Abstract

This project focuses on the challenge of enhancing image generation through the integration of adaptive augmentations within the latent space of Autoencoders (AEs). While the scarcity of diverse data often hinders model performance and scalability, data augmentation techniques offer a promising solution by generating diverse training examples from limited datasets. However, conventional augmentation methods may not fully meet the complex requirements of image generation tasks.

Recent progress in latent space enhancements has inspired this study to explore the possibility of utilizing acquired feature representations for conducting transformations that can capture essential data attributes while staying unaffected by irrelevant variations. Through implementing augmentations within the latent space of autoencoders, the objective of the framework is to enhance model robustness, generalization, and computational efficiency.

Key objectives include investigating existing Autoencoder architectures and latent space techniques, implementing a prototype framework incorporating adaptive augmentations, experimenting with augmentation strategies, optimizing the framework for efficiency and scalability, and validating its performance through quantitative metrics and qualitative assessments.

The project utilizes a dataset containing 200,000 images from the celebA dataset, with computational constraints limiting the analysis to 70,000 images. Training is conducted over 500 epochs, resulting in the generation of 128x128 images with a Fréchet Inception Distance (FID) score of 34.5. Through this work, we aim to contribute to the advancement of adaptive latent space augmentation techniques in AI image generation, paving the way for more robust, generalizable, and interpretable image generation models.