2021 IEEE International Symposium on Circuits and Systems (ISCAS) | 978-1-7281-9201-7/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISCAS51556.2021.9401522

# MorphoNet: a Deep Image Super Resolution Network using Hierarchical and Morphological Feature Generating Residual Blocks

Alireza Esmaeilzehi, *Student Member, IEEE,* M. Omair Ahmad, *Fellow, IEEE,* and M.N.S. Swamy, *Fellow, IEEE*
Department of Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada
Email: {a_smaei,omair,swamy}@ece.concordia.ca

***Abstract*—Morphological operations are nonlinear mathematical operations that are capable of performing signal processing tasks based on the structures and textures of the signals. With this motivation of the capability of morphological operations, in this paper, a novel residual block that can generate morphological features of images and fuse them with the conventional hierarchical features has been proposed. The proposed residual block is then used to design a light-weight deep neural network architecture in a residual framework for the task of image super resolution. It is shown that a fusion of morphological features of images with the conventional hierarchical features can improve the super resolution capability of a deep convolutional network. Experiments are performed to demonstrate the effectiveness of the proposed idea of using morphological operations and the superiority of the network designed based on this idea in super resolving low quality images.**

***Index Terms*—Image Super Resolution, Deep Learning, Morphological Image Processing, Residual Learning.**

## I. Introduction

Since the process of image acquisition by CCD cameras inherently involves subsampling of spatially continuous visual information, it yields degraded digital images, and therefore, cannot be used in many real-world applications, such as mobile devices, medical image analysis systems and robotic systems, in which the availability of high quality images is a prerequisite. Therefore, image super resolution of the low quality image is a crucial first step in these real-world applications. Emergence of deep neural networks has revolutionized the task of image super resolution[1]-[21]. The networks EDSR [1], RDN [2], RCAN [3], SAN [4] and DRN [5] employ a cascade of large number of residual blocks to map a degraded low resolution image to the high resolution one. Even though these networks provide very good performance, the use of large number of parameters employed by these heavy-weight networks and the large number of operations required by them precludes their deployment in many applications requiring the low power consumption and convenience of portability. In view of this limitation of the heavy-weight deep super resolution convolutional neural networks, design of a light-weight deep network that employs small numbers of parameters and operations is an important issue in the community. Keeping this crucial requirement of such applications, the networks CARN [6], IMDN [8] and OISR [7] have designed light-weight super resolution architectures as a cascade of small number of residual feature generating blocks in a non-recursive framework, but at the expense of slightly reduced performance. The richness of the features generated by a convolutional neural network has a direct bearing on its representational capability, and therefore, on its super resolution performance. Hence, when designing a super resolution network, it is important to focus on the richness of the features generated by it without making it deeper by increasing the number of convolutional layers either physically or virtually with more recursions. Most of the image super resolution networks existing in the literature employ residual blocks that generate features only at different hierarchical levels, which may not be adequately rich.

The quality of an image is very much dependent on the texture and image representation of the image. Hence, the success of an image super resolution process can be judged from its capability in enhancing the textures and structures in the image super resolved by it. Morphological operations [28]-[30] are the nonlinear mathematical operations that while performing signal processing aim at the textures and structures of the signal.

In this paper, we present, for the first time, a deep image super resolution architecture in a residual framework by proposing a novel residual block that is capable of producing features of the image based on its morphology, as well as the conventional convolutional features. The morphological features are learned by using the erosion and dilation operations and fused with the other hierarchical features to produce very rich set of features.

The paper is organized as follows. In Section II, the architecture of the proposed residual block is developed and the architecture of a low complexity network using this residual block for the task of image super resolution is presented. In Section III, the performance and complexity of the proposed super resolution network are presented and compared with those of the state-of-the-art light-weight super resolution networks that exist in the literature. Finally, some concluding remarks on the work of this paper are made in Section IV.

This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada and in part by the Regroupement Stratégique en Microélectronique du Québec (ReSMiQ).

