Implementation of SRGAN

1st M. Ahabb Sheraz
2021327
Ghulam Ishaq Khan Institute
Topi, Pakistan

2nd M. Ahmed
2021331
Ghulam Ishaq Khan Institute
Topi, Pakistan

***Abstract*—Super-resolution (SR) involves upsampling a low-resolution image into a higher resolution with minimal information distortion. In this paper, we tried to implement the popular Super-Resolution Generative Adversarial Network (SRGAN), an ingenious super-resolution technique that combines the concept of GANs with traditional SR methods.**

***Index Terms*—GANs, SR, SRGAN, VGG19**

I. Introduction

Super-resolution (SR) refers to increasing an image's resolution to obtain a higher-quality version. Traditional SR methods often rely on interpolation techniques, which may result in lossy detail and produce blurry images. Generative Adversarial Networks (GANs) offer a promising alternative for SR by generating high-resolution images that are perceptually realistic and visually appealing [1].

II. Background

A. Generative Adversarial Networks (GANs)

Goodfellow, I., et al. [2] introduced Generative Adversarial Networks (GANs) in 2014. GANs consist of two neural networks: a generator and a discriminator. The generator learns to generate synthetic data samples from random noise, while the discriminator learns to distinguish between real and synthetic samples. Through adversarial training, the generator learns to produce realistic samples that are indistinguishable from real ones.

B. Single Image Super-Resolution (SISR) Techniques (Prior to 2017)

Prior to the introduction of SRGANs, traditional single-image super-resolution (SISR) techniques mainly relied on interpolation methods such as bicubic interpolation. These methods often resulted in blurry images and did not capture fine details. Researchers then began exploring deep learning techniques, particularly convolutional neural networks (CNNs), for image super-resolution. Models such as Super-Resolution Convolutional Neural Network (SRCNN) showed promising results in enhancing the resolution of images. However, they tend to produce artifacts and over-smoothed regions in the super-resolved images, especially in areas containing high-frequency details [3].

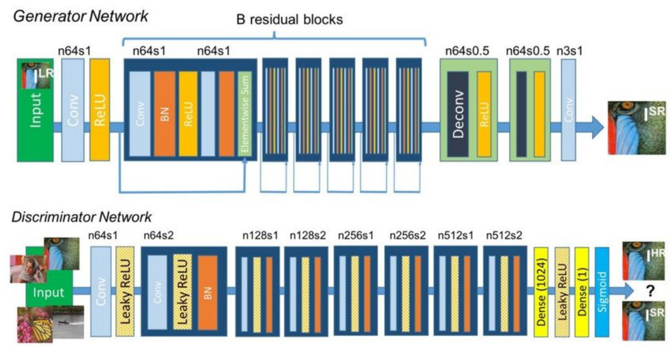

Fig. 1. Architecture of Generator and Discriminator Network with corresponding kernel size (k), number of feature maps (n) and stride (s) indicated for each convolutional layer [3].

C. Introduction of SRGAN (2017)

The SRGAN model as proposed in C. Ledig et al. [3], introduced a novel approach to super-resolution using the adversarial training framework of GANs. SRGAN offers superior perceptual quality, sharper edge and texture recovery, and better realism compared to SRCNN, making it a preferred choice for single-image super-resolution tasks in various applications.

III. SRGAN Architecture

The generator network employs residual blocks to keep information from previous layers alive and allow the network to choose from more features adaptively. Instead of adding random noise as the generator input, we pass the low-resolution image. The discriminator network is standard and is similar to the ones used in other GAN models [1].

Another noteworthy feature is the SRGAN's two loss functions: GAN loss and perceptual/content loss. The loss function helps in figuring the perceptually relevant characteristics. So not only is the adversarial loss helping adjust the weights, but the content loss is also doing its part. The content loss is defined as VGG loss, which means a pretrained VGG network output is compared pixel-wise [1].

IV. Training and Testing

Following C. Ledig et al. [3] and Bhattiprolu, S. [4] we implemented the SRGAN on Python using the Tensorflow Keras functional API. We used the VGG19 model for the content loss. We trained our SRGAN model on an NVIDIA GTX1660 Ti GPU using a random sample of 5000 images

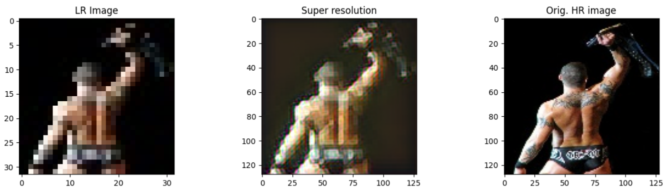

Fig. 2. Result from our implementation of SRGAN. A 32 × 32 image is regenerated to a 128 × 128 image. Despite a small training set and ten epochs, one can still observe improvements in the generated image compared to the low-resolution (LR) image.

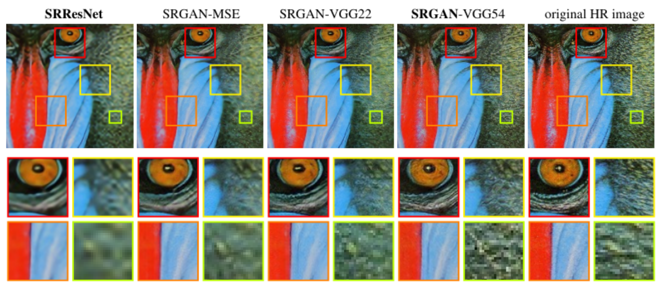

Fig. 3. SRGAN results from C. Ledig et al. [3].

from the MIRFLICKR25k dataset. The High-Resolution (HR) images had to be downsampled to 128 × 128 to be processable by our hardware. The Low-Resolution (LR) images were downsampled to 32 × 32. The model was then trained for 10 epochs with a learning rate of 0.001. Despite the small training set and low number of epochs, our model was able to show the effectiveness of the SRGAN architecture by producing better-quality images compared to LR images.

The original authors trained the model on an NVIDIA Tesla M40 GPU using a random sample of 350 thousand images from the ImageNet database. Their SRGAN variants were trained with 105 update iterations at a learning rate of 0.0001 and another 105 iterations at a lower rate of 0.00001 [3].

V. Conclusion

After gaining a brief knowledge of the concepts of image and video resolutions, we understood the concept of SRGANs in further detail. We then explored the architecture of this network in detail by looking at the generator and discriminator blocks accordingly. Finally, we developed a project to understand the significance of these generative neural networks.

Acknowledgment

We would like to extend our thanks to our course instructor, Sir Ahmer Rashid from Ghulam Ishaq Khan Institute, for enriching and challenging our understanding of deep neural networks. We also extend our thanks to Digital Sreeni, the YouTuber, for his insightful content and valuable contributions. His videos have been a source of inspiration and knowledge, enriching our understanding of GANs and SRGANs.

References

[1] A. Rosebrock, "Super-Resolution Generative Adversarial Networks (SRGAN)," PyImageSearch, Jun. 2022. [Online]. Available: https://pyimagesearch.com/2022/06/06/super-resolution-generative-adversarial-networks-srgan/. [Accessed: 2024-05-02].

[2] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).

[3] C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. [Online]. Available: https://arxiv.org/abs/1609.04802

[4] Bhattiprolu, S. (2022, February 09). 256 - Super resolution GAN (SRGAN) in keras. YouTube. https://www.youtube.com/watch?v=1HqjPqNglPct=640s

[5] Bhattiprolu, S. (2023). python for microscopists. GitHub.