1 Introduction and Acknowledgements

This report compares the results of three density estimation algorithms: Variational Autoencoders (VAE), Generative Adversarial Networks (GAN) and Wasserstein-GANs (WGAN).

Models are built using the MNIST and CIFAR10 data sets.

As a preprocessing step, the words images are converted to tensors and normalized.

A module for loading the data (*load_data.py*) encapsulates the DataLoader functions of PyTorch. Thankfully, loaders for both of these datasets are available in the Torchvision package.

The VAE code is based on `https://github.com/pytorch/examples/blob/master/vae/main.py`

The GAN code is based on `https://pytorch.org/tutorials/beginner/dcgan_faces_tutorial.html`

Finally, the WGAN is based on `https://wiseodd.github.io/techblog/2017/02/04/wasserstein-gan/`

For each of these algorithms, I contrast generated images with real sample images and examine the plot of the loss functions during training.

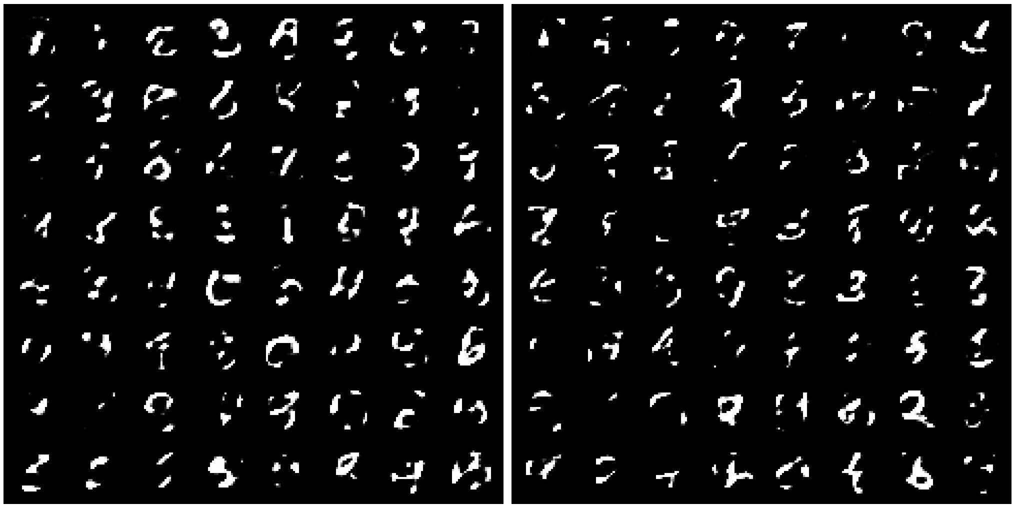
2 Variational Autoencoder

For this first VAE model (*mnist_vae.py* and *cifar10_vae.py*), the encoder and decoders are simple fully-connected structures with a bottleneck of dimension 20.

Figure 1: MNIST VAE Reconstructions After 20 Training Epochs

For the MNIST dataset, the reconstructions are fairly good after 20 training epochs. Sampling from the latent space, however, produces only vaguely numberlike shapes:

Figure 2: MNIST VAE-Generated Sample After 20 (left), 50 (right) Training Epochs

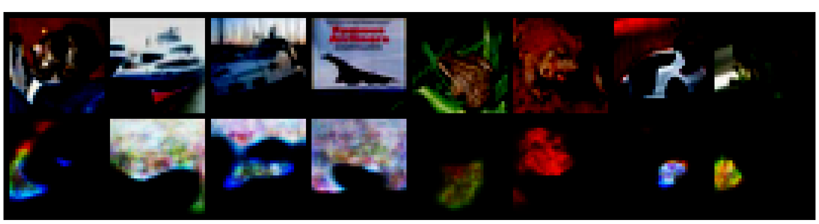


After an additional 30 training epochs, the reconstructions are slightly better, but again the generated images are not great.

The technique fared even worse on the CIFAR10 dataset, which has a greater variety and colour.

Figure 3: MNIST VAE Reconstructions After 50 Training Epochs

Figure 4: CIFAR10 VAE Reconstructions After 20 Training Epochs



Even after 50 training epochs, the VAE is unable to reconstruct the CIFAR10 images with any fidelity.

