

Variational Autoencoders with Decremental Information Bottleneck for Disentanglement



People-Centred AI
UNIVERSITY OF SURREY

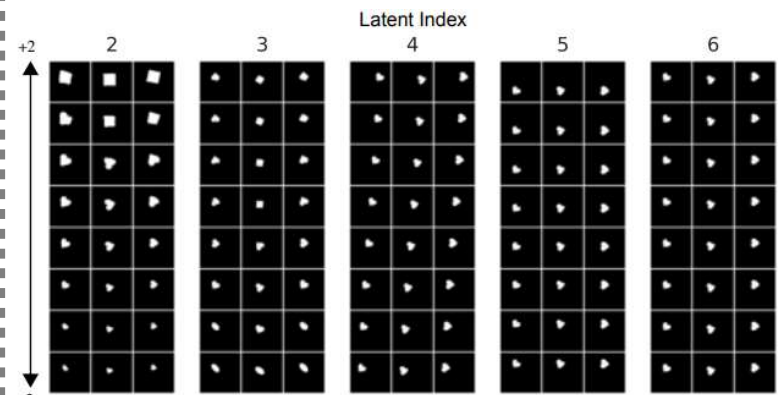
Jiantao Wu, Shentong Mo, Xingshen Zhang, Muhammad Awais Sara Atito, Zhenghua Feng, Lin Wang, Xiang Yang

We propose DeVAE with Hierarchical Latent Spaces (HiS) connected with Disentanglement-Invariant Transformations (DiT) to improve disentanglement and reconstruction simultaneously in a decremental information bottleneck principle.

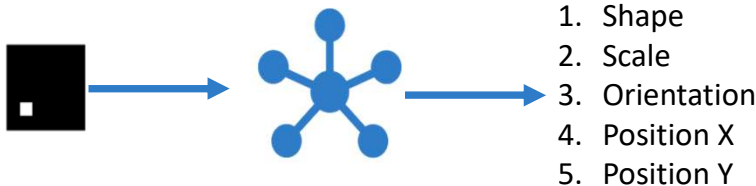
Experiment

dataset	model	MIG	DCI dis.	FactorVAE	Recon.
dSprites	DeVAE	0.34±0.02	0.53±0.02	0.80±0.03	48.31±27.98
	DynamicVAE	0.35±0.01	0.53±0.01	0.82±0.05	19.25±1.85
	β -TCVAE(12.0)	0.29±0.09	0.47±0.08	0.73±0.08	73.04±3.41
	β -VAE(6.0)	0.17±0.05	0.30±0.07	0.74±0.05	48.75±2.84
shapes3D	DeVAE	0.53±0.11	0.71±0.02	0.79±0.02	46.81±13.97
	DynamicVAE	0.54±0.04	0.68±0.03	0.87±0.10	31.02±3.56
	β -TCVAE(12.0)	0.49±0.11	0.73±0.07	0.78±0.01	44.53±5.69
	β -VAE(6.0)	0.42±0.18	0.68±0.06	0.82±0.06	34.95±2.34

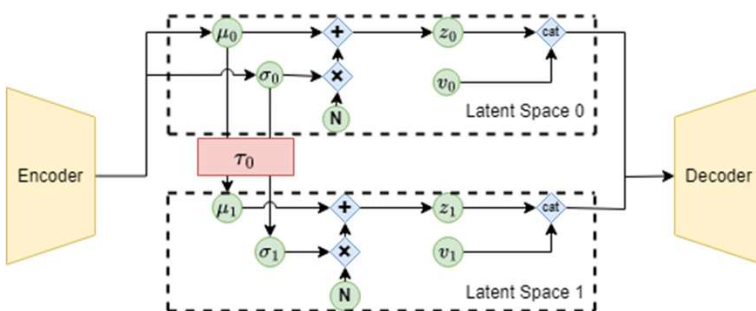
Table 1: Quantitative benchmarks on dSprites and shapes3D.



Disentangle Factors of Variance

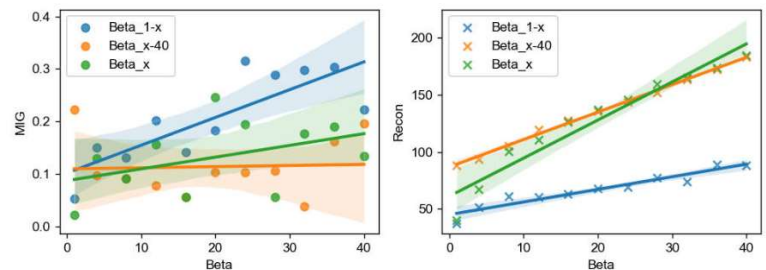


DeVAE



- ❖ Hierarchical latent spaces
- ❖ Disentanglement-invariant Transformation
- ❖ Decremental information bottleneck

Analysis



Conclusion

1. We propose DeVAE with a hierarchical structure to improve disentanglement and reconstruction fidelity simultaneously.
2. We develop a disentanglement-invariant transformation to guarantee that the representations in these latent spaces disentangle the same factors.
3. Our comprehensive experiments demonstrate the effectiveness of DeVAE.