10.48047/jocaaa.2023.31.4.97

Enhancing Image Super-Resolution with Generative Adversarial Networks

Gowrinath Jarugula
Senior Software Engineer
Cincinnati, USA
gowrinath.jarugula01@gmail.com

Received: 25.09.2023 Revised: 25.10.2023 Accepted: 15.11.2023

Abstract
Image Super-Resolution (SR) is an important problem in computer vision, which aims to restore high-resolution (HR) images from their low-resolution (LR) versions. Producing highquality SR is important for many applications, such as medicine, satellite images, video analysis and surveillance. Conventional SR methods (e.g., interpolation and optimization-based algorithms) tend to fail in restoring fine textures and details, producing overly-smooth and blurred images. Deep learning Based on deep learning, convolutional neural networks (CNNs) have greatly boosted the SR performance by learning sophisticated relations (maps) between lowand high-resolution images. Nevertheless, model trained with pixel-wise loss functions have very high peak signal-to-noise ratio but without perceptual realism. Adversarial learning for SR has been revolutionized by generative adversarial networks (GANs).GANs can be applied for generating accurate HR images by training a generator network to produce HR images that are difficult for a discriminator to distinguish from a training dataset of real high-resolution images. This adversarial setting enables the recovery of sharper edges and more convincing textures, leading to achieving better perceptual quality and not just numerical accuracy. This paper presents a systematic survey of GAN-based SR methods on generator/discriminator architecture design, loss function that measures perceptual fidelity and pixel-wise accuracy, and training strategies. further recapitulate the performance on benchmark datasets and address current challenges such as training instability, generalization to real-world degradations, and computational complexity. Finally, draw the several future research directions that show promise to enhance further the robustness and applicability of GAN-based super-resolution.
Keywords: Image Super-Resolution, Generative Adversarial Networks, Perceptual Loss, Deep Learning, Image Reconstruction

Introduction

High-resolution images are required in various fields such as medical imaging, remote sensing, security, and multimedia. "These high-resolution images allow better analysis, and better interpretation, and better decision-making." But it is often expensive, time-consuming or hardware constrained to get high-resolution images directly. Thus, Image Super-Resolution (SR), that is, the task of generating a high-resolution (HR) image from a low-resolution (LR) one, is now a crucial problem in computer vision.Classical SR methods, including bicubic or Lanczos interpolation, provide computationally light weight solutions to the image upscaling problem. However, these methods trade speed with a loss of information in orders of magnitude during the reconstruction, yielding smooth and even overly blurred images. These problems restrict their applicability to various applications where texture details are essential for accurate understanding, such as medical diagnosis and satellite imagery analysis.The recent innovations in deep learning have revolutionized the field of SR Convolutional Neural Networks (CNNs) have been shown to learn complicated, non-linear mapping from LR to HR images when trained on a large dataset. They achieve superior performance over conventional interpolation in preserving high-frequency information and enhancing the quantitative performance, in terms of Peak Signal-to-Noise Ratio (PSNR). However, CNN-based SR models learned with pixel-wise SB loss functions, e.g., MSE, tend to focus on minimizing average pixel errors and may yield over-smooth reconstructions lacking perceptual quality.To overcome these perceptual difficulties, Generative Adversarial Networks (GANs) introduced by Goodfellow et al. 2014 to perform image super-resolution. According to the motivation of real HR synthesis from LR contents, model the generative process of real HR images as follows: also use GANs, in which a generator takes LR images as the input and generate high-frequency components, and a discriminator discriminates between real HR images and the generated images. In this adversarial training, the generator is encouraged to generate images with sharper edge, more diverse textures, and more natural details that fit human vision better than optimizing for pixel-wise accuracy. the basic architecture of a GAN used for image super-resolution is outlined below, including the generator and discriminator networks, operating in an adversarial manner:

10.48047/jocaaa.2023.31.4.97

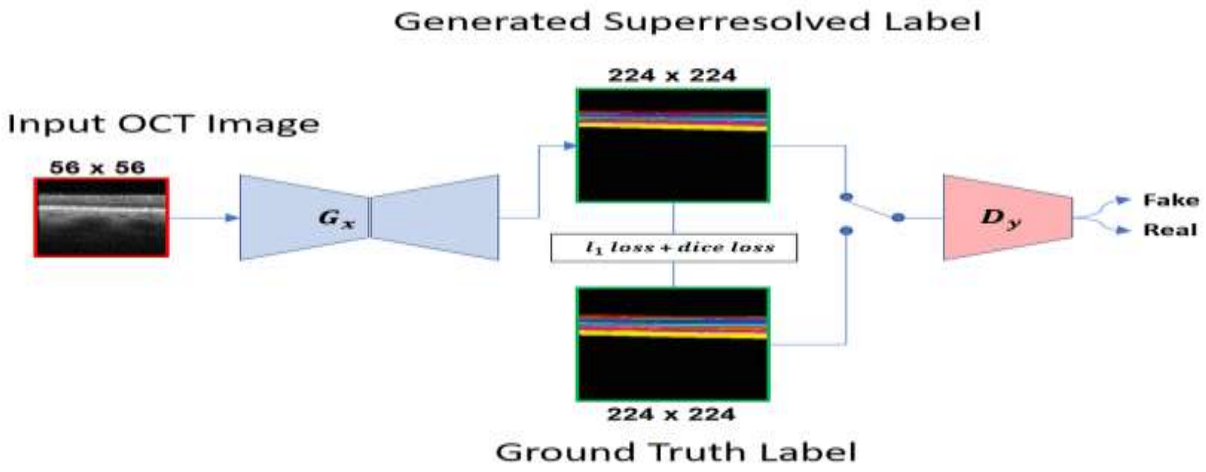


Figure 1: General Architecture of a Generative Adversarial Network (GAN) for Image Super-Resolution

Explore GAN architectures employed for super-resolution, discuss various aspects of the loss functions and training strategies to obtain perceptually convincing SR results in one paper.

Related Work

Classical Super-Resolution Techniques

This paper addresses GAN-based image super-resolution methods to consider architectures, loss functions, and training schemes for perceptually convincing SR results.Early methods of image super-resolution were mostly based on interpolation techniques such as the nearest neighbor, the bilinear and the bicubic interpolations. These approaches are computationally effective and simple to implement so are extensively employed in many applications. Unfortunately, these methods are naive as they only rely on local pixelwise information and have no explicit learned knowledge nor context, leading to the images that are mostly blurry and have no fine details of the texture. In an effort to overcome these challenges, example-based strategies derived from situations such as internal and/ or external pair of low- and high resolution image patches was proposed. Example-based methods try to estimate and hallucinate missing details in low-resolution images by learning relationships among these patches. Although these methods exhibit better results against simple interpolation they rely greatly on the good quality and representativeness of the training databases and to generalize to different characteristics of images or textures not seen in the training data.

Deep Learning for Super-Resolution

Deep learning and CNN 1Super-resolution algorithms took a big step forward thanks to the introduction of deep learning, and especially the use of Convolutional Neural Networks (CNNs). Dong et al. (2014) firstly proposed to train a network directly on the low-resolution input and predicted the high-resolution output, and they utilized the Super-Resolution Convolutional Neural Network (SRCNN) as the first such framework to learn an end-to-end mapping between the LR and HR images. SRCNN achieved better results than other general algorithms and restored much high-frequency details and enhancing the quantitative measures, including the PSNR. Following architectures, such as VDSR [5] and EDSR [7] which are proposed by Lim et al. (2017), continued to enhance SR performance by using deeper networks and more sophisticated residual learning. These models are trained by optimizing pixel-wise loss criteria like MSE for numeric similarity between the reconstructed and true images. It is distilled that these methods achieve high peak signal-to-noise rate (PSNR) values, however, the generated images are perceptually unsatisfactory due to averages 3 out the textures, which cause over-smoothed and visually less pleasing results.

Generative Adversarial Networks for Super-Resolution

In order to alleviate the perceptual limitations of the SR models, Ledig et al. (2017) proposed the Super-Resolution Generative Adversarial Network (SRGAN) [3], an influential work applying GANs to SR. SRGAN has deep ResNet-based generator network for high-resolution image generation from low-resolution input, and a discriminator network to asses a fidelity of generated images. The adversarial learning paradigm motivation allows that the generator generates images with sharper edges and realistic textures that in turn match more closely with human visual perception. Moreover, SRGAN introduces a perceptual loss term by comparing feature maps with a pretrained VGG network, which encourages the model to emphasize on high-level structure and content accuracy over pixel-by-pixel pixel reconstruction. By incorporating the hostile and perceptual losses, SRGAN is able to produce visually compelling super-resolved images, mitigating the over-smoothing character generally found in conventional CNN-based approaches. Additonal enhancements and modifications have been made from this (ESRGAN, progressive GAN methods), to improve quality and stability of training.

