Generative Models: Project 1

Hongli Lin (kag52rer), Michele Paterson (mipat104),
Sandro Jose Leal Miudo (salea100), Sven Klein (svkle100)

November 27, 2024

1 Methods

In this project, we learned to generate annotated images of terminal islands. We started by splitting the data into test and training sets. We kept Brain 20 as the test brain since it has the least amount of data, and a further investigation into the data revealed that the distribution of image sizes is roughly equal. The rest was then used for training, guaranteeing that we have as much training data as possible. We then trained a VAE on this dataset. For sampling, we developed an algorithm that ensures that each pixel of an image has the same probability of getting seen by our model to make the most use out of our training data. For the latent space, we choose a normal Distribution with mean 0 and variance 1. The KL divergence between two normal distributions is given by $KL(n_1, n_2) = \log \frac{\sigma_2}{\sigma_1} + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2}$. Since we have $\sigma_2 = 1$ and $\mu_2 = 0$ this simplifies to $KL(n_1, n_2) = -\frac{1}{2}(1 + \log(\sigma_1^2) - \sigma_1^2 - \mu_1^2)$. To avoid issues where $\log(\sigma_1)$ might not be defined, we decide to learn $\log(\sigma_1^2)$ instead of σ_1. For the model, we use convolutional layers with 1 fully connected layer to obtain σ_1 and μ_1. The model is also inspired by the architecture used in Higgins et al. [2017], but was made smaller for faster training.

2 Results

The images generated by the model look like plausible brain structures, at least to us non-experts. However, the generated brain structures seem to be a bit smoother than the ones from our dataset. The model struggles to generate sharp edges, and there seems to be less detail in the generated images. Therefore, the generated images look a bit like they are a zoomed-in version of the real images. We then analyzed the latent space by trying to interpolate between images. The interpolation between tiles looks very clean, and you can see the cell structures slowly deforming to transition from one image to another. This suggests that the latent space is meaningful and organized. We have also analyzed the latent space by looking at the PCA in 2 dimensions. In 2 Dimensions the latent space looks kind of like an ellipse. This also suggests that the latent space is at least somewhat organized, where different images are in different regions of the ellipse.

References

Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-VAE: Learning basic visual concepts with a constrained variational framework. In *International Conference on Learning Representations*, 2017. URL `https://openreview.net/forum?id=Sy2fzU9gl`.