Middle East Research Journal of Engineering and Technology
ISSN: 2789-7737 (Print) 2958-2059 (Online)
Frequency: Bi-Monthly
DOI: 10.36348/merjet.2023.v03i01.001

KUWAIT SCHOLARS PUBLISHER
Website: http://www.kspublisher.com/
Email: office.kspublisher@gmail.com

A Review Super Resolution Using Generative Adversarial Network-Applications and Challenges

Arun Agarwal[1*], Swatishee Chhotaray[1], Niraj Kumar Roul[1], Saurabh N. Mehta[2]

[1]Department of ECE, ITER, Siksha O Anusandhan Deemed to be University, Bhubaneswar, Odisha, India

[2]Department of Electronics and Telecommunication Engineering, Vidyalankar Institute of Technology, Mumbai- 400037, Maharashtra, India

Abstract: Artificial Neural Networks (ANNs), Deep Learning, and AI Deep learning is a subset of machine learning. In order to create pictures with a greater resolution, a high-resolution GAN combines a deep network with an opponent network. An approach to generative modelling that uses deep learning techniques like convolutional neural networks is known as generative adversarial networks, or GAN. In the Super-Resolution Generative Adversarial Network (SRGAN), a Generative Adversarial Network (GAN) may transform low-resolution images into super resolution images that are more finely detailed and of higher quality. In the past, CNNs were used to produce incredibly precise and detailed images. However, they could have trouble recalling specifics and usually draw hazy visuals which can be overcome by SRGAN.

Keywords: SRGAN, GAN, CelebA, Keras, ML, DL.

Review Paper

***Corresponding Author:**
Arun Agarwal
Department of ECE, ITER, Siksha O Anusandhan Deemed to be University, Bhubaneswar, Odisha, India.

How to cite this paper:
Arun Agarwal *et al.;* "A Review Super Resolution Using Generative Adversarial Network-Applications and Challenges". Middle East Res J. Eng. Technol, 2023 Jan-Feb 3(1): 1-6.

Article History:
| Submit: 11.11.2022 |
| Accepted: 25.12.2022 |
| Published: 12.01.2023 |

1. INTRODUCTION

GAN was first introduced in 2014 by Ian Good fellow and his colleagues. The basic idea was two constant models competition, meaning one's loss is another's gain. This approach leads them to produce realistic, undifferentiable protected entry. For example, if we entered a video, the output would also be in the form of a video.

GAN Models are trained to identify patterns or similarities inputs and can create items that are very close related to input. GAN has proven itself in a variety of difficult tasks such as resolution enhancement, face generation expressions and much more.

High-Definition films and high-resolution images are currently some of the most in-demand requirements for individuals to relax and unwind. The more enjoyable a picture or video is, the better its quality should be. Moreover, the audience's general watching experience is notable. Every updated visual innovation created in the present the globe strives to satisfy the demands of high-quality audio and video performance. With these photographs and videos' continuously improving quality whereas both the supply and demand for these goods are rapidly rise. However, it's probable that you won't always be able to reach or produce the finest quality pictures or videos given the limits of technology confronted within the proper process.

1. Technology restrictions or any other factors that result in these limits can be solved using super-resolution generative adversarial networks (SRGANs). With these as a guide. We can convert a large portion of the low-resolution photos or video information you can obtain using huge GAN structures into high-resolution image.

2. Related Work

One of the first methods that enable the model to reach an upscaling factor of approximately 4x for the majority of picture visualizations is the notion of SRGANs. Estimating and creating a high-resolution picture from creating low- resolution images is a really difficult process. CNN's previous was used to create high-resolution visuals that accelerate learning and accomplish superior accuracy.

However, there are times when they are unable to retrieving more minute details and frequently produce hazy photos. The suggested Most of these problems are overcome by the SRGAN architecture, which produces high-quality, cutting-edge pictures.

The majority of supervised methods for super-resolution use the high-resolution picture that uses the mean squared error loss is gained, and that image's underlying reality. This approach shows to presents a loss that uses the recently invented loss known as perceptual loss to battle increasingly perceptually focused characteristics.

