

TISR: Twin Image Super-Resolution using Deep Convolutional Neural Networks

Wazir Muhammad[†], Zuhaibuddin Bhutto^{††}, Jalal Shah^{††}, Murtaza Hussain Shaikh^{†††}, Syed Ali Raza Shah^{††††},
Shah Muhammad Butt^{†††††}, Salman Masroor^{†††††} Ayaz Hussain[†],

[†]Department of Electrical Engineering, Balochistan University of Engineering & Technology, Pakistan.

^{††}Department of Computer Systems Engineering, Balochistan University of Engineering & Technology, Pakistan.

^{†††}Department of Information Systems, Kyungshung University, Busan, South Korea.

^{††††}Department of Mechanical Engineering, Balochistan University of Engineering & Technology, Pakistan.

^{†††††}Sindh Madressa-tul-Islam University, City Campus, Karachi, Pakistan.

Abstract

Image super-resolution aims to reconstruct the visually pleasing high-resolution (HR) image from the degraded version or low-resolution (LR) ones. Under the remarkable improvement in the field of image and computer vision tasks. Despite its improvement in accuracy and performance, these models used convolutional neural network (CNN) layers side by side to increase the depth of the network, which is not a suitable way of design and it creates the vanishing gradient problems during the training. Furthermore, previous deep CNN methods rely on a single channel to reconstruct the HR output image, but later end layers cannot receive the proper information and work as a dead layer. In this paper, we are used two parallel deep convolutional neural networks with the same size and order of filters, known as Twin Image Super-Resolution using Deep Convolutional Neural Networks (TISR). Additionally, proposed method used two parallel branches for extracting the low, mid, and high-level features simultaneously. For multi-level feature extraction purposes latest Xception block is employed from GoogLeNet architecture. Our method is evaluated on three different benchmark test datasets including SET5, SET14, and BSDS100. Experimental results are demonstrated that our proposed method (TISR) outperforms then the existing state-of-the-art methods.

Keywords: Super-resolution, Convolutional neural networks, Depth wise Separable convolution, Xception block, Deconvolution.

1. Introduction

Single image super-resolution (SISR) is a technique for reconstructing a high-resolution image (HR) with superior visual quality from a single low-resolution (LR) input image [1-3]. Access to the original data, that is unavailable, and the quality of the received image cannot be estimated properly. If the image quality is poor and restoring the information becomes very difficult. As a result, it is critical to restore the original image in super-resolution. The application of image super-resolution (SR) is being used in a variety of applications right now, including medical imaging [4], atmospheric monitoring [5], closed-circuit television surveillance [6], security systems [7], satellite remote sensing [8], and robotics [9].

Before developing the deep learning-based approaches, various

image processing-based techniques were employed in the image super-resolution tasks. Li et al. [10] and Nasrollahi et al. [11] classified image super-resolution methods into three different groups, namely, interpolation-based method, reconstruction-based method, and learning-based method. Several interpolation-based methods such as linear [12], bilinear [13], or bicubic [13, 14] interpolations can be found in image super-resolution applications. These methods are straightforward, but they may not restore the image's high-frequency information, therefore a more complex understanding of the image may be necessary to recover it [10, 11]. Although, many single image super-resolution based deep learning methods [15-18] obtained the superior performance. All deep learning-based approaches used the back-propagation algorithms [19] to train the model. These strategies outperform prior statistics-based [20-23] and patch-based [24-29] models. Dong et al. suggested a deep learning-based method called super-resolution CNN (SRCNN) for image super-resolution [15]. SRCNN achieves superior performance as compared to earlier learning-based techniques because it uses more contextual information to restore lost image details. In general, the receptive field size of CNN is connected with the effective size of visual context for reconstruction [30]. SRCNN's receptive field size is determined by the convolutional kernel size in each layer as well as the CNN depth. Kim et al. constructed a very deep CNN [18], which stacks more convolutional layers to increase the receptive field size. Yamamoto et al. [31] were the first to apply deep CNN-based super-resolution algorithm in agriculture. Larger kernel sizes and deeper networks, on the other hand, add more parameters and demand more computational resources. Furthermore, once the kernel scale and depth are set, CNN only gives single scale contextual information for image reconstruction, ignoring the fact that real-world images are multi-scale in nature. Even though they produce superior results, but deeper networks require more parameters, which increases the computational cost and more memory usage. The training process for a deeper model is difficult to converge, and the testing process is time-consuming. To address these problems, we introduce a new concept about fast and accurate super-resolution network called Twin Image Super-Resolution using Deep Convolutional Neural Networks (TISR). We use a very deep twin network with Xception block network architecture and significantly improved the performance as well as reduce the computational cost in terms of model parameters.

