ECS795P Deep Learning and Computer Vision, 2021

Course Work 2:
Unsupervised Learning by Generative Adversarial Network

1. **What is the difference between supervised learning & unsupervised learning in image classification task? (10% of CW2)**
   - In supervised image classification training stage is required, which means that first we need to select some pixels from each class which is called training pixels. Finding the characteristics of training pixels and also finding other pixels which have same characteristics the image classification can be done. Whereas in unsupervised image classification, no training stage is required, but different algorithms are used for clustering. In real world scenario we sometimes don't have much information about the data. So in that case we can use unsupervised image classification as here information is not required for classification.

2. **What is the difference between an auto-encoder and a generative adversarial network considering (1) model structure; (2) optimized objective function; (3) training procedure on different components. (10% of CW2)**

   1) Model Structure:
      Generative Adversarial Network has two neural networks, one for the generator and second for the discriminator which competes against each other in a min-max game.
      Whereas The Auto-Encoder has four components, The Encoder, Bottleneck, Reconstruction Loss and the Decoder. The architecture is an Artificial Neural Network which is a feedforward Network, LSTM or a Convolutional Neural Network.

   2) Optimized objective function:
      In Generative Adversarial Network, the Generator has the goal to maximize the probability of D making a mistake, whereas the goal of the discriminator is to minimize this probability using gradient descent.
      For example in LSGAN least squared error function is used instead of the standard loss function. The overall final goal is to reduce the cross-entropy.
      Whereas in an auto encoder, Reconstruction loss is used which is the difference between the original data and the image recreated using the

decoder. The goal of the auto-encoder is to minimise the Reconstruction loss using gradient descent.

3) Training Procedure:
In Generative Adversarial Network, the Generator model G receives noise as input, then the Generator produces the fake data which is then sent to the Discriminator model D. Then D tries to classify the input as fake or the real data. The accuracy of the Discriminator D with which it differentiates between the real and fake data is the misclassification rate. Then according to the classification made by the discriminator D back propagation is done to update the weights of the Generator model G and the discriminator model D.

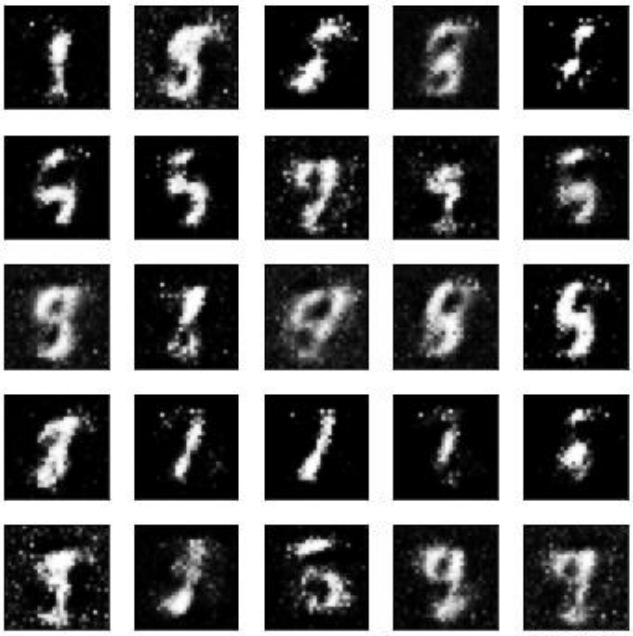
Auto-Encoder is a Sequential Neural Network that consists of two components Encoder and Decoder. The Encoder extracts features from the image, which is by reducing the image in height and width but simultaneously growing in depth. A latent representation of the image is made by the Encoder. Now the decoder decodes this latent representation and builds an image that satisfies our given criterion. The difference between the original and the recreated image is the Reconstruction loss. And then the network uses backpropagation to reduce this reconstruction loss.

3. **How is the distribution $p_g(x)$ learned by the generator compared to the real data distribution $p_{data}(x)$ when the discriminator cannot tell the difference between these two distributions? (15% of CW2)**

- Generator receives random noise as input, so $p_g(x)$ is a random distribution and as we continue training our generator it gets better at generating image close to the real data distribution. Finally, the global optimum exists when the discriminator cannot distinguish between real and generated data and at this global optimum $p_g(x) = p_{data}(x)$

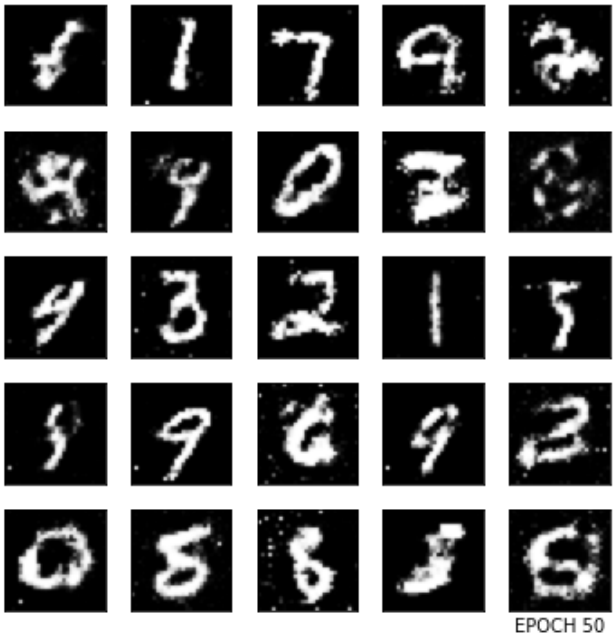
4. **Show the generated images at 10 epochs, 20 epochs,50 epochs,100 epochs by using the architecture required in Guidance. (15% of CW2)**

Generated image at 10 epochs.



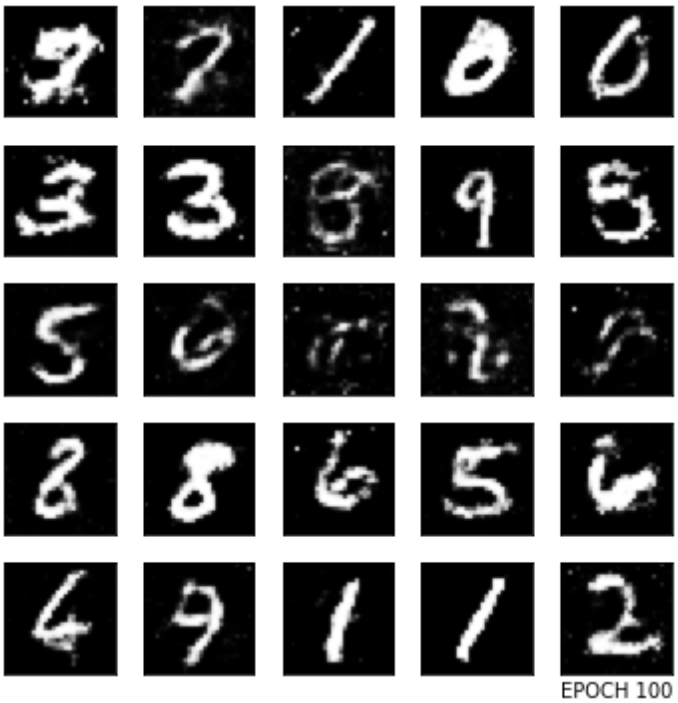
Generated image at 20 epochs.

Generated image at 50 epochs.



EPOCH 50

Generated image at 100 epochs.



EPOCH 100