JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

IMAGE RESTORATION USING RESIDUAL GENERATIVE ADVERSARIAL NETWORKS

P. Narasimha Rao[1], P. Akhila[2], P. Sai Praneeth[3], P. Sreeja[4], D. Ravali[5], Mrs. Soppari Kavitha[6]

[1,2,3,4,5]*B.Tech(IV-CSE)Students, Department of Computer Science and Engineering, ACE Engineering College, Hyderabad, India*

[6]*Associate Professor, Department of Computer Science and Engineering*
Ace Engineering College, Hyderabad, Telangana, India

Abstract : ***Due to increase in the crime rates in the major cities, proof like surveillance cameras feed and audio recordings become very crucial in the investigation. But majority of the Indian cities or households do not contain a CCTV camera or a surveillance camera, even if it does, the captured images or video is not up to the mark as it would be highly pixelated and the culprits would get away due to this loophole. Even in the field of medical research and diagnosis, the MRI or brain tumor scans would contain a lot of noise and the scanned images are not clear which hinders an effective diagnosis by the medical practitioner. The common enemy in the above-mentioned scenarios is the availability of low-resolution image which is not very helpful in our work. So, to tackle the above mentioned and many other problems, we have developed a robust deep learning based Single Image Super Resolution model which try to tackle the problem of low-resolution images by converting the obtained low-resolution images into high resolution versions. In this work we try to develop deep learning-based models that try to take in the low-resolution images and provide a high-resolution version of that particular image. In this work we try to retain very crucial features of the low-resolution image like edges, color textures, shadows etc. We start of by discussing the first ever models proposed in super resolution domain that is interpolation techniques, then we look over how the convolutional neural network-based model overshadowed the interpolation techniques, and to further enhance the image and give better results we study how Generative Adversarial Networks helps us in generating a high-resolution image with rich features in it.***

Keywords: Generative Adversarial Networks (GAN), Convolution Neural Network, Generator, Discriminator, Parametric Rectified Linear unit (PReLU), Leaky ReLU, Residual Blocks, Batch-Normalization, Dense.

I. INTRODUCTION

The goal of single-image super-resolution (SISR) algorithms is to generate a higher solution (HR) image from a low-resolution (LR) image input. Reconstructing a high-resolution image from its low-resolution image is a long challenging task in the field of computer vision. This task becomes even more difficult when a single low-resolution image available as input to recreate its high-resolution image. The super-resolution techniques are divided into two categories, which are multi-image super-resolution and single image super-resolution. In this paper, the single image super resolution techniques are discussed.

The single super-resolution approaches that are classified into three categories based on the application point of view are, interpolation-based, reconstruction based and example based. In this paper, the interpolation-based and example-based approaches are reviewed and compared. The different interpolation techniques, presented in the literature for the super-resolution of the images are the nearest neighbor interpolation, bi-cubic interpolation, super resolution conventional neural network, super-resolution residual network, and super resolution generative adversarial network. The nearest neighbor interpolation approach is the simple and easiest method among them.

The problem of super-resolution is solved by CNN, SR-Nets, VGG based transfer learning and GANs. There are many traditional techniques available in the market which promise to solve the problem like Bi-cubic Interpolation, which uses mean squared error and the information from the neighboring pixels to fill the gaps, but these techniques do not produce proper output as the output we get from these techniques are very blurry, pixelated and not up to the mark. We are using Generative Adversarial Networks which uses the best of both the worlds by making use of two neural networks one known as Generator and the other known as Discriminator.

II. LITERATURE SURVEY

Interpolation Techniques are the very first techniques introduced to tackle the problem of Single Image Super Resolution. Interpolation is nothing but the process of using the known data to estimate the values at unknown location, it is mainly used for zooming, shrinking, geometric corrections. In these techniques the low-resolution image is converted into high resolution image by upscaling the LR image by some upscaling factor. So, when upscaled there would be gaps formed in the image. These gaps are then filled with the information that is already available in the low-resolution image. There are various techniques that are introduced in the fields of interpolation like: 1. Nearest Neighbor, 2. Bilinear, 3. Bicubic

Here we mainly focus on Nearest Neighbor and Bicubic. The algorithm remains same for both of the techniques just the number of reference points would differ. For Distance Measure we will be using Euclidean distance.

