Image Super-Resolution Using GAN – A study

Darshan Parekh   
Department of Computer Science  
School of Engineering & Technology  
Sharda UniversityGreater Noida, India  
darshparekh03@gmail.com  
  
  
 Ankita Maiti   
Department of Computer Science  
School of Engineering & Technology  
Sharda UniversityGreater Noida, India  
maiti.ankita21@gmail.com  
  
  
Dr. Vishal Jain   
Department of Computer Science  
School of Engineering & Technology  
Sharda UniversityGreater Noida, India  
vishal.jain@sharda.ac.in

*Abstract*— Reconstructing low-resolution images to high-resolution images by building a neural network is quite challenging but can be used in many applications like medical imaging, public surveillance, or old photo recovery. Compared to previous methods, deep learning has a breakthrough in high-resolution accuracy and speed. By applying a deep network with Generative Adversarial Networks, this model aims to enhance low-resolution images to produce high-resolution images. The major focus is to reconstruct the image with high resolution by developing the image with low-resolution in order to preserve the main details in the reconstructed images.

Keywords— High-resolution, Low-resolution, Super-Resolution, GAN, Generative Adversarial Network

# Introduction

Generating high-resolution photos from low-resolution photos is a challenging task and has huge applications in the real world, including photo reconstruction, surveillance cameras, and computer-aided design. Recently, generative adversarial networks have gained extensive recognition [29] from the computer group because of their promising results in synthesis. Real-world images. Generative Adversarial Network (GAN) is a type of Artificial Intelligence algorithm used for unsupervised ML(Machine Learning). GAN is a deep NN (Neural Network) [20] architecture made of two networks, the first network being a generator and another being discriminator, which compete with each other (hence the name "adversarial"). GAN is about creation, such as drawing portraits or creating symphonies. The principal focal point of GAN is to produce information without any preparation.

In the Generative Adversarial Network (GAN) suggested by Goodfellow et al. [7] which was introductory in the field of Super-Resolution, the generative model is pitted to confront the opponent which is to fool the discriminator and train two models at the same time: the generator G that catches the distributed data and discriminator D evaluates that the sample comes from training dataset instead of generated image by the generator G. The process training of G is to improves the likelihood of D mistake. The proposed frame correlates with a very small and huge game with two-player.

Super-Resolution (SR) is the procedure of reconstructing one or more low-resolution images to high-resolution images of the same instance [26]. The difficult task of evaluating high-resolution (HR) images only from their corresponding low resolution (LR) is called Single Image Super-Resolution (SISR). For high magnification factors, the compositional nature of the undetermined super-resolution is particularly obvious. The texture details in the reconstructed SISR image are usually absent. The enhancement mission of Supervised Single Image Super-Resolution algorithms is usually (a) minimizing the MSE i.e. mean square error between the restored HR image and also maximizing the PSNR [28] i.e. peak signal-to-noise ratio or (b) Maximizing the SSIM i.e., Structural Similarity Index or both.

Many Super Resolution techniques use information from different images to create one upscaled image. The model tries to extract details from other images and reconstruct them in one single image. The most common technique is Mean Square Error (MSE) to find how far the image’s pixels are from the ground truth image’s pixels. For example, if we take two images A and B, we take the square [25] of the difference between every pixel in A and the corresponding pixel in B, combined and divided by the total number of pixels. MSE function is paired with the Peak signal-to-noise ratio which is normally used to determine the quality loss in an image, where this signal is ground truth image and noise is an error not recovered by the model.

The ultimate target of the Super-resolution algorithm is to minimize the (MSE) mean square error between Reconstructed Image and Real Image. The underdetermined SR problem is pronounced in Upscaling factors, due to which texture details in most of the images are absent. This minimized MSE also gives a high Peak Signal-to-noise Ratio (PSNR) which is a very famous method for evaluating Super Resolution models. The chances of MSE to generate detailed texture images is very limited and a high PSNR value doesn’t mean a better super-resolution image.

Inspired by the success of previous methods of Super Resolution, this work differs from previous methods in that we are using Deep Residual Network (ResNet) with skip connections and developing Super-Resolution Generative Adversarial network (SRGAN). For this network, we used perceptual loss using feature maps of the VGG network (trained on BSD200, and Urban100 dataset) which then combined with Discriminator Network stir solution hard to distinguish Super from High-Resolution Images.

