A Review on Self Learning based Methods for Real World Single Image Super Resolution

Yogesh J. Gaikwad

Dr. Vishwnath Karad Mitworld Peace University, School of Polytechnic & Skill Development, Pune, India

Corresponding author: Yogesh J. Gaikwad, Email: yogesh.gaikwad@mitwpu.edu.in

One of the active research issues in image processing is super resolution, which is used to boost picture resolution. The super resolution of a single image is obtained by rebuilding high-resolution (HR) pictures from low-resolution (LR) damaged photos (RSISR). This research examines publicly accessible datasets, RSISR assessment measures, and self-learning RSISR techniques. Comparisons are made in terms of reconstruction quality and computing efficiency utilising the self-learning RSISR technique and datasets.

Keywords: Self learning method, Real world image, Super resolution, Deep learning, Datasets.

2021. In Rahul Srivastava & Aditya Kumar Singh Pundir (eds.), *New Frontiers in Communication and Intelligent Systems*, 1–10. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.52458/978-81-95502-00-4-1

1 Introduction

Image super resolution is having limitations such as unknown downgrading, LR-HR images missing paired. Real world images do have problem of downgrading like blurring, additive noise and compression artefacts. Compression artefacts are nothing but distortion of media due to lossy compression application. In real-world image scenarios often performs poorly with models trained on datasets conducted manually. To overcome these limitations some work [2-5] has been proposed. Still there are some drawbacks in these studies, which will result in difficulty in training and over-perfect assumptions. In fact, it is propitious decision for specific domains, like intelligent surveillance, remote sensing, object tracking, scene rendering and medical imaging to apply SR.

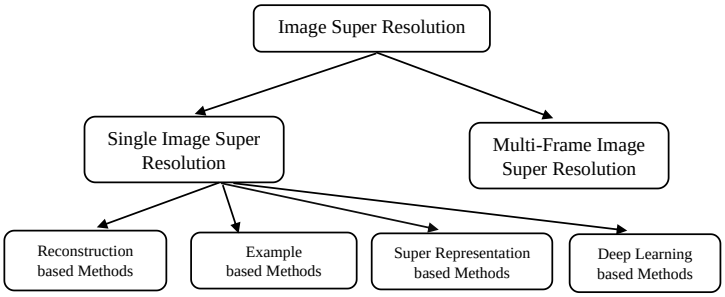

Fig. 1. Existing Super-Resolution Techniques

Images with higher resolution need upgraded hardware. With recent development in imaging devices and techniques we can achieve required high resolution images but with limitations. (i) The cost is quite high since the demand for applications is volatile. ii) We can receive fresh high-resolution photographs but not current low-resolution images in high-resolution. This is why super resolution is more flexible and inexpensive.

LR images are generated by downscaling RGB images manually with the help of various techniques such bicubic down sampling to train SR models.

12-bit or 14-bit raw images were captured by camera in reality but due to image signal processors of cameras produces 8-bit RGB images which losses lot of original signals and they have different features as compared to original images taken by the camera like demosaicing, denoising and compression. This is the main reason to use manually downscaled RGB images for SR. Some researcher's doing research to solve this problem. Chen et al. [6] discovered a relationship between imaging system field of view and picture resolution (R), using real-world dataset City100 offered data collecting methodologies, and got improved results in his proposed image synthesis model. Zhang et al. [7] created the real-world picture dataset SR RAW, which consists of paired RAW photographs and LR RGB images obtained using camera optical zoom to overcome the misalignment problem given by contextual bilateral loss. Xu et al. [8], on the other hand, are creating realistic training data by image processing simulation and developing dual CNN to use originally obtained radiance information in RAW photos.

2 Background

Reverse engineering is done to obtain reconstructed HR images from multiple LR images. This is achieved by original Hr image is down-sampled. To get LR images they are Warped, blur, and noise added. Super resolution will do exactly opposite up sample, de-blur, then it will add LR images to reconstruct HR images.

The following equation depicts the degradation of an LR picture (Y) from its matching HR image (X).

$$Y = DP(X, \theta_{DP}) \quad (1)$$

where DP () denotes the degradation process, which is described by the parameter set θ_{DP}.

