# Fine Tuning a Diffusion Model For Doodles Generation

Yuanhe Guo, Yiling Cao

## Introduction

Text-driven image generation AI has become a hot spot again since 2022, with the emergence of Dall-E2[1], stable diffusion[2], midjourny[3] and so much more. Other than using GAN, these projects open their arms to diffusion models. As is explained in this paper[4] by researchers in stability AI, they combine CLIP[5] with guided diffusion and create stunning results. Most state of the art diffusion models are integrated into the Hugging Face diffusers pipeline[6], and there's an open source webui for image generation[7].
In our project, we want to dive deep into image generation with diffusion models and try different fine-tuning methods to generate images in doodle styles.

## Objectives

We will use the quickdraw dataset by google creative lab, "a collection of 50 million drawings across 345 categories"[8]. We probably need to convert the data from ndjson format to jpg or png, and append more text labels.
With this dataset, we plan to implement the pipeline on the Hugging Face, make some improvements if possible, and train our model checkpoint.
We expect to get a fine-tuned diffusion model that specializes in generating doodles.

## References

[1] "Dall-e2." [Online]. Available: https://openai.com/product/dall-e-2

[2] "Stable diffusion public release." [Online]. Available: https://stability.ai/blog/stable-diffusion-public-release

[3] "Midjourney." [Online]. Available: https://www.midjourney.com/

[4] A. Nichol, P. Dhariwal, A. Ramesh, P. Shyam, P. Mishkin, B. McGrew, I. Sutskever, and M. Chen, "Glide: Towards photorealistic image generation and editing with text-guided diffusion models," 2022.

[5] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," 2021.

[6] "Diffusers." [Online]. Available: https://huggingface.co/docs/diffusers/main/en/index

[7] AUTOMATIC1111, "stable-diffusion-webui." [Online]. Available: https://github.com/AUTOMATIC1111/stable-diffusion-webui

[8] "Quickdraw-dataset." [Online]. Available: https://github.com/googlecreativelab/quickdraw-dataset