

Survey Existing Research and Reproduce Available Solutions

GANs:

Initially looked into using a GAN based on Pokemon GAN video from course:
<https://www.youtube.com/watch?v=yz6dNf7X7SA>

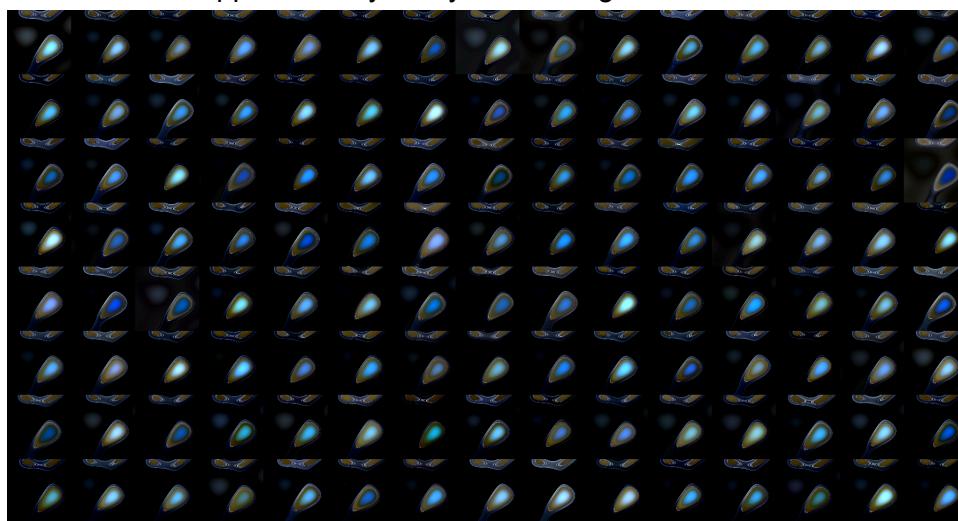
Which led to the following project:
<https://github.com/Zhenye-Na/pokemon-gan>

After surveying more recent research into GANs, a more modern architecture was discovered: StyleGAN. The most recent being StyleGAN3:
<https://github.com/NVlabs/stylegan3>

After testing for a 20 generations, results looked promising:



However, after approximately 4 days of training, results at 531 looked less promising:



DALL·E

With DALL·E, using the prompt “Create an image of a perfume on a white background” was able to achieve good results:



However, DALL·E is a closed system and so has no access to use for custom training dataset.

Stable Diffusion

In order to test out SD (Stable Diffusion), had to install an application to use an SD model on local machine. Chose to use DiffusionBee due to support for local machine architecture (M1 Mac).

Results were somewhat promising with the same prompt, “Create an image of a perfume on a white background”:

A screenshot of the DiffusionBee application. On the left, there is a sidebar with icons for home, library, settings, and others. The main area has a text input field containing the prompt "Create an image of a perfume on a white background". Below it is another text input field labeled "Enter your negative prompt here". Underneath are "Advanced Options" with a toggle switch turned on, "Model" set to "Default_SD1.5", "Resolution" set to 512x512, "Number of images" set to 4, "Seed" set to -1, and "Sampling Steps" set to 25. At the bottom is a blue "Generate" button. To the right, there are four generated images arranged in a 2x2 grid. The top-left image shows a hand holding a dark perfume bottle with white flowers floating around it. The top-right image shows a green leafy branch next to a small orange bottle. The bottom-left image shows a black perfume bottle on a purple floral background. The bottom-right image shows a clear perfume bottle with a pink label on a purple background.

After training with a small subset of the dataset (5 images), results were very promising with the following prompts:

Positive prompt: "a single bottle of perfume on a white background"

Negative prompt: "lowres, text, error, cropped, worst quality, low quality, normal quality, jpeg artifacts, signature, watermark, username, blurry"