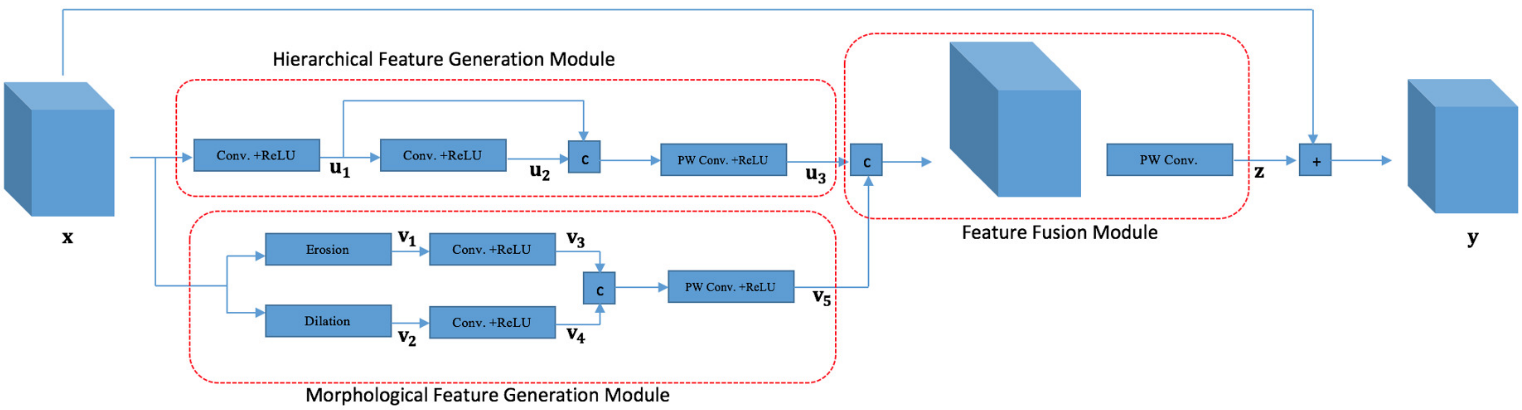


Fig. 1: Architecture of the proposed residual block. Conv. and PW Conv., respectively, denote the convolution and point-wise convolution operations.

## II. Proposed Deep Super Resolution Network

In this section, we first describe the architecture of a residual neural network for the task of image super resolution and then develop the architecture of the proposed residual block that is used in this network. We also explain the training details of the proposed network.

### A. Network Overall Architecture

The proposed network consists of four parts, namely, a feature extraction part that generates features of the low resolution input image, a nonlinear mapping part that maps the low level features to hierarchically higher level features through a cascade of the new residual blocks and upsampling and image reconstruction parts that provide the estimated high resolution image. In the feature extraction part, the features of the input image are extracted using a convolution operation employing 64 filters each with kernel size $3 \times 3$ followed by a ReLU activation operation. The nonlinear mapping part is composed of a cascade of 11 residual blocks. The architecture of the proposed residual block is developed in the next subsection. The upsampling part uses a depth-to-space transpose operation [22] with a scaling factor equal to that of the super resolution scaling factor. Finally, the high resolution image is constructed using 3 convolutional filters each with kernel size $3 \times 3$ by the image reconstruction part. We refer the proposed super resolution network to as MorphoNet.

### B. Proposed Residual Block

The proposed residual block is shown in Fig. 1. This block consists of three modules, a hierarchical feature generation module, a morphological feature generation module and a feature fusion module. In the hierarchical feature generation module, the feature tensor $\mathbf{x}$ input to the residual block is made to undergo a cascade of two convolution operations each followed by a ReLU activation operation yielding, respectively, two feature tensors $\mathbf{u}_1$ and $\mathbf{u}_2$ given by

$$\begin{aligned}\mathbf{u}_1 &= ReLU(W_1(\mathbf{x}))\\ \mathbf{u}_2 &= ReLU(W_2(\mathbf{u}_1))\end{aligned} \tag{1}$$

where each of the convolution operations $W_1$ and $W_2$ employs 64 filters with kernel size $3 \times 3$. The two feature tensors are then concatenatively fused as

