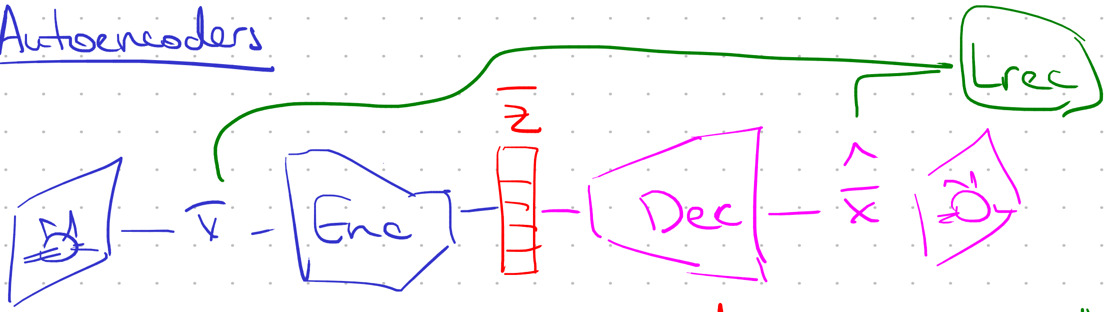
Autoencoders



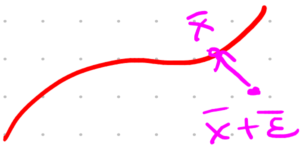
reconstruction loss

$\|\hat{\bar{x}} - \bar{x}\|^2$

L_{perc} $L_{texture}$

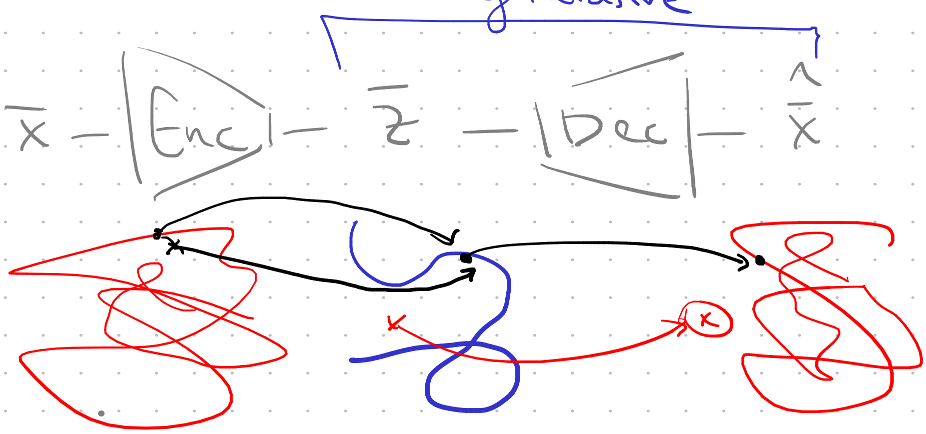
$\|VGG_2(\hat{\bar{x}}) - VGG_2(\bar{x})\|^2$

Denoising autoencoder



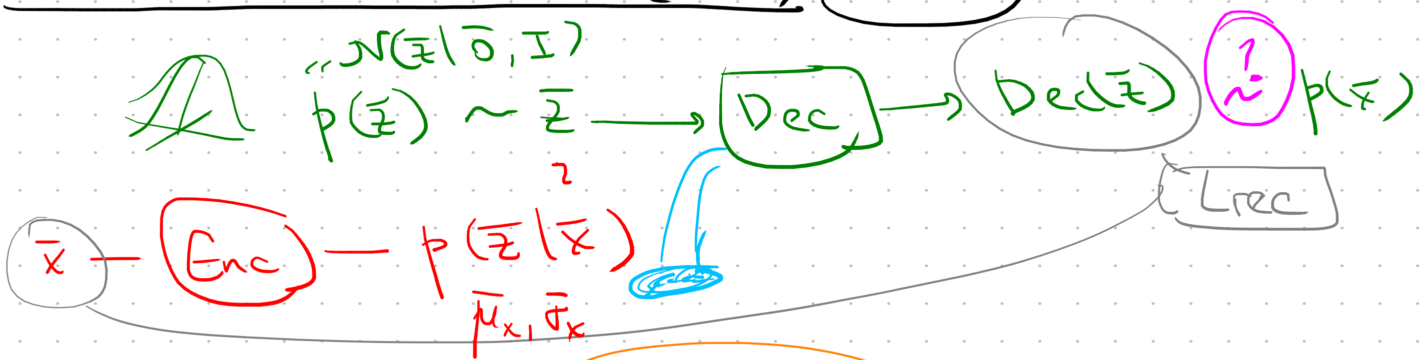
$d(\bar{x}, \hat{\bar{x}}) = \cancel{d(\bar{x}, Dec(Enc(\bar{x})))} \to min$

