VAE

- z is assumed to have a prior probability distribution

$$pr(z) = \mathcal{N}(0,1)$$

- The training loss for VAE becomes,

$$L = \frac{1}{N}(L_{reconstruction} + L_{divergence})$$

$$L = \frac{1}{N}((\hat{X} - X)^2 + \beta \sum_{j}^{z_{dim}} KL(q_j(z|X)||p(z)))$$

- Training Objective :

$$\theta_e, \theta_g = argmin\{\frac{1}{N}((\hat{X} - X)^2 + \beta \sum_{j}^{z_{dim}} KL(q_j(z|X)||p(z)))\}$$

Input Recognition and Preconditioning