ECS795P Deep Learning and Computer Vision, 2020

Course Work 2:
Unsupervised Learning by Generative Adversarial Network

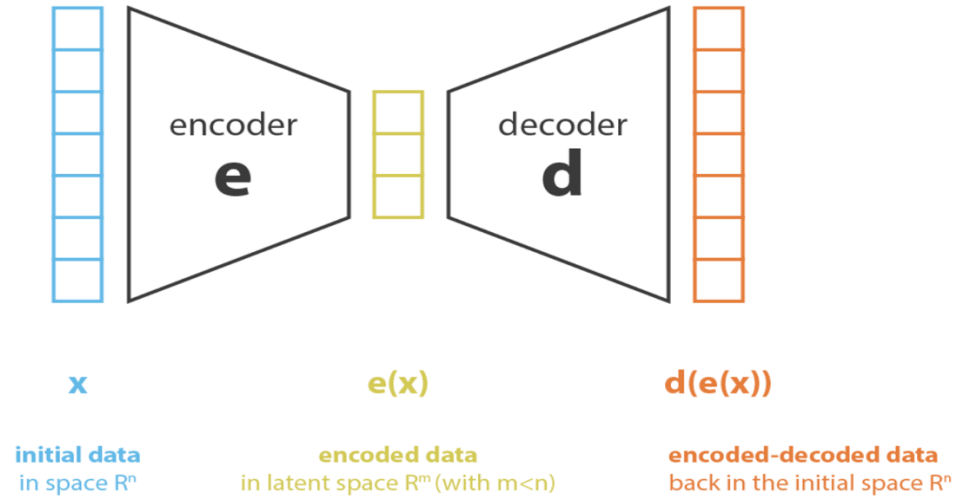
1. **What is the difference between supervised learning & unsupervised learning in image classification task? (10% of CW2)**

   In image classification task, supervised learning technique deals with the labeled data where the expected output data patterns are known to the system. On the contrary, the unsupervised learning works with unlabeled data in which the output is just based on the collection of perceptions.

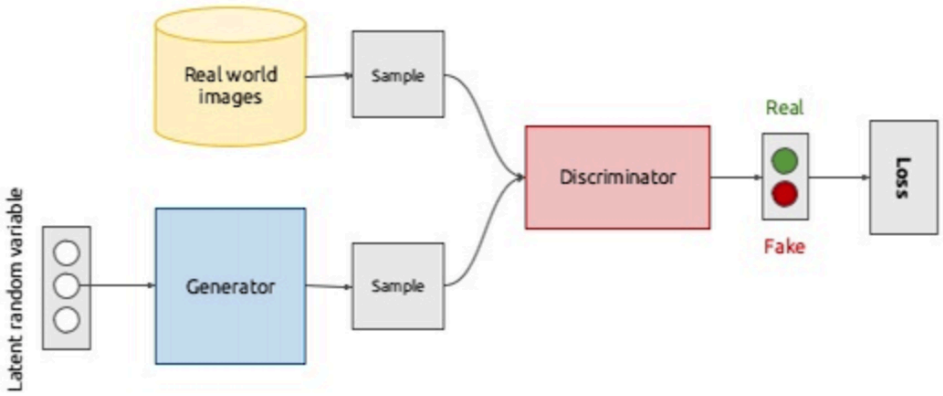
2. **What is the difference between an auto-encoder and a generative adversarial network considering (1) model structure; (2) optimized objective function; (3) training procedure on different components. (10% of CW2)**

   (1) Model structure:

   Auto-encoder is comprised of encoder and decoder. As shown in the figure below (the figure sources from Internet), the input data is encoded to low-dimensional representation $e(x)$ and then $e(x)$ is decoded to reconstructed input $d(e(x))$.



   GAN consists of generator and discriminator. As shown in the figure below (sourcing from the Internet), both the samples from the real data and those generated by generator are imported into the discriminator to discriminate.



(2) Optimized objective function:

For auto-encoder, the objective is to minimize reconstruction errors (such as squared errors).

$$L(x, x') = \| x - x' \|^2 \quad (1)$$

where the x is usually averaged over some input training set, and x' is the corresponding reconstruction output.

