To Generate High resolution Deep Space Images using Real-ESRGAN

Varad Joshi
dept. of computer engineering
VESIT
Mumbai, India
2019varad.joshi@ves.ac.in

Rushabh Rijhwani
dept. of computer engineering
VESIT
Mumbai, India
2019rushabh.rijhwani@ves.ac.in

Mr. Richard Joseph
dept. of computer engineering
VESIT
Mumbai, India
richard.joseph@ves.ac.in

Abstract.

As the space technology industry is burgeoning in the field of deep space exploration, researchers are compelled to test their psyche in analysis of the celestial bodies wrapping our universe. Eventually giving rise to machinations of new and sophisticated algorithms and systems that help to distil and reveal more about deep space that is yet abstruse to many of us. The current devices in this industry include some of high tech telescope and observatories are capable of taking incredible pictures of the deep space, which although are fortuitous in helping the space scientists to study the deep space entities, but the obscurity of those images impede the elucidation of the outer bound space entities;rendering the scientists unable to summarise much from the input. Thus, to overcome this cumbersome task of scratching the heads to get information from already low resolution image, we are aiming to build a model that enhances the low resolution of such space objects to high resolution data that will ease the task of space scientists in scrutinising the deep space bodies, eventually expediting the research process. The model we will be using to aid our problem is, Super Resolution General Adversarial Network(SRGAN)

Index Terms—High Resolution, Low Resolution, Deep Space data, Super Resolution General Adversarial Network, GAN.

I. INTRODUCTION

General Adversarial Networks or GAN in short: are becoming more cohesive which in turn has flourished its popularity in the research sphere as compared to the obsolete Convolutional Neural Networks. These GAN models provide more accurate results and also their cost function is near peripheral after being trained for sufficient epochs. The prodigious GAN architectures can help in upscaling of the low-resolution images and video content into the high-resolution entities.SRGAN is able to Upsample a muffled input image to recover a high resolution output of the same image.The SRGAN has glossified that extensive mean-opinion-score (MOS) test has had huge remarkable gains in regards to perceptual quality. SRGAN's MOS scores are more comparable to the primordial high-resolution pictures' scores than any other cutting-edge approach. Traditional super resolution networks and existing algorithms even though were less byzantine, they were not able to display pellucid images that would have helped significantly in deep space research if they were cogent enough. Hence, to decode these barriers, we decided to go with SRGAN model, that would eventually augment the image quality and obviate the possibility of studying the low resolution, downsampled images.

.II. OVERVIEW

A. Problem Statement

The practice of boosting an image's resolution from low-resolution (LR) to high-resolution is known as "image super resolution" (HR).It mostly alludes to using various methods to convert downsampled photos into upsampled, high definition images. With the aid of its generative adversarial nature, SRGAN exploits perceptual function loss to synthesize high resolution images from low resolution photos.The primary objective of the paper is focused on those far away, millions of light years distant inter galactic bodies, whose data isn't readily accessible in image format.Although some high-end telescope maybe able to capture these bodies aesthetically but are somewhat obtainable only in scant proportion and have distorted images.Thus, to speed up the task of research and bolstering the analysis phase in department of space science, Our team is driven to create a cutting-edge model that makes use of sophisticated real-time algorithms to provide us with the best results possible.Eventually decreasing the peril in the approach and giving results simultaneously.

B. Our Inspiration

The images procured from the below mentioned space devices are contorted.The scientific importance of super resolution is that it is used to enhance the warped images that are obtained from deep space observatories devices.The traditional algorithms and deep learning models are able to restore the enhanced version of distorted images to a certain extent,but the results are not persuasive enough. Hence, it becomes very perplexing to process the images for scientific purposes. Our model thus helps in overcoming these travails in the path of novelty and exploration.

C. Objective

1. To simplify the indigation procedure in the field of space science.

2. To popularise the use of more advanced and complex models in order to get results with high precision.

3. Recovery of enhanced images from convoluted input space images.

4. System that is affordable, convenient, and easy to use.

5. To galvanize and lionize the neophytes in the field of astronomy to come up with such fastidious and lucrative solutions.

