

# Universal Guidance for Diffusion Models

Dhruv Misra  
Roll Number: 210341

October 22, 2024

## 1 Overview

The paper "Universal Guidance for Diffusion Models" introduces an algorithm that enables diffusion models to be controlled by various guidance functions without retraining. It uses pre-trained models and applies guidance during the image generation process, allowing for flexibility with different input types like segmentation maps and face recognition. The approach uses forward and backward guidance to adjust the generated images and stepwise refinement to improve quality. The official GitHub repository link is: <https://github.com/arpitbansal297/Universal-Guided-Diffusion>

## 2 Implementation Details

The GitHub repository for the paper is inconsistent in terms of versions given in the requirements of the installations recommended, one particular example being the version requirements for '**typing-extensions**'. Some installations needed a 'typing-extensions' version greater than 4.10 while some required less than 4.6, which mandated the downscaling of certain versions of the requirements to at least try and get the code to run.

I began by cloning the repository on my terminal since it required a conda environment that doesn't run on kaggle/colab.

Instead, if you don't want to create a conda environment like when I tried running on colab/kaggle, download the requirements file. Otherwise, one can proceed with the README.md file instructions. The code as provided in the repository does not directly run. I have added the changes I have made in the README file as well, such as changing the '**num\_workers**' parameter to adapt to the laptop GPU requirements and decreasing the number of '**ddim\_steps**' and '**optim\_forward\_guidance\_wt**' to check if the code was stopping due to excessive computation requirements. But, that does not seem to be the case.

## 3 Dataset Description

The code runs on the dataset provided already in the repository, it is present in the GitHub link forked in its **/data** folder, and the python scripts provided have in-built functions to extract the same. The face detection and other algorithms run on standard **512 X 512 .png** images.

## 4 Results

The code could not run properly and, hence, did not generate the complete results. However, the code does create the result folder and loads the model, as will be evident upon running it, and malfunctions on one of the next steps. I am continuing to investigate the reasons behind this and will bring it to the evaluator's notice if some concrete evidence is found.

Alternatively, I also attempted to build a code for this **from scratch** and built the functions for forward as well as backward guidance, but it runs into errors on pipeline integration. The work is also attached separately in the '**Additional**' folder of the repository.