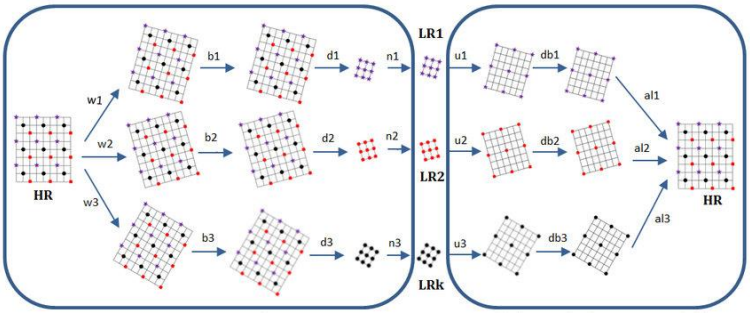

Fig. 2. (a) Generating LR images from HR (b) Basic premise for SR

We only have Y i.e., LR image and the degradation parameterθ_{DP} is not known,

SISR recovers desired HR image by inversing the degradation process done in Eq. (1), to get super resolved image from Y which is represented by $\hat{Y}$ which is an estimated real HR image Y as follows,

$$\hat{Y} = SR(Y, \theta_R) \quad (2)$$

where, SR () is the SR function defined by the parameter set θ_R,

Degradation process DP () and SR process SR () were inverses of each other's. SR (Y, θ_R) must be transformed to the degradation DP(X,θ_{DP})in order to achieve superior reconstruction performance.

Simulated degradation process mathematically obtained by following equation,

$$Y = SBX + n \quad (3)$$

Where B denotes the blurring operation and s denotes the down sampling procedure. In general, blurring is achieved by combining the HR picture with a Gaussian kernel. [1] Assume that n is white Gaussian noise.

Some studies use the basic degradation model with the "bicubic" kernel directly to downscale an HR picture to obtain an LR image. SR reconstruction performance on synthetic LR pictures is relatively excellent when learning-based SISR techniques such as RCAN [9], SAN [10], and RFANet [11] are used. Degradation process is complicated and unstable as it is influenced by various factors when it is compared with common degradation models in simulation. The difference between synthetic LR photos and genuine LR observations is enormous, resulting in a significant decline in the reconstruction performance of most available SISR algorithms using real-world images. To address this main limitation, academics have been working on RSISR for several years in a variety of approaches, including the provision of realistic datasets, SR performance assessment, and SR model development.

3 Datasets

This section provides a quick overview of publicly available datasets. Very few datasets consist of HR image along with LR images almost all datasets consist of only HR image to train and test models. To overcome these challenges more datasets for RSISR have been developed and they are listed below.

Table 1. DATASETS FOR RSISR

S. No	Datasets	Synthetic / Realistic	Scale factors
1	DIV2KRK[12]	Synthetic	*2, *4
2	Real SR [14]	Realistic	*2, *3, *4
3	DReal SR [15]	Realistic	*2, *3, *4
4	City100 [16]	Realistic	*2.9, *2.4
5	SR-RAW [17]	Realistic	*4, *8
6	TextZoom [18]	Realistic	*2
7	SupER [13]	Realistic	*2, *3, *4
8	ImagePairs [18]	Realistic	*2

3.1) DIV2KRK [12]: Bell-Kligler et al. [12] built this synthetic testing dataset for blind SR derived from DIV2K [61]. It consists of diverse images of 2Kresolution. From the validation set of DIV2K [61] 100 HR images were blurred; with V2KRK is more complex and random.

3.2) RealSR [14]: This real-world dataset was created by Cai et al. [14] for training and evaluating RSISR models. It contains 595 LR & HR picture pairings, which were created with two DSLR cameras. Cai et al. [14] offer a progressive image registration framework to accomplish pixel wise registration of photos taken at 28mm, 35mm, 50mm, and 105mm. After cropping lens distortion and interesting portions of rectified photographs in Photoshop, real-world LR-HR image pairings may be obtained.

3.3) DRealSR [15]: Wei et al. [15] built real world dataset DrealSR [15] which is having larger scale than RealSR [14]. To capture indoor and outdoor images 5 DSLR cameras were used with different resolutions; for alignment these images SIFT [57] algorithm is used. DRealSR [15] consists of 884(*2), 783(*3), 840(*4) image pairs of LR & HR.