$$\mathbf{u}_3 = ReLU(W_3(CONC(\mathbf{u}_1, \mathbf{u}_2))) \tag{2}$$

where $W3$ represents a point-wise convolution operations using 64 filters. In the hierarchical feature generation module, the features are learned solely though the convolution operation. In contrast, in the morphological feature generation module, features are also naturally learned but guided by morphological operations. In this module, the input feature tensor $\mathbf{x}$ first undergoes in parallel through the streams of the morphological erosion and dilation operations and then the resulting tensors $\mathbf{v}_1$ and $\mathbf{v}_2$ are convolved to produce morphologically guided features $\mathbf{v}_3$ and $\mathbf{v}_4$, respectively. Let $\mathbf{x}^k$ represent the $kth$ channel of the feature tensor $\mathbf{x}$. Then, the $kth$ channel of the feature tensor $\mathbf{v}_1$ resulting from the erosion operation is given by

$$\mathbf{v}_1^k[m,n] = (\mathbf{x}^k \ominus \mathbf{b})[m,n] = \min_{(i,j)\in B} \mathbf{x}^k[m+i, n+j] \tag{3}$$

where $\mathbf{b}$ is the structuring element defined over a neighborhood $B$. Similarly, the $kth$ channel of the feature tensor $\mathbf{v}_2$ resulting from, using the same structuring element $\mathbf{b}$ and defined over the same neighborhood $B$ as for the erosion operation, is given by

$$\mathbf{v}_2^k[m,n] = (\mathbf{x}^k \oplus \mathbf{b})[m,n] = \max_{(i,j)\in B} \mathbf{x}^k[m+i, n+j] \tag{4}$$

The feature tenors $\mathbf{v}_3$ and $\mathbf{v}_4$ are then obtained, respectively, by applying convolution operations to the feature tensors $\mathbf{v}_1$ and $\mathbf{v}_2$ as

$$\begin{aligned}\mathbf{v}_3 &= ReLU(W_4(\mathbf{v}_1))\\ \mathbf{v}_4 &= ReLU(W_5(\mathbf{v}_2))\end{aligned} \tag{5}$$

where each of the convolution operations $W_4$ and $W_5$ uses 64 filters with kernel size $3 \times 3$. The two morphological feature

