Conference Paper V1

by Wazir Muhammad

MCSNet: Multi-scale Channel Sharing based network for Single Image Super-resolution

[1]Wazir Muhammad, [2]Supavadee Aramvith

[1]Electrical Engineering Department, Faculty of Engineering Chulalongkorn University, Bangkok, 10330, Thailand and Electrical Engineering Department BUET, Khuzdar, Pakistan
[2] Multimedia Data Analytics and Processing Unit, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand.
Email: [1]wazir.laghari@gmail.com, [2]supavadee.a@chula.ac.th
Contact: [1]+923332634843., [2]+6622186911.

Abstract:. The performance and processing speed of single-image super-resolution have been increased every day due to the rapid advancement in convolutional neural network (CNN) architectures. However, CNN based architectures have a high computational cost, more processing time and reconstructed results are still not satisfactory. Furthermore, CNN architectures depend on the single channel type network architecture to reconstruct the high-resolution (HR) image, which is also create problem during the training and extracted features are not received at the later end layer. To report these problems, we propose a Multi-scale Channel Sharing based network for Single Image Super-resolution known as MCSNet. In addition, 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. Extensive quantitative and qualitative experimental results show that our proposed approach achieves a better trade-off against other state-of-the-art methods in terms of peak signal-to-noise ratio (PSNR), computational complexity, and processing speed. Especially on challenging enlargement scale factor 8×, our proposed method improves overall average PSNR on all test datasets including as Set5, Set14, Urban100, and Manga109 by 2.04dB with the base line method.

Index terms: Single image super-resolution, Dense channel sharing block, LeakyReLU, Convolutional Neural Network.

2. INTRODUCTION

The main target of single image super resolution (SISR) is to reconstruct the high quality / high-resolution (HR) output image from its degraded low quality / low-resolution (LR) input image [1] . The SISR is still challenging task to solve it due to its ill-posed problem. To report such problem, various image super-resolution (SR) methods have been developed, including bicubic interpolations, reconstruction, and learning-based approaches. Recently, deep CNN learning based image SR methods have been achieved tremendous performance due to the rapid improvement in the convolutional neural network (CNN) design. Recently, available approaches deep learning based CNN methods. Among them, Dong et al. [2] initially proposed the shallow based architecture using three CNN layers to develop the direct relationships between an original LR input image and high-resolution reconstructed output image known as super-resolution convolutional neural network (SRCNN). These three layers performs three basic operations, such as feature extraction, non-linear mapping, and restoration, respectively. The main flaw of this approach are to use the bicubic interpolation as a pre-processing step to upscale the LR image as an input and reconstruct the interpolated version of HR output image. The author of SRCNN [2] modified the architecture and replaced the pre-upsampling bicubic interpolation technique by post-upsampling learnable layer (transpose convolution) named as Fast SR Convolutional Neural Network (FSRCNN) [3]. This modification leads to decrease the number of parameters as well as memory consumption.

To improve the shortcomings observed during the SRCNN [2] and FSRCNN [3] models, Kim et al. used the concept of Visual Geometry Group (VGG) model which is used in the ImageNet classification task and proposed the 20-layer network known as Very Deep Super Resolution (VDSR) [4]. The performance of VDSR [4] is better than earlier approaches such as SRCNN [2] and FSRCNN [3] with the large margin gap. Furthermore, VDSR [4] employed the residual learning technique to avoid the vanishing gradient problem as well as increase the speed of convergence. The relationship between SRCNN [2] and the sparse-SCN methods confirms its acceptable performance. Wang et al. [5] replaced layers of mapping with the set of sparse based coding type networks and commonly known as SCN (sparse coding network). Deeper is the better follow this concept Kim et al., [6] proposed image SR method using a deeply recursive CN (DRCN). In this architecture used a very deep recursive type layers (16 times of recursions). DRCNN [6] method same convolutional layers applies repeatedly as several times as desired, but the computational cost does not increased although more times recursions are applied.

Likewise, Tai et al. suggested the concept of two-way novel network, one for a deep recursive residual network (DRRN) [7] [24], and next is used for a persistent memory network (MemNet) [8]. The earliest designed is employed for 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 the testing as well as training process. A 27-CNN layers with robust Charbonnier loss function network architecture proposed by Lai et al. presented the Laplacian pyramid SR network (LapSRN) [9] to gradually reconstruct the HR images.