3.4) City100 [16]: To characterize resolution of field of view FoV with the use of DSLR and smart phones Chen et al. [25] proposed City100 dataset which includes City100, NikonD5500 and iPhoneX. There is a counterbalance between the FoV and resolution for imaging system. If we zoom out the lens, we will get larger FoV but it is with low resolution. But if we zoom in the lens, we can increase the resolution of an image. This is the reason behind adjusting focal length or shooting distance by Chen et al. [25]. This lens length is 55mm, with 18mm reserved for HR-LR photography. For picture alignment, the SIFT [57] and RANSAC [58] algorithms were applied once more. It is done to enhance the accuracy of the intensity and colour correction.

3.5) SR-RAW [17]: Zhang et al. [17] suggested the SR-RAW dataset, which comprises of several levels of optical zoom RAW photos collected for the same scene at various resolutions by varying focal length. Seven photographs of each location were captured using a 24-240mm zoom lens. With 500 sequences, these seven image sequences were recorded in outdoor and indoor situations.

3.6) TextZoom [18]: The TextZoom dataset was created by Wang et al. [18] using RealSR [14], and SR-RAW [16] is the first real scene text SR dataset. The text images in this collection were cut from RealSR [14] and SR-RAW [16] photos, which included stores, automobiles, gardens, and building interiors. TextZoom [18] is divided into three difficulty levels: easy, medium, and severe. TextZoom [18] may be used to examine text image SR as well as text recognition.

3.7) SupER [13]: K¨ohler et al. [24] developed the SupER dataset by hardware binning. More than 80000 photographs were acquired from fourteen lab situations with four imaging resolutions and five compression levels using a Basler acA2000-50 gramme CMOS camera with f-1.8, 16mm fixed focus lenses. Three levels of resolution and binning factors were employed to create LR pictures matching to HR photos in order to achieve precise alignment between HR-LR images.

3.8) ImagePairs [18]: ImagePairs [18] was suggested by Joze et al. [18], which comprises 11421 LR-HR image pairs (LRHRIP) of various scenes acquired by a 5 mega pixel camera (LR) and a 20.1 mega pixel HR camera.

To capture same scene images simultaneously with two different cameras a beam splitter cube is used. But due to differences in focal length Joze et al. [26] proposed pixel based aligned LRHRIP with following 4 steps

i. ISP Process: In these process first images captured by LR-HR cameras were converted to colour images.
ii. Distortion: Using camera calibrations tangential and radial distortions were reduced.
iii. Alignment of LR-HR images are done globally and locally.
iv. Matching accuracy is improved of image pairs 10% of border is removed. As ImagePairs [18] includes raw images it should be used for ISP and other tasks.

4 Assessment Metrics of Image Quality

4.1) Peak Signal to Noise Ratio (PSNR): Different metrics are employed in image restoration for quality evaluation, however PSNR is the most often used statistic for Super resolution, denoising, deblocking, and deblurring.

$$\mathbf{PSNR = 10.\ Log_{10}(\frac{L2}{MSE})} \qquad \mathbf{(4)}$$

MSE is defined as follows,

$$\mathbf{MSE} = \frac{1}{HWC} \left\| \mathbf{Y} - \hat{\mathbf{Y}} \right\|_2^2$$

When employing 8-bit representations, L = 255 in most circumstances.

PSNR is most popular evaluation metric in pixel level MSE. It focuses only on differences between corresponding pixels instead of visual perceptions.

4.2) Structural Similarity Index (SSIM): For quantifying image quality conceptual metric is SSIM which measure the difference between two similar images conceptually.

4.3) Information Fidelity Criterion (IFC) [20]: Based on natural scene statistics the quality of images may be assessed by the information fidelity criterion (IFC) [20]. Characterization of natural images formed by statistics of the space can be done using models like Gaussian Scale Mixture is shown by researchers. Statistics of natural images will be disturbed by distortion and it will make unnatural images. Using natural sceneries and distortion models, quantify the mutual information between the test picture and the reference image to determine image visual quality. Overall, the IFC [20] does a good job of assessing the quality of super-resolved pictures [23].