[1] Single-image super-resolution: European Conference on Computer Vision

Single-image super resolution algorithms can be classified into four groups based on their image priors: prediction models, edge-based methods, image statistical methods, and patch based (or example-based) methods. The example-based approaches, for example, achieve state-of-the-art efficiency. Internal example-based methods take advantage of the self-similarity property of the input image to generate exemplar patches. It was proposed for the first time in Glasner's work, and several modified versions are proposed to speed up implementation. The methods that use external examples to learn a mapping between low and high-resolution patches from external datasets.

[2] Single-image super-resolution using sparse regression and natural image prior

The authors of Kim and Kwon stress the propensity of neighborhood approaches to over fit and use kernel ridge regression to create a more general map of example pairs. Gaussian process regression, trees, and Random Forests can all be used to solve the regression problem.

[3] Image super-resolution using deep convolutional networks

To achieve state-of-the-art SR output, Dong used bicubic interpolation to upscale an input image and trained a three-layer deep completely convolutional network end-to-end. It was then demonstrated that allowing the network to learn the upscaling filters directly would improve efficiency even further in terms of accuracy and time.

III. EXISTING SYSTEM

The modern image zooming and enhancing tools like Adobe Illustrator and Photoshop uses the techniques like Nearest Neighbors and BiCubic interpolation to enhance the quality of the image. The problem with this interpolation is, these interpolation techniques use the information already present in the image to predict or enhance the image, which is not an efficient system as using the already present pixel values and manipulating those values in the image would not produce a decent quality of the image, and there might be highly pixelated output or the crucial features might get displaced, with positioning errors, and the blurry and curvy edges. So, there is a desperate need for a more sophisticated yet an efficient predictive model that does not just copies or manipulated the pre-existing data.

