Text-guided Image Generation and Manipulation Based on Generative Adversarial networks

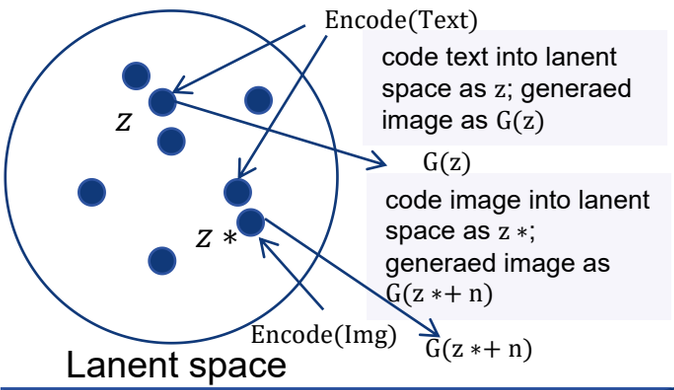
Yiqi LIANG, Yang ZHOU
Northeastern University

Poster
Outstanding Graduates

Introduction

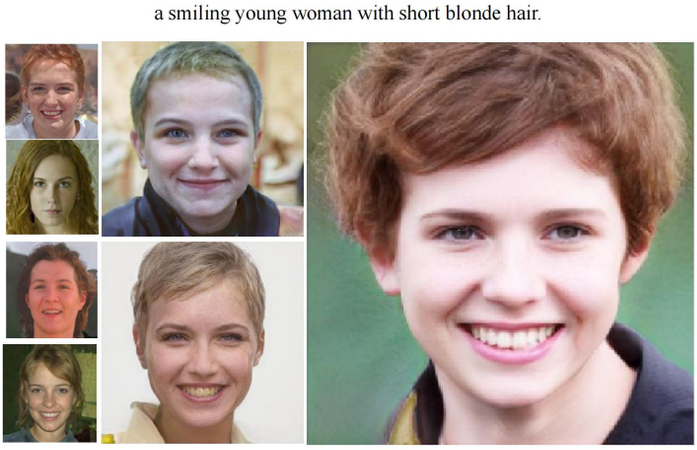
We build a novel framework based on text description to generate and manipulate portrait images. Our method uses GAN Inversion to convert a given image back to the latent space of a pre-trained GAN model, so that the image can be reconstructed by the reverse encoding of the generator

Method

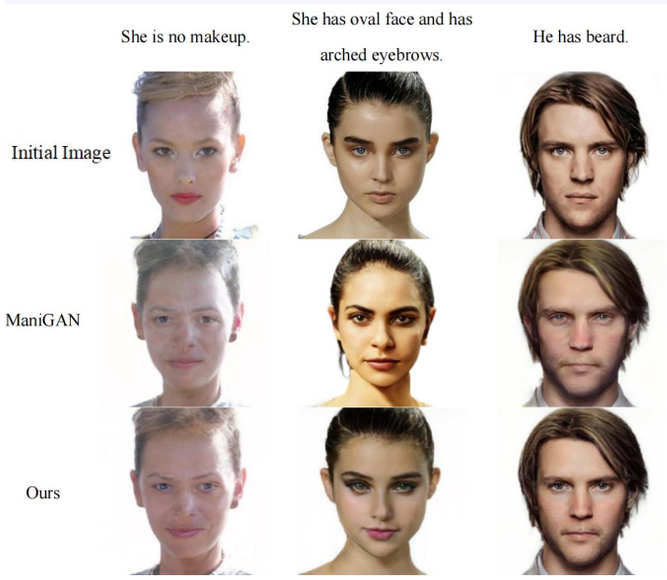


Results

The result of text-guided image generation is shown in Figure. It can be seen that the features of the generated image are basically aligned with the text, and the relationship between text and image is one-to-many, so one sentence of descriptive text can generate multiple corresponding face images.



In text-guided image editing, our work retains more details that are the same as the input image, that is, the areas not related to the text have not changed too much. This verifies that our model has good robustness.



Comperation

Medthod	FID⬇	Acc.(%)⬆
AttnGAN	110.62	19.7
ControlGAN	105.35	20.1
DM-GAN	112.78	17.7
DF-GAN	117.13	15.8
Ours	104.23	23.2

Conclusion

We build a framework that could tackle two different tasks (text-guided image generation and manipulation) and achieves high accessibility, diversity, controllability for portrait image generation and manipulation.