Recently, deep CNN based single image SR is to attained the improved performance as compared to conventional methods, but still faces some challenging issues [10]. Generally, deep CNN model used a convolutional layer's side-by-side to design a deeper framework, which introduce the vanishing gradient problem and later end layers work as a dead layer. Similarly, some well-known deep CNN methods [2, 4] used the interpolation technique as a pre-processing stage to upscale the LR image and then extract the CNN features from the upscaled version of HR image. These approaches are not successful, because it introduces the new noise in the models and used more burden on the model during the training. Additionally, deep CNN rely on two-way multi-scale channel sharing type network architecture, which is not extracts the complete feature information through the previous layers.

To address these problems, we have suggested the Multi-scale Channel Sharing based network for Single Image Super-resolution architecture, named as MCSNet, which use the Dense channel sharing ResNet blocks (DCSRB) with local as well global skip connection. In summary, we establish a novel Multi-scale Channel Sharing based network for Single Image Super-resolution, which performs a better improvement in terms of PSNR/SSIM, and perceptual quality point of view.

The overall main contribution of our MCSNet method can be summarize as follow:

• Inspired from Multi-scale Channel Sharing with ResNet architecture, we proposed a Multi-scale Channel Sharing based network for Single Image Super-resolution network for the image SR method with post-upsampling technique to reconstructing the high-quality HR output image.

• We proposed a new multi-channel path scheme to excellently boost the representations of features of the HR image. Multi-channel path scheme consists of dense channel sharing ResNet block (DCSRB) which facilitate the features extraction through each earlier and subsequent CNN layers.

• Conventional deep CNN methods used Rectified Linear Unit (ReLU) activation function, which is introduce the vanishing-gradient problems in a training, in case of deeper network architecture. In this regard, we changed the ReLU with leaky ReLU activation function to avoid the vanishing-gradient problems and perform the training of the model more efficiently. The remaining parts of paper is divided as follows: Proposed Method presented in Section II. The experimental calculations are discussed in a Section III. Finally, the conclusion is reported in the last section.

II. PROPOSED METHOD

In this paper, we propose the concept of multi-scale channel sharing type network architecture with the support of dense channel sharing Resnet block (DCSRB). First, we discuss the proposed network architecture of MCSNet in detail, and then we discuss the DCSRB block with different CNN layers used in the main architecture as shown in the Figure 1. The proposed network framework of MCSNet used two CNN layers followed by Leaky Rectified Linear Unit (LeakyReLU) activation function. These CNN layers extracted the initial features and then fed the four DCSRB blocks side by side as well as timely update features information received through local as well as global skip connections. Finally, the cumulative LR features pass through the deconvolution layer to upscale and reconstructed the HR output image. The design of LReLU depends on the ReLU and it provides two outputs one for (0, x) and other for (x,0). Additionally, earlier approach depends on maximum one or two channel sharing residual block to extract the image features of different scales [11]. The reconstructed features information satisfactory, but in case of deeper and denser network architecture information is lost and cannot extract the detailed features as well as sometimes information is stuck. In our proposed DCSRB block design in such a way to split the original features information with five different kernel size multiscale channels in the order of 1×1, 3×3, 5×5, 7×7, and 9×9. First branch used 1×1, convolutional neural network layers to reduce the dimension of the input data as well reduce the computational cost of the model [12] and other branches of different kernel size are used to extract the multi-channel information as shown in Figure 2. Finally, in the reconstruction phase, upsampling operation is used by 3×3 deconvolution to generate the high-quality or high-resolution (HR) output image. The 1×1 convolution operation is used before the deconvolution layer to reduce the size of dimension.