# RELATED WORK

Author [1] has suggested the method called SSSR (self-supervised super-resolution) which is PET dependent on dual GAN that trains the combinedly to generates SR PET images from unpaired [19] PET inputs in a self-supervised manner. For supervised training, the images have been used from the BrainWeb database. The proposed network receives low- resolution PET. The research shows the result that, SSSR appears to be weaker in comparison with VDSR it is better than classic deblurring. The last is true in light of the fact that VDSR is fully supervised and based on paired training data sets.

Author [2] has suggested the model conditional GAN for developing high-resolution images from linguistic label maps. It generates for combining 2048×1024 images as a result with new adversarial loss. It also produces the architecture of a new multi-scale generator and discriminator. By incorporating information from object instance segmentation which empowers object control, for example, eliminating/adding objects and changing the classification of the object. Also, the suggests a way to generate diversification given similar info, permitting users to alter the display of the objects. The outcomes recommend that conditional GANs [24] can combine high-resolution images with no user-created loss or pre-trained network.

Author [3] uses a deep residual network (ResNet) to develop a Super-Resolution GAN (SRGAN). They created a discriminator network to differentiate frames of super-resolution from HR reference. The mentioned GAN urges the recreations to go towards areas of the desired investigation with a likely higher number of photo-realistic frames hence producing more natural images. The upscaling factor that makes it unique video SR [27] is estimated by PSNR deep ResNet (SRResNet). They developed a generator network that enhances new loss which is determined on the VGG network. They used a self-video dataset to train the SRGAN.

Author [4] has suggested the model, super-resolution generative adversarial network (SRGAN) with the help of deep residual network (ResNet), for generating an image of super-resolution using GAN. The proposed network can generate photo-realistic images by upscaling factor ×4. They developed a function to determine loss which comprises both the content and an adversarial loss. The target of the proposed SRGAN is to overlook connection and skip diverge from MSE. The dataset used to be experimented on are Set5, Set14, and BSD100, and the testing part was done on BSD300.

Author [5] develops a network that uses attribute-embedded upscaling which comprises two types of networks, one being upsampling network and the other one being a discriminative network. The purpose of the upsampling network is to skip connection with the help autoencoder. Meanwhile, the goal of the discriminator is to check if the image of faces after being super-resolved consists of the expected attributes, and after that, the loss is used by the upsampling network for updating. With the upscaling factor × 8, the unaligned images of faces can be super-resolved the extremely low resolution (16 × 16 pixels). To train the network, the dataset being used is CelebA (Celebrity Face Attributes). As a result, the network can not only super resolve face images with low resolution but also can change the attributes info of the image resulting in the up-sampling of the end image.

Author [6] has suggested the method Gradient Map Generative Adversarial Network (GMCAN) which generates images in relation to Human Vision System (HVS) to create a layout of a loss function by combining Image Quality Assessment (IQA). The type of GAN used in this method is Wasserstein GANs(WGAN-GP) which is an improved version to control the instability of the initial GAN. It consists of two parts, the first being network architecture and the second being the loss function. It also focuses majorly on training the network with different datasets which are Part of the ImageNet dataset, DF2K (DIVIK + Flickr2K) dataset, DF2K + OST (OutdoorSceneTraining) dataset, and DIVIK dataset. Also to test the network, experiments were performed on different datasets including Set5, Set14, BSD100, and Urban100. But GMCAN is unable to show adequate outcomes when handling strong repetitive shapes.

Author [7] focuses on training two models the evaluating generative models by an adversarial process. The generative model is the first model which captures the data and the discriminative model, the second model, whose purpose is to evaluate that the data was sent by the training data, not a generator. The process of training is for the generator is to magnify the chance of committing a mistake by the discriminative model. The generator can be considered as indistinguishable from a group of counterfeiters who seeks to make counterfeit currency and use it without identification, while the discriminator is equivalent to the police, trying to distinguish counterfeit currency. The range of datasets on which the adversarial network was trained includes, the CIFAR-10, Toronto Face Database (TFD), and MNIST dataset.

Author [8] has suggested an algorithm that accesses a dataset of the image to train to generate a high-resolution image when the image is being zoomed in. To create the training data, high-resolution images are used and breakdown every image in such a way that fixing the breakdown of the images can be done later in the procedure. Normally, the image is been blurred and subsample them to make a low-resolution image of ½ the number of existing pixels in every dimension. The algorithm performs its best to evaluate the low-resolution image of vocabulary or letters of a text i.e. when text is zoomed in high-resolution characteristic is formed.