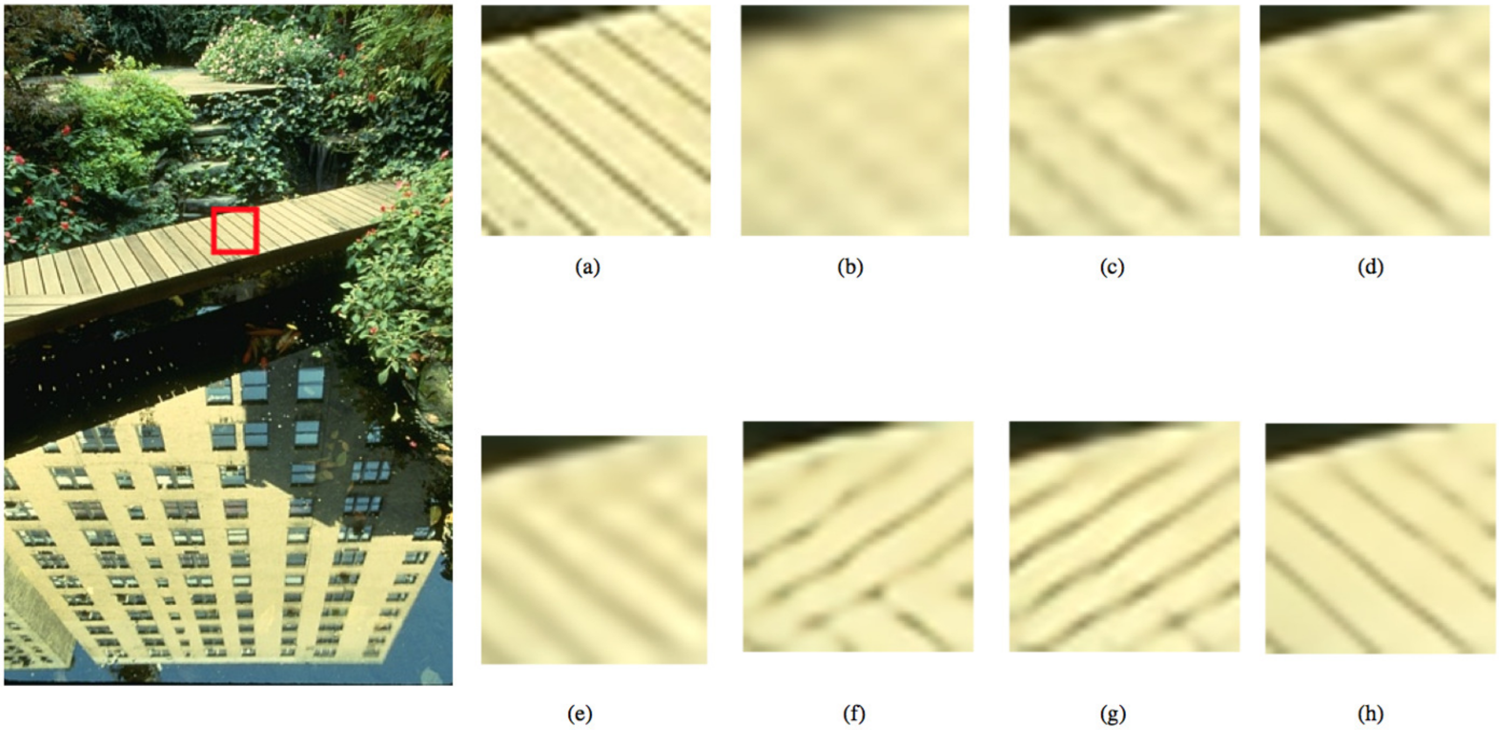


Fig. 2: Visual quality of images *img021* super resolved by various schemes. (a) Ground truth. (b) Bicubic. (c) SRCNN. (d) VDSR. (e) DRCN. (f) CARN. (g) IMDN. (h) Morphonet.

TABLE I: Results on the ablation study of the proposed residual block.

| Network with | *Set5* | *Set14* | *BSD100* | *Urban100* |
|---|---|---|---|---|
| *Variant* | 34.38 (0.9277) | 30.44 (0.8446) | 29.15 (0.8080) | 27.98 (0.8502) |
| *Proposed* | 34.52 (0.9284) | 30.53 (0.8455) | 29.20 (0.8085) | 28.05 (0.8510) |

tensors $\mathbf{v}_3$ and $\mathbf{v}_4$ are concatenatively fused to yield the feature tensor:

$$\mathbf{v}_5 = ReLU(W_6(CONC(\mathbf{v}_3, \mathbf{v}_4))) \tag{6}$$

where the point-wise convolution operation $W_6$ uses 64 filters. Next, the feature tensors $\mathbf{u}_3$ and $\mathbf{v}_5$ obtained, respectively, from the hierarchical and morphological feature generation modules, are fused to obtain the block's residual feature tensor given by

$$\mathbf{z} = W_7(CONC(\mathbf{u}_3, \mathbf{v}_5)) \tag{7}$$

where $W_7$ is a point-wise convolution operation using 64 filters. Finally, the feature tenor $\mathbf{x}$ input to the residual block is added to the residual feature tensor $\mathbf{z}$ to yield the output feature tensor $\mathbf{y}$ of the residual block.

### C. Training Details

For training the proposed super resolution network, sub-images of size $48 \times 48$ are extracted from the 800 training images of the *DIV2K* [23] dataset. The parameters of the proposed network are updated using the $\ell 1$ norm of the loss between the ground truth and estimated high resolution training samples. The $\ell 1$ norm loss is minimized using the stochastic gradient descent optimizer with the initial learning rate of 0.1. The batch size and weight decay parameter of the convolution operations are set to 64 and $10^{-4}$, respectively.

## III. Experimental Results

In this section, we first perform an ablation study on the proposed residual block in order to show the effectiveness of the morphological feature generation module in the residual block on the network performance. We then present and compare the performance of the proposed network on four benchmark datasets, namely, *Set5* [24], *Set14* [25], *BSD100* [26] and *Urban100* [27] and its complexity with that of the state-of-the-art light-weight image super resolution networks existing in the literature.

