The Euclidean Space is Evil: Hyperbolic Attribute Editing for Few-shot Image Generation

Lingxiao Li[1], Yi Zhang[2], Shuhui Wang[3]*

[1]Columbia University, [2]Oxford University, [3]Institute of Computing Technology, Chinese Academy of Sciences

ICCV23 PARIS

Motivation

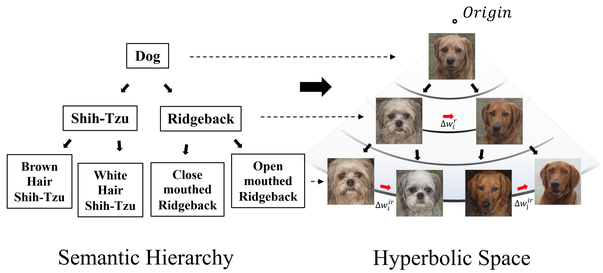

Figure 1: Illustration of hierarchical attribute editing in hyperbolic space.

- Existing methods suffer from the trade-off between the quality and diversity of generated images.
- As Figure 1 shows, like the ubiquity of hierarchies in language, the semantic hierarchy is also common in images.
- Hyperbolic space can naturally and compactly encode hierarchical structures, which is illustrated in Figure 2.
- HAE can generate new images of unseen categories by moving the latent code in the proposed latent space.

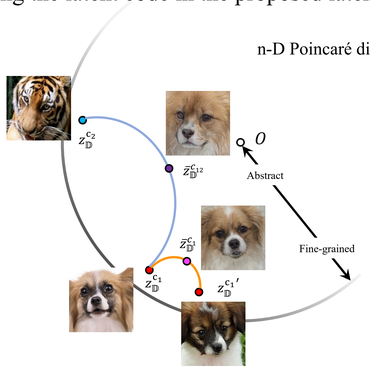

Figure 2: Illustration of the property of hyperbolic space on the Poincaré disk.

Methodology

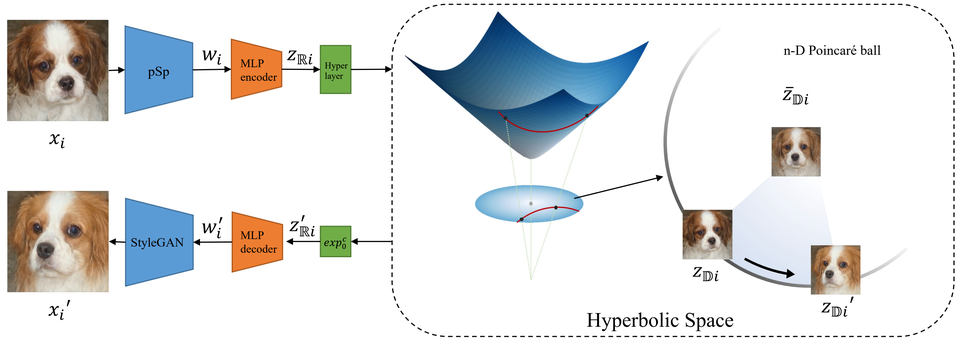

Figure 3: The overview of HAE.

The Hyper layer is a hyperbolic feedforward layer called Möbius linear layer which is used to project the latent code from Euclidean space $\mathbb{R}^n$ to hyperbolic space $\mathbb{D}^n$. $\bar{z}_{\mathbb{D}i}$ can be viewed as the "parent" or average code of $z_{\mathbb{D}i}$ and $z_{\mathbb{D}i}{}'$. One can generate diverse images without changing the category by moving the latent code from one child to another of the same parent in the hyperbolic space.

Experiments

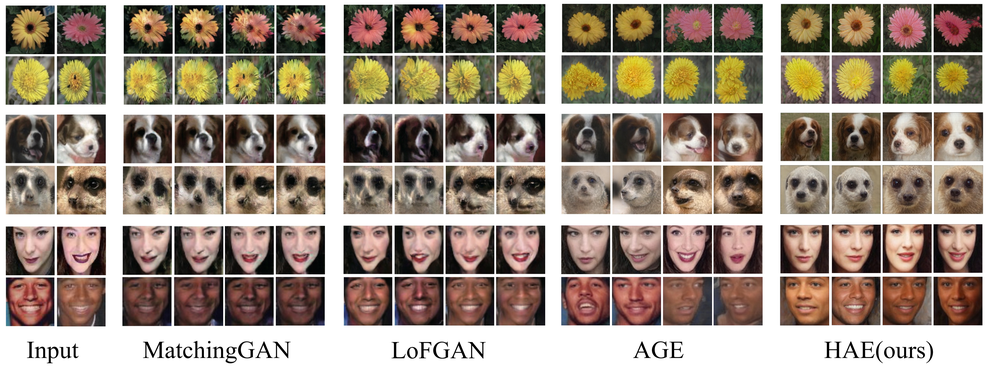

Figure 4: Comparison between images generated by MatchingGAN, LoFGAN, AGE, and HAE on Flowers, Animal Faces, and VGGFaces.

Figure 5: Images generated by HAE by adding the same perturbation on the latent code of a given image with different hyperbolic radii.

Method	Settings	Flowers		Animal Faces		VGG Faces*	
		FID(↓)	LPIPS(↑)	FID(↓)	LPIPS(↑)	FID(↓)	LPIPS(↑)
DAWSON [39]	3-shot	188.96	0.0583	208.68	0.0642	137.82	0.0769
MatchingGAN [28]	3-shot	143.35	0.1627	148.52	0.1514	118.62	0.1695
F2GAN [30]	3-shot	120.48	0.2172	117.74	0.1831	109.16	0.2125
LoFGAN [23]	3-shot	79.33	0.3862	112.81	0.4964	**20.31**	0.2869
DeltaGAN [27]	1-shot	109.78	0.3912	89.81	0.4418	80.12	0.3146
Disco-FUNIT [29]	1-shot	90.12	0.4436	71.44	0.4511	-	-
AGE [15]	1-shot	45.96	0.4305	28.04	0.5575	34.86	0.3294
SAGE [14]	1-shot	**43.52**	0.4392	27.43	0.5448	34.97	0.3232
HAE (Ours)	1-shot	50.10	**0.4739**	**26.33**	**0.5636**	35.93	**0.5919**

Table 1: FID(↓) and LPIPS(↑) of images generated by different methods for unseen categories on three datasets.

Summary

- We propose a simple yet effective method for few-shot image generation. We use hyperbolic space as the latent space to capture the hierarchy among images. HAE is the first attempt to use hyperbolic latent spaces for few-shot image generation.
- We show that in our designed hyperbolic latent space, the semantic hierarchical attribute relations among images can be reflected by their distances to the center of the Poincaré disk.
- Extensive experiments and visualization suggest that HAE achieves stable few-shot image generation with state-of-the-art quality and diversity.

Code: https://github.com/lingxiao-li/HAE **Contact: ll3504@columbia.edu**