Dataset Selection and Key Findings in VAE Image Generation Study

Dataset Selection Rationale

This study employed two complementary datasets to investigate how data characteristics influence VAE performance:

Primary Dataset - CIFAR-10: Selected for its complexity as a benchmark containing 60,000 32×32 color images across 10 classes. CIFAR-10 presents challenges including varied backgrounds, complex object shapes, and RGB color information, making it an ideal testbed for evaluating VAE capabilities on natural images.

Secondary Dataset - Fashion-MNIST: Chosen as a contrasting dataset with 70,000 28×28 grayscale images of fashion items across 10 classes. This dataset offers cleaner backgrounds, more consistent foreground-background separation, and simpler grayscale representation, allowing for comparative analysis against the more complex CIFAR-10.

Key Findings

1. **Dataset Complexity Impact:**
    - Fashion-MNIST produced better reconstructions than CIFAR-10 due to its simpler grayscale format and more uniform structure
    - CIFAR-10's RGB data and complex backgrounds required significantly greater model capacity to reconstruct effectively
2. **Latent Space Properties:**
    - Larger latent dimensions improved reconstruction quality while reducing compression efficiency
    - Fashion-MNIST classes exhibited better separation in latent space, suggesting more distinct visual features
    - Silhouette scores provided quantitative confirmation of latent space organization differences
3. **β-VAE Experimentation:**
    - Increasing β values from 0.2 to 5.0 demonstrated the classic reconstruction-regularization trade-off
    - Higher β values produced more organized latent spaces at the cost of slightly degraded reconstruction quality
    - Unexpectedly, KL loss decreased with higher β values while reconstruction loss increased

4. **Visual Analysis:**

    - Both datasets produced blurry reconstructions typical of standard VAEs
    - Fashion-MNIST reconstructions retained more recognizable structural information
    - CIFAR-10 reconstructions captured color information but struggled with fine details

This comparative study highlights the fundamental importance of dataset characteristics in generative modeling. The findings suggest that generative model architectures should be tailored to specific dataset properties, with considerations for color complexity, background variation, and class distinctiveness. Future work should explore architecture modifications including deeper networks, attention mechanisms, and alternative loss functions to address the limitations identified.