E4T

- □□□□□Encoder-based Domain Tuning for Fast Personalization of Text-to-Image Models
- □□□□□https://arxiv.org/abs/2302.12228
- TOG 2023

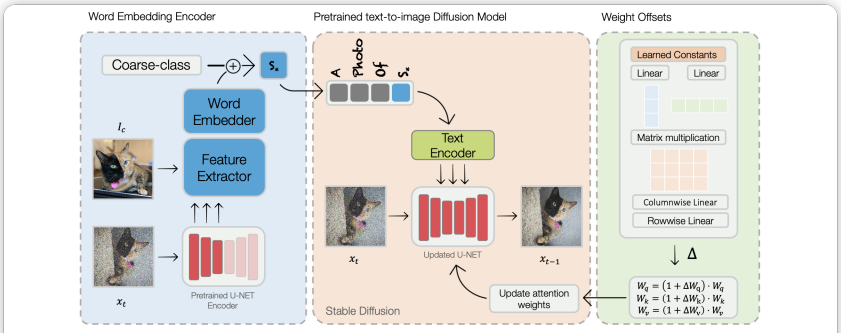

Figure 2: Method overview. Our method efficiently adapts pre-trained diffusion models (middle) to handle novel concepts. We train two components, a word-embedding encoder (left) and weight offsets (right). The encoder predicts a new code in the diffusion model's embedding space which best describes the input concept. In addition, we optimize learned weight offsets to specialize the text-to-image model to the target domain (*e.g.* cats). Our offsets are learned constants, regularized through a neural network. During training, we use $x_0 = I_c$. Here, we show inference, where x_T is sampled from the Gaussian prior as usual and hence x_t is no longer a noisy version of I_c.

□□□□□□□□□□□

□□□□□□□□□□□