IV. PROPOSED SYSTEM

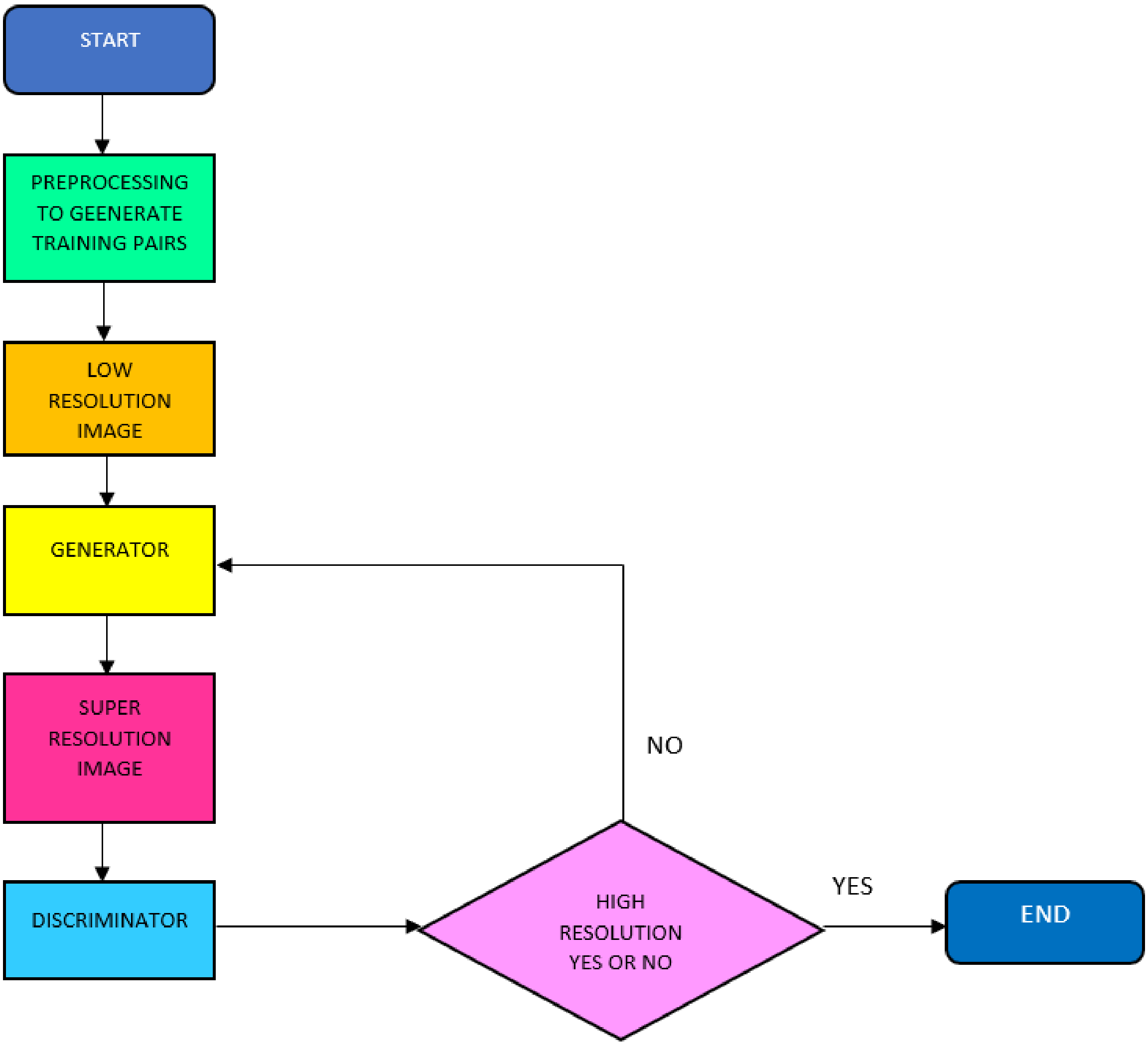
We first take a look at Interpolation Techniques and how they work. In our project scope we would be looking at two of the interpolation techniques named Nearest Neighbors interpolation and Bicubic Interpolation. As the necessity for a more robust and efficient predictive model that can overshadow the interpolation techniques and produce a very convincing results by retaining crucial features with enhanced quality has increased, we build a Convolution Neural Network based model named SRCNN which is the first ever predictive model. Then we checkout how the residual model named SRResNet easily overtook the SRCNN model in terms of image quality and feature retention. Then we also developed our main model named SRGAN which works on the principle of Adversarial training to produce even more convincing results in terms of image restoration named SRCNN (Super Resolution Convolution Neural Network) which is a result of the work by Chao Dong et. al. namely “Image Super-Resolution Using Deep Convolutional Networks”. Then we checkout the GAN Based model named SRGAN as mentioned in the work of Christian Ledig et al. We are using Generative Adversarial Networks which uses the best of both the worlds by making use of two convolutional layers one known as Generator and the other known as Discriminator, here the generator network is used to generate the images based on the input image, and the discriminator network takes this generated image and tells whether is true or fake, so here the generator tries to by-pass the discriminator and with the feedback of the discriminator the generator would try to generate a better acceptable image, and this method also combines two loss function for booth Generator and Discriminator which results in better refinement of image. In the proposed system we have implemented Nearest Neighbor Interpolation, Bicubic Interpolation, Super Resolution Convolutional Neural Network, Super Resolution Generative Adversarial Network (powered by residual networks). Even though the topics of SRCNN and other interpolation techniques are mentioned in related work and literature survey we still would implement them and check how those models work.

V. METHODOLOGY

Super Resolution Residual Generative Adversarial Network

Even though there has been a breakthrough in the efficiency, accuracy, and speed of the single image super resolutions which have been achieved through faster and deeper convolutional neural networks one caveat still lingers, that is the loss of finer texture details and minute important parts of the image when upscaled with a bigger scaling factor. The mainstream traditional CNN’s have been trying to improve the resolution by minimizing loss functions when upscaled like mean squared error, but this lacked high-frequency details and would not yield satisfactory results when being upscaled. The generative models have been bringing a breakthrough in the deep learning problems, So in Residual Generative Adversarial Network-based super-resolution model has been proposed to tackle the problem mentioned above. This SR-GAN is the first-ever architecture to achieve the state-of-the-art results even when the image is upscaled by the factor of 4. Extending the previous discussions of loss functions and their importance further, a new loss function named perceptual loss is proposed which is a combination of adversarial loss and content loss. Initially an SRResNet model has been developed from optimizing the previous SRCNN model which generates realistic images, to improve this further another neural network is added to the SRResNet model together to form a GAN, which outperforms the SRResNet and SRCNN models.

