Diffusion Models (Overview)

Dr. Alireza Aghamohammadi

Overview

- A diffusion model has two parts: an *encoder* and a *decoder*.
- The encoder takes a data sample, x, and maps it to a series of latent variables, $z_1, \ldots, z_T$.
- During this process, the data is gradually mixed with noise until only noise remains. At this point, both the conditional distribution, $q(z_T \mid x)$, and the marginal distribution, $q(z_T)$, approximate the standard normal distribution.
- The decoder reverses this process, starting with z_T and working backward through $z_{T-1}, \ldots, z_1$, removing noise at each step.
- After training, new data samples are generated by sampling a noise vector, z_T, and passing it through the decoder.
- The encoder is predefined; all learnable parameters are in the decoder.

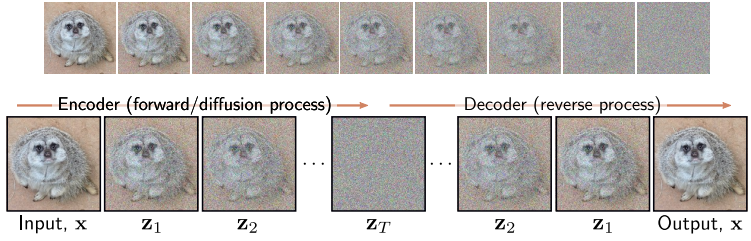

Image credit: Understanding Deep Learning

- ❖ Diffusion models are probabilistic models that define a nonlinear mapping from latent variables to observed data, where both have the same dimension.
- ❖ These models are easy to train and can produce very high-quality samples. They also scale well on parallel hardware.
- ❖ However, generating new samples can be computationally expensive due to the need for multiple forward passes through the decoder network.