219013044, Buys Enock Onkarabile
Proposed Topic: Hybrid VAE-GAN for low-resolution face generation

Problem statement:
Face generation using generative models has gained significant attention in the tech scale, yet most research focuses on high-resolution outputs that need substantial computational resources. Generative Adversaria networks (GANs) do good at producing these realistic images, but then they suffer from training instability and mode collapse(Goodfellow et al., 2014). While Variational autoencoders (VAEs) provide stable training but then generate blurry images because of their reconstruction-based objective(Kingma and Welling, 2013).
Hybrid approaches that combine these two phenomena architecture show promise in addressing its these limitations by leveraging the stability of VAEs with the high-quality generation of GANs(Larsen et al., 2016).

Existing literature lacks comprehensive studies on the hybrid performance for low-resolution face generation (32x32 pixels) as most focus on natural images/high h resolution faces(VanRullen and Reddy, 2019), leaving a gap in knowing how they behave when computational resources are limited, and facial features must be preserved at low resolution. This research aims to fill this gap by evaluating hybrid VAE-GAN architectures for low resolution face generation.

Research Objectives:
I will evaluate hybrid VAE-GAN architectures for low-resolution face generation to address this gap. Primary objectives are:

1. To implement and compare hybrid VAE-GAN architectures against standalone VAE and GAN baselines for 32x32 face generation.
2. To evaluate different fusion strategies (early, late, and progressive fusion)
3. Provide quantitative analysis of the quality diversity trade-off in hybrid models for resource constrained face generation, filling the gap in performance benchmarks

Hypothesis: Hybrid VAE-GAN architectures will demonstrate superior performance compared to standalone VAE and GAN models for low-resolution face generation, specifically achieving better quality diversity trade offs measured by improved FID scores while maintaining training stability.

Methodology:
I will use three core components using Numpy and Pandas:

1. a VAE module with encoder-decoder architecture
2. a GAN module with generator-discriminator networks
3. hybrid fusion mechanism

The VAE encoder will compress 32x32x3 images to a 128-dimensional latent space, while the decoder reconstructs images from this presentation. The GAN generator transforms 100-dimensional noise vectors to 32x32x3 images, with a discriminator providing adversarial feedback. The CelebA dataset will be processed to 32x32 resolution for training and evaluation(Liu et al., 2015).

Expected Contributions:

Addressing the empirical gap in hybrid VAE-GAN literature by providing

1. systematic performance benchmarks for hybrid VAE-GAN architectures on low-resolution face generation
2. comparison of fusion strategies with quantitative results for the face generation domain
3. filling a specific gap in understanding how these models behave when computational resources are limited

References:

1. GOODFELLOW, I. J., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDE-FARLEY, D., OZAIR, S., COURVILLE, A. & BENGIO, Y. 2014. Generative adversarial nets. *Advances in neural information processing systems,* 27.
2. KINGMA, D. P. & WELLING, M. 2013. Auto-encoding variational bayes. Banff, Canada.
3. LARSEN, A. B. L., SØNDERBY, S. K., LAROCHELLE, H. & WINTHER, O. Autoencoding beyond pixels using a learned similarity metric. International conference on machine learning, 2016. PMLR, 1558–1566.
4. LIU, Z., LUO, P., WANG, X. & TANG, X. Deep learning face attributes in the wild. Proceedings of the IEEE international conference on computer vision, 2015. 3730–3738.
5. VANRULLEN, R. & REDDY, L. 2019. Reconstructing faces from fMRI patterns using deep generative neural networks. *Communications biology,* 2**,** 193.