Diffusion Modded Image Generator

—

By Tariq A.

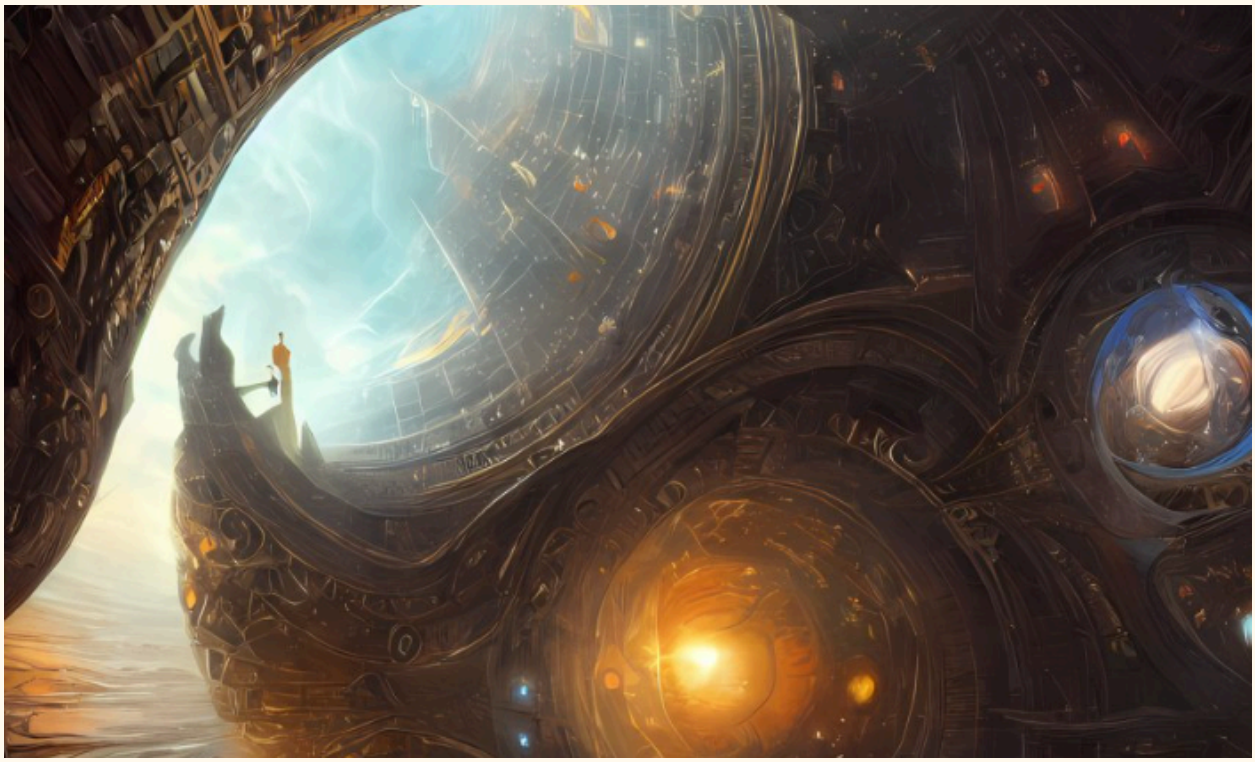
Building a standalone custom queue-based, scalable, multi-user product for stable diffusion finetuning and image generation.

What is this DMIG project?

The Diffusion Modded Image Generator is a cutting-edge, standalone, custom-built product designed to facilitate scalable, multi-user stable diffusion fine-tuning and high-quality image generation. This project encompasses the complete lifecycle of model development, from data preprocessing and model training to deployment and documentation, ensuring robust performance, transparency, and accessibility.

What was the DMIG project's goal?

1. Develop a scalable, multi-user image generation system using stable diffusion.
2. Ensure a high-quality and diverse image outputs through meticulous model fine-tuning.
3. Document the entire development process for transparency and reproducibility.
4. Deploy the model on a widely accessible platform to ensure robust performance and ease of use.



How did you get the data for DMIG?

The initial dataset comprised a diverse collection of high-resolution images sourced from various public image repositories, such as Unsplash, ImageNet, and COCO. The dataset included a wide range of categories, such as nature, urban landscapes, portraits, and abstract art, to ensure the model's ability to generate varied and high-quality images.

Some key insights

This project often downscale images during preprocessing but trained on high-resolution data to retain fine details.

How did you clean the data?

Surprisingly there is no magic here, just cold hard steps to perform over and over again, till you get the *cleaner* data. Steps like:

1. **Removing Duplicates:** Identified and removed duplicate images to prevent the model from overfitting to redundant data.
2. **Filtering Out Low-Quality Images**: Applied automated filters to remove images with poor resolution, excessive noise, or irrelevant content.
3. **Normalization:** Standardized the images to a consistent size and aspect ratio to ensure uniformity across the dataset.
4. **Augmentation:** Used data augmentation techniques like random cropping, rotation, flipping, and color jittering to artificially expand the dataset and improve model generalization.

Some key insights

- We employed automated scripts to detect and filter out low-quality images. This approach is used by projects like Google’s DeepDream.
- We implemented a wide range of augmentation techniques to help models learn from more varied data, such as StyleGAN.

Which Evaluation Metrics were used?

1. **Inception Score (IS):** Measures the quality and diversity of generated images by evaluating how well the generated images can be classified by a pre-trained Inception model.
2. **Frechet Inception Distance (FID):** Assesses the similarity between the generated images and real images by comparing the statistics of the generated images to those of the real images. Lower FID scores indicate better quality.
3. **Human Evaluation:** Conducted qualitative assessments where human evaluators rated the visual quality and relevance of the generated images.

Which models were used on this data?

1. CNN
2. Stable Diffusion
3. Transfer Learning