Author [9] has suggested a network called coupled GAN (CoGAN) which focuses on learning about the joint distribution of multi-domain images without any tuple alike images. Learning of joint distribution that only draws samples from the marginal distribution can be completed by applying the weight sharing constraint that limits the network limit. The CoGAN network is based on the possibility that deep neural networks get familiar with a progressive element portrayal. By implementing the layers that decode significant level semantics in the GANs to share the weights, it powers the GANs to disentangle the high-level semantics similarly. The layers that decode low-level information then, at that point map the common portrayal to images in singular spaces for confusing the separate discriminative models. CoGAN includes a pair of GAN, each is in charge of incorporating images in a single domain. For training the network, four datasets were used including the MNIST dataset, Celeb Faces Attributes dataset, RGBD dataset, and NYU dataset.

Author [10] has suggested the model, conditional generative adversarial network (CGAN) states that methods have no control over the information being created in an unconditioned generative model. Though, by molding the model on extra data, it is feasible to coordinate the process of information generation. These types of conditioning can be found on class names, on some piece of information for inpainting, or even on information from various methodologies. If both the generative model and the discriminative model are adjusted based on some other information named y which can be any additional data (such as class labels, etc.), then GAN can be further extended to a conditional model. For training the unimodal, the datasets which are been used MNIST. For training multimodal, the datasets which are been used are Flickr, ImageNet dataset, YFCC100M2 dataset, and MIR Flickr 25000 dataset. The outcomes displayed in this paper are incredibly starter, yet they exhibit the capability of conditional adversarial nets and show a guarantee for intriguing and helpful applications.

Author [11]proposed an extremely precise super-resolution method with the help of a deep convolutional network which is been used for ImageNet Classification motivated by VGG-Net. Relatable data over an enormous image domain is taken advantage of productively by dropping small filters repeatedly in a deep network. For training, high-resolution images are directly modeled by SRCNN. An image that is of high resolution can deteriorate into an image with low-frequency information which is correlated to a low-resolution image and also into a piece of high-frequency information that corresponds to image details. The input and output images share similar information of low-frequency. This shows that SRCNN fills both needs, the first is to convey the input to the last layer which is somewhat similar to what an autoencoder does and the second is to recreate residuals. For training purposes, images from Berkeley Segmentation Dataset are being used and for testing purposes, four datasets were used which include Set5, Set14, Urban 100, and B100 datasets.

Author [12]proposed a method of super-resolution with the help of a deeply-recursive convolutional network (DRCN). Expanding recursion depth can further develop execution without presenting new boundaries for extra convolutions. Two methods have been suggested to facilitate the struggle of training. Firstly, supervision of all the recursion is required. The reconstruction method is something very similar for all recursions. As every recursion prompts an alternate HR forecast, we join all predictions coming about because of various levels of recursions to convey a more precise last prediction. The subsequent proposition is to utilize a skip-connection from the input to the reconstruction layer. In the super-resolution, an LR image (input) and an HR image (output) share similar data overall. A version of the input, nonetheless, is probably going to be lessened during many forward passes. We expressly interface the input to the layers for output reconstruction. This is especially viable when information and yield are exceptionally associated. The network proficiently reuses weight boundaries while taking advantage of an enormous picture domain. To facilitate the trouble of training the model, we utilize recursive-supervision and skip connection. This strategy works best on sharping the edges of the picture, particularly to the example. For training purposes Set 5, Set 14, B100, Urban100 datasets is been used.

Author [13] has suggested the method called SimGAN which is Stimulated and Unsupervised [23] learning in which the job is to get familiar with a model to work on the authenticity of a stimulator’s output utilizing unlabeled genuine information while saving the comment data from the stimulator. The method utilizes an adversarial network much like GAN but with input as a synthetic image. Many main modifications are done to keep intact annotation and equalize training including a local adversarial, self-regularization term, and updating the discriminative model using a history of refined images. For training purposes, the datasets been used are UnityEyes, NYU hand pose dataset, and MPIIGaze datasets for testing. The SimGAN’s output beats the model by 8.8% prepared on genuine images with supervision,