4.4) LPIPS [21]: An image's quality is measured by the distance between two MVG models that fit natural pictures and assessed images. The learnt metric LPIPS [21] is used for referenced based picture quality evaluation. LPIPS is done by discriminating between the reference and test pictures in deep feature space, according to human judgments. To fit quality-aware features collected from photos, the MVG model is employed. Features such as the Generalised Gaussian Distribution (GGD) and Asymmetric Generalised Gaussian Distribution (AGGD) parameters are used to characterise the behaviour of image patches.

4.5) PIQE [22]: The perception-based quality evaluator (PIQE) is a no-reference metric for assessing picture quality [22]. By splitting the test picture into non-overlapping chunks, block level analysis is utilised to detect distortion and grade quality. The combined block level quality scores are used to assess the image's quality.

4.6) NRQM [19]: This is a no-reference quality metric (NRQM) that has been learnt for evaluating super-resolved pictures [19]. NRQM predicts super-resolved picture perception ratings, encompassing global frequency characteristics and spatial details.

5 Methods and Technologies

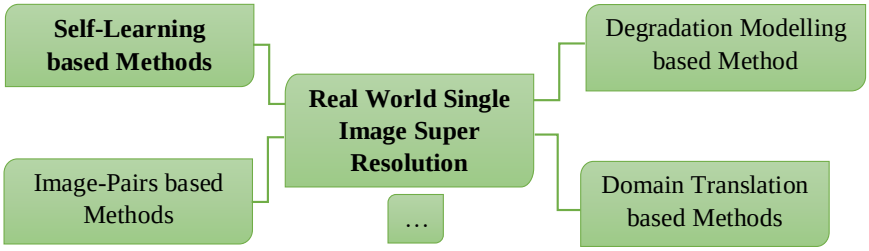

Fig. 3. Existing RSISR Methods

More focus is given to RSISR as the SR performance on synthetic data is giving better results. Fig 2. Shows existing RSISR techniques grouped into four categories based on their principles and characteristics as degradation modelling- based methods [27]-[37], image pairs-based methods [38]-[49], and self-learning-based methods [33], [50]-[56].

Self-Learning-based Methods

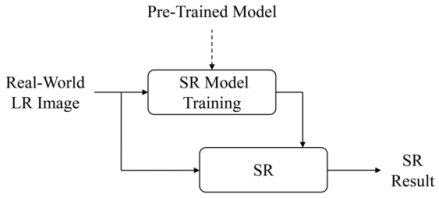

Fig. 4. Self-learning-based SR method

For training SR models existing RSISR methods use paired or unpaired training data which is external dataset. Viscosity between testing and training data results in tightly bound SR performance. The characteristics of real-world image training data are not always consistent. To reduce the impact of training testing variance on SR performance, information from the LR input is used to build an image-specific SR model, as shown in fig. 4.Shocher et al. [50] created the Zero-Shot SR (ZSSR), which is based on a common property of natural images, namely cross scale internal recurrence of information. During the testing phase, example pairs were extracted from the LR test image and its degraded images and used to train image-specific LR HR relations with an 8-layer CNN. Because the single test picture provides insufficient training data, data augmentation is used when extracting image-specific LRHR pairs. ZSSR [50] adapts itself to different testing images with unknown and unideal degradation processes in order to achieve excellent SR performance on real-world images. Bell-Kligler et al. [33] proposed using the cross-scale recurrency property to train an image-specific GAN (Kernel GAN) to model the degradation process (blur kernel) of the input. To achieve a fully self-supervised image-specific RSISR framework, the blur kernel estimation model KernelGAN [39] was plugged into the reconstruction model ZSSR [50].Kim et al. [51] created the DBPI, a unified internal learning-based SR framework that consists of an SR network and a downscaling network to train the image-specific degradation and SR networks simultaneously. The downscaling network is used in the self-supervised training phase of DBPI to

reconstruct the LR picture from its downscaled counterpart, while the SR network is optimised. A down scaling network is trained to recover the LR input image from its super resolved form. Similarly, Emad et al. [52] proposed DualSR [52], which optimises both the image-specific down sampler and the relative upsampler. The DualSR [52] is trained with three losses cycle-consistency, masked interpolation, and the adversarial loss using patches from the test image, resulting in [51], [52] complementary training of the image specific degradation and SR network that is beneficial to the reconstruction framework.

