Vector Quantized Variational Autoencoders on Novel Datasets

Shivanshu Gupta **Kolby Nottingham** **Preethi Seshadri**

Abstract

We explore using the Vector Quantized Variational Autoencoder (VQ-VAE) to generate discrete representations for the Kaokore dataset, which contains images of facial expressions from traditional Japanese illustrations (`https://github.com/rois-codh/kaokore`). The framework VQ-VAE is built on, Variational Autoencoders (VAE), learn continuous latent representations. While continuous representations are flexible, many real world attributes are better defined discretely, and some current state-of-the art model architectures, like transformers, only work with discrete data. Additionally, VAEs have been shown to exhibit posterior collapse, which means that latent codes are ignored. In this project, we experiment with VQ-VAEs on a novel dataset and design experiments to test the advantages and disadvantages of this approach in terms of generation quality and learned latent structure. Since the VQ-VAE paper uses CIFAR10 and 128x128 ImageNet images, there might be additional experimentation and training required to achieve good performance on the Kaokore dataset. We demonstrate the strengths and weaknesses of VQ-VAEs over traditional VAEs.

1 Introduction

Generative machine learning models learn...

One such generative machine learning model is the variational autoencoder (VAE). VAEs learn...

While traditional VAEs learn a continuous latent space, vector quantized variational autoencoders (VQ-VAE) learn discrete latent variables. Discrete variables have the advantage when it comes to...

In this work, we compare traditional VAEs to thier discrete counterpart, VQ-VAEs. We implement each of these methods and run experiemnts with each on the Kaokore dataset. To the best of our knowledge, this is the first time VAEs have been used to learn a latent space for this dataset.

2 Background

Generative models...

Autoencoders...

3 Variational Autoencoders

4 Vector Quantized Variational Autoencoders

5 Experiments

In a series of experiments, we compare traditional VAEs to VQ-VAEs. We first...

Preprint.

	Loss	Recon. Loss
VAE testset	0	0
VQ-VAE testset	0	0

Table 1: Training loss and reconstruction loss on the Kaokore testset for both algorithms.

Figure 1: Testset images (top) comared with images generated by VAE (middle) and VQ-VAE (bottom).

5.1 Kaokore Dataset

The Kaokore dataset consists of...

5.2 Results

As indicated in table...

6 Conclusion

The authors of the VQ-VAE paper presented promising results. However, bsaed on our experiments, we conclude that it is not strictly better than VAE on every dataset...