III. EXPERIMENTS

In the experimental section, we explain the training and testing datasets strategy for reconstructing the HR image. procedures. In our MCSNet used the 3 different training datasets, such as Yang et al., [13], BSDS100 [14] and DIV2K [15] (including 50 images) for training purposes. The complete training dataset is split into two sections, like 80% for the purpose of training and 20% for validation. Our proposed method is tested on four standard publicly available datasets, including Set5 [16], Set14 [13], Urban100 [17], and Manga109 [18]. Set5 [16], Set14 [13] used the images in the order of 5 and 14. In Urban100 and Manga109 [18] available images 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. Our model performs training as well as testing under the environment of Windows 10 operating system with NVIDIA GeForce RTX20270 GPU model having 16 GB RAM. Training performs on enlargement factor 8x in Keras 2.15 and MATLAB2018 a framework. The quantitative experimental results present in the Table 1, which shows that proposed method attains higher average quantitative values.

CONCLUSION

In this paper, we proposed a Multi-scale Channel Sharing based network for Single Image Super-resolution (MCSNet) based on the convolution, LeakyReLU and dense channel sharing Resnet block (DCSRB) for reconstruct the visually pleasing HR output image. To increase the quantitative as well qualitative evaluation of LR image, we employed the four DCSRB bocks followed by LeakyReLU activation function. Furthermore, increase the training convergence process and fixed the dying ReLU problem in the dense CNN model, we replaced the ReLU activation function with LeakyReLU non-linear activation function. Moreover, whole network architecture adopts a local as well global skip connections to capture the more high-frequency feature information for single image super-resolution reconstruction. Quantitative as well as qualitative performance evaluated on four benchmark test dataset demonstrates that proposed MCSNet can superior performance in terms of reconstructed results.

Additionally, we evaluate the performance of our proposed method in terms of perceptual quality with different publicly available methods as shown in Figure 3. The proposed and other methods in Figure 3 used 3 images of butterfly, zebra, and PrayerHaNemurenai are obtained from Set5, Set14, and Manga109 with scale factor 8×. In the case of the bicubic interpolation technique texture generated are blurry output. Other methods such as SRCNN, VDSR, and DRCN produce improved results then baseline method. In our proposed method reconstructed output is a visually pleasing texture details with high value of PSNR/SSIM.

ACKNOWLEDGEMENTS

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), and Thailand Science research and Innovation Fund Chulalongkorn University CU_FRB65_soc-(1)_001_27_01”.

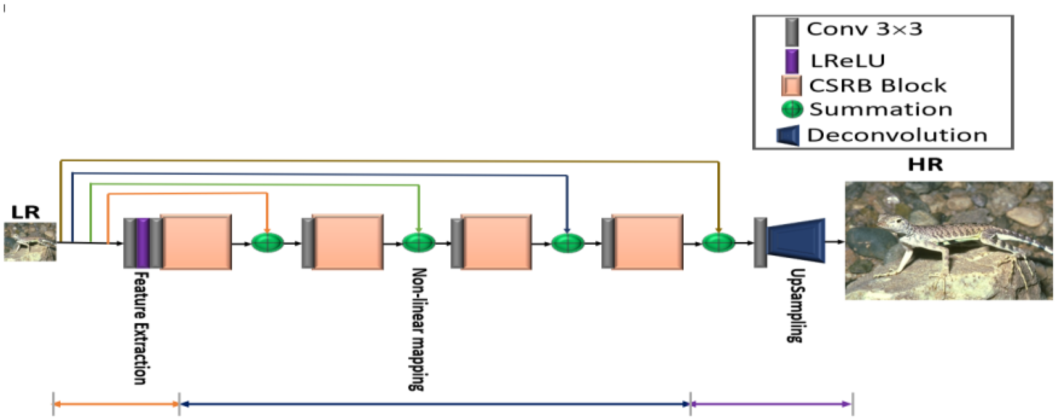

Figure 1: Proposed Multi-scale Channel Sharing based network for Single Image Super-resolution (MCSNet)