### A. Ablation Study

In order to investigate the effectiveness of the morphological feature generation module on the network performance, we form a variant of the proposed residual block by removing this module from the residual block. Table I gives the performance of the super resolution network employing the proposed residual block and its this variant on the four benchmark datasets with the scaling factor 3. Our objective in the design of the proposed residual block is to generate morphological features in addition to the conventional hierarchical features that are generated solely through the convolutional operations in order to provide a very rich set of features. It is seen by

TABLE II: PSNR (SSIM) values resulting from applying MorphoNet and various state-of-the-art methods to images of four benchmark datasets.

| Dataset | Scaling | Bicubic | SRCNN | VDSR | DRCN | LapSRN | MemNet | IDN | SRFBN | CARN | IMDN | OISR | MorphoNet (Proposed) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Set5 | ×2 | 33.66 (0.9299) | 36.66 (0.9542) | 37.53 (0.9587) | 37.63 (0.9588) | 37.52 (0.959) | 37.78 (0.9597) | 37.83 (0.9600) | 37.78 (0.9597) | 37.76 (0.9590) | 38.00 (0.9605) | 38.02 (0.9605) | 38.04 (0.9614) |
| | ×3 | 30.39 (0.8682) | 32.75 (0.9090) | 33.66 (0.9213) | 33.82 (0.9226) | N/A | 34.09 (0.9248) | 34.11 (0.9253) | 34.20 (0.9255) | 34.29 (0.9255) | 34.36 (0.9270) | 34.39 (0.9272) | 34.52 (0.9284) |
| | ×4 | 28.42 (0.8104) | 30.48 (0.8628) | 31.35 (0.8838) | 31.53 (0.8854) | 31.54 (0.885) | 31.74 (0.8893) | 31.82 (0.8903) | 31.98 (0.8923) | 32.13 (0.8937) | 32.21 (0.8948) | 32.14 (0.8947) | 32.23 (0.0.8951) |
| Set14 | ×2 | 30.24 (0.8688) | 32.42 (0.9063) | 33.03 (0.9124) | 33.04 (0.9118) | 33.08 (0.913) | 33.28 (0.9142) | 33.30 (0.9148) | 33.35 (0.9156) | 33.52(0.9166) | 33.63 (0.9177) | 33.62 (0.9178) | 33.77 (0.9196) |
| | ×3 | 27.21(0.7385) | 29.28 (0.8209) | 29.77 (0.8314) | 29.76 (0.8311) | N/A | 30.00 (0.8350) | 29.99 (0.8354) | 30.10 (0.8372) | 30.29 (0.8407) | 30.32 (0.8417) | 30.35 (0.8426) | 30.53 (0.8455) |
| | ×4 | 26.00 (0.7027) | 27.49 (0.7503) | 28.01 (0.7674) | 28.02 (0.7670) | 28.19 (0.772) | 28.26 (0.7723) | 28.25 (0.7730) | 28.45 (0.7779) | 28.60 (0.7806) | 28.58 (0.7811) | 28.63 (0.7819) | 28.77 (0.7855) |
| BSD100 | ×2 | 29.56 (0.8431) | 31.36 (0.8879) | 31.90 (0.8960) | 31.85 (0.8942) | 31.80 (0.895) | 32.08 (0.8978) | 32.08 (0.8985) | 32.00 (0.8970) | 32.09 (0.8978) | 32.19 (0.8996) | 32.20 (0.9000) | 32.32 (0.9025) |
| | ×3 | 27.21 (0.7385) | 28.41 (0.7863) | 28.82 (0.7976) | 28.80 (0.7963) | N/A | 28.95 (0.8004) | 28.96(0.8001) | 28.96 (0.8010) | 29.06 (0.8034) | 29.09 (0.8046) | 29.11 (0.8058) | 29.20 (0.8085) |
| | ×4 | 25.96 (0.6675) | 26.90 (0.7101) | 27.29 (0.7251) | 27.23 (0.7233) | 27.32 (0.728) | 27.40 (0.7281) | 27.41 (0.7297) | 27.44 (0.7313) | 27.58 (0.7349) | 27.56 (0.7353) | 27.60 (0.7369) | 27.65 (0.7391) |
| Urban100 | ×2 | 26.88 (0.8403) | 29.50 (0.8946) | 30.76 (0.9140) | 30.75 (0.9133) | 30.41 (0.910) | 31.31 (0.9195) | 31.27 (0.9196) | 31.41 (0.9207) | 31.92 (0.9256) | 32.17 (0.9283) | 32.21 (0.9290) | 32.06 (0.9288) |
| | ×3 | 24.46 (0.7349) | 26.24 (0.7989) | 27.14 (0.8279) | 27.15 (0.8276) | N/A | 27.56 (0.8376) | 27.42 (0.8359) | 27.66 (0.8415) | 28.06 (0.8493) | 28.17 (0.8519) | 28.24 (0.8544) | 28.05 (0.8510) |
| | ×4 | 23.14 (0.6577) | 24.52 (0.7221) | 25.18 (0.7524) | 25.14 (0.7510) | 25.21 (0.756) | 25.50 (0.7630) | 25.41 (0.7632) | 25.71 (0.7719) | 26.07 (0.7837) | 26.04 (0.7838) | 26.17 (0.7888) | 26.01 (0.7830) |

The values in the red font indicate the best performance and those in the blue font represent the second best performance.