In summary, the main contributions of this paper are threefold:

Manuscript received January 5, 2022

Manuscript revised January 20, 2022

<https://doi.org/10.22937/IJCSNS.2022.22.1.57>

- We introduce a lightweight network architecture for image SR and replace the single path network architecture with twin networks architecture.
- Our proposed Xception blocks for extracting the multi-scale features information efficiently. Each block depends on 3 types of kernel sizes of the order 3×3 , 5×5 and 7×7 .
- We are adopting a competitive strategy to replace the traditional bicubic upsampling strategy with deconvolution layer and pre-processing upsampling step is replaced with post upsampling step to further reduce the computational burden of the model.

The rest of this paper is laid out as follows. Section 2 goes overview of the related works. In Section 3, we discuss the network architecture in detail. The experimental results on image super-resolution are presented in Section 4. Section 5 contains the conclusions.

2. Related works

Over the last five years, there have been numerous research articles published in the field of deep convolutional neural network-based single image super-resolution. Initially, stepped up into the field of SR by Dong et al. [32] and proposed the concept of super-resolution convolutional neural network abbreviated as SRCNN. This approach extends the idea of sparse representation with convolutional neural network. Deep convolutional neural has basically three categories of network architectures i.e., shallow, deep, and dense network architecture. SRCNN is in under the category of shallow type network architecture and used three CNN layers followed by ReLU activation function except the last layer. SRCNN has three stages such as feature extraction, non-linear mapping, and image reconstruction. The performance of SRCNN is improved as compared to earlier approaches and design is very simple, but it has some issues. Firstly, the SRCNN used the bicubic interpolation as a pre-processing step to upscale the LR image into HR image. Secondly, it uses larger kernel size and features are not extracted directly from the original LR input image. To resolve these issues and accelerate the performance, same author introduced the concept of FSRCNN [33]. In this approach removing the bicubic interpolation pre-processing step with post-processing step by deconvolution layer. Furthermore, FSRCNN used shrinking, expanding, and deconvolution layers for extracting the HR image. After that, research community shift from shallow to deeper network architecture to further increase the performance of the CNN model. In this race, first time very deep 20 layers based convolutional neural network for image super-resolution is introduced by Kim et al. [18], known as Very Deep Super-Resolution (VDSR). The designed concept of VDSR is followed by a popular VGG-net architecture having a fixed CNN kernel size of the order of 3×3 with 64 number of filters. Main issue with VDSR is to use the bicubic interpolation as a pre-processing step, which is the extra burden on the model and increase the computational cost during the training. Furthermore, Kim et al. [17] suggested the concept of deeply recursive convolutional neural network (DRCN) for image super-resolution purposes. In this architecture 16 times used the same layer recursively. The performance of DRCN is improved, but its testing time is too slow and requires more memory space.

Shi et al., suggested an efficient sub-pixel convolutional neural network (ESPCN) [3], which included an efficient sub-pixel convolution layer to upscale the final LR feature maps into the HR output. Almost all later SR algorithms, such as SRResNet [16], EDSR [34], and others, incorporated the efficient sub-pixel convolution layer. These approaches are used the transposed or sub-pixel convolution layer to extract the features from the original LR input image. These networks can attain real-time performance or be designed to be exceedingly deep or wide with the use of deconvolution and sub-pixel convolution layers. Furthermore, such approaches simply stack building blocks in a serial manner and ignoring the hierarchical information contained in each building block. Lai et al., [35] use the concept of a pyramid and designed a new architecture known as a Deep Laplacian Pyramid Super-Resolution Network abbreviated as LapSRN. LapSRN resolve the issue of outliers with the help of L_1 loss also known as Charbonnier loss.

3. Proposed Method

In this section, we present our proposed Twin Image Super-Resolution using Deep Convolutional Neural Networks (TISR) for reconstructing the high-quality output image from the low-quality input image. Fig 1 shows the overall network architecture comprising two paths (also known as Twin networks). The TISR network architecture is split into four stages: low, mid, and high-level feature extraction, and upsampling and reconstruction. For upsampling strategy, we used a post-processing technique to reduce the computational cost as well as training time.