Figure 5: CIFAR10 VAE Reconstructions After 50 Training Epochs

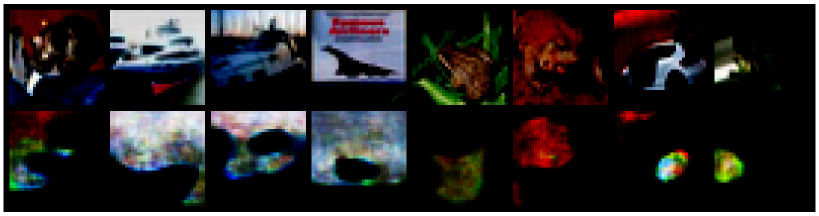
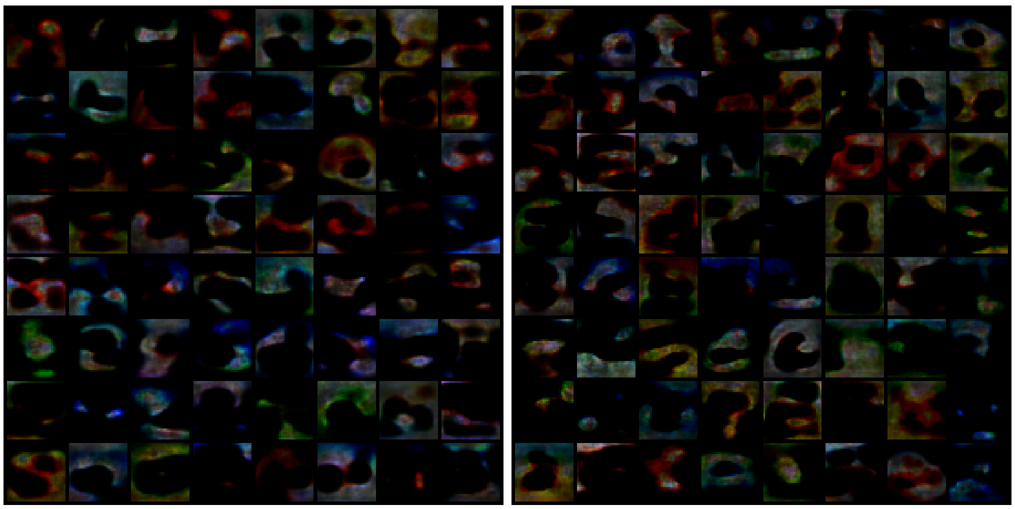
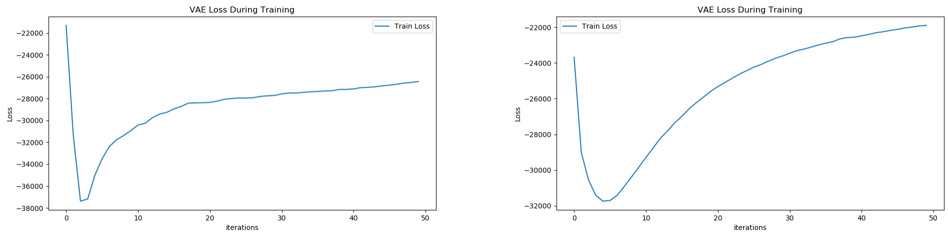


Figure 6: CIFAR10 VAE-Generated Sample After 20 (left), 50 (right) Training Epochs



For VAE, the loss function for consists of a cross-entropy term and the KL divergence term. Since both terms should be non-negative, I suspect that the KL term (which is negated in the function) is dominating the cross-entropy, leading to these negative values. The values do seem to be increasing, so perhaps additional training is warranted.

Figure 7: Plotting the Loss Function for VAE With MNIST (left) and CIFAR10 (right)



3 Generative Adversarial Network[1]

In my first GAN attempt, I modeled the generator and discriminator networks using fully-connected layers but the results were unimpressive.

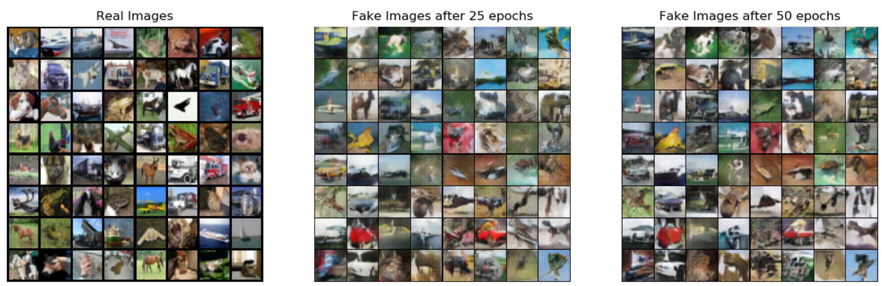
For my second attempt, I followed the tutorial linked in the intro which is based on [2]. In this paper, the authors implement the discriminator and generator using convolutions (and so-called transpose convolutions). Below are sample images generated after 25 and 50 training epochs.

Figure 8: DCGAN Generated MNIST Images After 25, 50 Epochs

With both datasets, the quality of images does seem to improve with longer training. The MNIST fakes, in particular, are fairly convincing.

However, even after 50 epochs the CIFAR10 generated images are of disappointing quality.

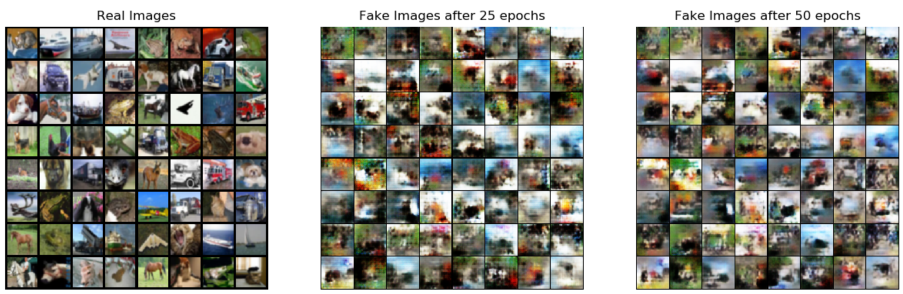
Figure 9: DCGAN Generated CIFAR10 Images After 25, 50 Epochs



4 Wasserstein GAN Variant

Figure 10: WGAN Generated MNIST Images After 20 Epochs

Figure 11: WGAN Generated CIFAR10 Images After 50 Epochs



5 Conclusion

I should have allotted more time for building the models, testing them and trying different parameters.

References

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[2] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.