© 2017 IJNRD | Volume 2, Issue 6 June 2017 | ISSN: 2456-4184

ENHANCING IMAGE QUALITY THROUGH GAN-BASED SUPER-RESOLUTION WITH INNOVATIVE QUALITY METRICS

[1]Narendra Kumar D N , [2]Anand Bhamashetti

[1,2] Lecturer , Department of Electronics and Communication ,Government Polytechnic Zalaki

Abstract : This "Advancing Single Image Super-Resolution (SISR) with GMGAN: Bridging the Gap in Visual Quality" In recent years, Single Image Super-Resolution (SISR) has emerged as a captivating research frontier. Remarkable progress in SISR owes its success to deep learning and the power of Generative Adversarial Networks (GANs). However, despite these advancements, the generated images still exhibit undesirable artifacts.This paper introduces a novel approach named GMGAN to tackle SISR challenges. Our method is designed with the goal of producing images that align more closely with the capabilities of the Human Visual System (HVS). To achieve this, we introduce a quality loss component by seamlessly integrating an Image Quality Assessment (IQA) metric known as Gradient Magnitude Similarity Deviation (GMSD). Notably, this represents the first instance of truly incorporating an IQA metric into the SISR framework.Furthermore, we address the stability concerns inherent in conventional GANs by adopting a variant known as WGAN-GP. Through a series of experiments, we demonstrate that GMGAN, enhanced by quality loss and WGAN-GP, delivers visually compelling results and establishes a new benchmark in the field of Single Image Super-Resolution.

1. INTRODUCTION

"Advancing Single Image Super-Resolution (SISR) for Photorealistic Results" Single Image Super-Resolution (SISR) is a challenging task aimed at enhancing low-resolution (LR) images to recover high-resolution (HR) counterparts. However, it's inherently complex as there are multiple HR images that could correspond to a single LR image. Despite significant progress in the field, the key challenge remains achieving photorealistic results with natural textures while minimizing unwanted artifacts.

Solutions have evolved from traditional methods to learning-based approaches. Learning-based methods employ deep neural networks, trained with a sophisticated loss function—typically, the mean squared error (MSE) loss, closely related to the Peak Signal-to-Noise Ratio (PSNR). However, MSE tends to produce overly smooth outcomes, and PSNR, while widely used for evaluation, relies solely on low-level pixel differences and may not suit SISR, which deals with more complex image transformations.

To address these issues, we introduce a novel loss term called "quality loss," moving away from the pursuit of state-of-the-art PSNR values. This quality loss considers higher-level image attributes and visual fidelity.
Moreover, many learning-based methods employ Generative Adversarial Networks (GANs) to generate visually appealing results. However, the original GAN often faces training instability. We address this challenge by adopting a more stable variant.Additionally, the importance of training datasets should not be underestimated. While ImageNet is a popular choice, it may not be ideal for SISR due to image quality variations.Single Image Super-Resolution is a critical task with applications in medical imaging, surveillance, astronomy, and face recognition. Deep learning, particularly Convolutional Neural Networks (CNNs) and Fully Convolutional Neural Networks (FCNs), shows remarkable potential for tackling this task. These networks are capable of learning intricate features and solving specific tasks efficiently.In the context of assisting the visually impaired, technology plays a crucial role in providing access to information and communication. Deep learning systems, including CNNs, offer innovative solutions for various applications, overcoming challenges posed by moving objects and improving accessibility.

In conclusion, our research aims to push the boundaries of Single Image Super-Resolution to achieve photorealistic results, address challenges related to GAN training stability, and emphasize the importance of tailored training datasets.