Proposed Architecture



GAN Intuition

GAN are generative model architecture introduced by Ian Goodfellow in 2014. Generative Adversarial Networks are the latest trending topic in deep learning which managed to solve many complex problems that seemed very complex for the traditional deep learning techniques. Here the basic working of the GAN is explained. GAN is a neural network architecture where 2 neural networks are used instead of one, where one is the generator and the other is the discriminator. As the name suggests the Generator tries to generate random samples and gives them to the discriminator, the job of the discriminator is to check and tell whether the generated image is real or fake.

Generator

The generator part of a GAN learns to create fake data by incorporating feedback from the discriminator. It learns to make the discriminator classify its output as real. Generator training requires tighter integration between the generator and the discriminator than discriminator training requires. The portion of the GAN that trains the generator includes:

- random input
- generator network, which transforms the random input into a data instance
- discriminator network, which classifies the generated data
- discriminator output
- generator loss, which penalizes the generator for failing to fool the discriminator

The Super Resolution GAN, like other GAN architectures, has two parts: a generator and a discriminator. The generator generates data based on a probability distribution, while the discriminator attempts to guess weather data from the input dataset or generator. After that, the generator attempts to optimise the produced data in order to deceive the discriminator. In its most basic form, a GAN takes random noise as its input. The generator then transforms this noise into a meaningful output. By introducing noise, we can get the GAN to produce a wide variety of data, sampling from different places in the target distribution.

Discriminator

The discriminator in a GAN is simply a classifier. It tries to distinguish real data from the data created by the generator. It could use any network architecture appropriate to the type of data it's classifying. The discriminator's training data comes from two sources:

- Real data instances, such as real pictures of people. The discriminator uses these instances as positive examples during training.
- Fake data instances created by the generator. The discriminator uses these instances as negative examples during training.

The discriminator connects to two loss functions. During discriminator training, the discriminator ignores the generator loss and just uses the discriminator loss. We use the generator loss during generator training, as described in the next section. During discriminator training:

- The discriminator classifies both real data and fake data from the generator.
- The discriminator loss penalizes the discriminator for misclassifying a real instance as fake or a fake instance as real.
- The discriminator updates its weights through backpropagation from the discriminator loss through the discriminator network.

SRGAN Pre Processing

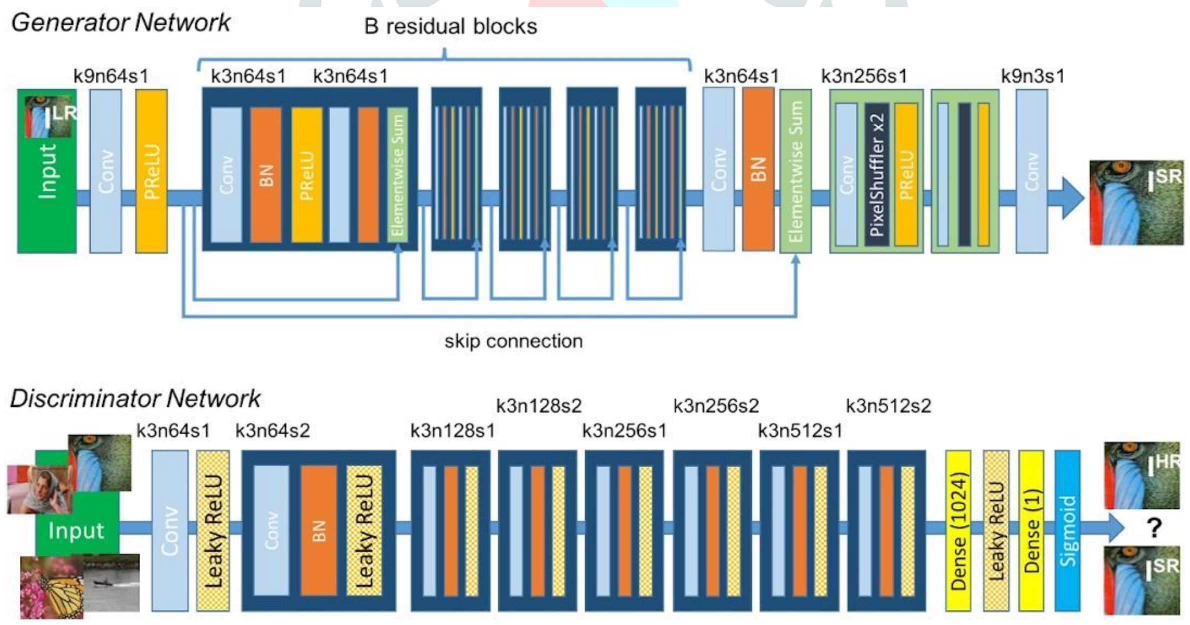
As mentioned already in the previous explicitly have to prepare such a dataset. This is the main pre-processing step in the SRGAN model. In the SRGAN model due to our GPU constraints and training time, we fix the input size as 64x64 pixels and the output as 128x128 pixels, and here we can observe that the upscaling factor of GANs are 4x. We get the output image improved 4 times compared to input. Our pre-processing involves 2 steps, they are:

