TIJER || ISSN 2349-9249 || © March 2024, Volume 11, Issue 3 || www.tijer.org

Image Resolution Enhancing Using ESRGAN Models.

Prabhu M Rokade, Sahil Nikam, Shreevats Nandan, Vaibhav Chouhan, Prof. Sheetal Shimpikar

[1]Student, [2]Student, [3]Student, [4]Student, Faculty

Department of Computer Engineering,

Pillai College of Engineering New Panvel, Navi Mumbai, India

Abstract - Image Enhancing System Using ESRGAN Model represents the increasing demand for high-quality video content in fields like medical imaging, satellite imaging, scientific research, etc. There is a growing need for solutions to enhance the quality of the existing video footage and images. Our project aims to do so using Deep learning algorithms on Enhanced Super Resolution Generative Adversarial Networks (ESRGAN) to upscale the resolution of images. This project focuses on enhancing the visual quality of images by upscaling them while preserving details and minimizing artifacts. The primary goal is to develop innovative algorithms and methods for increasing the resolution of images such as Conv2DTranspose, Cropping2D and customize loss functions. The ESRGAN model is a generative adversarial network that consists of two parts: a generator network that produces high-resolution images and a discriminator network that evaluates the generated images. The generator and discriminator are trained in a competition to produce high quality images that are indistinguishable from real images.

Index Terms - Generative Adversarial Networks (GANs), Enhanced Super Resolution Generative Adversarial Networks (ESRGAN), Convolutional neural networks (CNNs), artificial neural networks (ANNs),deep neural networks (DNNs), Peak signal-to-noise ratio (PSNR).

I. INTRODUCTION

We love high-quality images that are clear and detailed, but sometimes we come across pictures that are a bit blurry or not as sharp as we would like. This project is all about making these blur images look much better. like this when we take a small picture and make it bigger it usually becomes blurry and fuzzy. We want to change to make these pictures bigger while keeping them sharp and clear.
The task of Enhancing Upscaled image resolution is a challenging task in image processing. Generative Adversarial Networks (GANs) have shown immense potential in this area. GAN is a subset of Deep Learning that is based on artificial neural networks (ANNs) and convolutional neural networks (CNNs)with multiple layers, also known as deep neural networks (DNNs). Among them ESRGAN have shown the greatest potential.

The objectives of the Image Upscaling and Enhancement are as follows:

1. Enhancing Image Quality: The primary objective is to improve the quality and resolution of digital images. This involves developing techniques to make images look sharper, clearer, and more visually appealing.
2. Image Upscaling: The project aims to upscale images, increasing their size while preserving or enhancing their detail and quality, especially for images with low resolution.
3. Detail Preservation: Preserving the essential details, textures, and structures within images during the upscaling and enhancement process is a key focus, ensuring that the enhanced images maintain their original characteristics.

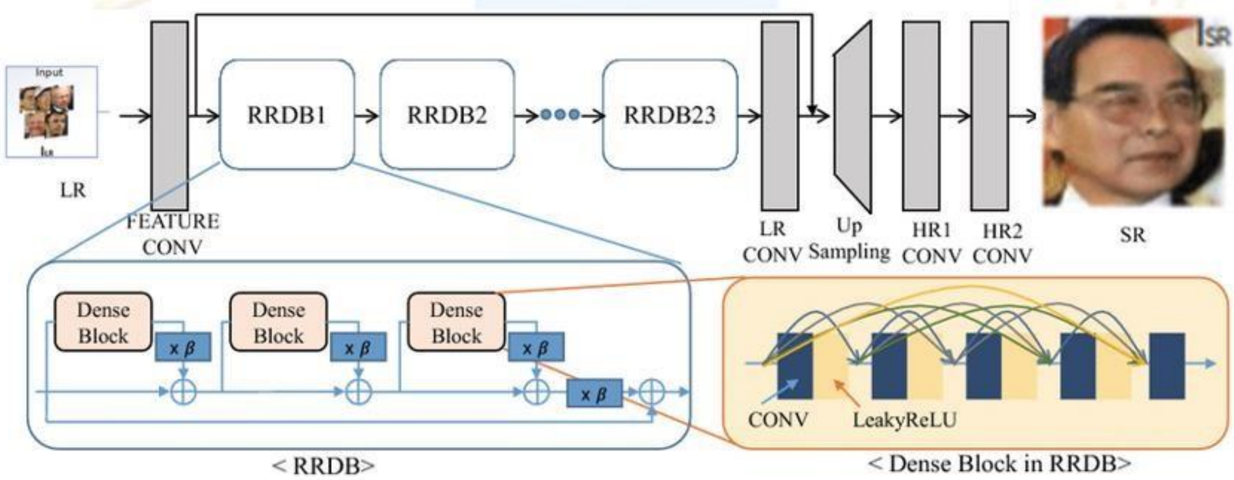

Fig.1 ESRGAN Structure

II. LITERATURE SURVEY

J. Goodfellow et al. [14] proposed adversarial nets framework, the generative model is pitted against an adversary: a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The generative model can be thought of as analogous to a team of counterfeiters, trying to produce fake currency and use it without detection, while the discriminative model is analogous to the police, trying to detect the counterfeit currency. Competition in this game drives both teams to improve their methods until the counterfeits are indistinguishable.