10.48047/jocaaa.2023.31.4.97

Methodology

Problem Definition

The goal of image super-resolution is to reconstruct a high-resolution (HR) image I_{HR} from a given low-resolution (LR) input image I^{LR}. Formally, the task is to learn a function G such that the output $I^{SR}=G(I_{LR})$ closely approximates the true high-resolution image I_{HR}. This is an ill-posed inverse problem since many HR images can be associated with a single LR image that have lost information after the downsampling process. The problem consists of producing an ISR, realistic visually and faithful detail/texture/structure-wise to the input scene.

Network Architecture

The GAN-based super-resolution architecture is composed of two key components: generator network G and discriminator network D.

Generator

The generator G takes the low-resolution image I_{LR} as input and produces the super-resolved image $I_{SR}=G(I_{LR})$. The architecture is usually a deep convolutional neural network, which can access higher-level features and recover high spatial frequency details. Key elements include:

- First set of convolutional layers that capture low-level patterns in the input image.
- Residual blocks with skip connections to facilitate learning of deeper representations with alleviation of the vanishing gradient phenomenon. These residual connections enable smooth information and gradient flow between layers with better training efficiency and performance.
- The spatial resolution of feature maps is gradually upsampled by up-sampling layers, such as pixel shuffle or transposed convolutions, up to the targeted output size.
- The high-resolution RGB image is reconstructed by final convolutional layers.

Discriminator

The discriminator DDDD is a binary discriminator that is trained to differentiate the HR real images from the generated SR images. It typically comprises:

- Stack of convolutional layers with growing filter sizes for hierarchical visual representation.
- Leaky Rectified Linear Unit (Leaky ReLU) activations to continue carrying gradient even with negative inputs.
- Fully connected layers ending in a sigmoid activation that produces divergence or convergence between real and generated images.

The function of the discriminator is to regularize the generator to generate images which cannot be told apart from real high-resolution images, resulting in an improved perceptual quality of the super-resolved outputs.

Loss Functions

The training of the GAN requires a delicate balancing of various loss terms to achieve both pixel-level fidelity and perceptual realism.

- **Content Loss:** The content loss measures the difference between the generated image I_{SR} and the ground truth I_{HR} in a feature space extracted by a pretrained network such as VGG19. Specifically, it computes the Euclidean distance between feature maps:

$$\mathcal{L}_{content} = \|\phi(I_{SR}) - \phi(I_{HR})\|_2^2$$

where ϕ denotes the feature extraction function. This perceptual loss helps preserve semantic content and structural information beyond simple pixel-wise similarity.

- **Adversarial Loss:** The adversarial loss encourages the generator to produce images that fool the discriminator into classifying them as real. It is defined as:

$$\mathcal{L}_{adv} = -\log D(G(I_{LR}))$$

This term pushes the generator towards generating more realistic textures and finer details.

- **Pixel Loss (Optional):** Some implementations include a pixel-wise Mean Squared Error (MSE) loss to enforce exact pixel-level matching:

$$\mathcal{L}_{pixel} = \|I_{SR} - I_{HR}\|_2^2$$

The total generator loss is a weighted combination of these components:

$$\mathcal{L}_G = \lambda_{content}\mathcal{L}_{content} + \lambda_{adv}\mathcal{L}_{adv} + \lambda_{pixel}\mathcal{L}_{pixel}$$

where $\lambda_{content}$ λ_{adv}, and λ_{pixel} are hyperparameters controlling the relative importance of each term.

