

# Dreambooth Fine-tuning for Stable Diffusion

Monil Lodha

October 22, 2024

## 1 Introduction

This report documents the process of fine-tuning Stable Diffusion using Dreambooth. Dreambooth allows for teaching new concepts to the model using a small set of images.

## 2 Pretrained Model Used

Stable Diffusion is a latent text-to-image diffusion model capable of generating photo-realistic images given any text input. It combines an autoencoder with a diffusion model that is trained in the latent space of the autoencoder. I used the Stable Diffusion v1-4 model available freely on the web as `CompVis/stable-diffusion-v1-4`.

## 3 Dataset

The dataset I have used is one of the subjects from the official provided by the paper Dreambooth Dataset. The dataset used for fine-tuning consists of images downloaded from the URLs provided in the code. The images represent a mini yellow alarm clock, which is the new concept being introduced to the model. A total of 4 images are used. Prior preservation is not used.

## 4 Implementation Details

The implementation is based on the Hugging Face Diffusers library. The Stable Diffusion v1-4 model is used as the base model. The code includes steps for:

- Installing necessary libraries.
- Setting up the `DreamBoothDataset` and `PromptDataset` classes.
- Generating class images (if prior preservation is enabled).
- Loading the Stable Diffusion model.
- Setting up training arguments.
- Defining the training function.
- Running the training process using `accelerate.notebook_launcher`.
- Setting up the pipeline for inference.
- Running the Stable Diffusion pipeline to generate images using various prompts.

Key hyperparameters used in the training process include a learning rate of  $5e-06$ , a maximum of 100 training steps, a batch size of 2, and gradient accumulation steps of 2.

## 5 Testing and Results

The results of the fine-tuning process are presented in the form of images generated by the model. The images are generated based on a list of prompts, including:

- a mini yellow alarm clock in the jungle
- a mini yellow alarm clock in the snow
- a mini yellow alarm clock on the beach
- a mini yellow alarm clock on a cobblestone street
- a mini yellow alarm clock on top of pink fabric
- a mini yellow alarm clock on top of a wooden floor

The generated images, along with the prompt used can be viewed in the final cell of the python notebook provided.



Figure 1: a mini yellow alarm clock in the jungle



Figure 2: a mini yellow alarm clock in the snow