Alec Radford et al. [17] proposed and evaluated a class of architectures named Deep Convolutional GANs (DCGAN)methods, and deep learning methods. They used the trained discriminators for image classification tasks, showing competitive performance with other unsupervised algorithms and visualized the filters learnt by GANs and empirically show that specific filters have learned to draw specific objects.

Christian Ledig et al. [17] proposed a generative adversarial network (GAN) for image super resolution (SR) is presented. It's the first framework capable of inferring photo-realistic natural images for 4× upscaling factors. To achieve this, they proposed a perceptual loss function which consists of an adversarial loss and a content loss. The adversarial loss pushes our solution to the natural image manifold using a discriminator network that is trained to differentiate between the super-resolved images and original photo-realistic images. In addition, they used a content loss motivated by perceptual similarity instead of similarity in pixel space.

III. METHODOLOGY

To improve the quality and resolution of digital images dimension, several steps are performed as follows.

1. Normalize Each HR image of the dataset into RGB format for training the model and making batches of the images.
2. Create a low-resolution image from the HR image.
3. Pass that LR image to Generator, and it will give you an SR image, which is not perfect yet.
4. The discriminator will get HR images and SR images as input, and it will try to tell which are real and which are fake and give each a percentage value of the probability of the reality of the image.
5. Generator and Discriminator both will calculate their losses and gradients and pass that value to the optimizer. The loss functions we are using are as follows:
   - **Binary cross entropy loss function**:

     $$-\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1-y_i) \cdot \log(1-p(y_i))$$

     Here Yi represents the actual class and log(p(yi)is the probability of that class.

   - **Perceptual loss function**: (= *tf.reduce_mean(tf.square(hr_features - sr_features))*)
     It is a loss function that measures the difference between the high-level features of two images, we extracted from a pre-trained VGG neural network.

   - **Color loss function:** (= *tf.reduce_mean(tf.square(hr_image_color_for_loss - sr_image_color_for_loss))*) It is used to maintain the color integrity of the generated images.

   - **adversarial loss function**: (= *-tf.reduce_mean(discriminator(sr_image))*)
     It is used to maintain the discriminator loss which is adverse to generator. Basically, we use it to see how good our discriminator is doing.
6. The optimizer function will help to modify weights to get the minimum loss of both the generator and discriminator by finding the minima. We use Adam optimizer as we need a variable learning rate for smooth training of the model.
7. This process will be repeated for the number of epochs specified.
8. Generator will try to fool Discriminator every time, and Discriminator will also learn how to distinguish between real and fake. Both will improve simultaneously.
9. After the training, the generator will be able to generate the image as real as possible, and loss is also minimal.
10. We freeze the weights of the generator, and thus the model is trained.
11. So, we pass a blurry image to the generator, which is trained in that process. The image gets upscaled and its resolution will be increased by the generator.
12. We finally used PSNR score for evaluation purpose and tuned the hyperparameters of the model to improve it.

Terms used:

1. Image Processing: Normalize Each HR image of the dataset into RGB format for training the model and making batches of the images.
2. Data Batches: Make batches for training, like 100 images in one batch.
3. LR Image: Convert High resolution image into low resolution image for training of generator.
4. SR Image: SR image is output by generator after taking input as a LR image at start of training generator has random weights and noises so it will generate random output.
5. Loss: Generator and Discriminator both will calculate their losses and gradients and pass that value to the optimizer.
6. optimizer: The optimizer function will help to modify weights in order to get the minimum loss of both the generator and discriminator.
7. Weights and Bias updating to minimize loss: generator and discriminator both will change their weights in order to minimize the loss. So, change will happen in order to improve output.
8. Epoch: This process will be repeated for the number of epochs specified.