A new kind of content loss called VGG Loss [Fig 1] was added in the Perceptual Losses for Style Transfer in Real-Time and Super-Resolution superior style and resolution transfer structure Both types of perceptual loss are present. Content loss as well as antagonistic loss. The potential formulation of this loss is interpreted in accordance with the next interpretation.

$$l^{SR} = \underbrace{\underbrace{l^{SR}_{X}}_{\text{content loss}} + \underbrace{10^{-3} l^{SR}_{Gen}}_{\text{adversarial loss}}}_{\text{perceptual loss (for VGG based content losses)}}$$

Fig. 1

This loss is favoured over the mean-squared error loss since we are not concerned with comparing the pictures pixel-by-pixel. The enhancement in image quality is what we are most worried about. Hence, by we can do more with the SRGAN model by employing this loss function favorable outcomes.

2.1 GAN

Deep learning is being employed in the field of artificial intelligence, and studies have found it to be effective. Although it can't make things, deep learning is capable of recognizing them. Ian Good fellow came up with the notion, "What if you pitted two neural networks against each other?" The most popular machine learning algorithm is GAN [Fig. 2], which was developed over the course of multiple tries [1]. GANs were referred to as "the trendiest innovation in Machine Learning in the last 20 years" in the seminar "Unsupervised learning: The Next Frontier in AI" held in September 2016.

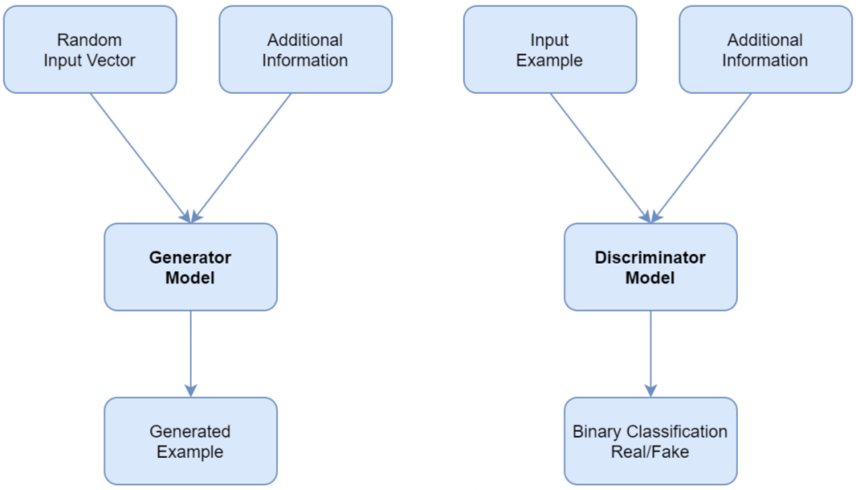

Fig. 2

2.2 SRGAN

For single picture super-resolution, SRGAN is a generative adversarial network. It employs an adversarial loss and a content loss component of a perceptual loss function. Using a discriminator network trained to distinguish between the original photo-realistic pictures and the super-resolved images, the adversarial loss drives the solution to the natural image manifold. A content loss motivated by perceptual similarity rather than similarity in pixel space is also used by the authors. Remaining blocks for feature extraction make up the majority of the real image.

2.3 Generative Modeling

In machine learning, generative modeling is a type of uncorroborated learning that lacks labels. Sample creation and density estimation are both capabilities of generative models. While implicit density prohibits assessment of this model, explicit density does. Different uses include image-to-image translation, single-image super resolution, and artistic creation. In order to manipulate large dimensional probability distributions, incorporate them into reinforcement learning, and provide missing data, generative models should be researched Finally, GANs allows machine learning to use multimodal outputs.