2. LITERATURE REVIEW AND OUR RESEARCH CONTRIBUTIONS

In this section, we provide an overview of the existing research in the field and highlight our unique contributions. GuiminLina et al .,[1] In this study, we address Image Super-Resolution (SR) using a 7-layer Dilated Convolutional Neural Network (DCNN) with skip-connections, harnessing the power of deep learning, specifically CNNs. Our approach excels at reconstructing high-resolution images from their low-resolution counterparts by leveraging the flexibility of dilated convolutions to control the FOV effectively.Bee Lim et al.,[2] Recent advances in super-resolution leverage DCNN, with a focus on residual learning techniques, yielding superior performance. Our work introduces the Enhanced Deep Super-Resolution Network (EDSR), surpassing current state-of-the-art SR methods in performance Yifan Wang et al.,[6] Recent deep learning methods for single image super-resolution have delivered impressive results in error measures and perceptual quality. Nonetheless, achieving high-quality results for substantial upsampling factors remains a persistent challenge Chao Ma et al .,[7] Single-image super-resolution (SR) techniques strive to generate high-quality, high-resolution (HR) images from low-resolution (LR) inputs. These approaches leverage various priors, including edges, gradients, neighboring interpolation, regression, and patches, to enhance image quality.RaduTimofte et al.,[9] Single image super-resolution (SR) focuses on restoring high-frequency details in high-resolution (HR) images from single low-resolution (LR) inputs. It tackles the inherently ill-posed nature of the problem through interpolation-based, reconstruction-based, and learning-based techniques.Tong Tong et al.,[11] The network's dense skip connections counter the vanishing-gradient problem in deep networks by establishing short paths from output to each layer. Deconvolution layers enhance upsampling and reconstruction speed while significantly reducing parameter count, boosting computational efficiency. Our evaluation on four benchmark datasets establishes a new state-of-the-art performance.Wei-Sheng Lai et al.,[15] Introduce Laplacian Pyramid Super-Resolution Network (LapSRN), leveraging convolutional neural networks to progressively reconstruct sub-band residuals of high-resolution images through transposed convolutions and coarse-resolution feature maps.AssaShocher et al.,[16] Deep Learning has significantly boosted Super-Resolution (SR) performance, yet supervised methods rely on predefined LR-HR data, lacking robustness to real-world artifacts like sensor noise, compression, and non-ideal PSF.YulunZhang et al.,[17] Introduce the Residual Dense Network (RDN) for image super-resolution, optimizing the utilization of hierarchical features from all convolutional layers through Residual Dense Blocks (RDBs) that capture rich local features. While traditional methods offer partial solutions to SISR, the advent of learning-based approaches has significantly outperformed them in terms of both performance and efficiency. As a result, our emphasis in this discussion centers on learning-based methods. Our research contributes Single image super-resolution (SISR) using generative adversarial network (GAN) , a novel approach that optimizes the enhanced generative model in gradient map space, yielding superior visual quality, as confirmed by experimental results . We incorporate a quality loss by integrating the IQA metric Gradient Magnitude Similarity Deviation, chosen for its meaningful derivative and consistent ability to predict perceptual quality compared to human subjective evaluation. Additionally, we address training instability by substituting the original GAN with WGAN-GP, promoting smoother decision boundaries for improved generator training and output quality. To assess the impact of diverse training datasets, we conducted a comprehensive series of experiments using various training datasets. Our findings underscore the positive influence of a substantial number of high-quality training images with diverse textures on the overall outcome.

3. METHODOLOGY ENHANCING IMAGE SUPER-RESOLUTION THROUGH GAN WITH INNOVATIVE QUALITY ENHANCEMENT

In this section, we provide a comprehensive overview of our innovative approach,GAN with map gradinat. GAN synergizes the advantages of an Image Quality Assessment (IQA)-based loss function with advanced adversarial training. Our presentation covers two core components: network architecture and the detailed exploration of loss functions. We commence with an in-depth examination of the GAN architecture and subsequently delve into the intricacies of our novel loss terms. Notably, we introduce a groundbreaking quality loss, representing the first genuinely IQA-based loss term.

3.1 ARCHITECTURE

Inspired by SRGAN, the architecture of GAN with map gradient maintains the same discriminator structure as SRGAN. However, to further elevate the perceptual quality of the reconstructed images, we introduce three key modifications to the generator:

We enhance the generator's capabilities by replacing the original residual blocks with Residual-in-Residual Dense Block (RRDB) blocks , allowing for dense connections and a multilevel residual network.

Addressing training instability, we replace the original GAN with Wasserstein GAN with Gradient Penalty (WGAN-GP) , fostering more stable training and the generation of realistic results.

For improved memory efficiency and reduced computational complexity, we follow the approach suggested in EDSR by removing batch normalization (BN) layers.

Figure [1] illustrates the comprehensive network architecture of GAN with gradient map.