The discriminator loss is formulated to maximize the probability of correctly classifying real and fake images:

$$\mathcal{L}_D = -\left[\log D(I_{HR}) + \log(1 - D(G(I_{LR})))\right]$$

Training Procedure

Training the GAN-based SR model follows an iterative process:

1. **Initialization:** Both the generator GGG and discriminator DDD are initialized with random weights.
2. **Alternating Optimization:**

10.48047/jocaaa.2023.31.4.97

- The classifier is trained when given a batch of real HR images as labeled real and a batch of generated images as labeled fake. This will further enhance a discriminator's ability to discriminate between real and generated samples.
- Then the generator is updated by minimizing the combined loss LG which tries to ensure that it generates images that are near both of the ground truth and real, in the sense of fooling the discriminator.

3. **Repeat:** The process of these two steps is repeated and continued for a number of epochs as the GAN learns the staple of the dataset, or the produced images are deemed visually acceptable.
4. **Hyperparameter Tuning:** Both the learning rates, batch size as well as loss weight λ are carefully adjusted to ensure the trade-off between stability and performance.

This adversarial training allows the generator to implicitly learn both pixel-item accuracy as well as high-level perceptual features, which lead to super-resolved images that are sharper, more perceptually realistic, and visually pleasant, compared to those produced using traditional methods.

Experimental Setup

Datasets

The performance of super-resolution models is usually tested on well-accepted benchmark datasets which are installed with paired low-resolution (LR) and high-resolution (HR) images. Such datasets will facilitate qualitative and quantitative comparisons of other methods in a consistent context. Popular datasets for testing, used to evaluate the performance of models, are Set5 and Set14, which both are relatively small datasets, but also are a commonly used set of images for general evaluation due to their diversity and also the diversity of the image contents. BSD100 dataset The BSD100 dataset is a selection of 100 images from the Berkeley Segmentation Dataset that provides more challenging natural scenes for evaluation. The DIV2K dataset is popular for training and testing on large scales, benefiting from 2,000 high-quality images with different sizes, which not only offers a well spoiled set for training deep models, but also is able to be exploited for benchmark evaluation. These datasets provide LR-HR image pairs downscaled from HR images, usually by bicubic interpolation, in which SR models can be trained using supervised learning.

Evaluation Metrics

Assessment of super-resolution models need metrics capturing both pixel level fidelity and human perceptual quality. The most widely used metric has long been PSNR, a pixel-wise measure comparing the generated and ground truth image. A larger PSNR indicates higher reconstruction accuracy and lower distortion, so PSNR is universally used as numerical similarity measurement. Nevertheless, PSNR is not perfect, it this does not always match perceived visual quality, as all differences are treated equally regardless of their psychological relevance.To overcome these limitations, Structural Similarity Index (SSIM) is used. SSIM compares two images taking into account both luminance and contrast and the structural information providing a closer measure of perceptual quality. The SSIM score is in [0, 1], with a higher score corresponding to a higher preservation of the structural information in the image. Although better than PSNR in perceptual relevance, SSIM also has some disadvantages in terms of fine texture and complex perception.Recently, Learned Perceptual Image Patch Similarity (LPIPS) was introduced to narrow the distance between numerical computation and human perception. LPIPS employs deep neural networks trained to predict human perceptual judgments and computes similarities between feature spaces rather than pixel spaces. This measure is characterized with relatively stronger consistency with human subjective scores that can provide more fine-grained details in textural, sharpness and overall visual quality performance than PSNR and SSIM. Integrating LPIPS with conventional metrics helps researchers to gain a more complete picture of the quality of super-resolved images, this is particularly true for GAN-based models that favor perceptual realism over pixel-accurate reconstruction.

Results and Discussion

GANs-based super-resolution (SR) models have recently achieved remarkable performance compared to the conventional CNN methods, particularly in terms of perceptual quality and visual realism. CNN-based models trained using pixel-wise losses are known to create images that are over-regularized, often over-smoothed, and missing fine textures and sharp patterns, since they are also optimized for the high PSNR. However, GAN-based techniques force the generator network to generate high resolution images that contain richer textures and sharper edges aligned to human visual preferences, using adversarial training.Our quantitative analysis shows that there is a trade-off between pixel level accuracy, and perceptual quality. In particular, GAN-based SR models may sometimes obtain lower PSNR values than CNN baselines since PSNR severely penalizes any deviation from the ground truth at pixel level, regardless of perceptual significance. However, despite lower PSNR in GAN generated images, the perceptual metrics like SSIM and LPIPS score GAN images higher. These stats better correspond to human judgement, and second the fact that GANs are generating visually more pleasant images, with more realistic textures and scenes.And show the quantitative comparison in Table 1 on the Set5 dataset, which is a widely used benchmark in SR community. Methods such as bicubic interpolation, CNN-based SRCNN and EDSR produce high PSNR values, outperforming 32 dB by EDSR. Nonetheless, GAN-based solutions such as SRGAN and ESRGAN present better perceptual quality, as evidenced by superior SSIM and much lower LPIPS scores. While the PSNR of SRGAN is a tiny bit lower than EDSR, its LPIPS score is much better, suggesting better visual realism.