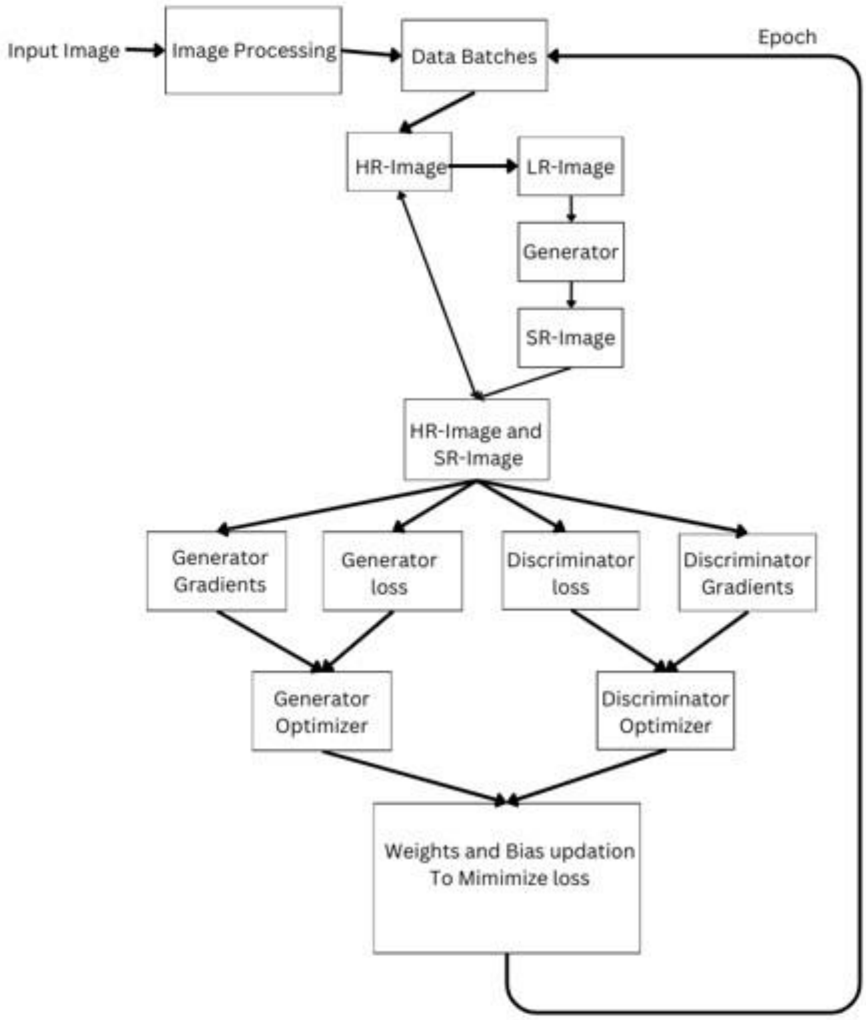

Fig.2 Activity (Model training flow)

Evaluation used:

PSNR is most used to measure the quality of reconstruction of images. The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs, PSNR approximates human perception of reconstruction quality equation.

$$MSE = \frac{\sum_{M,N} [I_1(m,n) - I_2(m,n)]^2}{M * N}$$

In the previous equation, M and N are the number of rows and columns in the input images.

$$PSNR = 10 \log_{10}\left(\frac{R^2}{MSE}\right)$$

In the previous equation, R is the maximum fluctuation in the input image data type. For example, if the input image has a doubleprecision floating-point data type, then R is 1.

IV. CONCLUSIONS

In this report, we introduced a framework for generating high resolution images from provided low resolution images, we can conclude that ESRGAN model approach would give better performance results. The results, as evidenced by metrics such as PSNR and SSIM, indicate that ESRGAN outperforms conventional upscaling techniques, offering sharper and more realistic images. This has promising implications for various fields, including medical imaging, entertainment, and computer vision, where high-quality image resolution is essential. However, we also acknowledge some limitations and challenges, such as the need for substantial computational resources. As a result, future work may involve optimizing ESRGAN and exploring its potential in real-time applications.

V. REFERENCES

[1] Chenyang Zhu, “Improved Image Super-resolution Real ESRGAN”, school of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China IEEE 2022.

[2] Hongying Liu1, Peng Zhaol , Zhubo Ruan1, Fanhua, “Large Motion Video Super-Resolution with Dual Subnet and Multi Stage Communicated Upsampling”, Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, School of Artificial Intelligence, Xidian University AAAI 2022.

[3] Xintao Wang, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks”, CUHK-SenseTime Joint Lab, The Chinese University of Hong Kong, [cs.CV] 17 Sep 2018.

[4] Christian Ledig, “Photo-Realistic Single Image Super Resolution Using a Generative Adversarial Network ”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017.

[5] Bee Lim, Sanghyun, Son, Heewon Kim, Seungjun Nah, Kyoung Mu Lee, “Enhanced Deep Residual Networks for Single Image Super Resolution ”, Department of ECE, ASRI, Seoul National University, 08826, Seoul, Korea 2017.

[6] Christian Ledig, “Generative Adversarial Networks (GAN)A Gentle Introduction”, Department of Statistics and Data Science University of Texas at Austin 2016.

[7] Alec Radford, Luke Metz, “UNSUPERVISED REPRESENTATION LEARNING WITH DEEP CONVOLUTIONAL GENERATIVE ADVERSARIAL NETWORKS ”,Under review as a conference paper at ICLR 2016 arXiv:1511.06434v2 [cs.LG] 7 Jan 2016 .

[8] J. Goodfellow, Jean Pouget Abadie , Mehdi Mirza, Bing Xu, David WardeFarley, Sherjil Ozair , Aaron Courville, Yoshua Bengio, “Generative Adversarial Nets ”,Departement d’informatique et de recherche op ´ erationnelle ´ Universite de Montr ´ eal ´ Montreal, QC H3C 3J7 2014.