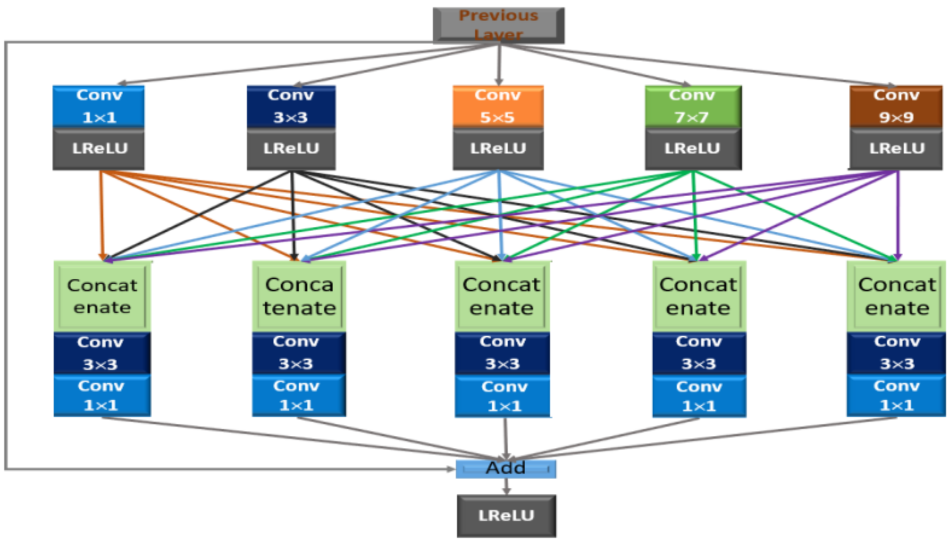

Figure 2: Dense channel sharing ResNet Block (DCSRB)

Table 1

The quantitative comparison of average PSNR/SSIM with existing image super-resolution algorithms on four test datasets (Set5, Set14, Urban100, and Manga109) on challenging enlargement factor ×8.

Datasets	Scale	Bicubic PSNR SSIM	SRCNN [2] PSNR SSIM	FSRCNN [3] PSNR SSIM	SCN [5] PSNR SSIM	VDSR [4] PSNR SSIM	DRCN [6] PSNR SSIM	MCSNet PSNR SSIM
SET5 [16]	×8	24.40 0.6580	25.33 0.6900	20.13 0.5520	25.59 0.7071	25.93 0.7240	25.93 0.7230	25.95 0.7231
SET14 [13]	×8	23.10 0.5660	23.76 0.5910	19.75 0.4820	24.02 0.6028	24.26 0.6140	24.25 0.6141	25.12 0.6152
URBAN100 [17]	×8	20.74 0.5160	21.29 0.5440	21.32 0.5380	21.52 0.5571	21.70 0.5710	21.71 0.5710	23.38 0.5821
MANGA109 [18]	×8	21.47 0.6500	22.46 0.6950	22.39 0.6730	22.68 0.6963	23.16 0.7250	23.20 0.7240	23.43 0.7331
Average	×8	22.43 0.5975	23.21 0.6300	20.90 0.5613	23.45 0.6408	23.76 0.6585	23.77 0.6580	**24.47** **0.6634**