Self-learning RSISR techniques like ZSSR [50], KernelGAN [33], and DBPI [51] have two major drawbacks due to their self-supervised training strategy. I) Despite the fact that bigger scale external data is accessible, it is ignored since optimised SR models only employ internal data. II) These procedures were time intensive due to the online instruction. Meta-learning is included into self-learning-based SR techniques to overcome these limits. Based on ZSSR [50], Soh et al. [53] introduced the MZSR (meta-transfer learning for zero-shot SR), which consists of three steps: large-scale training, meta-transfer learning, and meta test.

1) On the big-scale dataset DIV2K [55], large scale training step one trains an 8-layer SR network with pixel wise l1 loss to make SR network and meta-learning training easier.
2) The goal of meta-learning is to establish a general starting point for internal learning by using Model-Agnostic Meat-Learning [56], which allows the model to easily adapt to changing visual situations with only a few gradient updates.
3) The input picture is degraded in the met-test phase to provide example pairs for updating model parameters, and then it is sent to the modified model to generate the SR result.

Meta-transfer learning for zero-shot super-resolution [53] outperforms meta-learning-based SR techniques in terms of super resolved picture quality and running time, reconstruction quality, generalisation capacity, and processing efficiency.

6 Challenges and Future Scope

As we seen in section 3 and 4 research on RSISR are positively done still there are some problems need further exploration. In this section we discuss some of the challenges and future work.

6.1) Image Datasets

Dataset is essential when it comes to self-learning equally as SR techniques for any research. In this field of research several datasets were designed but still it is required to develop more datasets focused on realistic image with more accuracy, images captured with different resolutions on same scene.

6.2) SR Algorithms

Still, it is not possible to apply RSISR algorithms to practical applications even though performance was increasing. As there are two major limitations of real-world images suffers from degradation problem therefore it is necessary to adapt RSISR models with ever changing real-world images. Other major limitation is with resources required are very highly configured for large model which is time consuming and also requires more storage space. As a result, lightweight SR model design and implementation are required.

6.3) Evaluation Criteria

PSNR and SSIM are the two most widely utilised SR assessment measures because to the difficulty of accurately measuring visual quality of super-resolved pictures. Due to this these metrics were unfit for implementing in practical applications. For RSISR, which is a critical and urgent research subject, more precise evaluation criteria should be devised.

Preservation of smoothness in flat regions, textures should be enhanced in details, sharpening of edges all these points should be considered while evaluation. The challenge is to more accurately and simply measure visual quality.

7 Conclusion

Researchers are now focusing on real-world picture super resolution, but self-learning based image super resolution should be given greater attention. This paper looks at recent super resolution approaches (self-learning-based algorithms) for realistic pictures, datasets, and assessment metrics for RSISR model training and evaluation. Some challenges should be addressed immediately as discussed in previous section. This review will help to understand existing studies on RSISR, datasets and self-learning method with challenges.