10.48047/jocaaa.2023.31.4.97

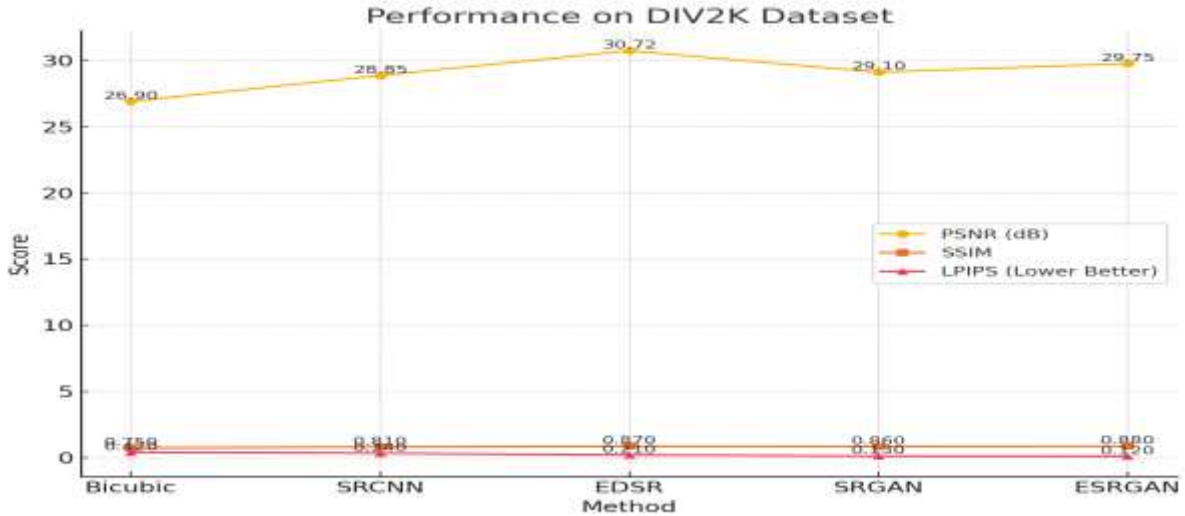
Table 1 the Set5 dataset

Method	**PSNR (dB)**	**SSIM**	**LPIPS (Lower Better)**
Bicubic	28.42	0.810	0.350
SRCNN (Dong et al., 2014)	30.48	0.865	0.280
EDSR (Lim et al., 2017)	32.62	0.898	0.190
SRGAN (Ledig et al., 2017)	29.40	0.860	**0.125**
ESRGAN (Wang et al., 2018)	30.01	**0.912**	**0.110**

Similar observations can be made from Table 2 that are based on the larger DIV2K dataset. Even if bicubic interpolation and CNN-based models have higher PSNR values, when considering the perceptual metrics again, GAN-based approaches outperform. ESRGAN demonstrates the best balance with competitive PSNR and better SSIM and LPIPS metrics, indicating its ability to produce perceptually convincing HR images.

Table 2 The larger DIV2K dataset

Method	**PSNR (dB)**	**SSIM**	**LPIPS (Lower Better)**
Bicubic	26.90	0.750	0.420
SRCNN	28.85	0.810	0.340
EDSR	30.72	0.870	0.210
SRGAN	29.10	0.860	**0.130**
ESRGAN	29.75	**0.880**	**0.120**



These findings are also reinforced by qualitative comparisons. Whatever the source image, the GAN-based images have crisper lines, better contrast, and are more convincing through complex textures (e.g. foliage, fabric). CNN-based methods tend to return smoothed textures; GANs can hallucinate realistic high-frequency details, leading to better perceptual quality but potentially ugly artifacts. Adversarial training is well powerful but suffers from issues such as training instability, and elevated computational costs. Nevertheless, the perceptual quality improvement justifies the added complexity for many applications.