Author [14] has suggested the network called Style and Structure GAN (S2-GAN) has two main part which includes- the Structure GAN that creates a surface normal map and secondly, the Style GAN which receives input as the surface normal map and creates the 2D image. Both GANs are first trained independently and then they are combined by joint learning. Structure GAN encodes the underlying geometry which includes voxel, mesh characterization, etc of the scenario. Style GAN enciphered the texture on the illumination and the objects. S2-GAN resolves the procedure of generation of the image into various parts. The model is more interpretable and produces more practical pictures contrasted with the guidelines. The author also depicts that the proposed technique can discover RGBD portrayals in an unaided way. For training purposes, the dataset that is used is the NYUv2 dataset, Places, ImageNet dataset, RGBD dataset, SUN RGBD dataset.

In Table 1, the summarised literature review of various method used by authors is mentioned .

TABLE 1. COMPARISON OF METHODS GIVEN BY VARIOUS AUTHOR

| Method | Description | Dataset Used | | Ref No |
| --- | --- | --- | --- | --- |
| Training Dataset | Testing Dataset |
| Super-resolution generative adversarial network (SRGAN) using ResNet | This network can generate photo-realistic images by upscaling factor × 4 using the deep residual network. | Set5, Set14, BSD100 | BSD300 | [4] |
| Gradient Map GAN (GMCAN) | It generates images about HVS (Human Vision system) to design a loss function by combining Image quality Assessment. | DF2K (DIVIK + Flickr2K), DF2K + OST (OutdoorSceneTraining), and DIVIK. | Set5, Set14, BSD100 and Urban100 | [6] |
| GAN | It focuses on training two models the evaluating generative models by an adversarial process. | MNIST, the Toronto Face Database  (TFD), and CIFAR-10. | MNIST, the Toronto Face Database  (TFD) | [7] |
| Coupled GAN (CoGAN) | It is constructed to study the joint distribution in two different domains which include a pair of GANs. | MNIST, Celeb Faces Attributes, RGBD, and NYU dataset |  | [9] |
| Conditional GAN (CGAN) | Construction of GAN by adding some data to condition on both the generator and discriminator. | MNIST, Flickr, ImageNet, YFCC100M2, and MIR Flickr 25000 dataset | MNIST | [10] |
| Super-resolution method with deep CNN | It overcomes the drawback of SRCNN by increasing models network depth and achieving better accuracy. | Berkeley Segmentation Dataset | Set5, Set14, Urban 100, and B100 | [11] |
| Super-resolution with a deeply-recursive convolutional network (DRCN) | The super-resolution model with a deep recursive layer improves the overall performance without the need of adding a new parameter. |  | Set 5, Set 14, B100, Urban100 | [12] |
| SimGAN | Stimulated and Unsupervised learning where the job is to get familiar with a model to work on the authenticity of a stimulator’s output utilizing unlabeled genuine information while saving the comment data from the stimulator | UnityEyes, NYU hand pose dataset and MPIIGaze | NYU hand pose and MPIIGaze | [13] |
| Style and Structure GAN (S2-GAN) | It includes- the Structure GAN that creates a surface normal map and secondly, the Style GAN which takes input as the surface normal map and creates the 2D image. | NYUv2, SUN RGBD, Places, RGBD, and ImageNet dataset. | NYUv2, SUN RGB-D, | [14] |
| Generative multi-adversarial network (GMAN) | It is a framework for extending GAN to various discriminators to produce higher quality samples in a small number of iterations. | MNIST, CelebA and CIFAR-10 |  | [15] |
| Conditional adversarial network | Image-to-image conversion problem which is effective in synthesizing pictures from label maps, coloring images, and recreating objects. | Cityscape, UT Zappos50K, CMP Facades, Google Maps, and Paris Street View dataset. | UT Zappos50K | [16] |
| Perceptual GAN | It improves detection performance of small through reducing the difference of representation of small objects with the large objects. | Tsinghua-Tencent 100K benchmark and Caltech benchmark | Tsinghua-Tencent 100K benchmark and Caltech benchmark | [17] |
| Stacked GAN | With Conditioning Augmentation to come up with 256×256 photo-realistic images constrained on text representation. | CUB, Oxford-102, and MS COCO | CUB and Oxford-102 | [18] |
| Energy-based GAN (EBGAN) | It consists of two method- GANs and auto-encoders in which the discriminator act as the energy function that allocates the low energy to the area around the data while the high energy goes to the remaining areas. | MNIST, ImageNet, LSUN bedroom, and CelebA |  | [22] |
| GAN | With techniques like Feature matching, historical averaging, minibatch discrimination virtual batch normalization, and one-sided label smoothing to improve the convergence of the GANs | MNIST, CIFAR-10, ImageNet and SVHN |  | [21] |

