PreciseControl

- □□□□□PreciseControl: Enhancing Text-To-Image Diffusion Models with Fine-Grained Attribute Control
- □□□□□https://arxiv.org/abs/2408.05083
- ECCV 2024

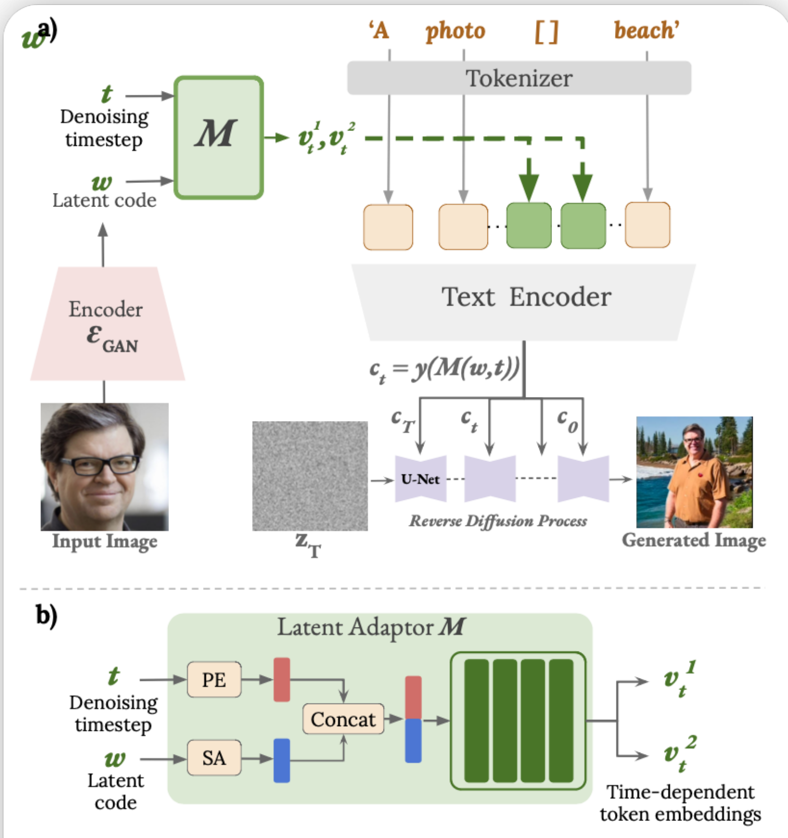

Fig. 2: Framework for personalization. Given a single portrait image, we extract its w latent representation from encoder $\mathcal{E}_{GAN}$. The latent w along with diffusion timestep t are passed through the latent adaptor $\mathcal{M}$ to generate a pair of time-dependent token embeddings (v_t^1, v_t^2) representing the input subject. Finally, the token embeddings are combined with arbitrary prompts to generate customized images.

□□□□□□w+ Adapter□□□□□□□StyleGAN□□w+□□w+□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□w+□□□□□□□□□□□□□□□□□□□□t□□□□□□□□□embedding□□□□□□□□□□□□□□□□□□t□□□□□□□□□embedding□

□□□□□□□□□□□□□□W□□□□□□□□□□□□□□□□□□□□□□W□□□□□□UNet□□LoRA□□□□□□□□□□□□□□□□□□□□□□□□

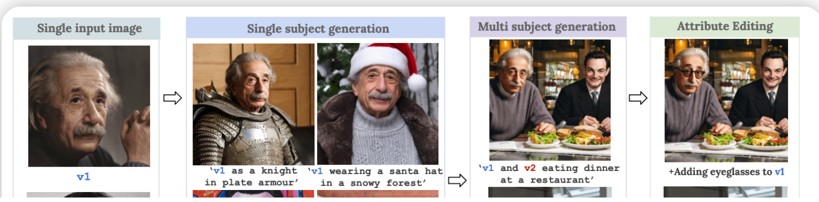

Fig. 1: Given a single portrait image, we embed the subject into a text-to-image diffusion model for personalized image generation. The embedded subject can then be transformed or placed in a novel context using text conditioning. The proposed method can also compose multiple learned subjects with high fidelity and identity preservation. To obtain precise inversion of face, we condition the T2I model on the rich $\mathcal{W}+$ latent space of StyleGAN2. This enables our method to additionally perform fine-grained control over the generated face with continuous control over facial attributes such as age and beard.

- □□□□□□□□□□□□StyleGAN2□FFHQ□□□□W□□□□□□□□
- □□□CLIP□□□□□□□
- □□□□□□□
- □□□https://rishubhpar.github.io/PreciseControl.home/

Newer Older

2024-11-07
AnyDoor

2024-10-28
CosmicMan