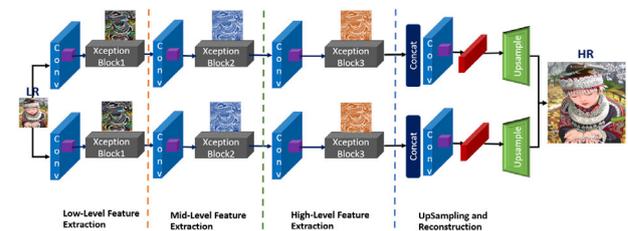


Figure 1. The proposed Network architecture of Twin Image Super-Resolution using Deep Convolutional Neural Networks (TISR).

3.1 Feature Extraction

For feature extraction purposes state-of-the-art methods such as SRCNN and VDSR used a pre-processing step as a bicubic interpolation to upscale the LR image into an HR image. The total features are extracted from the upscaled version, this way of feature extraction is to add new noises in the features and increase the extra burden on the model during the training. In this paper, we have not used a pre-processing technique for feature extraction but extracted the features directly from the original LR input image. Furthermore, earlier approaches are used a single straight path for extracting the feature, we are used two paths. For low-level feature extraction purposes, we used 2 convolution layers followed each by Xception block. All CNN layers are used

the kernel size of the order of 3×3 with 64 feature maps. Similar way, we extract the mid and high-level features simultaneously. A convolutional layer is a basic layer of deep learning model and it contains several filters whose number of parameters can be learned. Convolution filters have a smaller height and weight than the input volume. To create a neuron-based activation map, each filter is convolved with the input volume. Another method is to slide the filter across the width and height of the input, computing the dot products between the input and filter at each spatial position. The output volume of the convolutional layer is calculated by stacking the activation maps of all filters along the depth dimension. Each neuron in the activation map is only connected to a limited local region of the input volume since the width and height of each filter are designed to be less than the input. The local connection of the convolutional layer allows the network to train filters that respond optimally to a specific region of the input. Furthermore, by convolution the filter and the input, the activation map is formed, and the filter parameters are shared across all local positions. For efficient expression, learning, and generalization, the weight sharing method reduces the number of parameters required.

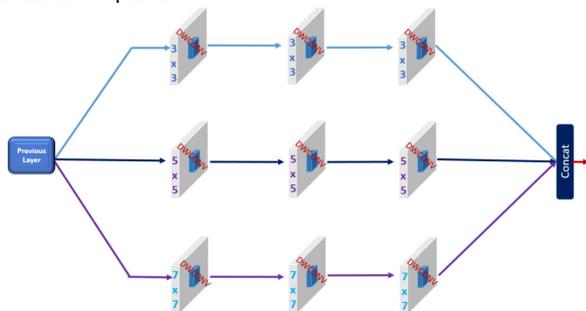


Figure 2. Proposed multi-scale Xception block.

3.2 Proposed Multi-scale Xception Block

Xception [36] is a Depthwise Separable Convolutions-based deep convolutional neural network architecture proposed by Google. Furthermore, Google also introduce the concept of inception architecture. Our proposed multi-scale Xception block idea is borrowed from Xception architecture [36], which was initially proposed by Google engineers and stands as Extreme version of Inception with addition of depthwise separable convolution. In our Xception block remove all 1×1 convolution layers and used a 3 convolution layers with different kernel sizes. For low-level feature extraction purposes, we are used small kernel size of the order 3×3 with 64 number of filters. For Medium level feature extraction purposes, we used kernel size of the order 5×5 and similarly for high-level feature extraction purposes we apply 7×7 kernels with same 64 number of filters. Finally, all branches are concatenated to extract the hierarchical rich features information simultaneously as shown in Fig 2.

3.3 UpSampling and Reconstruction

UpSampling is the final stage of our model for generating the high-resolution (HR) output image. There are many ways to upscale the LR features into HR features, such as bicubic

interpolation, nearest-neighbor interpolation, or bilinear interpolation. The implementation of these techniques is very simple, but the results are not satisfactory and create the jagged ringing artifacts and introduce the new noises in the reconstructed output image. To handle such types of issues, we used a deconvolution layer to upscale the LR features into HR features. Deconvolution operation is the inverse convolution operation. To reduce the computational cost of the model, we cannot direct fed the LR features into the Deconvolution layer, so we add two bottleneck layers of the kernel size is 1×1 . The use of the bottleneck layer is to reduce the number of parameters and then we fed the LR features to the deconvolution layer to reconstruct the upscale HR features. Finally, both branch features are concatenated to reconstruct the visually pleasing high-quality HR output image.

