Exploration of Super Resolution Image Enhancement

Alexander Nishio

Introduction

Super-resolution is a image processing technique in the field of image resolution, where the resolution of a low-quality image is enhanced. It plays a critical role in various fields in medical, satellites, and space backgrounds for scientists to analyze microscopic or distant areas at a higher level. As technological and scientific advancements continue to improve, there is an increase in demand

Background

I explored two models in this project: Real-ESRGAN and SUPIR, which are two of the best super resolution models in their own domain. Below is a general overview of the two models:

Real-ESRGAN (GAN Model)

- Trained using a Generator and a Discriminator that competes with each other
- Generator tries to create fake images that can fool the Discriminator
- Discriminator tries to guess the real image between the generated image and the actual image
- Runs into issues if the Discriminator or the Generator is not effective enough

SUPIR (Diffusion Model)

- Adds noise to the original image and slowly denoises the image to learn complex patterns
- Adjusts loss to make sure the image is as high quality as possible
- There is an attached text prompt option to help guide the model to the correct output
  - Can plug in a Natural Language Processing Model for nice results
- Extremely computationally expensive, and the model is very large

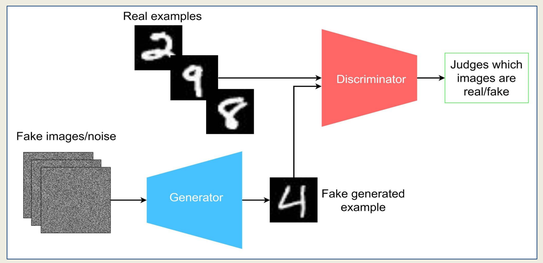

Figure 1: Diagram of a Generative Adversial Network (GAN) Model

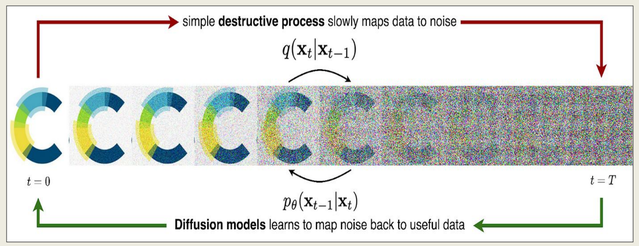

Figure 2: Diagram of a Diffusion Model

Results

Below are the results highlighting the advantages SUPIR has over Real-ESRGAN:

Advantages of SUPIR:

- The image enhancement is much more effective compared to Real-ESRGAN, as shown in the bus stop, the mountain scenery, the road, and the grass
- Renders images relatively fast (~1 minute) for the quality generated

Disadvantages of SUPIR:

- Extremely large model (30-40 GB) and requires GPU
- The compute time is extremely long if the resolution is too high
- Does not render text very well
- Can easily produce unrealistic results if no prompt is provided

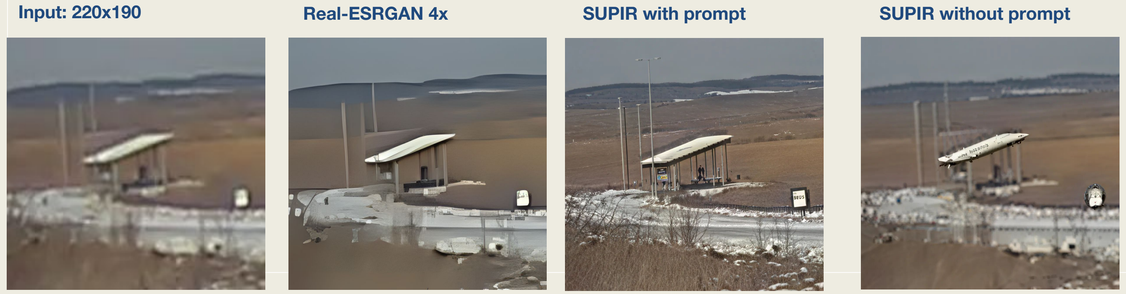

Figure 3: Comparison of Results between the Two Models

Summary and Conclusions

We explored the advantages and disadvantages of

This research was supported in part by the U.S. Department of Energy (DOE), Omni Technology Alliance Internship Program. The program is championed by the DOE's Office of Chief Information Officer (OCIO) and represents a partnership with the leadership of the Office of Economic Impact and Diversity, the Office of Science, the Office of Nuclear Energy, and the National Nuclear Security Agency. The program is administered by the Oak Ridge Institute for Science and Education.