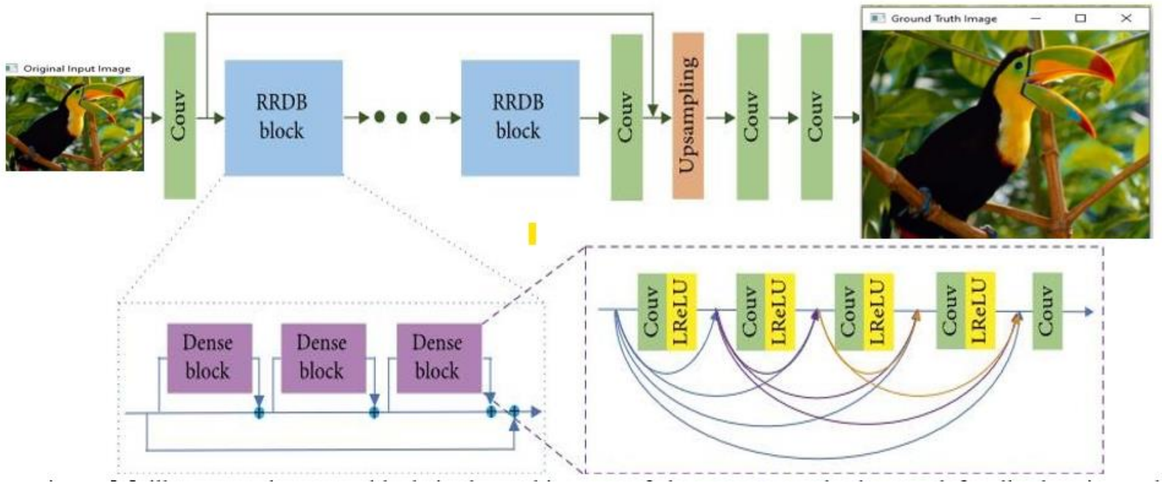

Figure [1] illustrates The upper block is the architecture of the generator. The bottom left tells the internal architecture of each RRDB block and the downright informs the inner structure of each dense block.

In Figure 1, the overall architectural framework of SRGAN remains intact, while significant enhancements are introduced in the low-level architecture. Specifically, the conventional residual blocks are substituted with Residual-in-Residual Dense Block (RRDB) structures, each featuring a more intricate design. This innovation establishes a hierarchical residual-in-residual approach, enabling enhanced feature exploration as outlined in Section 2.1.1.

Within each dense block, the integration of short paths connecting every layer enhances information flow, exemplifying the concept of dense connections. Our configuration incorporates 23 RRBD blocks, elevating the network's capability to capture intricate details.

In addition to architectural improvements, the removal of Batch Normalization (BN) layers brings several advantages. First, BN layers normalize features, reducing network's range flexibility. Second, they consume significant computational resources and GPU memory. Consequently, the elimination of BN layers enhances network performance, particularly in resource-constrained environments. Lastly, when deep networks undergo adversarial training, BN layers can introduce undesired artifacts. For these compelling reasons, we have chosen to exclude BN layers from our design.

4. LOSS FUNCTIONS

This section delves into the loss function, which serves as the optimization objective for learning-based Single Image Super-Resolution (SISR) techniques. We provide a comprehensive breakdown of the individual loss terms:

- Perceptual loss (lP).

To address the limitations of the MSE loss and enable a more precise assessment of semantic and perceptual disparities between images, we introduce and optimize a perceptual loss. This perceptual loss is rooted in high-level features extracted from a pretrained network. MES loss id defined as

$l_{\mathrm{P}} = \|\varnothing\left(G_{\theta}\left(I_{\mathrm{LR}}\right)\right) - \varnothing\left(I_{\mathrm{HR}}\right)\|_2^2$ where θ refers to the 19-layer VGG network

- Mean Squared Error loss (MSE), denoted as lMSE.
  MSE loss function $l_{\mathrm{MSE}} = \|G_{\theta}(I_{\mathrm{LR}}) - I_{\mathrm{HR}}\|_2^2$ where the parameter of the generator is denoted by θ; the generated image, namely, ISR, is denoted by Gθ(ILR); and the ground truth is denoted by IHR.
- Quality loss (lQ).
  Drawing inspiration from the Image Quality Assessment (IQA) metric known as Gradient Magnitude Similarity Deviation (GMSD), we introduce a novel loss term called quality loss. Our choice of GMSD is driven by two key considerations. Firstly, full-reference IQA metrics are essential for assessing generated image quality, as they rely on reference-based evaluation—an essential criterion for use as a loss function. While metrics like Mean Squared Error (MSE), equivalent to PSNR, fall short in aligning with human perceptual quality, we prioritize reference-based metrics known for their strong performance. These include Visual Information Fidelity (VIF) , measuring shared information between reference and distorted images, and Feature Similarity Index (FSIM) , capturing dissimilarity based on local phase congruency and gradient magnitude.