4. Experiments

The quantitative and qualitative evaluations are present to validate the performance of our proposed method. We use Yang 91 image dataset and extra 200 images are obtained from BSDS200 for our training purposes. The Adam was used as an optimizer with an initial learning rate is 0.0003 including 16 as a mini-batch size. Our model fully converges on 200 epochs. For training purposes, we run our code on NVIDIA GPU RTX 2070. Keras TensorFlow library is used for designing the model architecture.

4.1 Computational Complexity of Proposed Method.

The computational complexity of any model plays a vital role in the practical application point of view. The model has a higher computational cost it takes more testing time as compared to the lower computational cost of the model. All deep learning communities measure the computational cost in terms of network parameters. The network has more parameters it means having a more computational cost and vice versa. Our model has fewer parameters as compared to recent deep learning models such as VDSR, LapSRN, and DRCNN as shown in Fig 3 and 4.

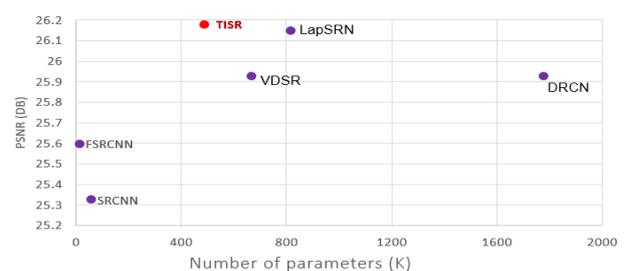


Figure 3. Quantitative evaluation of the network depth/ computational cost in terms of k parameters versus PSNR on image SR dataset SET5 enlargement factor $4 \times$.

4.2 Quantitative Comparison with existing state-of-the-art methods

The quantitative comparison of our proposed TISR method with 9 standard methods including base-line bicubic method,

such as Bicubic, A+, RFL, SelfExSR, SRCNN, FSRCNN, SCN, VDSR, DRCN and LapSRN were experimentally compared with our proposed method. Table 1 present the PSNR (dB) / SSIM comparison result with the existing deep CNN-based image SR methods on three main benchmark datasets SET5, SET14 and BSDS100 with challenging factor 3×, 4× and 8×. Results are observed from Table 1 and 2, that our method attained the improved performance as compared to other state-of-the-art methods as well as has a smaller number of parameters. Furthermore, graphical representation a shown in Fig: 5 and 6.

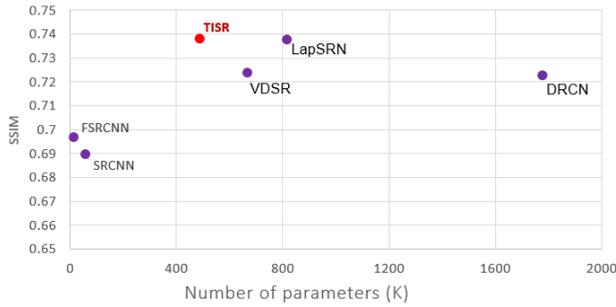


Figure 4. Quantitative evaluation of the network depth/ computational cost in terms of k parameters versus SSIM on image SR dataset SET5 enlargement factor 4×.

Table 1. Experimental evaluation of our proposed method with other image SR methods with scale factor 3×, 4× and 8×. First-best values are indicated in red color with bold and second-best in blue-colors with underline.

