LDEdit: Towards Generalized Text Guided Image Manipulation via Latent Diffusion Models
Paramanand Chandramouli, Kanchana Vaishnavi Gandikota
University of Siegen

LDEdit - Introduction

Goal: To develop a fast and flexible approach to open domain image manipulation from text prompts.

Our Solution: Adapt pretrained text-to-image latent diffusion model to perform text guided manipulations using DDIM sampling.

Advantages:
- Faster manipulation in lower dimensional latent space.
- DDIM sampling ensures a near cycle-consistency between source and target.

Overview of LDEdit

Visual Comparisons

Face: Tanned, Zuckerberg
Dog: Bear, Fox
Tennisball: Baseball, Orange
Strike: van Gogh, Picasso
VGGAN+CLIP [1] Ours

Local Editing with Mask Inputs

Input: girl + dog
Mask: girl + monkey
Output: girl + dog + monkey

Simultaneous editing of multiple attributes

Left to right: (i) Input (ii) girl with a dog (iii) woman with a dog (iv) old woman with a cat (v) boy with a tennisball

References