For GAN, it is a two-player minimax game, and its objective function is shown below.

$$\min max\, V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[logD(x)] + \mathbb{E}_{z \sim p_z(z)}\left[log\left(1 - D\big(G(z)\big)\right)\right] \quad (2)$$

where we train D to maximize the probability of assigning the correct label to both training examples and samples from G. We simultaneously train G to minimize $log(1 - D(G(z)))$.

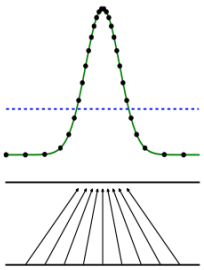
(3) Training procedure on different components:

For auto-encoder, encoder is used to learn a representation for a set of data (typically for dimensionality reduction) and the decoder uses reduced encoding to reconstruct an output as close as possible to its original input.

For GAN, generator is used to learn generating plausible data and discriminator learns to distinguish the generator's fake data from real data.

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

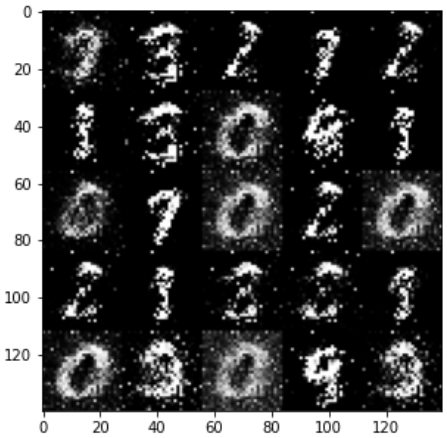
When the discriminator cannot tell the difference between these two distributions, the distribution $p_g(x)$ learned by the generator is the same as the real data distribution $p(x)$. In other words, both of the generator and discriminator have enough capacity (i.e. they will reach a point at which both cannot improve because $p_g(x) = p(x)$) and thus the discriminator is unable to differentiate between the two distributions, i.e. $D(x) = \frac{1}{2}$. As shown in the figure below, the figure references from original GAN paper.



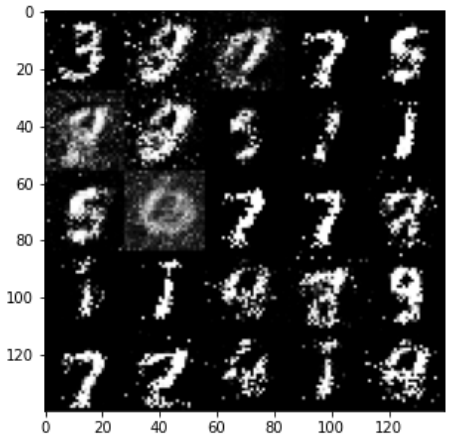
(d)

Where the black dotted line denotes the real data distribution $p(x)$, the green solid line denotes the generative distribution $p_g(x)$, and the blue dashed line denotes the discriminative distribution $D(x)$.

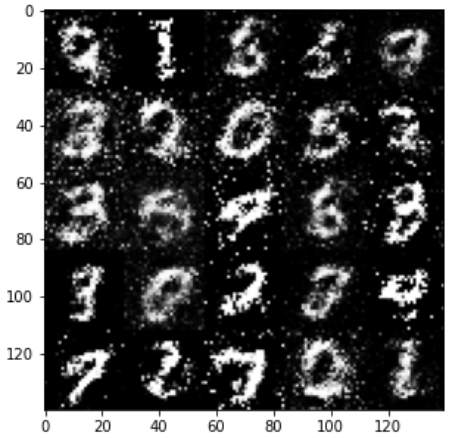
4. **Show the generated images at 10 epochs, 20 epochs,50 epochs,100 epochs by using the architecture required in Guidance. (15% of CW2)**



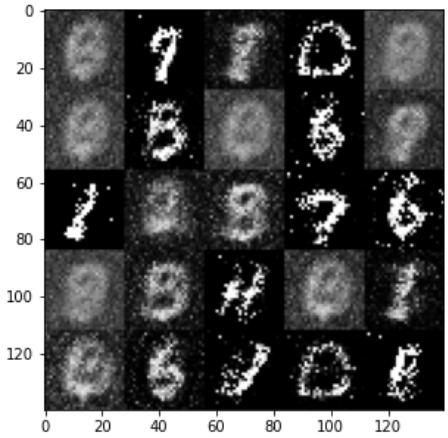
Epoch 10



Epoch 20



Epoch 50



Epoch 100