III. THEORY

Super Resolution General Adversarial Network or also known as SR-GAN; is a deep learning model proposed by researchers at twitter.SR-GAN is generally used for SISR(single image super resolution). The primary objective of this model is to reclaim the quality textures of the ground truth data of the image that had been distorted due to upsampling of the input image or noise or any other factor. SR-GAN uses a deep network with an adversary network in order to give HR(high resolution) images as the output. SR-GAN uses a perpetual loss function which is made up of adversarial loss and content loss. The solution to the natural image is propelled by the adversarial loss function, multitudinous where a discriminator is used to differentiate between the original input image i.e. the ground truth data and the generated super resolution image from the generator.

A. Survey of the Existing systems

To augment the resolution of the deep space photos, we used a Super Resolution General Adversarial Network model. This model belongs to the GAN family and is one of the most well- known and popular ones.According to our research, SinGAN, SR-CNN, ESRGAN, deeply recursive convolutional network (DRCN), and efficient subpixel convolutional neural network were the most widely utilised models for this application (ESPCN). For the purpose of reconstructing the Compressed Sensing Magnetic Resonance Imaging, a new model based on conditional generative adversarial network (DAGAN) (CS-MRI) was presented. According to the literature survey, our new model provides better accuracy and efficiency than other convolutional neural networks like ESRGAN and few architectures like Enhanced Deep Super-Resolution network(EDSR) which supervises specific super-resolution scale and Multi-scale Deep Super-Resolution system(MDSR)that creates a single model from multiple scales of high-resolution images. SinGAN model as well as conditional GAN models showed accuracy less than 75 percent which is less as compared to our "Real ESRGAN" model. Some of the Convolutional neural networks that use single image dehazing process had higher perceptual loss function and take more processing as well training time than SRGAN model used in this system. While studying the other models, we found out their experimental results in usage and then compared it with SRGAN model. For reference purposes,we are providing picturesque results for better visualisation.

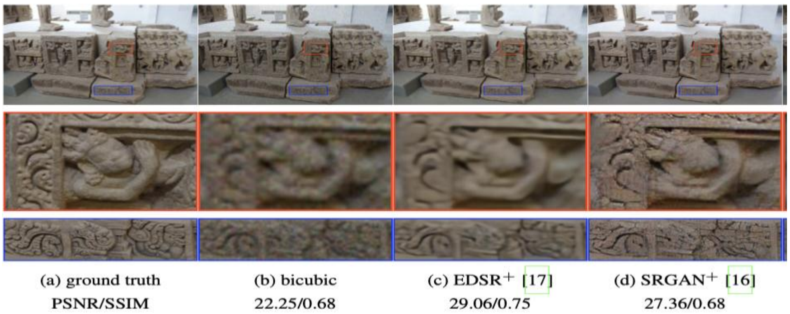

Fig.1.Result comparison between different image resolution models

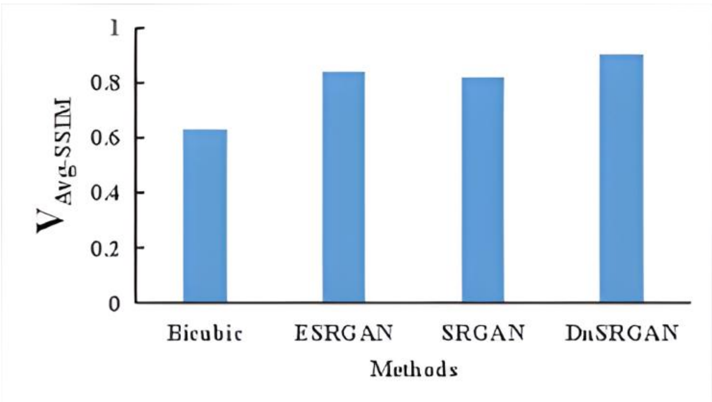

Fig. 2. Graphical comparison of different image resolution models.

