



- home
- papers
- study
- life
- About me

Custom Diffusion

- Multi-Concept Customization of Text-to-Image Diffusion
- <https://arxiv.org/abs/2212.04488>
- CVPR 2023

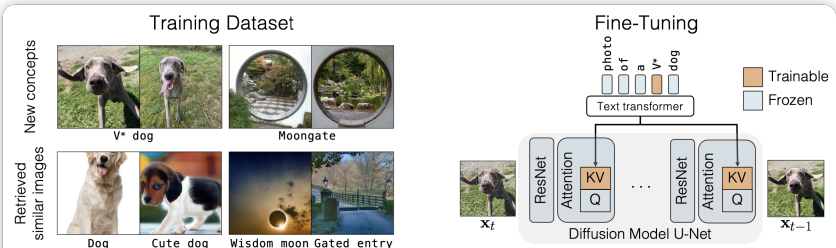


Figure 2. **Custom Diffusion.** Given images of new concepts, we retrieve real images with similar captions as the given concepts and create the training dataset for fine-tuning, as shown on the left. To represent personal concepts of a general category, we introduce a new modifier token V^* , used in front of the category name. During training, we optimize key and value projection matrices in the diffusion model cross-attention layers along with the modifier token. The retrieved real images are used as a regularization dataset during fine-tuning.

Custom Diffusion

cross-attention KV

cross-attention KV token token embedding

CLIP caption

KID

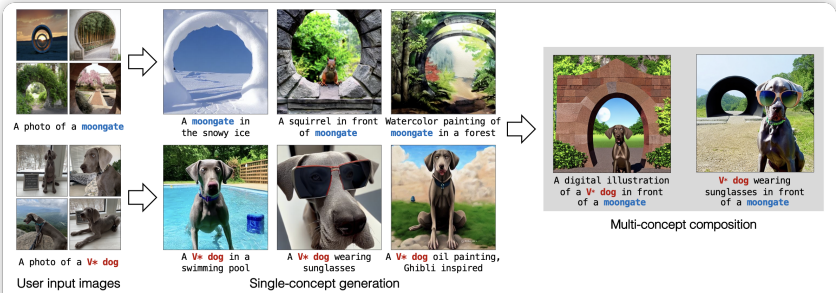


Figure 1. Given a few images of a new concept, our method augments a pre-trained text-to-image diffusion model, enabling new generations of the concept in unseen contexts. Example concepts include personal objects, animals, e.g., a pet dog, and classes not well generated by the model, e.g., moongate (a circular gate [83]). Furthermore, we propose a method for composing multiple new concepts together, for example, V* dog wearing sunglasses in front of a moongate. We denote personal categories with a new modifier token V^* .

Newer

Older

2024-08-23
ELITE

2024-08-21
DreamBooth

leicheng © 2022-2025

Archive RSS feed GitHub Email QR Code

Made with Montaigne and bigmission