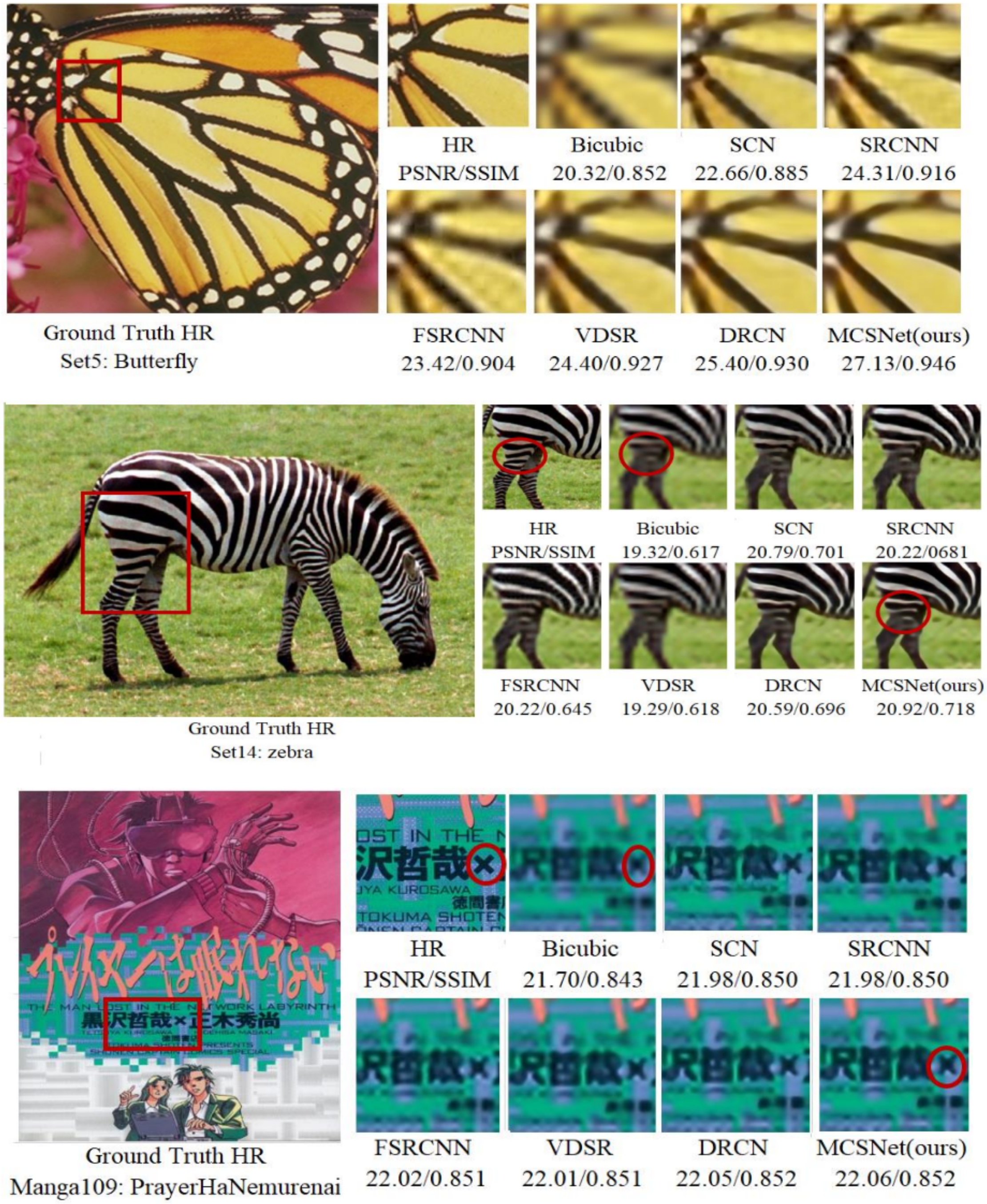

Figure 3: Presents the perceptual quality comparison on enlargement factor 8× of different image super-resolution methods using test datasets of Set5, Set14 and Manga109.

REFERENCES

1. Freeman, W.T., E.C. Pasztor, and O.T. Carmichael, *Learning low-level vision.* International journal of computer vision, 2000. **40**(1): p. 25-47.
2. Dong, C., et al., *Image super-resolution using deep convolutional networks.* IEEE transactions on pattern analysis and machine intelligence, 2015. **38**(2): p. 295-307.
3. Dong, C., C.C. Loy, and X. Tang. *Accelerating the super-resolution convolutional neural network*. in *European conference on computer vision*. 2016. Springer.
4. Kim, J., J.K. Lee, and K.M. Lee. *Accurate image super-resolution using very deep convolutional networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
5. Wang, Z., et al. *Deep networks for image super-resolution with sparse prior*. in *Proceedings of the IEEE international conference on computer vision*. 2015.
6. Kim, J., J.K. Lee, and K.M. Lee. *Deeply-recursive convolutional network for image super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
7. Tai, Y., J. Yang, and X. Liu. *Image super-resolution via deep recursive residual network*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
8. Tai, Y., et al. *Memnet: A persistent memory network for image restoration*. in *Proceedings of the IEEE international conference on computer vision*. 2017.
9. Lai, W.-S., et al. *Deep laplacian pyramid networks for fast and accurate super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
10. Cai, Y., et al., *Image Reconstruction of Multibranch Feature Multiplexing Fusion Network with Mixed Multilayer Attention.* Remote Sensing, 2022. **14**(9): p. 2029.
11. Li, J., et al. *Multi-scale Residual Network for Image Super-Resolution*. in *Computer Vision – ECCV 2018*. 2018. Cham: Springer International Publishing.
12. Lan, R., et al., *MADNet: a fast and lightweight network for single-image super resolution.* IEEE transactions on cybernetics, 2020. **51**(3): p. 1443-1453.
13. Yang, J., et al., *Image super-resolution via sparse representation.* IEEE transactions on image processing, 2010. **19**(11): p. 2861-2873.
14. Arbelaez, P., et al., *Contour detection and hierarchical image segmentation.* IEEE transactions on pattern analysis and machine intelligence, 2010. **33**(5): p. 898-916.
15. Agustsson, E. and R. Timofte. *Ntire 2017 challenge on single image super-resolution: Dataset and study*. in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
16. Bevilacqua, M., et al., *Low-complexity single-image super-resolution based on nonnegative neighbor embedding.* 2012.
17. Huang, J.-B., A. Singh, and N. Ahuja. *Single image super-resolution from transformed self-exemplars*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
18. Matsui, Y., et al., *Sketch-based manga retrieval using manga109 dataset.* Multimedia Tools and Applications, 2017. **76**(20): p. 21811-21838.