Method	Scale	SET5 PSNR/SSIM	SET14 PSNR/SSIM	BSDS100 PSNR/SSIM
Bicubic	3×	30.41/0.869	27.55/0.775	27.22/0.741
A+ [25]	3×	32.62/0.909	29.15/0.820	28.31/0.785
RFL [37]	3×	32.47/0.906	29.07/0.818	28.23/0.782
SelfExSR	3×	32.66/0.910	29.18/0.821	28.30/0.786
SRCNN	3×	32.78/0.909	29.32/0.823	28.42/0.788
FSRCNN	3×	33.18/0.914	29.37/ <u>0.824</u>	28.53/0.791
SCN	3×	32.62/0.908	29.16/0.818	28.33/0.783
VDSR	3×	33.67/ <u>0.921</u>	29.78/ 0.832	28.83/0.799
DRCN	3×	33.83/0.922	29.77/ 0.832	28.80/0.797
LapSRN	3×	<u>33.82/0.922</u>	<u>29.87/0.832</u>	<u>28.82/0.798</u>
TISR (ours)	3×	<u>33.82/0.922</u>	29.89/0.832	28.83/0.798
Bicubic	4×	28.43/0.811	26.01/0.704	25.97/0.670
A+ [25]	4×	30.32/0.860	27.34/0.751	26.83/0.711
RFL [37]	4×	30.17/0.855	27.24/0.747	26.76/0.708
SelfExSR	4×	30.34/0.862	27.41/0.753	26.84/0.713
SRCNN	4×	30.50/0.863	27.52/0.753	26.91/0.712
FSRCNN	4×	30.72/0.866	27.61/0.755	26.98/0.715
SCN	4×	30.41/0.863	27.39/0.751	26.88/0.711
VDSR	4×	<u>31.35/0.883</u>	28.02/0.768	27.29/0.726
DRCN	4×	31.54/0.884	28.03/0.768	27.24/0.725

LapSRN	4×	31.54/0.885	<u>28.19/0.772</u>	<u>27.32/0.727</u>
TISR (ours)	4×	31.54/0.885	28.21/0.773	27.33/0.728
Bicubic	8×	24.40/0.658	23.10/0.566	23.67/0.548
A+ [25]	8×	25.53/0.693	23.89/0.595	24.21/0.569
RFL [37]	8×	25.38/0.679	23.79/0.587	24.13/0.563
SelfExSR	8×	25.49/0.703	23.92/0.601	24.19/0.568
SRCNN	8×	25.33/0.690	23.76/0.591	24.13/0.566
FSRCNN	8×	25.60/0.697	24.00/0.599	24.31/0.572
SCN	8×	25.59/0.706	24.02/0.603	24.30/0.573
VDSR	8×	25.93/ <u>0.724</u>	<u>24.26/0.614</u>	24.49/ <u>0.583</u>
DRCN	8×	25.93/0.723	24.25/0.614	24.49/0.582
LapSRN	8×	<u>26.15/0.738</u>	<u>24.35/0.620</u>	24.54/0.586
TISR (ours)	8×	26.18/0.738	24.35/0.623	<u>24.53/0.586</u>

Table 2. Network depth in k number of parameters of different single image SR methods.

Method	#Param k	SET5 PSNR/SSIM	SET14 PSNR/SSIM	BSDS100 PSNR/SSIM
Bicubic	-/-	30.41/0.869	27.55/0.775	27.22/0.741
A+ [25]	-/-	32.62/0.909	29.15/0.820	28.31/0.785
RFL [37]	-/-	32.47/0.906	29.07/0.818	28.23/0.782
SelfExSR	-/-	32.66/0.910	29.18/0.821	28.30/0.786
SRCNN	57	32.78/0.909	29.32/0.823	28.42/0.788
FSRCNN	12	33.18/0.914	29.37/ <u>0.824</u>	28.53/0.791
SCN	42	32.62/0.908	29.16/0.818	28.33/0.783
VDSR	665	33.67/ <u>0.921</u>	29.78/ 0.832	28.83/0.799
DRCN	1,775	33.83/0.922	29.77/ 0.832	28.80/0.797
LapSRN	812	<u>33.82/0.922</u>	<u>29.87/0.832</u>	<u>28.82/0.798</u>
TISR (ours)	488	<u>33.82/0.922</u>	29.89/0.832	28.83/0.798

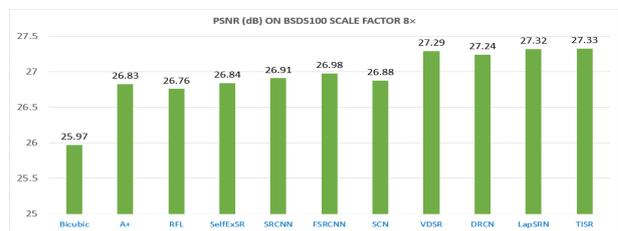


Figure 5. Quantitative evaluation of PSNR on BSDS100 dataset enlargement factor 8 × with other state-of-the-art methods. Our proposed method obtained the highest PSNR as compared to other methods.

