
Abstract: Implementing the Introspective Variational Autoencoder

Nicholas Alonso

Department of Cognitive Sciences
University of California-Irvine
nalonso2@uci.edu

Hao-Che Hsu

Department of Economics
University of California-Irvine
haoche.hsu@uci.edu

Edward King

Department of Developmental and Cell Biology
University of California-Irvine
eking2@uci.edu

We will re-implement the introspective variational autoencoder (IntroVAE) [1] to synthesize realistic images. Current approaches for generating images often rely on generative adversarial networks (GAN), pure VAE models, or VAE-GAN hybrids. GANs are noted for their training difficulties and unstable gradients, VAEs produce blurry images due to their inexpressive loss functions, and hybrid models require complicated architectures that often perform worse than regular GANs. IntroVAE repurposes the inference model to additionally act as a discriminator, enabling the model to self-estimate differences between generated and real images in an adversarial manner without the inclusion of a separate discriminator network or sacrificing training stability. The model combines VAE and GAN loss functions to bypass the training difficulties associated with GANs, while also encouraging the creation of less blurry, more realistic images than are produced by VAEs. We hope to replicate and deliver comparable image quality to those presented in the research, and confirm the advantages of this model over standard VAEs and GANs.

References

- [1] Huaibo Huang, zhihang li, Ran He, Zhenan Sun, and Tieniu Tan. Introvae: Introspective variational autoencoders for photographic image synthesis. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 52–63. Curran Associates, Inc., 2018.