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)**

   Supervised Image Classification: This method involves the user defining certain features, pixel groups and label classes to which the images would belong. The algorithms used tries to find the similarity of the input image and the class it belongs to.

   Unsupervised Image Classification: In this, each image in the dataset belongs to an inherent class in the dataset without the use of labelling data. Two major classes of algorithms; clustering algorithms and dimension reduction algorithms are used for the purpose of classification based on the similarity of features identified by the algorithms.

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:
      - The Generative Adversarial Network has two neural networks, a generator and a discriminator, which compete against each other in a minimax game.
      - The Auto-Encoder has four components: Encoder, Bottleneck, Decoder and Reconstruction Loss. The architecture is using an Artificial Neural Network, which can be a FeedForward Network, LSTM or a Convolutional Neural Network.

   2. Optimized objective function:
      - The Generator has the goal to increase the misclassification rate of the Discriminator, while the goal of the Discriminator is to decrease the misclassification rate, using gradient descent (eg LSGAN replace the standard loss function of the G and D with a least squared error function). The final goal of the entire network is to reduce the cross-entropy.

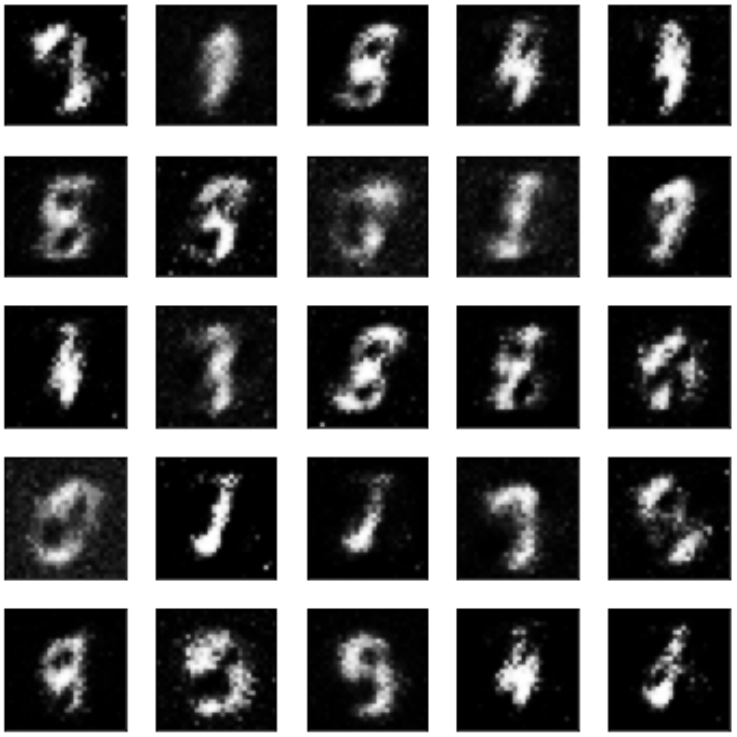
- In an autoencoder, we have Reconstruction Loss, which is the difference between the original data and the recreated image using the Decoder. The goal of the autoencoder is to minimise the Reconstruction Loss using gradient descent.

3. Training procedure:
   - **GAN**: The Generator network G receives random noise as input, the G network then produces the fake data, this is then sent forward to the Discriminator network D. The task of D is to try and classify if the input it received is fake or real data. The accuracy with which D is able to differentiate between the real and fake data is called the misclassification rate. Depending on the classification made by D, backpropagation is used to update the weights of both Generator G and discriminator D.
   - **AutoEncoder:** It is an unsupervised ANN that compresses the input data and then reconstructs an image using the reduced representation that is as close as possible to the original input. The difference between the original and the recreated image is the Reconstruction Loss. The training of the network uses backpropagation to reduce 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)**

   Initially, the generator receives random noise as input and hence $p_g(x)$ is a random distribution, as we continue training the generator gets better at mimicking the real data distribution. Ultimately, the global optimum exists when the discriminator cannot tell the difference between real and generated data, at this global optimum $p_g(x) = p_{data}(x)$.

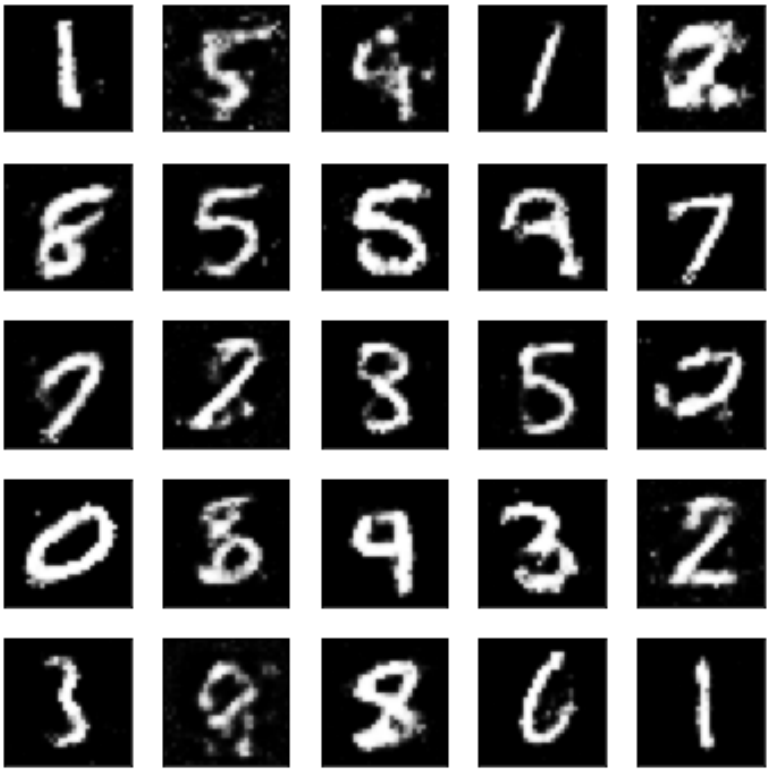
4. **Show the generated images at 10 epochs, 20 epochs,50 epochs,100 epochs by using the architecture required in Guidance. (15% of CW2)**



EPOCH 10

EPOCH 20

EPOCH 50



EPOCH 100