# PROPOSED METHOD

In this work for generating super-resolution images, we used GANs as the framework. HR pictures are produced by combining GAN's deep network with an adversarial network, which is more attractive since they are more detailed than SRResNet's comparable architecture. These are measured with the amplification factor measured by PSNR ResNet. We propose a generator network with optimized perceptual loss. This loss is based on the feature map of the VGG network rather than the content loss based on MSE.

For Super-resolution Generator Network is fed with an interpolated low-resolution image ILR as input. Here ILR is the low-resolution image of its High-resolution image IHR. Downsampling photos using different approaches such as Bicubic, bilinear, and others with downscaling operation factor r yields LR images. High-resolution images IHR *R* are only present for training. Images with C color channels, *ILR* can be stated as *W X H X C* and *IHR*, *ISR*by *rW X rH X C*.

The primary aim is to train a Generator G that accepts LR images and outputs corresponding HR images. We train the generator as a feedforward neural network . For training IHR, n= 1,…N with corresponding to *ILR* , n= 1,…N

(1)

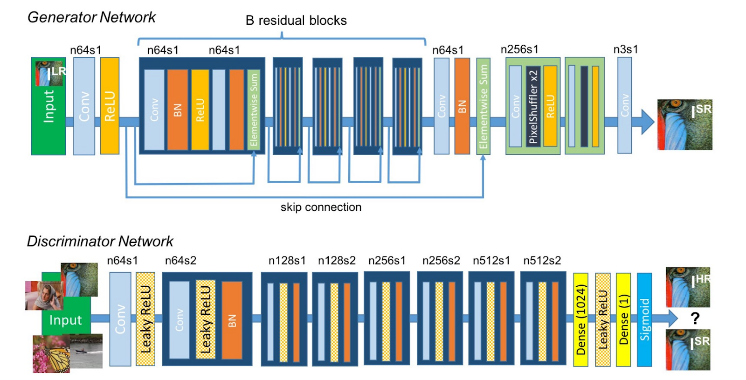
we use the perceptual loss ISR as a weighted sum of most loss components.

## Adversarial network architecture

After setting up the generator, we define a discriminator model with a generator networkto resolve the minimum/maximum problem.

(2)

The goal of this design is to train a generator network to trick a discriminator network, which has been specifically trained to distinguish between actual and super-resolution pictures. Now the discriminator decision generator is used to learn to create an image of HR image similar to a real image, which makes it more difficult for the discriminator to determine which is a super-resolution or a real image.

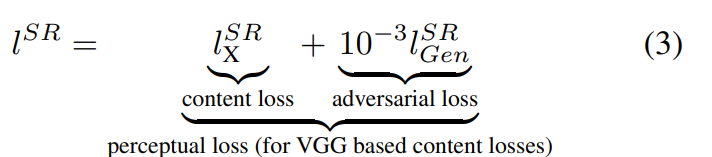
The generator network shown in figure 1 has B residual blocks of the same design. 64 feature maps with batch normalization layers along with two 3x3 kernels, and input image parameters. The resolution of the ReLu activation function is increased by two sub-pixel convolutional layers. 

*Figure 1: (Source* *- Ledig et al.[4]) Architecture of the model*

A trained discriminator was used to tell the difference between real HR images and generated super-resolution images. Use Leaky ReLu activation instead of max pooling because it helps the gradient to pass through the architecture more easily. It contains 8 convolutional layers, and the number of 3x3 cores keeps increasing. When the features are doubled, 64 to 512 kernels in the VGG network. The stride convolution reduces the frame resolution, so the generated feature map is changed by 2 dense layers to ensure.

## Perceptual loss

This is crucial for the execution of the generator network. The loss function ƖSR examines the solutions associated to related characteristics. Then the perceptual loss is formulated as a weighted sum of the adversarial loss and content loss components.

 (3)

## Content loss

In previous implementations, this method is widely used and it gives high PSNR values but usually the lack of detail in the image leads to a smooth texture, instead of relying on pixel-to-pixel loss, the loss function we use is based on a weighted combination. This loss is implemented in ReLu activation layers. VGG loss feature representation of super-resolution images and the reference image.