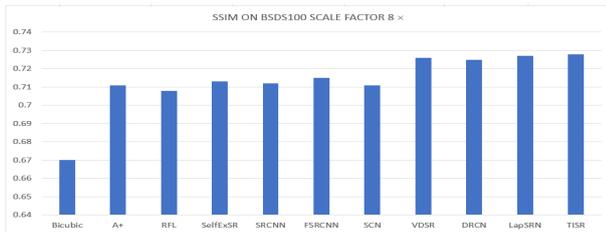
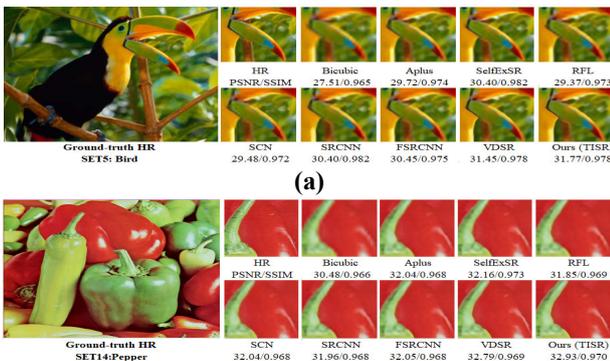


Figure 6. Quantitative evaluation of SSIM on BSDS100 dataset enlargement factor $8 \times$ with other state-of-the-art methods. Our proposed method obtained highest PSNR as compared to other methods.

Additionally, qualitative evaluation of TISR as shown in Fig. 7, with scale factor $4 \times$. In Figure 7(a) Bird image is obtained from SET5 dataset and test the results of patch on 8 standard methods. Our TISR model reconstruct the fine edges smoothly as compare to base-line bicubic method. Similarly, in Figure 7(b), and (c) our proposed method visually pleasing, because other methods generated result are blurry and jagged artifacts.



References

1. Sheikh, H.R., A.C. Bovik, and G. De Veciana, *An information fidelity criterion for image quality assessment using natural scene statistics*. IEEE Transactions on image processing, 2005. **14**(12): p. 2117-2128.
2. Zhou, W., et al., *Blind binocular visual quality predictor using deep fusion network*. IEEE Transactions on Computational Imaging, 2020. **6**: p. 883-893.
3. Shi, W., et al. *Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
4. Cherukuri, V., et al., *Deep MR brain image super-resolution using spatio-structural priors*. IEEE Transactions on Image Processing, 2019. **29**: p. 1368-1383.
5. Dudczyk, J., *A method of feature selection in the aspect of specific identification of radar signals*. Bulletin of the Polish Academy of Sciences. Technical Sciences, 2017. **65**(1).

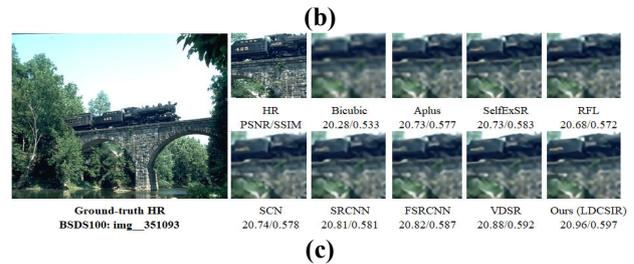


Figure 7. Qualitative as well as quantitative comparisons of TISR approach with publicly available image SR methods on sale factor $4 \times$.

5. Conclusions

In this paper, Twin Image Super-Resolution using Deep Convolutional Neural Networks (TISR) is proposed. The TISR model used the two way architecture with stacking Xception blocks for extracting the multi-scale feature information through the original LR input image. Furthermore, apply the different kernel sizes in the Xception blocks such as 3×3 , 5×5 , and 7×7 to extract the low, medium, and high-level feature information simultaneously. Additionally, we have replaced the traditional bicubic upsampling pre-processing step into deconvolution layer as a post-processing technique to further decrease the computational burden of the network during the training. Experimental quantitative, as well as qualitative results, validate that the proposed lightweight TISR network achieved favorable results against the state-of-the-art methods.