(i) Framework

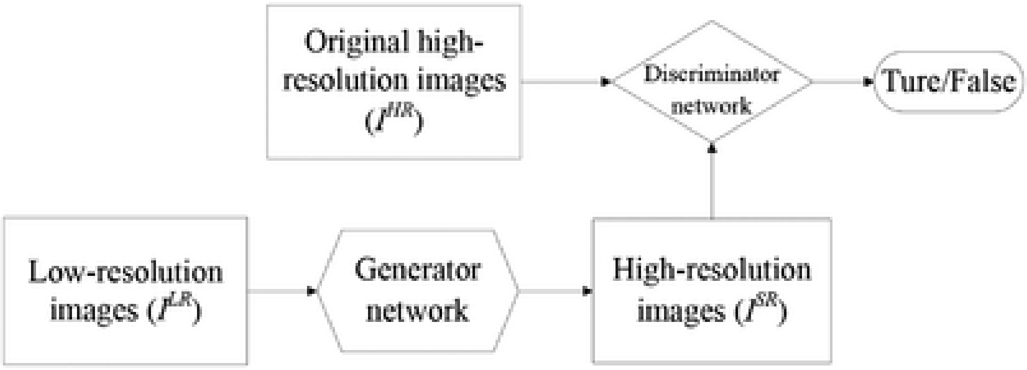

Fig. 3

(ii) Training

On the other hand, supervised training settings were used to try to regulate the output more precisely. By presenting a picture to the trainee during the (as opposed to random noise z) network as an input, and one-to-one mapping between the input and output image. This region is the focus of our efforts. The one is SRGAN [2].

One of the well-known one-to-one mapping efforts. Network situation and attempts to produce an output with particular characteristics In particular, closely related to our work, Image-to-image translation using cGANs has been shown by Isola *et al.,* [3]. As such, SRGAN is a member of the Conditional GANs is a more general word.

Authors: Liang Gong and Yimin Zhou (IEEE 2019).

This document emphasizes provenance GAN, its recent applications, extension variants, existing problems and some other applications. Mathematical preparation was also explained in detail. GAN classification was one of the main ones main points of the article Application of GAN in an image and video as image generation, image by image translation, image super-resolution, video generation, video frame prediction, etc. It also has listed advantages and disadvantages. Key Findings: Classification and structure of GANs which is the main thing we need before we start with implementation.

Authors: Chaoyue Wang, Chang Xu, Xin Yao, Dacheng Tao (IEEE 2018)

There are some cases where GAN does not serve its purpose. Sometimes because of the poor training model or sometimes due to instability data. This particular article suggests how GANs the model can be improved. The authors took a completely different path than that traditional method and introduced a generator and discriminator. The qualities of each of them the generated material is evaluated and only preserved generators are used for further use operation. Key Takeaways: The end result of this EGAN (Evolutionary GAN) approach was the foundation the main applications we focused on. EGAN plays an integral role in the super resolution of the image.

Authors: Han Wang, Wei Wu and Yang Su, Yongsheng Duan and Pengze Wang (IEEE 2019)

SRGAN is used to increase resolution image. This is the most proven method create high-resolution photorealistic images. However, this paper focuses on how SRGAN works and how the model can be modified to make it better Result. An encoder block is fixed in the generator model to get much more clarity images. It also extracts essential features image and can synthesize it better resolution. The encoder block uses a simple encoder network to extract important information as it then connects to the existing CNN. The CNN parameters are reduced for simplicity distortion. Key information: This method shows that super images can be obtained with better resolution clarity and higher resolution with the proposed parallel encoder connection model.

GPU

A specialized electrical circuit known as the GPU is used in conjunction with a CPU to create 2D and 3D graphics. The term "GPU" or "graphics card" is occasionally used in the gaming world. The usage of GPUs to accelerate computer workloads is growing, especially in industries like financial modeling, cutting-edge research, deep learning, analytics, oil and gas exploration, etc. GPUs are capable of processing tens of thousands of operations each cycle.

TPU

Tensor Processing Unit abbreviation The Tensor Processing Unit (TPU) is an integrated circuit created specifically for Google's open-source machine learning framework TensorFlow. TPUs have been powering Google's data centers since 2015, however Google still uses CPUs and GPUs for other types of machine learning. TPU can operate at a cycle rate of up to 128000. A co-processor designed with TensorFlow (a programming framework) to accelerate deep learning tasks; because general purpose programming on TP requires a lot of work, general purpose programming on TPU has not yet seen the creation of any practical compilers.

INFOGAN

Instead of employing a single unstructured noise vector, INFOGAN splits the input noise into two components known as the latent code and in compressible noise, which targets the significant structured semantic aspects of the real data distribution that has incompressible noise. For the creation of informal infogan, infogan was employed.

CycleGAN

The headway of unpaired data is essential, and cycle consistency sets an upper limit on conditional entropy. It is also possible to derive it using the framework of variational inference.

F-GAN

F-gan, which may be used to train gan, is based on f divergence. The benefits of the divergence are in the complexity of training as well as the quality of the generative models.