(4)

Where,

is VGG19 feature map

is High resolution image

## Adversarial loss

Along with content loss, authors combined the generative network to perceptual loss. The support network tries to trick the discriminator network to make the solution close to the real image. The loss of the generator ƖSR is explained on the basis of the probabilities of the discriminator DΘD(GΘG(ILR)) is the Probability that the reconstructed image GΘG(ILR)) is a original HR image.

(5)

Where,

is Low resolution image

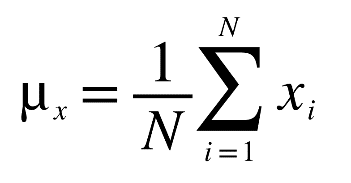
## Structural Similarity Metric

To perform comparative analysis quality between two images relies on estimated errors between truth image and super resolute image. The most common metric is to estimate the difference between values of each corresponding pixel between truth image and super resolute image. Human visual perception can easily identify dissimilarities between images hence, Structural Similarity Metric replicates this behavior of differentiating images.

Structural Similarity Metric calculation ranges between -1 to 1. +1 value indicates both images are very similar and -1 indicates images are every different.

This is based on three key features-

* *Luminance:* which can be defined as average of all pixels values and formula,

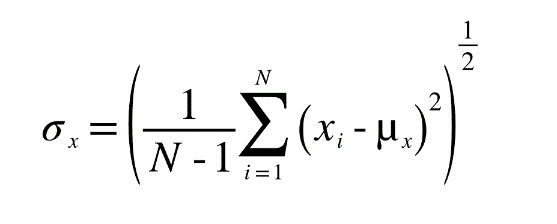
 (6)

*Where,*

*x is image   
N the is total number of pixels  
 is the ith pixel value*

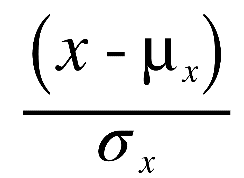
Luminance is calculated on both images x and y, which can also be defined as SR and HR images comparison is function of *l(x,y)* is then function of and .

*Contrast:* which can be defined as the standard deviation of all pixel values and formula.

 (7)

Contrast comparison *c(x,y)* is then function of and .

* *Structure*: can be defined as unit standard deviation

 ()

# EXPERIMENTS

In the proposed research, we experimented on commonly used benchmark datasets Set5 and Set14. All implementation was executed with the scale factor of 2x between Low-Resolution images and High-Resolution images and for comparison analysis, we used SSIM (Structural Similarity Index Metric) and VIFM (Visual Information fidelity Measure). Super Resolved Images from reference models were received from online sources. Results of SRCNN models are available online.

# TRAINING DETAILS

The proposed models were trained on NVIDIA tesla GPU with different datasets mentioned in Table 2. BSD200 (Berkeley Segmentation dataset) includes natural images with human annotation. The Urban100 dataset contains 100 images of urban scenes. The training dataset and testing dataset are different for better evaluation of the model. For testing purposes, Set5 and Set14 dataset has been used. Images with low resolution are obtained by downsampling images with high resolution by a factor of 2x. A mini-batch of 18 random 50 x 50 HR images was used for training. For optimization we used Adam. For Feature extraction, we used the VGG19 model. We alternatively updated Discriminator and Generator which is equal to k =1, Generator model consists of 16 Residual blocks.

TABLE 2. DATASETS USED FOR TRAINING & TESTING

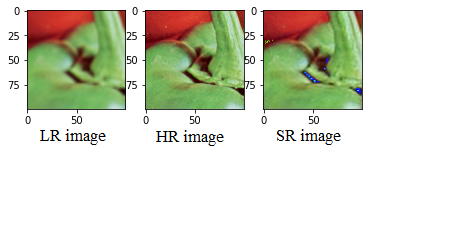
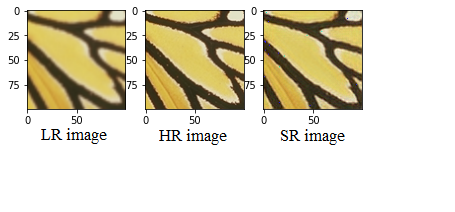
|  |  |
| --- | --- |
| Training | Testing |
| BSD200 | SET5 |
| URBAN100 | SET14 |