References

[1] H. Chen et al., "Real-World Single Image Super-Resolution: A Brief Review" in *arXiv,* 2103.
[2] Y. Yuan et al., "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks," in *CVPRW*, 2018.
[3] Y. Bei et al., "New techniques for preserving global structure and denoising with low information loss in single-image super-resolution," in *CVPRW*, 2018.
[4] A. Bulat, J. Yang and G. Tzimiropoulos, "To learn image super resolution, use a gan to learn how to do image degradation first," in *ECCV*, 2018.
[5] K. Zhang, W. Zuo and L. Zhang, "Learning a single convolutional super-resolution network for multiple degradations," in *CVPR*, 2018.
[6] C. Chen et al., "Camera lens super-resolution," in *CVPR*, 2019.
[7] X. Zhang et al., "Zoom to learn, learn to zoom," in *CVPR*, 2019.
[8] X. Xu, Y. Ma, and W. Sun, "Towards real scene super-resolution with raw images," in CVPR, 2019.
[9] Y. Zhang et al., "Image super resolution using very deep residual channel attention networks," in *Proc. Euro. Conf. Comp. Vision*, 2018, pp. 286–301.
[10] T. Dai et al., "Second-order attention network for single image super-resolution," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2019, pp. 11 065–11 074.
[11] J. Liu et al., "Residual feature aggregation network for image super-resolution," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2020.
[12] S. Bell-Kligler, A. Shocher and M. Irani, "Blind super-resolution kernel estimation using an internal-gan," in *33rd Conf. Neural Inform. Proc. Systems*, 2019, pp. 284–293.
[13] T. K¨ohler et al., "Toward bridging the simulated-to-real gap: Benchmarking superresolution on real data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 11, pp. 2944–2959, 2020.
[14] P. Wei et al., "Component divide-and-conquer for real-world image superresolution," in *Eur. Con. Comp. Vision*, 2020.
[15] C. Chen et al., "Camera lens super resolution," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2019, pp. 1652–1660.
[16] X. Zhang et al., "Zoom to learn, learn to zoom," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.,* 2019, pp. 3762–3770.
[17] W. Wang et al., "Scene text image super-resolution in the wild," in *Eur. Con. Comp. Vision*,, 2020.
[18] H. R. VaeziJoze et al., "ImagePairs: Realistic super resolution dataset via beam splitter camera rig," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops*, 2020, pp. 518–519.
[19]C. Ma et al., "Learning a no-reference quality metric for single-image super-resolution," *Comp. Vision Image Under.*, vol. 158, pp. 1–16, 2017.
20] H. R. Sheikh et al., "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Tran. Image Proc.*, vol. 14, no. 12, pp. 2117– 2128, 2005.
[21] R. Zhang et al., "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2018, pp. 586–595.
[22] N. Venkatanath et al., "Blind image quality evaluation using perception based features," in 21st *Nat. Conf. Commun.*, 2015, pp. 1–6.