B. Limitations and Research Gap

Some of the models like SRCNN, ESPCN, DRCN incorporate higher loss functions that can increase the complexity of the training phase. Additionally, it has been discovered that newer works use GAN or reinforcement learning before vision restoration. Blind SR can be categorised into two different groups, the first of which uses explicit degradation representations and typically has two parts: degradation prediction and conditional restoration. On the other hand, inaccurate degradation assessments will inevitably produce artefacts. Before training a unified network to overcome blind SR, another strategy is to collect/precipitate training pairs that are as near to real-world data as possible.

.IV. SR-GAN FOR SUPER RESOLUTION OF DEEP SPACE IMAGES

A. Architecture

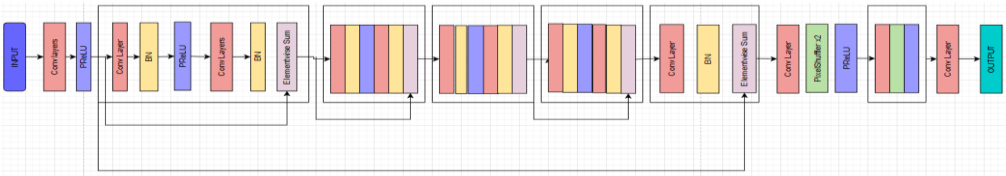

Fig. 3. SRGAN generator

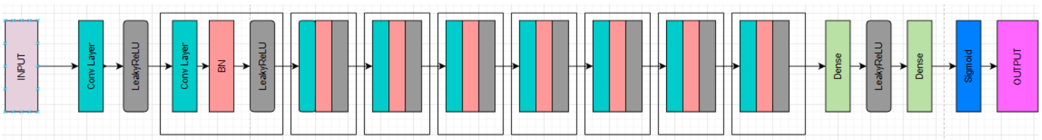

Fig. 4. SRGAN discriminator

The above figure represents the architecture of SR-GAN(super resolution general adversarial network).The model consists of two main blocks.

1. Generator.

The SR-GAN architecture is based on the ResNet architecture and consists of a generator, Batch Normalisation and Parametric ReLU. The main task of the discriminator is to check both the generated image of the generator and then try to distinguish between them. While training, a gaussian filter is applied to the high resolution image(IHR) which then gives us a low resolution image (DHR) of the same high resolution by applying a downsampling operation.

2. Discriminator.

The SR-GAN is based on the architecture of DC-GAN which is known as deep convolutional general adversarial networks. The primary task of the discriminator is to discriminate between the input images and those produced by the generator. After the 512 feature maps, two dense layers follow them along with a layer of activation function i.e. LeakyReLU.

•Input low resolution is interpreted as C x H x W

•Input high resolution image and input super resolution image is interpreted as C x rW x rH

1. The Generator function gives the HR(high resolution) image which is actually classified as a super resolved image of the corresponding low resolution image.

2.Discriminator D is trained to discern between genuine images and super resolved images

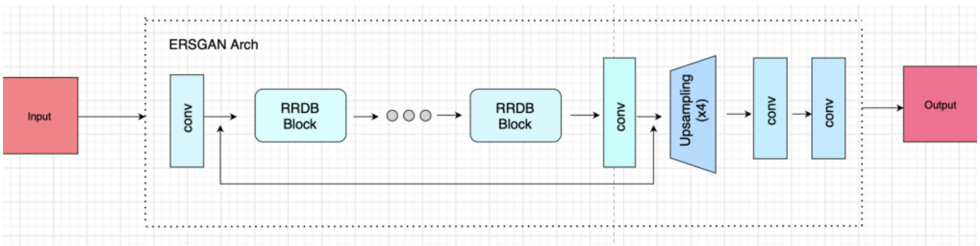

Fig. 5 Real-ESRGAN Architecture

The Real ESRGAN can do high computation on a mini batch of input images and of high degradation space than its predecessor. In order to solve this, we have used a u-net architecture for our discriminator. The U-Net may give the generator detailed per-pixel feedback and produce realness values for each pixel.