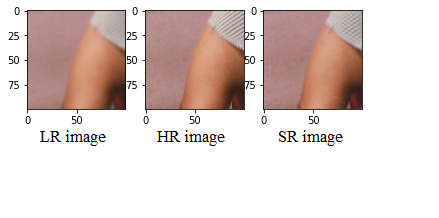
# RESULTS

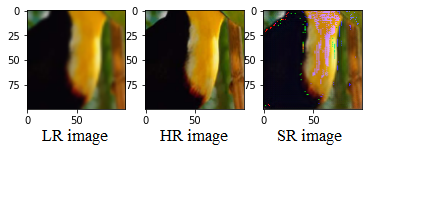
On evaluating generator network, on comparing PSNR values of SRCNN Super resolved images with our SRGAN Super Resolved images, SRCNN have high PSNR values with smooth images on the other hand SRGAN having less PSNR values with detailed images on some parts. We further noticed minor artifacts in Generated Super-Resolved images. Comparison between HR Images and SR images is done via SSIM values as show in Fig 2. Quantitative results are summarized in Table 3.

TABLE 3. PERFORMANCE OF LOSS FUNCTION ON THE ADVERSARIAL NETWORK ON SET5 AND SET14 DATASETS

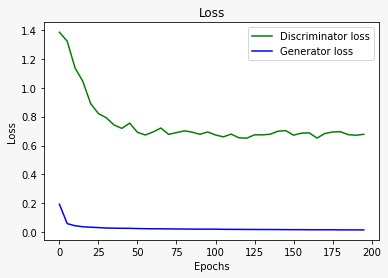
|  |  |  |
| --- | --- | --- |
|  | SET5 | SET14 |
| SSIM | 0.8268 | 0.8087 |





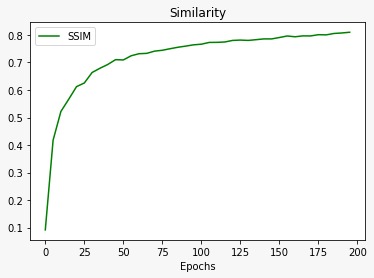


*Figure 2: Comparison between LR, HR, and SR image.*



*Figure 3: Graphical Representation of Discriminator and Generator loss*

Discriminator loss was calculated using adversarial loss equation shown in Fig. 3. Generator loss is combined using content loss equation and adversarial loss equation. The Loss graph signifies that the model has found some optimum and can’t improve anymore, which also means the model has learned thoroughly. In Fig 4. the structural similarity index measure is shown on different epoch which gives estimated errors between truth image and super resolute image.



*Figure 4: Structural Similarity Index Measure*

# CONCLUSION

In this research, the authors tried to compare the existing algorithm SRCNN with SRGANs by using publicly available methods as mentioned in the proposed method section which amplifies the adversarial loss and content loss by training GANs and further experimented with our SR images with SSIM measures. The main aim of this model is to improve the quality of images with super-resolution rather than computational efficiency. We found that deeper SRGAN models are challenging to train due to the display of different objects, we also found that the GAN model purely focuses on details of texture that is a major contrast between super-resolved images generated by SRGAN and super-resolved images generated by SRCNN.

# FUTURE WORK

For better results of SRGANs, we can use different VGG architecture such as VGG22 and VGG54 and with the use of different publicly available datasets we can increase SRGAN performance

ACKNOWLEDGMENT

First of all, we would like to take this chance to express our significant appreciation to the School of Computer Science & Engineering (SET), Sharda University for the support and environment provided to complete our system and implement it as a pilot project. We also extend our gratitude to Dr. Vishal Jain for his supervision, exemplary guidance, monitoring, and constant support.