  Our choice of GMSD balances feasibility and efficiency. While VIF and FSIM may offer superior quality prediction, their formulations are more complex and non-differentiable, rendering them impractical for optimization within neural networks. Hence, we opt for GMSD to define the quality loss, considering both its simplicity and consistent ability to predict perceptual image quality in harmony with Human Visual System (HVS).
- Adversarial loss for the generator (lGA).

The chosen generative model is the Generative Adversarial Network (GAN) , which, however, is susceptible to training instability. Wasserstein GAN (WGAN) mitigates this issue by introducing the Wasserstein distance and a value function, offering superior theoretical properties compared to the original GAN. Nevertheless, WGAN relies on weight clipping to ensure that the discriminator remains within the 1-Lipschitz space, which can lead to challenges like vanishing or exploding gradients unless the clipping threshold is meticulously fine-tuned.

To address this limitation of weight clipping, we adopt an alternative strategy known as gradient penalty from WGAN-GP to enforce the Lipschitz constraint. In contrast to WGAN, WGAN-GP calculates the Wasserstein distance between two distributions and penalizes the gradient of the discriminator concerning its input, eliminating the need for weight clipping. This approach not only encourages the discriminator to learn smoother decision boundaries but also accelerates the training phase and enhances the quality of generated images.

- Adversarial loss for the discriminator (lDA).

Collectively, these loss terms define the generator's loss function, which guides the learning proces. Loss function for genareter is

$lG = \alpha lMSE + \beta lp + \gamma lQ + \delta lGA$

The loss function fo the discriminator D is

$lD = \delta lDA$

lp=perceptual loss lQ=quality loss

lGA=adversarial loss for the generator lDA=adversarial loss for the discriminator

where α, β, c, and δ are the weights for each loss term.

5. RESULTS

5.1 Datasets: The proposed method is tested with standard datasets such as part of ImageNet,DF2K ,DF2K+OST,DIVIK Augmentation.

Part of Image Net: (short for ImageNet for convenience),consisting of 10000 relatively low-qualityimages. Note that ImageNet dataset is a benchmarkfor object category classification and detection onhundreds of categories originally, and thus thisdataset is characterized by large number but lowquality of contained images.

DF2K dataset: consisting of 3450 relatively high quality images in total. Here, DF2K is short forDIVIK [45] + Flickr2K [46], which contains 2650and 800 relatively high-quality images, respectively.Note that DIVIK dataset is a professional dataset for image restoration tasks.

DF2K + OST dataset: consisting of 3902 images totally.Here, 2650 relatively high- quality images arefrom DF2K dataset and 1342 images of rich texturesare from OutdoorSceneTraining (OST) [19] dataset.

DIVIK [45] augmentationdataset:consisting of 20480 relativelyhigh-quality images flipped or rotated fromDIVIK dataset. As suggested in [47], to enhance thefinal performance, we rotated each original imagefrom DIVIK dataset by 90°, 180°, and 270° and thenflipped them vertically to get 8 corresponding imagesincluding identity without altered content.