1. Take the input image and resize the image to 64 x 64 pixels and label them as Low-Resolution image.

2. (if the image is not 128 x 128 pixels) Take the input image and resize the image to 256 x 256 pixels and label them as High-Resolution image.

To verbally explain we have prepared two sets of images one set labels as Low-Resolution Images which contain all the images in 64 x 64 pixels resolution and the second one as High-Resolution Images where all the images are in 256 x 256 pixels resolution. Now, our dataset is prepared and we can give this Low Resolution and High-Resolution images as the input to the model and train the model to produce high resolution images as the output.

Residual GAN Architecture



The Super Resolution GAN, like other GAN architectures, has two parts: a generator and a discriminator. The generator generates data based on a probability distribution, while the discriminator attempts to guess weather data from the input dataset or generator. After that, the generator attempts to optimise the produced data in order to deceive the discriminator. The following are the architectural specifications for the generator and discriminator:

Generator Architecture

Since residual networks are simple to train and enable them to be significantly deeper in order to produce better performance, the generator architecture uses them instead of deep convolutional networks. This is due to the residual network's use of a link form known as skip connections.

16 residual blocks were created by ResNet. Two convolutional layers with small 3*3 kernels and 64 feature maps are used in the residual block, followed by batch-normalization layers and ParametricReLU as the activation function. Two qualified sub-pixel convolution layers boost the resolution of the input image. Instead of using a fixed value for a rectifier parameter (alpha) like LeakyReLU, this generator architecture uses parametric ReLU as an activation function. It learns the parameters of the rectifier over time and improves accuracy at a low computational cost.

A high-resolution image (HR) is down sampled to a low-resolution image during the preparation (LR). The picture is then up sampled from low resolution to super-resolution by the generator architecture. The image is then passed through the discriminator, which attempts to discriminate between a super-resolution and a high-resolution image.