REFERENCES

1. T.-A. Song, S. Chowdhury, K. Kim, K. Gong, G. El Fakhri, Q. Li, and J. Dutta, “Super-resolution PET using a very deep convolutional neural network,” in Proc IEEE Nucl Sci Symp Med Imag Conf. IEEE, 2018.
2. Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, Bryan Catanzaro, “High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs” in CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018.
3. Karthika Gopan, and Kumar G.S, “Video Super-Resolution With Generative Adversarial Network” in 2nd International Conference on Trends in Electronics and Informatics (ICOEI),2018.
4. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, “Photo-realistic single image super-resolution using a generative adversarial network,” CoRR, vol. abs/1609.04802, 2016.
5. Xin Yu, Basura Fernando, Richard Hartley, and Fatih Porikli, “Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes” in Conference on Computer Vision and Pattern Recognition, IEEE, 2018
6. Xining Zhu, Lin Zhang, Lijun Zhang, Xiao Liu, Ying Shen, and Shengjie Zhao, “GAN-Based Image Super-Resolution with a Novel Quality Loss” Mathematical Problems in Engineering, Hindawi, vol. 2020
7. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. “Generative adversarial nets” in Advances in Neural Information Processing Systems (NIPS), pages 2672–2680, 2014.
8. W. T. Freeman, T. R. Jones, and E. C. Pasztor. “Example-based super-resolution” in IEEE Computer Graphics and Applications, 22(2):56–65, 2002.
9. M.-Y. Liu and O. Tuzel. “Coupled generative adversarial networks”. In *Advances in Neural Information Processing Systems* *(NIPS)*, 2016.
10. M. Mirza and S. Osindero. “Conditional generative adversarial nets”. *arXiv preprint arXiv:1411.1784*, 2014.
11. J. Kim, J. Kwon Lee, and K. Mu Lee, “Accurate image super-resolution using very deep convolutional networks,” in *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, 2016, pp. 1646–1654.
12. J. Kim, J. K. Lee, and K. M. Lee. “Deeply-recursive convolutional network for image super-resolution” In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
13. A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb. “Learning from simulated and unsupervised images through adversarial training” In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
14. X. Wang and A. Gupta. “Generative image modeling using style and structure adversarial networks” In *European Conference on Computer Vision (ECCV)*, 2016.
15. I. Durugkar, I. Gemp, and S. Mahadevan. “Generative multiadversarial networks” In International Conference on Learning Representations (ICLR), 2016.
16. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. “Image-to-image translation with conditional adversarial networks”. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
17. J. Li, X. Liang, Y. Wei, T. Xu, J. Feng, and S. Yan. “Perceptual generative adversarial networks for small object detection”. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
18. H. Zhang, T. Xu, H. Li, S. Zhang, X. Huang, X. Wang, and D. Metaxas. “StackGAN: Text to photo-realistic image synthesis with stacked generative adversarial networks”. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
19. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. “Unpaired image-to-image translation using cycle-consistent adversarial networks”.In IEEE International Conference on Computer Vision (ICCV), 2017.
20. A. Radford, L. Metz, and S. Chintala. “Unsupervised representation learning with deep convolutional generative adversarial networks”. In International Conference on Learning Representations (ICLR), 2015.
21. T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. “Improved techniques for training GANs”. In Advances in Neural Information Processing Systems
22. (NIPS), 2016.J. Zhao, M. Mathieu, and Y. LeCun. “Energy-based generative adversarial network”. In International Conference on Learning Representations (ICLR), 2017.
23. Z. Yi, H. Zhang, P. T. Gong, et al. “DualGAN: Unsupervised dual learning for image-to-image translation”. In IEEE International Conference on Computer Vision (ICCV), 2017.
24. Guim Perarnau, Joost van de Weijer, Bogdan Raducanu,and Jose M. Álvarez “Invertible Conditional GANs for image editing” In Workshop on Adversarial Training, NIPS 2016, Barcelona, Spain.
25. X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. P. Smolley. “Least squares generative adversarial networks”. In IEEE International Conference on Computer Vision (ICCV), 2017.
26. Shakya, Subarna. “Unmanned Aerial Vehicle with Thermal Imaging for Automating Water Status in Vineyard.” Journal of Electrical Engineering and Automation 3, no. 2 (2021): 79-91.
27. Raj, Jennifer S., and Mr C. Vijesh Joe. "Wi-Fi Network Profiling and QoS Assessment for Real Time Video Streaming." IRO Journal on Sustainable Wireless Systems 3, no. 1 (2021): 21-30.
28. Darney, P. Ebby, and I. Jeena Jacob. “Rain Streaks Removal in digital images by Dictionary based sparsity process with MCA Estimation.” Journal of Innovative Image Processing 3, no.3 (2021): 174-189.
29. Hamdan, Yasir Babiker. "Construction of Statistical SVM based Recognition Model for Handwritten Character Recognition." Journal of Information Technology 3, no. 02 (2021): 92-107.