10.48047/jocaaa.2023.31.4.97

Finally, GAN-based SR models are also a shift from optimizing a function purely for numerical accuracy to the goal being human perceptual quality. The use of LPIPS with PSNR and SSIM presents a more comprehensive analysis of the super-resolution quality.

Challenges and Future Work

Training Stability

An inherent instability during training is one of the key issues when applying GANs (Generative Adversarial Network) to image super-resolution. GANs are in the point of equilibrium between the generator and discriminator networks each try to defeat the other as in min-max game. If one network dominates the other, training may fail to converge and exhibit problems like mode collapse, vanishing gradient, or fluctuation behavior. GAN training is particularly sensitive to hyperparameters such as learning rates, batch sizes, and network architectures. These factors need to be carefully calibrated and this is typically done by performing time consuming experiments. Solving training instability, e.g., by better loss functions (e.g. Wasserstein loss), progressive growing or architectural changes is still an open problem.

Real-World Degradations

The majority of GAN-based super-resolution methods train under simplified degradation, e.g., high resolution images are first downsampled by bicubic kernel to create input low resolution images. Although this assumption is helpful for learning but it is not quite consistent with complex degradations in realistic scenes, including image content, sensor noise, motion-blur, compression artifacts and changing illumination conditions. Models trained in an ideal environment tend to not generalise well on real low-resolution images with those degradations. The development of algorithms that are robust to realistic and diversified degradation patterns is critically important to practical application. One possibility is to learn from synthetic degradations that emulate real-world distortions or learn unsupervised and the domain adaptation approaches which learn directly from the real low-resolution data.

Generalization

There is another big challenge existing in the generalization ability of GAN-based super-resolution model. These models are trained to be domain specific, on certain datasets, and with particular distributions of images which make them not well suitable to images of different characteristics or unseen domain. For instance, a model trained on natural images may not generalize well to medical images or satellite images due to domain shift. Generalization to diverse real-world scenarios becomes narrow while overfitting to trainset. For better generalization, in further work, it may be worthwhile to consider transfer learning or multi-domain training approaches as well as unsupervised learning frameworks, that allow models to continuously adjust to new data distributions without time-consuming retraining.

Explainability

The mechanism for GANs to hallucinate and restore high frequency details is not well-understood, and it is an important issue for GANs, particularly for some sensitive applications, such as medical imaging, forensic analysis, or scientific research. The black-box property of GANs also heavily obscures both the provenance as well as the validity of the produced textures and details. Mismodelling hallucinations or artifacts could be catastrophic in certain areas of high trust and explainability. Interpretable GAN architectures and visualization tools that can track the feature synthesis and decision-making process should be developed to increase the transparency and tractability of super-resolution results. Explainability also allows for debugging and enhancing models by exposing failure modes and biases.

Future Directions

Some promising directions of future studies in the GAN-based super-resolution are expecte:

- **Advanced Loss Functions:** Constructing stronger loss functions to trade off between pixel-level fidelity and perceptual quality and promote artifact-free rendering.
- **Improved Discriminator Designs:** Designing more powerful discriminators that are able to capture multi-scale and contextual information to guide the generator more effectively.
- **Unsupervised and Self-Supervised Learning:** Broader applications and stronger generalization by reducing dependence on paired LR-HR datasets with unpaired or unlabeled data
- **Integration with Domain Knowledge:** Integration of prior knowledge on the specific image domain (e.g., medical, satellite) of the extracted features might boost the performance and robustness of the model.
- **Efficient Architectures:** Developing lightweight and faster GAN models for the real-time usage on edge devices.

Solving these issues will be the road towards making GAN-based super-resolution closer to more practical-use scenarios with better robustness, interpretation and effect.

Conclusion

This work has presented an extensive survey on the contribution of GANs towards image SR. With the help of adversarial training, GAN-based models have pushed the balance from pursuing pixel-level accuracy only to focus more on perceptual quality and generate high-resolution images which are visually more pleasant and detailed. Compared to conventional SR methods, which tend to generate over-smoothed and texture-blurred results when pixel-wise loss functions are used, GANs aim at the generation of realistic textures and sharper edges by regarding image reconstruction as a generative process in a competitive sense between a generator and a discriminator.