Least Squares-GAN

It has been suggested that this network be used to address the vanishing gradient issue with conventional GANs. In lieu of the cross entropy loss, LSGANs do this using the least square loss function.

WGAN

In addition to offering valuable learning curves for debugging and hyper parameter searches, WGAN aids in increasing the stability of learning. It limits the k-lipschitz constraints and is a lipschitz constraint with the use of the graduated penalty.

Unrolled-GAN

Rounds is a method for stabilizing the gan by describing how the generator goal in relation to an unrolled optimization of the discriminator is altered between it, it is typically unstable and results in hoar solutions. Additionally, it is not practical in actual applications.

Spectra Normalized GANs (SN-GAN)

The discriminator's training is steady as a result. These work well with existing techniques and are computationally efficient. Comparatively speaking to the earlier training stabilization approaches, this method produces images of comparable or greater quality.

LAPGAN

This network creates high-quality pictures by using a cascade of CNN within a laplacian pyramid structure. Because of this, it creates images with greater resolution when compared to the original GAN.

DCGAN

G and D are described in this way as MLP in the original gan. In terms of visuals, CNN is superior to MLP. Therefore, deep convolutional reward networks in DCGANs, which produce higher outcomes, are used to calculate G and D.

Progressive GAN

The progression GAN is based on a progressive neural network, and its major goal is the progressive, steady evolution of both the generator and the discriminator from a low resolution, adding new layers that simulate more fine features as training goes.

Self-alternative Advesial Network (SAGAN)

The main goal of a GAN is the subtle, firm development of both the generator and the discriminator from a low resolution, adding new layers that model finer and finer details as training goes on. Progressive neural networks serve as the foundation for the progressive GAN architectures.

BigGAN

It can produce pictures with a resolution of up to 512 by 512 pixels, which is fairly high. You need to have enough data to be able to duplicate the BigGAN result. It also modifies discriminators and adds an encoder for representation learning.

Hybrids of Autoencoder and GAN

There are encoders and decoders in this network. The encoder's goal is to learn a representation for a setoff of data in order to reduce the dimensionality of the data. The decoder's goal is to produce a representation from the redacted encoding that is as similar to the input as feasible. Adversarial Generator-Encoder (AGE) Network is one of the examples.

Multi-GAN Learning

The GAN1 and GAN2 pair, which make up the multi-gan or linked gan, each generate pictures in a specific domain.

3. Implementation Method

The Google Colab platform (https://colab.research.google.com) was used to implement the project. The following may be done by a programmer using Google Colab.

- Python code writing and execution.

- Record the code you use to support mathematical equations.
- Create/Upload/Share notebooks.
- Upload and save notebooks to/from Google Drive.
- Publish or import notes from GitHub.
- Import outside datasets, such as those from Kaggle.
- Combine Keras, TensorFlow, PyTorch, and OpenCV.
- Free Cloud storage and GPU.

3.1 Design Evolution

We also looked at how, visually, SRGAN generator network reconstructions change as the number of training iterations increases. Figure 8 shows the visual outcomes after varying the amount of training iterations. It is intriguing that the generator significantly deviated from the SRResNet initialization after just 20 000 training rounds. This generates noise-filled reconstruction with a lot of high frequency content with an increase in training. The baboons from Set14 reconstruction iterations resemble the reference picture more closely. However, there isn't much to see aesthetically change during the last 50–100k update rounds [8].

4. Performance Evaluation

The majority of supervised methods for super-resolution use the high-resolution picture that uses the mean squared error loss is gained, and that image's underlying reality. This approach shows to present a loss that uses the recently invented loss known as perceptual loss to battle increasingly perceptually focused characteristics. A new kind of content loss called VGG Loss was added in the Perceptual Losses for Style Transfer in Real-Time and Super-Resolution superior style and resolution transfer structure Both types of perceptual loss are present content loss as well as adversarial loss. The potential formulation of this loss is interpreted in accordance with the next interpretation.

$$l^{SR} = l_x^{SR}(content\ loss) + 10^{-3} l_{Gen}^{SR}(adversarial\ loss)$$

Equation-1(Perceptual loss)

This loss is favoured over the mean-squared error loss since we are not concerned with comparing the pictures pixel-by-pixel. The enhancement in image quality is what we are most worried about.