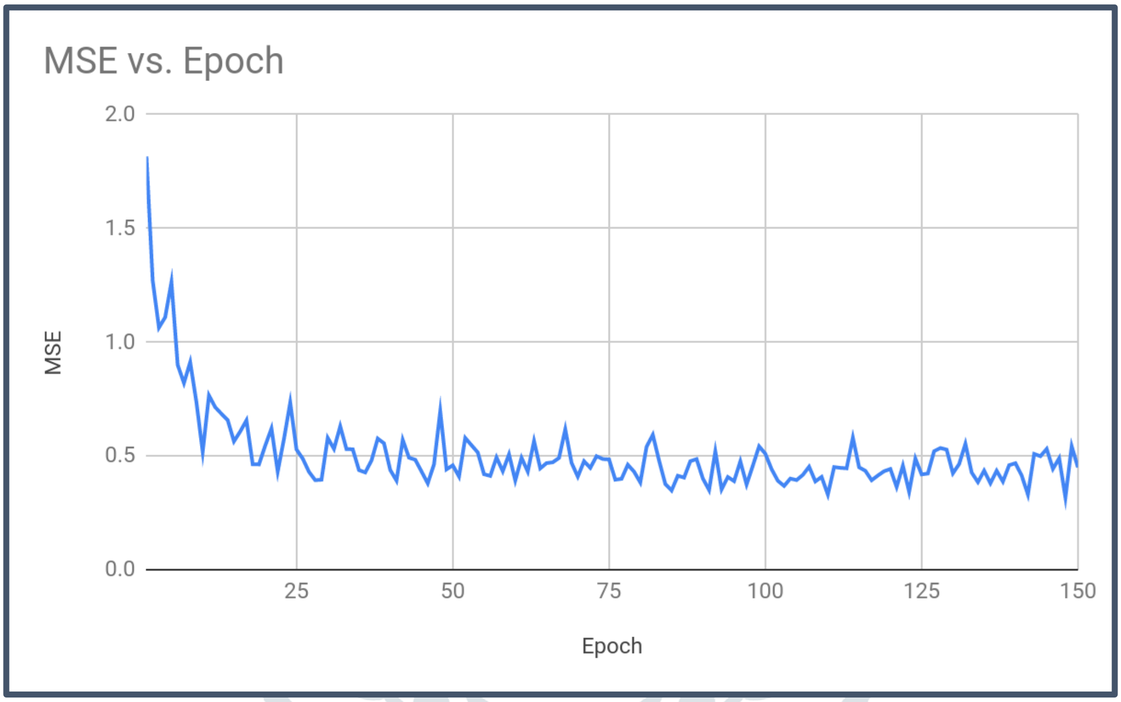
Discriminator Architecture

The discriminator's job is to tell the difference between real HR images and produced SR images. The network has eight convolutional layers, each with 3*3 filter kernels, with the number of kernels increasing by a factor of two from 64 to 512. Where the number of features is doubled, strided convolutions are used to minimize the image resolution. To obtain likelihood for sample classification, the 512 feature maps are followed by two dense layers with a leakyReLU added between them and a final sigmoid activation function.

VI. PERFORMANCE ANALYSIS

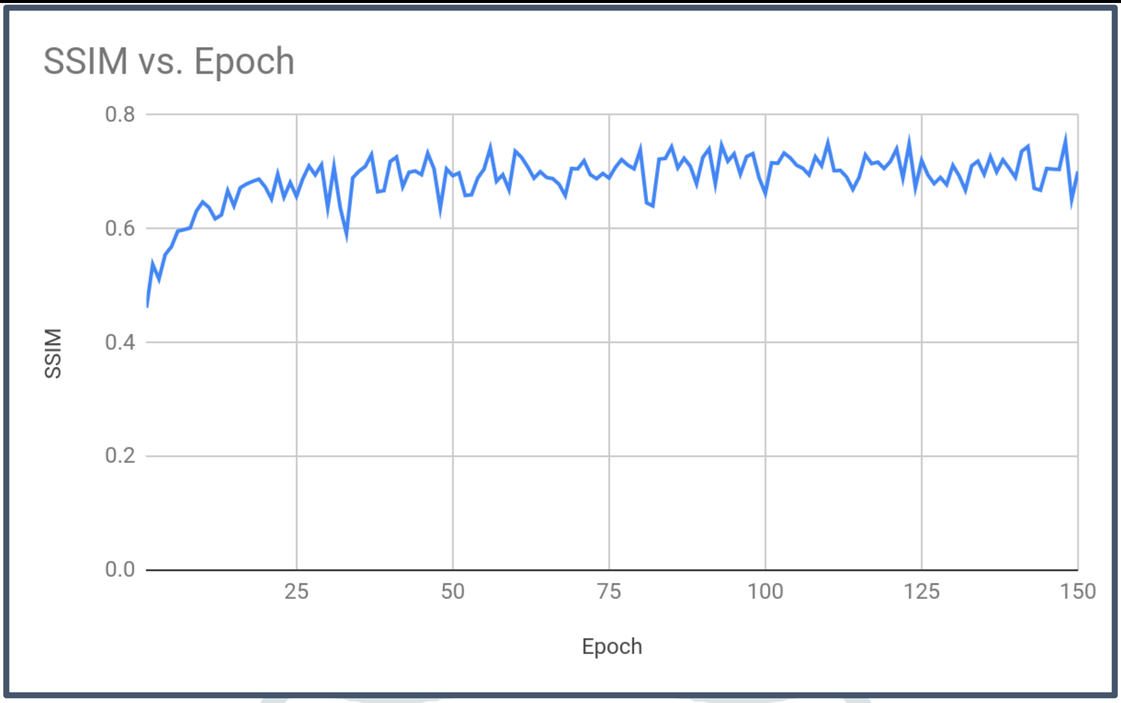
MSE of validation set images:

The MSE represents the cumulative squared error between the original and the modified image, whereas PSNR represents a measure of the peak error.



SSIM of Validation set images:

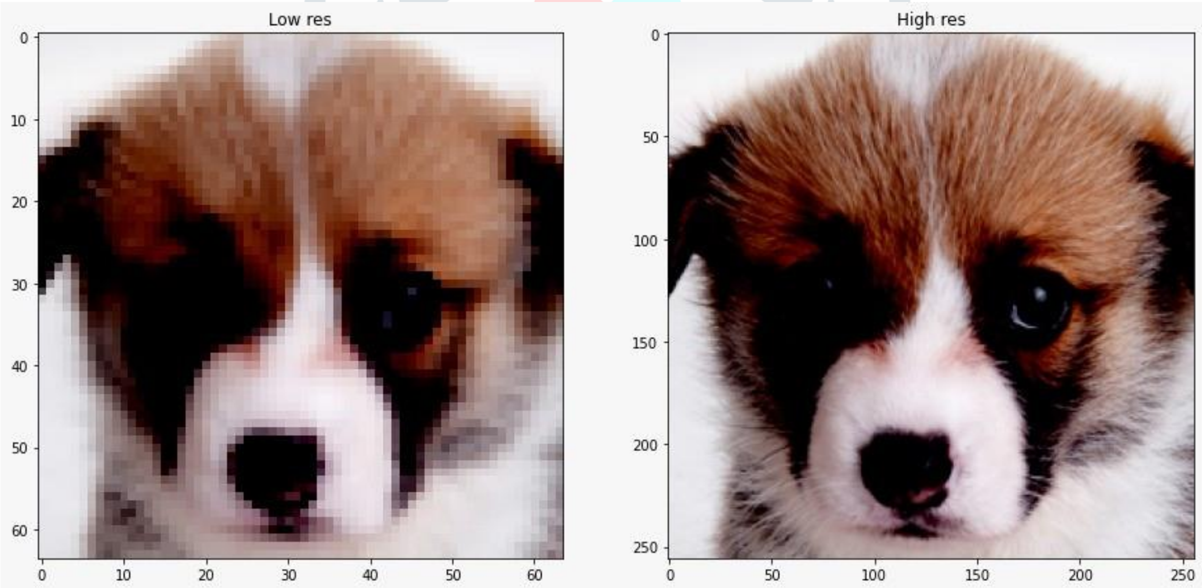
The Structural Similarity Index (SSIM) is a perceptual metric that quantifies image quality degradation* caused by processing such as data compression or by losses in data transmission. It is a full reference metric that requires two images from the same image capture—a reference image and a processed image.