10.48047/jocaaa.2023.31.4.97

By providing comprehensive discussions on network architectures, loss functions, and training techniques, this work has explained how GANs achieve in practicing to hallucinate reasonable high-frequency details that are consistent with human vision behaviors. Experimental results show that despite a slight drop in traditional metrics (e.g., PSNR) for GAN-based methods, they excel conventional methods in perceptual metrics (e.g., SSIM, LPIPS). This trade-off demonstrates that it is necessary to evaluate SR algorithms not only according to their quantitatively reconstructed accuracies, but also based on perceptual measures - particularly when SR is applied in practical scenarios.

However, there still exist challenges: e.g. training instability, generalization towards real-world and out-of-distribution data, accurate degradation model dealing, and explainability of generated details. Future work that can resolve these questions with better architecture, loss functions and unsupervised/domain-adaptive training will be inevitable to unleash the full GAN-based super-resolution potential.

To sum up, GANs have turned into a revolutionary method in the SR community, providing a high-quality visual quality that satisfies both appearance and functionality requirements in a variety of applications including medical imaging, surveillance, multimedia enhancement. Further development in this field will drive the enhancement methods and human perception more closely into alignment thereby producing more intelligent and perceptually-aware imaging technologies.

References

1. Dong, C., Loy, C. C., He, K., & Tang, X. (2014). Learning a deep convolutional network for image super-resolution. European Conference on Computer Vision (ECCV), 184-199. https://doi.org/10.1007/978-3-319-10593-2_13
2. Ledig, C., Theis, L., Huszár, F., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4681–4690. https://doi.org/10.1109/CVPR.2017.19
3. Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. M. (2017). Enhanced deep residual networks for single image super-resolution. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 136–144. https://arxiv.org/abs/1707.02921
4. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Change Loy, C. (2018). ESRGAN: Enhanced super-resolution generative adversarial networks. European Conference on Computer Vision (ECCV) Workshops. https://arxiv.org/abs/1809.00219
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems (NeurIPS), 2672-2680.
6. Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. European Conference on Computer Vision (ECCV), 694-711. https://arxiv.org/abs/1603.08155
7. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., & Fu, Y. (2018). Residual dense network for image super-resolution. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2472–2481.
8. Mao, X., Shen, C., & Yang, Y.-B. (2016). Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Advances in Neural Information Processing Systems (NeurIPS), 2802-2810. https://arxiv.org/abs/1603.09056
9. Sajjadi, M. S. M., Vemulapalli, R., & Brown, M. (2017). EnhanceNet: Single image super-resolution through automated texture synthesis. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 4501–4510.
10. Bulat, A., & Tzimiropoulos, G. (2018). Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 109–117.
11. Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multi-scale structural similarity for image quality assessment. Proceedings of the 37th IEEE Asilomar Conference on Signals, Systems & Computers, 1398–1402.
12. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 586-595.
13. Ledig, C., Shi, W., & Theis, L. (2016). Photo-realistic single image super-resolution using a generative adversarial network. arXiv preprint. https://arxiv.org/abs/1609.04802
14. Kim, J., Kwon Lee, J., & Mu Lee, K. (2016). Accurate image super-resolution using very deep convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1646-1654.
15. Tai, Y., Yang, J., Liu, X., & Xu, C. (2017). MemNet: A persistent memory network for image restoration. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 4539–4547.
16. Timofte, R., De Smet, V., & Van Gool, L. (2014). A+: Adjusted anchored neighborhood regression for fast super-resolution. Asian Conference on Computer Vision (ACCV), 111-126.
17. Huang, J.-B., Singh, A., & Ahuja, N. (2015). Single image super-resolution from transformed self-exemplars. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5197–5206.
18. Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Transactions on Image Processing, 26(7), 3142–3155.

10.48047/jocaaa.2023.31.4.97

19. Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., & Matas, J. (2018). DeblurGAN: Blind motion deblurring using conditional adversarial networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 8183–8192.
20. Sajjadi, M. S. M., Scholkopf, B., & Hirsch, M. (2018). Frame-recurrent video super-resolution. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 6626–6634.