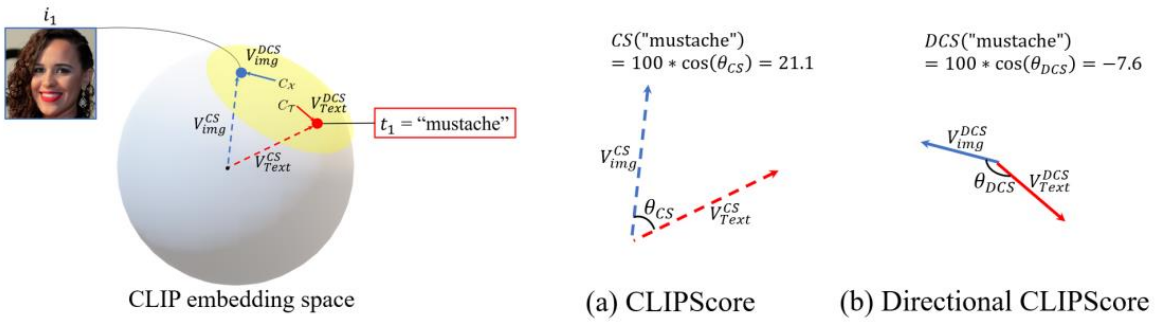
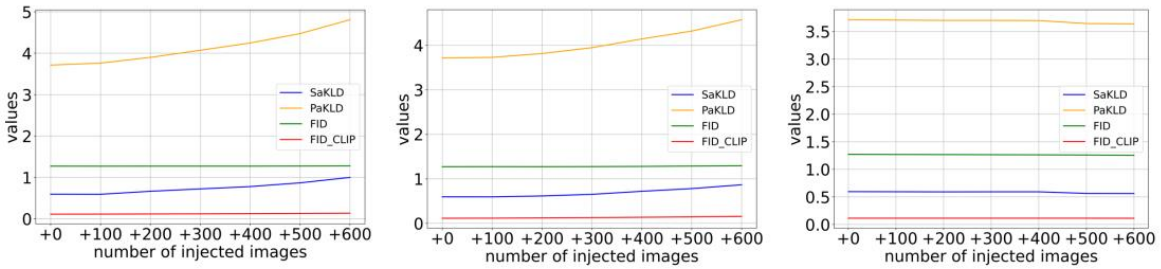
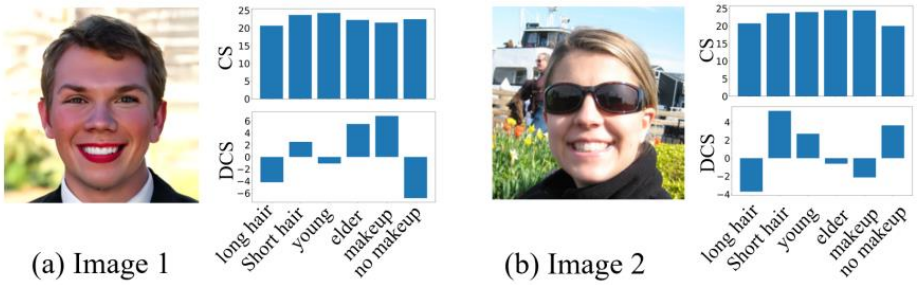


Figure 1: **Illustration of CLIPScore and Directional CLIPScore.** (a) CLIPScore evaluates the similarity between V_{img}^{CS} and V_{text}^{CS} from the coordinate origin, where the angle between the two vectors is bounded, resulting in a limited similarity value. (b) Directional CLIPScore gauges the similarity between V_{img}^{DCS} and V_{text}^{DCS} using the defined means of images $C_{\mathcal{X}}$ and texts $C_{\mathcal{T}}$ as the origin, the range of similarity is unrestricted.



(a) "man&bang" injection (b) "man&makeup" injection (c) normal image injection

Figure 2: **Validation of metrics by including correlated images.** We set two distinct subsets of the FFHQ dataset (30,000 images each), and linearly injected correlated images hard to find into a subset (a)(b), and normal images (c). While SaKLD and PaKLD captures distribution changes, FID and FID_{CLIP} score remained stagnent.



(a) Image 1 (b) Image 2

Figure 3: **Scores for opposite attributes.** We examined CS and DCS values for opposite attributes such as "long hair" and "short hair". In both images, Directional CLIPScore shows opposite directions for opposite attributes.

Figure 4: **Samples from ProjectedGAN.** While the FID of ProjectedGAN is comparable to major models, SaKLD and PaKLD of ProjectedGAN are significantly inferior because it fails to capture distributions of attributes, generating such a baby with beard.