

Analisis Kemampuan Beta-VAE pada Dataset yang Berbeda

Bramantya Purbaya¹, Bedy Purnama², Edward Ferdian³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹bramanp@student.telkomuniversity.ac.id, ²pembimbing1@telkomuniversity.ac.id,

³pembimbing2@telkomuniversity.ac.id

Abstract

Synthetic data has been a subject of several studies in machine learning cases, one of which involves generating new data due to the lack of existing data. However, producing and controlling various variations of input data distributions is still a matter of research. This research uses variation of the Variational Auto Encoder (VAE) method for generating synthetic data, namely Beta-Variational Auto Encoder (Beta-VAE). VAE itself is an unsupervised learning method capable of generating synthetic data, but the variation it produces are not as structured compared to Beta-VAE. In this research, the original Beta-VAE method utilized to generate synthetic data trained on four different datasets. PSNR, SSIM and FID score metrics are used to evaluate the Beta-VAE model. Comparing each Beta-VAE model trained with different datasets and analyzing each model. The result of the research found that the model trained with CelebA had the best result as seen from matrix evaluation.

Keywords: synthetic data, dataset, autoencoder, variational autoencoder