4.1. Loss Function

Perpetual loss function (LSR), which is used by the SRGAN, is the weighted sum of two types of loss: content loss and adversarial loss. For the generator architecture's performance, this loss is crucial.

4.2. Content Loss

For the SRResnet architecture, which is the most popular, we apply pixelwise MSE loss and MSE loss for picture Super Resolution in this work. The high frequency material in the image, however, is not handled by MSE loss, which led to the production of extremely smooth images. As a result, the paper's authors opted to employ VGG layer loss. Based on the 19-layer pre-trained VGG network's ReLU activation layers, this VGG loss. According to this definition, this loss:

$$l_{VGG|i.j}^{SR} = \frac{1}{W_{ij}H_{ij}} \sum_{x=1}^{W_{ij}} \sum_{y=1}^{H_{ij}} (\emptyset_{ij}(I^{HR})_{x,y} - \emptyset_{ij}(G_{\emptyset_G}(I^{LR})_{x,y})$$

Equation-2(simple content loss)

$$l_{VGG|i.j}^{SR} = \frac{1}{W_{ij}H_{ij}} \sum_{x=1}^{W_{ij}} \sum_{y=1}^{H_{ij}} (\emptyset_{ij}(I^{HR})_{x,y} - \emptyset_{ij}(G_{\emptyset_G}(I^{LR})_{x,y})$$

Equation-3(VGG content loss)

4.3. Adversarial Loss

A discriminator that has been taught to distinguish between high resolution and super resolution pictures is used in the adversarial loss, which drives the generator to produce images that are more resemblant of high-resolution images.

5. Current Research Directions

Presently GAN are widely researched for medical imaging and ophthalmology. Further they are used for high frequency predictions by design of optimal training model.

6. Applications

SR methods can be used to transform LR video pictures into high-definition images. Numerous medical imaging techniques can deliver both functional and anatomical details on the make-up of the human body. Enhanced satellite photos, picture restoration and recovery of vintage images, automatic improvement of camera image quality by creating software based on learning models.

7. Challenges and Future Scope

A successful project manager is aware of the rapid drop in morale and productivity that may happen when a team is asked to do activities that are not within their purview. Typically, assuming that a project will be successfully finished within a specific timeline is unrealistic.

Different GANs that may be utilized to create commercial applications have been developed by scientists and researchers. The use of many technologies, including artificial intelligence, machine learning, computer games, and computer simulation. Using 3D GANs, you can create forms, photorealistic pictures, paintings into photographs, translate images, create maps, and more.

8. CONCLUSION

SRGANs are a better way for processing ultra-high-resolution images and videos when compared to other techniques like bicubic, CNN, and deep neural networks. These data-driven networks may be utilized in a wide range of sectors to improve the technology that is now available and develop new technologies that are based on it. Advancements in computer vision, machine learning, deep learning, and AI.

REFERENCES

1. https://www.semanticscholar.org/paper/Nonlinear-Multi-scale-Super-resolution-Using-Deep-Tran-Panahi/8385720cf6c8562b09946a61f48c032e94e994f0
2. https://blog.ampedsoftware.com/2022/04/27/does-deep-learning-based-super-resolution-help-humans-with-face-recognition/
3. https://medium.com/what-is-artificial-intelligence/this-ai-makes-blurry-faces-look-60-times-sharper-7fcd3b820910
4. https://www.geeksforgeeks.org/super-resolution-gan-srgan/
5. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690).
6. Johnson, J., Alahi, A., & Fei-Fei, L. (2016, October). Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision* (pp. 694-711). Springer, Cham.
7. Loss functions https://www.geeksforgeeks.org/super-resolution-gan-srgan/
8. Yang, C. Y., Ma, C., & Yang, M. H. (2014, September). Single-image super-resolution: A benchmark. In *European conference on computer vision* (pp. 372-386). Springer, Cham.
9. Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. In International Conference on Learning Representations (ICLR), 6.
10. Emily, D., Soumith, C., Arthur, S., & Rob, F. (2015). Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks. *Advances in Neural Information Processing Systems (NIPS)* pp.1486–1494. https://arxiv.org/ftp/arxiv/papers/1903/1903.09922.pdf