[23] C. Y. Yang, C. Ma and M. H. Yang, "Single-image super-resolution: A benchmark," in *Eur. Con. Comp. Vision*, 2014, pp. 372–386.
[24] T. K¨ohler et al., "Toward bridging the simulated-to-real gap: Benchmarking super resolution on real data," *IEEE Tran. Pattern Anal. Mach. Intell.*, vol. 42, no. 11, pp. 2944–2959, 2020.
[25] C. Chen et al., "Camera lens super resolution," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2019, pp. 1652–1660.
[26] H. R. VaeziJoze et al., "ImagePairs: Realistic super resolution dataset via beam splitter camera rig," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.* Workshops (CVPRW), 2020, pp. 518–519.
[27] W. Z. Shao and M. Elad, "Simple, accurate, and robust nonparametric blind super-resolution," in *Int. Conf. Image and Graphics*, 2015, pp. 333–348.
[28] W. Z. Shao et al., "Nonparametric blind super-resolution using adaptive heavy-tailed priors," *J. Math. Imaging and Vision*, vol. 61, no. 6, pp. 885–917, 2019.
[29] J. Gu et al., "Blind super-resolution with iterative kernel correction," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2019, pp. 1604– 1613.
[30] V. Cornillere et al., "Blind image super-resolution with spatially variant degradations," *ACM Tran. Graphics*, vol. 38, no. 6, pp. 1–13, 2019.
[31] Y. Huang et al., "Unfolding the alternating optimization for blind super resolution," in *34th Conf. Neural Inform. Proc. Systems*), 2020.
[32] T. Michaeli and M. Irani, "Nonparametric blind super-resolution," in *Proc. IEEE Int. Conf. Compu. Vision*, 2013, pp. 945–952.
[33] S. B. Kligler, A. Shocher and M. Irani, "Blind super-resolution kernel estimation using an internal-gan," in *33rd Conf. Neural Inform. Proc. Systems*, 2019, pp. 284–293.
[34] A. Bulat, J. Yang and G. Tzimiropoulos, "To learn image superresolution, use a gan to learn how to do image degradation first," in *Proc. Eur. Con. Comp. Vision*, 2018, pp. 185–200.
[35] R. Zhou and S. Susstrunk, "Kernel modeling super-resolution on real low-resolution images," in *Proc. IEEE Int. Conf. Compu. Vision*, 2019, pp. 2433–2443.
[36] J. Xiao, H. Yong and L. Zhang, "Degradation model learning for real-world single image super-resolution," in *Proc. Asian Conf. Comp. Vision*, 2020.
[37] X. Ji, Y. Cao et al., "Real-world super resolution via kernel estimation and noise injection," *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops*, 2020, pp. 466–467.
[38] Y. Yuan et al., "Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops*, 2018, pp. 701–710.
[39] Y. Zhang et al., "Multiple cyclein- cycle generative adversarial networks for unsupervised image super resolution," *IEEE Tran. Image Proc.*, vol. 29, pp. 1101–1112, 2020.
[40] G. Kim et al., "Unsupervised real-world super resolution with cycle generative adversarial network and domain discriminator," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops*, 2020, pp. 456–457.
[41] S. Maeda, "Unpaired image super-resolution using pseudo supervision," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2020, pp. 291–300.
[42] K. Prajapati et al., "Unsupervised single image super-resolution network (USISResNet) for real-world data using generative adversarial network," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops*, 2020, pp. 464–465.
[43] T. Zhao et al., "Unsupervised degradation learning for single image super-resolution," *arXiv preprint arXiv*:1812.04240, 2018.
[44] C. You et al., "CT super-resolution gan constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE)," *IEEE Tran. Medical Imaging*, vol. 39, no. 1, pp. 188–203, 2020.
[45] M. Fritsche, S. Gu and R. Timofte, "Frequency separation for real-world super-resolution," in *IEEE Int. Conf. Comp. Vision Workshop*, 2019, pp. 3599–3608.
[46] R. M. Umer, G. L. Foresti and C. Micheloni, "Deep generative adversarial residual convolutional networks for real-world super-resolution," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops*, 2020, pp. 438–439.
[47] M. S. Rad et al., "Benefiting from bicubically down-sampled images for learning real-world image super-resolution," in *Proc. IEEE Winter Conf. Appl. Comp. Vision*, 2021, pp. 1590–1599.
[48] A. Lugmayr, M. Danelljan and R. Timofte, "Unsupervised learning for real-world super-resolution," in *IEEE Int. Conf. Comp. Vision Workshop*, 2019, pp. 3408–3416.
[49] S. Chen et al., "Unsupervised image super-resolution with an indirect supervised path," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops,* 2020,pp. 468–469.
[50] A. Shocher, N. Cohen and M. Irani, ""Zero-Shot" super-resolution using deep internal learning," in in *Proc. IEEE Conf. Comp. Vision and Pattern Recog.*, 2018, pp. 3118– 3126.

[51] J. Kim, C. Jung and C. Kim, "Dual back-projection-based internal learning for blind super-resolution," *IEEE Signal Proc. Lett*., vol. 27, pp. 1190–1194, 2020.
[52] M. Emad et al., "DualSR: Zero-shot dual learning for real-world super-resolution," in *Proc. IEEE Winter Conf. Appl. Comp. Vision*, 2021, pp. 1630–1639.
[53] J. W. Soh, S. Cho and N. I. Cho, "Meta-transfer learning for zero-shot super-resolution," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog*., 2020, pp. 3516–3525.
[54] S. Park et al., "Fast adaptation to super-resolution networks via meta-learning," in *Euro. Conf. Comp. Vision*, 2020.
[55] E. Agustsson and R. Timofte, "NTIRE 2017 challenge on single image super-resolution: Dataset and study," in *Proc. IEEE Conf. Comp. Vision and Pattern Recog. Workshops,* 2017, pp. 126–135.
[56] C. Finn, P. Abbeel and S. Levine, "Model-agnostic meta-learningfor fast adaptation of deep networks," in *Int. Conf. Machine Learning*, 2017.
[57] D. G. Lowe, "Distinctive image features from scale-invariantkeypoints," *Int. J. Comp. Vision*, vol. 60, no. 2, pp. 91–110, 2004.
[58] M. A. Fischler and R. C. Bolles, "Random sample consensus: aparadigm for model fitting with applications to image analysis andautomated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.