Lab 8 Exercise – Exploring Latent Spaces

1. Exploring the latent space of a VAE and the code space of a standard auto-encoder

1.1 Systematically sample a VAE and an Autoencoder

As seen below, following the instruction, I build up synthetic figures on the Autoencoder model created on Colab 8.1 and the VAE model on Colab 8.3. With latent space dimensions specified to be 2, I sample codes from $-4 \sim 4$ on either dimension, then decode them with their own decoder.
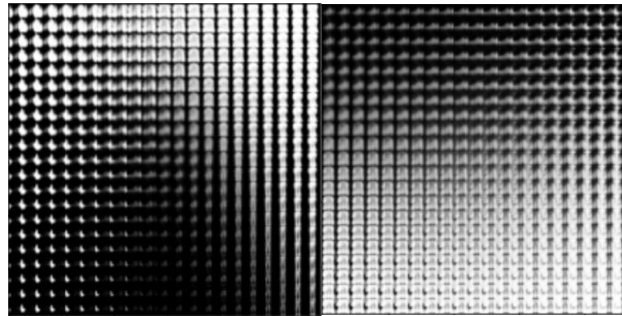

Figure 1 and 2: The samples systematically created from the VAE's latent space and the AE's one.

2. Compare the latent spaces of the VAE and autoencoder

It can be discovered that on VAE's figure, images are changing quite fast. By contrast, on AE's figure images almost remain unchanged. For instance, the figures below show two 4 by 4 figures on the central of VAE and AE. For VAE, the bottom left image is totally different from the top right one, while there is not much difference among all images. This character can be also found by the number of kinds of object shown on the range: VAE's figure has at least 8 kinds of object, while AE's one only includes 4 kinds.

AE seeks to do a best reconstruction work so that it has a certain mapping from an input image to its reconstruction image. For an AE model, it would better have a large distance between two different images, in order to better distinguish them (like SVM).

Then a VAE model is trained to do a reconstruction work and also to have generability (to be invariant to noise). There will be a trade-off, since they cannot be both done ideally. So, for generability, a VAE intend to restrict $p(Z|X)$ for all input image X to be standard Gaussian, which avoids σs to be 0 (a model could have best reconstruction ability but not invariant to noise). And for reconstruction, it will have different μs for different inputs to recognise them. Consequently, at the zero point there is high overlap of different $p(Z|X)$ (see the rightmost figure), making it the greatest point of $p(Z)$, and also making it most various.
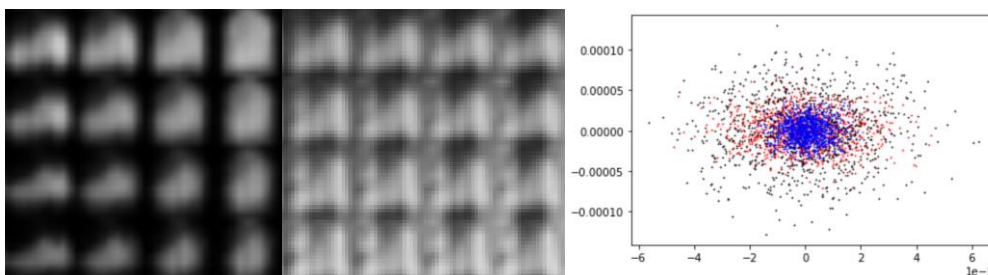

Figure 3,4: The centre patches on figure 1 and 2   Figure 5: Three distribution of $p(Z|X)$