TABLE III: Complexity of various super resolution schemes.

| Method | Number of Parameters |
|---|---|
| SRCNN [9] | 57K |
| VDSR [12] | 665K |
| DRCN [17] | 1770K |
| MemNet [15] | 677K |
| IDN [14] | 553K |
| SRFBN-S [20] | 483K |
| CARN [6] | 1592K |
| IMDN [8] | 715K |
| OISR [7] | 1550K |
| MorphoNet (Proposed) | 1414K |

comparing the results of this table corresponding to using the proposed residual block and its variant, that by removing the morphological feature generation module, the performance of the network gets reduced significantly.

### *B. Comparison with the State-of-the-Art Schemes*

Table II gives the performance in terms of the PSNR and SSIM metrics of the proposed super resolution network and those of ten other super resolution neural networks, namely, super resolution convolutional neural networks (SRCNN) [9], very deep super resolution network (VDSR) [12], deep recursive convolutional network (DRCN) [17], Laplacian pyramid super resolution network (LapSRN), memory persistent network (MemNet) [15], information distillation network (IDN) [14], super resolution feedback network (SRFBN) [20], cascaded residual network (CARN) [6], information multi-distillation network (IMDN) [8] and ODE-inspired super resolution network (OISR) [7]. It is seen from this table that the proposed network generally outperforms all the networks used in our comparison. Specifically, it is seen that the proposed network outperforms OISR, which is the best performing state-of-the-art super resolution network in the light-weight category employing the number of parameters in the neighborhood of $1.5M$ parameters or less, in 18 out of 24 cases of the PSNR and SSIM metric values.

Table III gives the complexity in terms of the number of parameters of the proposed and state-of-the-art light-weight super resolution networks. It is seen from this table that the proposed network employs $1.4M$ parameters, which is lower than that $1.55M$ parameters employed by the OISR network. Thus, considering the complexity and performance together, the proposed network can be regarded to be the best network among all the light-weight super resolution networks.

Fig. 2 shows the zoomed segments of the image *img021* from the *BSD100* dataset super resolved by various light-weight super resolution networks. It is seen from this figure that the image super resolved by the proposed network has the best visual quality. Specifically, the ridge textures of the pathway in the wooden bridge are recovered more accurately by the proposed network in comparison to that recovered by all the other networks. In particular, the ridge orientation in the recovered images by CARN [6] and IMDN [8] are completely altered from that of the ground truth image.

## IV. Conclusion

In this paper, a novel residual feature generating block has been proposed and used in a deep neural network for the task of single image super resolution. The proposed residual block generates morphological features using the morphological erosion and dilation operations in a learnable framework, and fuses them with those generated by the conventional convolutional operations in order to extract a very rich set of features. The proposed block is then used in constructing a light-weight residual deep convolutional neural network for super resolving poor quality low resolution images. The effectiveness of the proposed morphological residual block has been shown by performing experiments on the proposed network using challenging benchmark datasets and comparing the performance results with those of the existing light-weight super resolution networks. The proposed network has been shown to generally outperform all the light-weight networks used in our comparison. It has been conclusively demonstrated that an amalgamation of morphological features with the conventional features obtained solely through the convolution operations can improve the performance of super resolution networks.

# References

[1] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. "Enhanced deep residual networks for single image super-resolution", In *CVPR*, 2017.

[2] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu,"Residual dense network for image super-resolution", in *CVPR*, 2018.

[3] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong and Y. Fu,"Image Super-Resolution Using Very Deep Residual Channel Attention Networks", In *ECCV*, 2018.

[4] T. Dai, J. Cai, Y. Zhang, S-T. Xia and L. Zhang, "Second-order Attention Network for Single Image Super-Resolution", In *CVPR*, 2019.

