Ultra-Realistic: Generating Ultra-realistic Super-resolution Images with A-ESRGAN

Dushyant Singh Udawat (ds35) , Christopher Cai (cdcai2) , Pratikshit Singh (ps71)

Problem Definition

In our project, we are planning to implement A-ESRGAN, which is the state of the art GAN model for blind image super-resolution. This version of SR is built upon improvements over previous models such as SRGAN, ESRGAN, and RealSRGAN. It modifies the discriminator used in the aforementioned models by using a multiscale attention U-net discriminator that helps the generator generate more detailed images and reduce the blurring of edges apparent in the previous models. We will use the DIV2K dataset, which contains 800 images for training, 100 images for validation, and 100 images for testing. We also plan to test our implementation on two other datasets (more information in the datasets heading) and compare our results with those of other popular super-resolution models such as **SRGAN, ESRGAN, and Real-ESRGAN on** non-reference natural image quality evaluator(NIQE) [2] metric.

Motivation

Image super resolution is pursued for its practicality in various fields like medical imaging, surveillance, and remote sensing, and for enhancing user experience in applications like video streaming and gaming with clearer and more detailed visuals. Our primary motivation for this project is to gain hands-on experience working with a large volume of images, allowing us to explore and understand various data features and transformations in greater detail. Additionally, this project will provide us with an opportunity to utilize transformers, a cutting-edge machine learning technique, in the context of a real-world application. Through this project, we hope to gain a deeper understanding of convolutional neural networks and their practical applications, as well as develop our skills in machine learning and image processing.

Datasets

1. For training of our A-ESRGAN model , we will use DIV2K[9] Dataset containing 800 HR (high-resolution) images for use in image super-resolution research.
2. For testing we will use two of the following available datasets,
    a. Set5
    b. Set14
    c. BSD100
    d. Sun-Hays80
    e. Urban100

The selection among these five will be done on the basis of which combination of these two datasets will allow us to compare our results with other SR models.

Schedule

Week	Expected Progress
1	Take reference from the SRGAN paper and understand the layer architecture and implementation.
2	Explore the various parameters for each layer. Work on the research and development of the Image super-resolution block.
3	Attention Mechanism - Figure out the layer architecture and its implementation. Explore the various parameters for layers involved.
4	Work on the development of channel and spatial attention blocks.Merge both spatial and channel attention blocks with the super resolution model.
5	We will compare the results of object detection from low resolution images to super-resolution images, between our model and the previous work.
6	Explore further work including- additional layers, adding noises other than gaussian to the images.

References

1. Wang, Xintao, et al. "ESRGAN: Enhanced super-resolution generative adversarial networks." Proceedings of the European Conference on Computer Vision (ECCV). 2018.
2. Zhang, Y., Liu, X., Yuan, X., & Wu, S. (2021). A-ESRGAN: Training Real-World Blind Super-Resolution with Attention U-Net Discriminators. arXiv preprint arXiv:2112.10046.
3. Yuan, X., Liu, W., Zhang, X., & Tang, X. (2019). Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 0-0).
4. SRGAN: https://github.com/tensorlayer/srgan Authors: Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang
5. ESRGAN: https://github.com/xinntao/ESRGAN Authors: Xinntao Yuan, Yanmei Jin, and Shengfeng He
6. A-ESRGAN: https://github.com/xinntao/Adaptive-ESRGAN Authors: Xinntao Yuan, Ke Yu, and Shixiang Wu
7. RealSRGAN: https://github.com/yhongwei/RealSR-PSNR Author: Hongwei Yong
8. DIV2K Dataset, Agustsson, E., & Timofte, R. (2017). NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops.
9. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, and A. Tejani, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 105–114. (SRGAN)

Backgrounds

1. **Dushyant Singh Udawat**

I am an Industrial engineering graduate student specializing in Data Analytics. My passion lies in implementing complex Mathematical models and Machine Learning algorithms to solve real world problems. I chose this project because I wanted to learn how to work with Image data, how to test CNN models, and how to apply Attention models, which I believe I will be able to accomplish through the project.

2. **Pratikshit Singh**

I'm a Computer Science graduate student currently focusing on data mining and machine learning. This semester I'm working with computer vision and generative models through courses and relevant projects because of my interest in state-of-the-art technology and models, to learn about underlying implementations of GPTs, memory-networks, and diffusion models. This project and course will give me exposure and opportunity to work wih CNNs, attention networks, GANs and real world applications of techniques like super-resolution and object detection.

3. **Christopher Cai**

I am a Computer Science MCS student interested in computer vision and machine learning. I am familiar with building CNN models with PyTorch and have completed a few projects using them for image classification and semantic segmentation. However, I have limited knowledge on GAN networks and attention which are becoming increasingly popular, so working on this project will give me an opportunity to learn about how GANs work and how to apply them to real world problems.