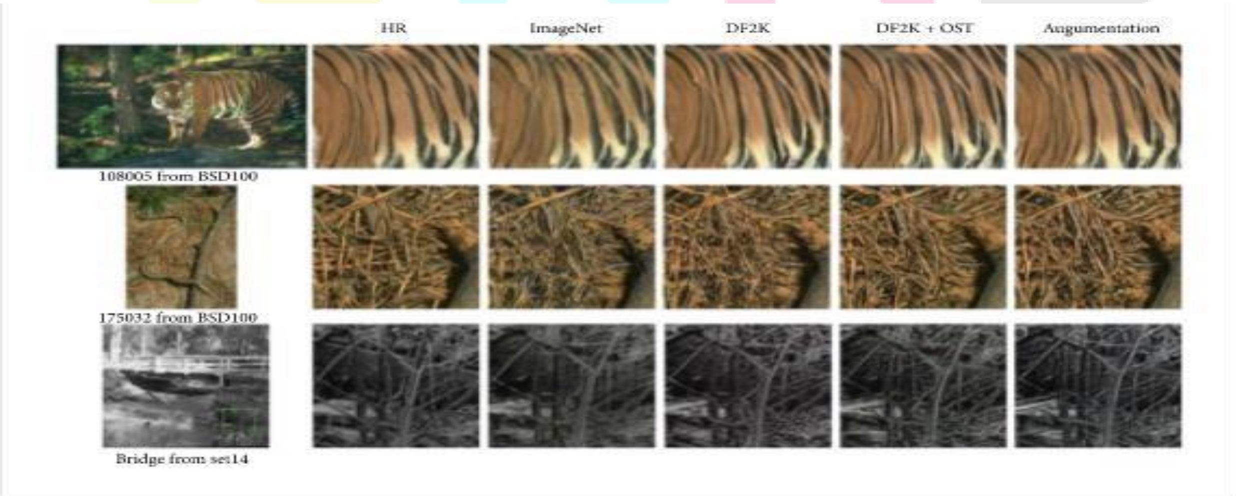

Figure [2] Visual comparisons on four different training datasets.

© 2017 IJNRD | Volume 2, Issue 6 June 2017 | ISSN: 2456-4184

For evaluation, experiments were conducted on fourpublic benchmark datasets: Set5 , Set14 , BSD100, and Urban100 . Set5, Set14, and BSD100 containnatural scenes such as woman, butterfly, and bridge, whileUrban100 contains urban scenes with details in differentfrequency bands such as high-rise buildings, brick wall, andretro church. Because the published results of some state-ofthe-art methods do not contain the case of Urban100,Urban100 was adopted only by our proposed GMGAN.

Below are result that converted in to super resolution image.

Original image [1] Bicubic [2] SRGAN[3] our module [4]

As depicted in the preceding Figure [3], Image 1 represents the original image with lower resolution and diminished quality. Image 2 showcases a high-resolution outcome achieved through the Bicubic module. Image 3 illustrates a super-resolution image produced by the SRGAN module. The final result, Image 4, represents a super-resolution image generated using the Gradient Map GAN method, offering superior visual quality discernible to the naked eye when compared to other techniques.

5.2 Qualitative Results

While visual perception is a valuable indicator, it alone does not serve as the definitive measure of image quality. To comprehensively assess the quality of images generated by our methods and other modules, we conducted additional analyses. Our evaluation included key metrics such as Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), and Structural Similarity Index (SSIM). We compared these metrics between the original image and the super-resolution images produced by our method and two other techniques.

Original image [1] Bicubic [2] SRGAN[3] our module [4]

(MSE/SSIM/PSNR) (0.0818/0.8341/22.69) (0.08212/0.9102/33.14)

(0.09213/0.9432/41.32)

Original image [1] Bicubic [2] SRGAN[3] our module [4]

(MSE/SSIM/PSNR) (0.0771/0.8421/21.96) (0.08321/0.891/24.321) (0.0911/0.9512/26.347)

Figure[4] . Comparison of parameters like MSE,SSIM and PSNR of original image along with other methods and our methods .

As shown figure[4] the qualitative results have better results in our module compare to other module.

6. CONCLUSION AND FUTURE SCOPE

Addressing the challenge of producing visually pleasing super-resolution images has remained a pivotal concern within the realm of Single Image Super-Resolution (SISR). In response, this paper introduces GAN, a novel approach that amalgamates an Image Quality Assessment (IQA)-based loss function with enhanced adversarial training.

Furthermore, we delve into an extensive exploration of the influence of diverse training datasets on the GAN's performance.Our comprehensive experimentation underscores the commendable capabilities of GMGAN, demonstrating its proficiency in generating super-resolved images that exhibit heightened photorealism while minimizing the presence of undesirable artifacts. A noteworthy observation gleaned from our research is the substantial advantage conferred by employing an extensive repository of high-quality training images, particularly those characterized by intricate textures, in enhancing the ultimate image quality.As we steer toward future developments, our trajectory involves refining the network architecture to equip GMGAN with the capability to effectively handle images that exhibit pronounced repetitive structures. This ongoing endeavor aligns with our commitment to advancing the state of the art in Single Image Super-Resolution techniques.