6. Turchini, F., et al., *Deep learning based surveillance system for open critical areas*. Inventions, 2018. **3**(4): p. 69.
7. Hazra, D. and Y.-C. Byun, *Upsampling Real-Time, Low-Resolution CCTV Videos Using Generative Adversarial Networks*. Electronics, 2020. **9**(8): p. 1312.
8. Das, M. and S.K. Ghosh, *Deep-STEP: A deep learning approach for spatiotemporal prediction of remote sensing data*. IEEE Geoscience and Remote Sensing Letters, 2016. **13**(12): p. 1984-1988.
9. Pierson, H.A. and M.S. Gashler, *Deep learning in robotics: a review of recent research*. Advanced Robotics, 2017. **31**(16): p. 821-835.
10. Li, X., et al., *Deep learning methods in real-time image super-resolution: a survey*. Journal of Real-Time Image Processing, 2020. **17**(6): p. 1885-1909.
11. Nasrollahi, K. and T.B. Moeslund, *Super-resolution: A survey*. IEEE Transactions on Image Processing, 2017. **26**: p. 3523-3541.

- a comprehensive survey*. Machine vision and applications, 2014. **25**(6): p. 1423-1468.
12. Tong, C. and K. Leung, *Super-resolution reconstruction based on linear interpolation of wavelet coefficients*. Multidimensional Systems and Signal Processing, 2007. **18**(2): p. 153-171.
 13. Sun, N. and H. Li, *Super resolution reconstruction of images based on interpolation and full convolutional neural network and application in medical fields*. IEEE Access, 2019. **7**: p. 186470-186479.
 14. Jing, L., G. Zongliang, and Z. Xiuchang, *Directional bicubic interpolation-a new method of image super-resolution*. Proceedings of ICMT, Atlantis Press, 2013: p. 470-477.
 15. 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.
 16. Ledig, C., et al. *Photo-realistic single image super-resolution using a generative adversarial network*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
 17. 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.
 18. 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.
 19. LeCun, Y., et al., *Handwritten digit recognition with a back-propagation network*. Advances in neural information processing systems, 1989. **2**.
 20. Fernandez-Granda, C. and E.J. Candes. *Super-resolution via transform-invariant group-sparse regularization*. in *Proceedings of the IEEE International Conference on Computer Vision*. 2013.
 21. Yang, J., Z. Lin, and S. Cohen. *Fast image super-resolution based on in-place example regression*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013.
 22. He, H. and W.-C. Siu. *Single image super-resolution using Gaussian process regression*. in *CVPR 2011*. 2011. IEEE.
 23. Efrat, N., et al. *Accurate blur models vs. image priors in single image super-resolution*. in *Proceedings of the IEEE International Conference on Computer Vision*. 2013.
 24. Dai, D., R. Timofte, and L. Van Gool. *Jointly optimized regressors for image super-resolution*. in *Computer Graphics Forum*. 2015. Wiley Online Library.
 25. Timofte, R., V. De Smet, and L. Van Gool. *A+: Adjusted anchored neighborhood regression for fast super-resolution*. in *Asian conference on computer vision*. 2014. Springer.
 26. Zhu, Y., Y. Zhang, and A.L. Yuille. *Single image super-resolution using deformable patches*. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
 27. Gao, X., et al., *Image super-resolution with sparse neighbor embedding*. IEEE Transactions on Image Processing, 2012. **21**(7): p. 3194-3205.
 28. Zhang, K., et al. *Multi-scale dictionary for single image super-resolution*. in *2012 IEEE conference on computer vision and pattern recognition*. 2012. IEEE.
 29. Wang, S., et al. *Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis*. in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012. IEEE.
 30. Jain, V. and S. Seung, *Natural image denoising with convolutional networks*. Advances in neural information processing systems, 2008. **21**.
 31. Yamamoto, K., T. Togami, and N. Yamaguchi, *Super-resolution of plant disease images for the acceleration of image-based phenotyping and vigor diagnosis in agriculture*. Sensors, 2017. **17**(11): p. 2557.
 32. Dong, C., et al. *Learning a deep convolutional network for image super-resolution*. in *European conference on computer vision*. 2014. Springer.
 33. Dong, C., C.C. Loy, and X. Tang. *Accelerating the super-resolution convolutional neural network*. in *European conference on computer vision*. 2016. Springer.
 34. Lim, B., et al. *Enhanced deep residual networks for single image super-resolution*. in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017.
 35. 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.
 36. Chollet, F. *Xception: Deep learning with depthwise separable convolutions*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
 37. Schuler, S., C. Leistner, and H. Bischof. *Fast and accurate image upscaling with super-resolution forests*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.