[5] Y. Guo, J. Chen, J. Wang, Q. Chen, J. Cao, Z. Deng, Y. Xuy and M. Tany, "Closed-loop Matters: Dual Regression Networks for Single Image Super-Resolution", In *CVPR*, 2020.

[6] N. Ahn, B. Kang and K.A. Sohn,"Fast, accurate, and lightweight super-resolution with cascading residual network", in *ECCV*, 2018.

[7] X. He, Z. Mo, P. Wang, Y. Liu, M. Yang and J. Cheng, "ODE-inspired Network Design for Single Image Super-Resolution", In *CVPR 2020*.

[8] Z. Hui, X. Gao, Y. Yang and X. Wang, " Lightweight Image Super-Resolution with Information Multi-distillation Network. In *ACMMM 2019*.

[9] 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.

[10] D. Liu, Z. Wang, B. Wen, J. Yang, W. Han and T. Huang, "Robust single image super-resolution via deep networks with sparse prior", IEEE Transactions on Image Processing, vol. 25, no.7, pp. 3194-3207, 2016.

[11] A. Esmaeilzehi, M.O. Ahmad and M.N.S. Swamy, "SRNMFRB: A Deep Light-weight Super Resolution Network using Multi-receptive Field Feature Generation Residual Blocks", In *ICME 2020*.

[12] J. Kim, J.K. Lee and K.M. Lee, "Accurate image super-resolution using very deep convolutional network", In Proc. Computer Vision and Pattern Recognition, 2016.

[13] Y. Tai, J. Yang, X. Liu, "Image super-resolution via deep recursive residual network", In *CVPR*, 2017.

[14] Z. Hui, X. Wang and X. Gao, "Fast and accurate single image super-resolution via information distillation network", in *CVPR*, 2018.

[15] Y. Tai, J. Yang, X. Liu, and C. Xu, "MemNet: A persistent memory network for image restoration", In *ICCV*, 2017.

[16] A. Esmaeilzehi, M. O. Ahmad and M.N.S. Swamy, "Compnet: A new scheme for single image super resolution based on deep convolutional neural network", IEEE Access, vol. 6, pp. 59963 - 59974, Dec. 2018.

[17] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution", In *CVPR*, 2016.

[18] W-S. Lai, J-B. Huang, N. Ahuja, and M-H. Yang, "Deep laplacian pyramid networks for fast and accurate super-resolution", in *CVPR*, 2017.

[19] A. Esmaeilzehi, M. O. Ahmad and M.N.S. Swamy, "SRSubBandNet: A New Deep Learning Scheme for Single Image Super Resolution Based on Subband Reconstruction", in *ISCAS*, 2019.

[20] Z. Li, J. Yang, and Z. Liu, and X. Yang and G. Jeon and W. Wu, "Feedback Network for Image Super-Resolution", In *CVPR*, 2019.

[21] A. Esmaeilzehi, M.O. Ahmad and M.N.S. Swamy, "MGHCNET: A Deep Multi-Scale Granular and Holistic Channel Feature Generation Network for Image Super Resolution", In *ICME 2020*.

[22] W. Shi, J. Caballero, F. Huszaa¡r, 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*CVPR*, 2016.

[23] E. Agustsson, R. Timofte, "Ntire 2017 Challenge on Single Image Super-Resolution: Dataset and Study", in *CVPR*, 2017.

[24] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. Alberi-Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding", in *BMVC* , 2012.

[25] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*, 2012.

[26] 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 *ICCV*, 2001.

[27] J.B. Huang, A. Singh and N. Ahuja, "Single image super-resolution from transformed self-exemplars", In Proc. Computer Vision and Pattern Recognition, 2015.

[28] G. Franchi, A. Fehri and A. Yao, "Deep morphological networks", Pattern Recognition, 2020.

[29] D. Mellouli, T.M. Hamdani, M.B. Ayed and A.M. Alimi, "Morph-CNN: a morphological convolutional neural network for image classification", in *International Conference on Neural Information Processing*, 2017.

[30] K. Nogueira, J. Chanussot, M.D. Mura, W.R. Schwartz and J.A. dos Santos,"An Introduction to Deep Morphological Networks", in *arXiv preprint*, arXiv:1906.01751 (2019).