B. Loss Function of SR-GAN

The total loss i.e. the perpetual loss of SRGAN is the sum of content loss and the adversarial loss. Content Loss: Two types of losses are used in the SRGAN architecture. Image super resolution models commonly use pixel wise MSE loss of the SRResnet architecture. This loss alone is not able to deal with the high frequency content and hence, loss function of different VGG layers are used which are based on the ReLU activation function layer of VGG-19 model. The loss is hence defined as:

$$l_{MSE}^{SR} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I^{LR})_{x,y})^2$$

Fig. 6. Content Loss

$$l_{VGG/i.j}^{SR} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

Fig. 7. VGG content loss in SR-GAN

Adversarial Loss: The Adversarial loss is given as output from the discriminator and input into the generator for the generator to learn to generate pragmatic high resolution images which are more similar to the ground truth data. It is a loss function that lets the generator win the war by helping it to generate a high resolution image making it difficult for the discriminator to distinguish between the original image and the generated image.

$$l_{Gen}^{SR} = \sum_{n=1}^{N} - \log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

Fig. 8. Adversarial loss of SRGAN

The sum of adversarial loss of the generator and the content loss is equivalent to the perpetual loss

$$l^{SR} = \underbrace{l_{X}^{SR}}_{\text{content loss}} + \underbrace{10^{-3} l_{Gen}^{SR}}_{\text{adversarial loss}}$$

Fig.9 Perpetual Loss

C. Loss Function of Real-ESRGAN

A more effective perpetual loss has been used in Real ESRGAN which is different than that used in SRGAN, where the features are constrained before the activation takes place unlike after which we can see is the case in SRGAN. In the previous model i.e. the SRGAN model the perpetual loss was defined at the activation layer in order to minimise the distance between two features. But in this model, the features are used before the activation function which will be able to overcome the following drawbacks. The first drawback that this model addresses is that the activated features are very sparse even after a deep neural network and the Second drawback it addresses is that use of features after activation has resulted into incompatible brightness in the output or the super resolved image when it is compared to the ground truth images.The perpetual loss or the total loss of Real ESRGAN is given in the figure.10.:

$$L_G = L_{\text{percep}} + \lambda L_G^{Ra} + \eta L_1,$$

Fig. 10 Perpetual Loss in Real-ESRGAN where,

$$L_1 = \mathbb{E}_{x_i} ||G(x_i) - y||_1$$

Fig.11. represents the content loss of the Real ESRGAN.

D. Algorithm and Process Design.

In figure 12, the first order layer represents the classical image degradation model to synthesize various input images. The first layer consists of blur (model blur is used as a convolution with linear blur filters), Resize (downsampling) and Noise (we usually consider gaussian noise and poisson noise).

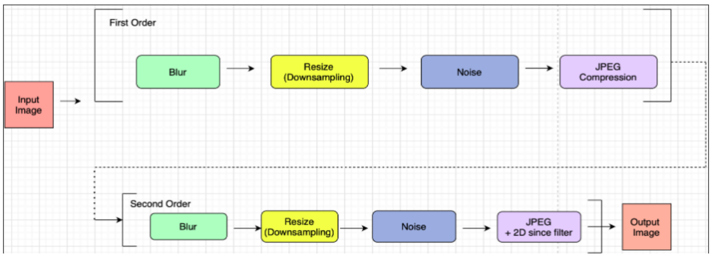

Fig. 12 Input Image Processing in Real-ESRGAN

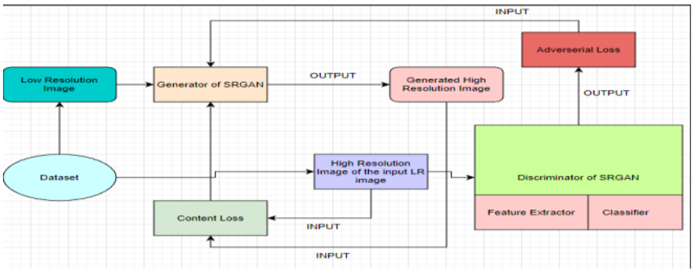

Fig. 13 Algorithm and Process Design

V. EXPERIMENTAL RESULTS

.A. Analysis

Fig .15. First image represents the low resolution image and the second image is the super resolved image i.e. the output image of the Real ESRGAN