Conference Paper V1

ORIGINALITY REPORT

21%	18%	20%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	res.mdpi.com Internet Source	5%
2	www.ncbi.nlm.nih.gov Internet Source	2%
3	www.journaltocs.ac.uk Internet Source	2%
4	Xinxia Fan, Yanhua Yang, Cheng Deng, Jie Xu, Xinbo Gao. "Compressed multi-scale feature fusion network for single image super-resolution", Signal Processing, 2018 Publication	1%
5	www.nature.com Internet Source	1%
6	paper.ijcsns.org Internet Source	1%
7	Lecture Notes in Computer Science, 2014. Publication	1%
8	www.hindawi.com Internet Source	1%

9	Yongliang Tang, Weiguo Gong, Qiane Yi, Weihong Li. "Combining sparse coding with structured output regression machine for single image super-resolution", Information Sciences, 2018 Publication	1 %
10	iis.bjtu.edu.cn Internet Source	1 %
11	journals.plos.org Internet Source	1 %
12	Katipot Inkong, Viphada Yodpetch, Santi Kulprathipanja, Pramoch Rangsunvigit, Praveen Linga. "Influences of different co-promoters on the mixed methane hydrate formation with salt water at moderate conditions", Fuel, 2022 Publication	1 %
13	Wazir Muhammad, Supavadee Aramvith, Takao Onoye. "Multi-scale Xception based depthwise separable convolution for single image super-resolution", PLOS ONE, 2021 Publication	1 %
14	export.arxiv.org Internet Source	1 %
15	Wazir Muhammad, Zuhaibuddin Bhutto, Arslan Ansari, Mudasar Latif Memon et al. "Multi-Path Deep CNN with Residual Inception	1 %

Network for Single Image Super-Resolution", Electronics, 2021
Publication

16	"Intelligent Data Engineering and Automated Learning – IDEAL 2017", Springer Science and Business Media LLC, 2017 Publication	<1 %
17	Sung Ye Kim, Preeti Bindu. "Realizing Real-Time Deep Learning-Based Super-Resolution Applications on Integrated GPUs", 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), 2016 Publication	<1 %
18	koreascience.kr Internet Source	<1 %
19	"Computer Vision – ECCV 2016", Springer Science and Business Media LLC, 2016 Publication	<1 %
20	Communications in Computer and Information Science, 2015. Publication	<1 %
21	arxiv.org Internet Source	<1 %
22	link.springer.com Internet Source	<1 %
23	www.ns2.thinkmind.org Internet Source	<1 %

24	"Computer Vision – ECCV 2018 Workshops", Springer Science and Business Media LLC, 2019 Publication	<1 %
25	Ding Qin, Xiaodong Gu. "Single-image super-resolution with multilevel residual attention network", Neural Computing and Applications, 2020 Publication	<1 %
26	Jianmin Jiang, Hossam M. Kasem, Kwok-Wai Hung. "A Very Deep Spatial Transformer Towards Robust Single Image Super-Resolution", IEEE Access, 2019 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		