VII. RESULTS

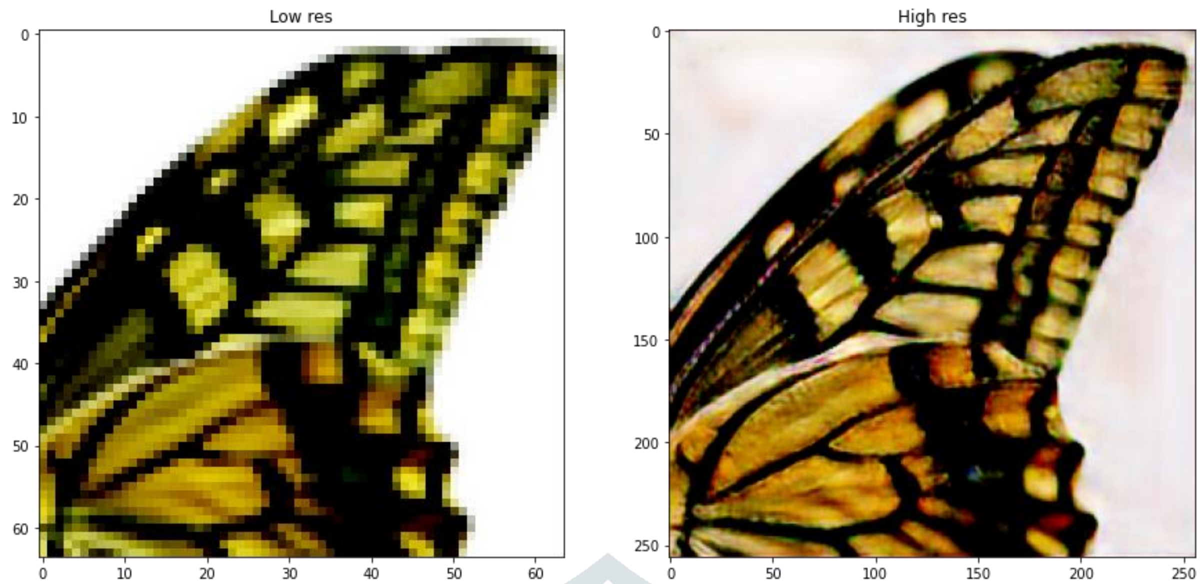
It is to be noted that this model has been trained with the 4x upscaling factor, and this SRGAN architecture works like the state-of-the-art algorithm. The results are as shown below.

Pre-processing Results



The left side of the image is the 4x downscaled image whereas the right-side image is the ground truth, here we are using 4x as upscaling factor.

SRGAN Model Predictions



VIII. CONCLUSIONS

The image super resolution task is one of the very sought out task as it can be applied in any field of the science and technology and it can improve the quality of research and investigation. In the SRGAN model, the model is also trained on various datasets. The generator and discriminator are trained separately and made to work in an adversarial environment. The model is trained for 5000 epochs. The pictorial results mentioned clearly tells us that the SRGAN outperforms all of the models, even having a huge task of upscaling by a factor of 4, it managed to obtain a very high PSNR values and a very low MSE rate. So, we can easily conclude that the SRGAN model outperforms all of the other models. Obtaining an average PSNR values of 60 and greater and having MSE value less than 1 clearly indicated that this is the state of art algorithm in the domain of single image super resolution.

ACKNOWLEDGEMENT

We would like to thank our guide Mrs. Soppari. Kavitha for her continuous support and guidance. Due to her guidance, we can complete our project successfully. Also, we are extremely grateful to Dr. M.V.VIJAYA SARADHI, Head of the Department of Computer Science and Engineering, Ace Engineering College for his support and invaluable time.

REFERENCES

[1] C.-Y. Yang, C. Ma, and M.-H. Yang. Single-image super-resolution: A benchmark. In European Conference on Computer Vision (ECCV), pages 372–386. Springer, 2014.

[2] Zhang, N., Donahue, J., Girshick, R., Darrell, T.: Part-based RCNNs for fine-grained category detection. In: European Conference on Computer Vision. pp. 834–849 (2014).

[3] Kim, K.I., Kwon, Y.: Single-image super-resolution using sparse regression and natural image prior. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(6), 1127–1133 (2010).

[4] K. I. Kim and Y. Kwon. Single-image super-resolution using sparse regression and natural image prior. IEEE Transactions on pattern analysis and Machine Intelligence, 32(6):1127–1133, 2010.

[5] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. IEEE Transactions on pattern analysis,38(2):295–307, 2016.

[6] C. Dong, C. C. Loy, K. He and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol:38, no:2, pp: -295-307, 1-Feb-2016, doi: 10.1109/TPAMI.2015.2439281.