References

[1] GuiminLin,Qingxiang Wu, Liang Chen, LidaQiu, Xuan Wang, Tianjian Liu, XiyaoChen,"Deep unsupervised learning for image super-resolution with generative adversarial network",signalprocessing:Image communication,68,88-10-2018.

[2] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, "Enhanced DeepResidual Networks for Single Image Super-Resolution,"*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,pp.136-144,2017.

[3] JoanBruna, Pablo Sprechmann, YannLeCun, "super resolution with deep convolutional sufficient statisics,"arXiv:1511.05666v4[cs.cv] ,2016.

[4] Mehdi S. M. Sajjadi Bernhard Sch¨olkopf Michael Hirsch,"EnhanceNet: Single Image Super-Resolution Through Automated Texture Synthesis,"*IEEE International Conference on Computer Vision (ICCV),* pp. 4491-4500,2017.

[5] Kwang In Kim and Younghee Kwon,"Single-Image Super-resolution UsingSparse Regression and Natural Image Prior,"pp.1-8,2010.

[6] Yifan Wang, Federico Perazzi , Brian McWilliams, "A Fully Progressive Approach to Single-Image Super-Resolution,"*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,pp.977-986,2018.

[7] Chao Ma, Chih-Yuan Yang, XiaokangYanga, Ming-Hsuan Yang,"Learning a No- Reference Quality Metric for Single-Image Super-Resolution,"arXiv:1612.05890v1[cs.CV]18 Dec 2016.

[8] Mohammad Saeed Rad, BehzadBozorgtabar, ClaudiuMusat, Urs-Viktor Marti, Max Basler, Hazim Kemal Ekenel, "Benefiting from multitask learning to improve single image superresolution,"[arXiv:1907.12488[cs.CV], 2019.

[9] RaduTimofte,RasmusRothe,Luc Van Gool, "Seven ways to improve example-based single image super resolution,"IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1865-1873,2015

[10] Christian Ledig, Lucas Theis, FerencHusz´ar, Jose Caballero, Andrew Cunningham, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,pp.4681-4690,2016.

[11] Tong Tong, Gen Li, Xiejie Liu, QinquanGao, "Image Super-Resolution Using Dense Skip Connections", *IEEE International Conference on Computer Vision (ICCV),* pp.4799- 4807,2017.

[12] WufengXue, Lei Zhang,XuanqinMou, Alan C. Bovik, "Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index," IEEE Transactions on ImageProcessing,pp.684-695,2013.

[13] EirikurAgustsson, RaduTimofte, "NTIRE 2017 Challenge on Single Image Super- Resolution: Dataset and Study," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*,pp.126-135,2017.

[14] Lin Zhang, Lei Zhang,XuanqinMou David Zhang, "A Feature Similarity Index for Image Quality Assessment,"IEEE Transactions on image processing,vol.20,no.8,2011.

[15] Wei-Sheng Lai, Jia-Bin Huang, NarendraAhuja, "Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution," IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp.624-632,2017.

[16] AssafShocher, NadavCohen,MichalIrani, "Zero-Shot Super-Resolution using Deep Internal Learning," IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp.3118-3126,2017.

[17] Yulun Zhang, YapengTian, Yu KongandBinengZhong,"Residual Dense Network for Image Super-Resolution," IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp.2472-2481,2018.

[18] Jiwon Kim, Jung Kwon Lee and Kyoung Mu Lee,"Accurate Image Super-Resolution Using Very Deep Convolutional Networks," IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp.1646-1654,2016.

[19] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. Lecun, OverFeat,"Integrated recognition localization andconvolutional networks," in: International Conference on Learning Representations, ICLR, 2014.

[20] J.-B. Huang, A. Singh, and N. Ahuja."Single image superresolution from transformed self-exemplars,"IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[21] R. Zeyde, M. Elad, and M. Protter."On single image scale-up using sparse- representations," In Curves and Surfaces, pp. 711–730. 2012.

[22] C. Dong, C. C. Loy, K.He, and X. Tang."Learning a deep convolutional network for image super-resolution," In ECCV, 2014.

[23] S. Baker and T. Kanade,"Limits on super-resolution and how to break them," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1167–1183, 2002.

[24] C. Ledig, L. -eis, F. Husz´ aret al., "Photo-realistic single image super-resolution using a generative adversarial network,"inProceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 105–114, IEEE, Honolulu, HI, USA, July 2017.