Our final model exhaustively includes sampling of the distorted images and produces highly elucidated space body images using proper resolution techniques. The suggested methodology saves significant time and is practical and close to certain. After testing the model on several images of deep space, the output images have sharper and richer texture when compared to the lower resolution images and the peak signal-to-noise ratio has improved..In the formula given below, MAXi is basically the maximum possible value of the pixel of the image.The average PSNR or the peak signal-to-noise ratio is 29.791 and structural similarity index turn out to equal to 0.8215 Where PSNR and SSIM are given by the formula which is explained in the figure .22

$$
\begin{aligned}
PSNR &= 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \\
&= 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \\
&= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE).
\end{aligned}
$$

$$
\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}
$$

Fig. 22 PSNR and SSIM Formula

VI. CONCLUSION

We Successfully implemented the model for generating high resolution deep space images captured by the telescope and with the help of this ”Super Resolution General Adversarial Network” model, we can easily enhance the resolution of any images of deep space objects up-to 1024X1024. Since the model in use is a bit convoluted, it requires high-end machines to run it.Also it requires enormous computation power and takes more time to train and show expected results as compared to other models in question. With the enhancement in the structure of hidden layers of Real ESRGAN we can expect to get more quality assured results in case of real world problems. In comparison to the classical degradation detection process for images and models Real ESRGAN has improved a lot but changes like adding an optimizer, more compatible activation.

1. REFERENCES

[1] Daniel Glasner, Shai Bagon, and Michal Irani. Super-resolution from a single image. IICCV, 2009.

[2] Demiray, B.Z., Sit, M. Demir, I. D-SRGAN: DEM Super-Resolution with Generative Adversarial Networks. SN COMPUT. SCI. 2, 48 (2021).

[3] Lian, S., Zhou, H., Sun, Y. (2019). FG-SRGAN: A Feature-Guided Super-Resolution Generative Adversarial Network for Unpaired Image Super-Resolution. In: Lu, H., Tang, H., Wang, Z. (eds) Advances in Neural Networks – ISNN 2019. ISNN 2019. Lecture Notes in Computer Science(), vol 11554. Springer, Cham.

[4] Liu, J., Chen, F., Wang, X., Liao, H. (2019). An Edge Enhanced SRGAN for MRI Super Resolution in Slice-Selection Direction. In: , et al. Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy. MBIA MFCA 2019 2019. Lecture Notes in Computer Science(), vol 11846. Springer, Cham.

[5] J. Li, L. Wu, S. Wang, W. Wu, F. Song and G. Zheng, ”Super Resolution Image Reconstruction of Textile Based on SRGAN,” 2019 IEEE International Conference on Smart Internet of Things (SmartIoT),2019, pp. 436-439, doi: 10.1109/SmartIoT.2019.00078.

[6] B. Liu and J. Chen, ”A Super Resolution Algorithm Based on AttentionMechanism and SRGAN Network,” in IEEE Access, vol. 9, pp. 139138-139145, 2021,

[7] Zhong, Z., Chen, Y., Hou, S., Wang, B., Liu, Y., Geng, J., Fan,S., Wang, D., Zhang, X.: Super-resolution reconstruction method of infrared images of composite insulators with abnormal heating based on
improved SRGAN. IET Gener. Transm. Distrib. 16, 2063– 2073 (2022).
[8] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single
image super-resolution: Dataset and study. In CVPRW, 2017.
[9] Google drive link:Google drive link

Biographies

Varad Joshi is currently pursuing a bachelor's degree in computer engineering from Vivekanand Education Society’s Institute of Technology,Chembur-Mumbai, Affiliated to University of Mumbai. He is in Final Year of Engineering and has worked in the field of Machine Learning, AI and Software Development. He’s intended to graduate in the year 2023.

Rushabh Rijhwani is currently pursuing a bachelor's degree in computer engineering from Vivekanand Education Society’s Institute of Technology,Chembur-Mumbai, Affiliated to University of Mumbai. .