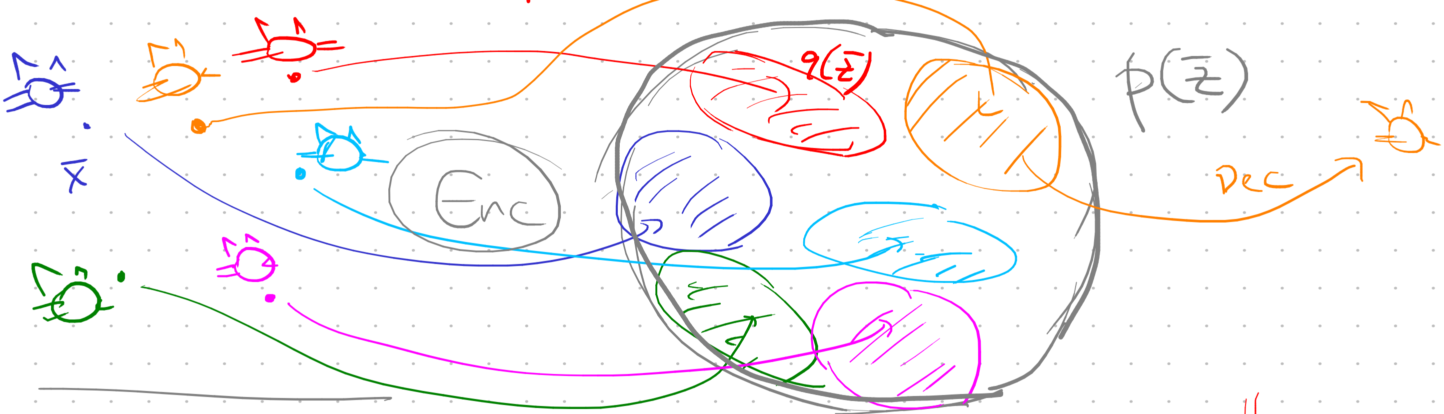
$d(\bar{x}, Dec(Enc(\bar{x}+\bar{\varepsilon}))) \to min$



AAE

Variational autoencoder (VAE) 2013





$$\underbrace{p(\bar z)}_{\uparrow \ \mathcal{N}(\bar z|\bar 0, I)}\ \underbrace{p(\bar x|\bar z)}_{\uparrow \ \text{Decoder}} = p(\bar x, \bar z) = p(\bar x)\ \underbrace{p(\bar z|\bar x)}_{\uparrow \ \text{Encoder}} \approx q(\bar z)$$

$$p(\bar x|\bar z) = \mathcal{N}(\bar x \mid f(\bar z), cI)$$

$$f(\bar z) = \text{Decoder}(\bar z)$$

$$\mathbb{E}_{q(\bar z)} \log p(\bar x) = \mathbb{E}_{q(\bar z)} \log p(\bar x, \bar z) - \log p(\bar z|\bar x)$$

$$\log p(\bar x) = \mathbb{E}_{q(\bar z)}[\log p(\bar x, \bar z)] - \mathbb{E}_q[\log p(\bar z|\bar x)] \pm \mathbb{E}_q[\log q(\bar z)]$$

$$\log p(\bar x) = \int q(\bar z) \log \frac{p(\bar x, \bar z)}{q(\bar z)} d\bar z + \underbrace{\int q(\bar z) \log \frac{q(\bar z)}{p(\bar z|\bar x)} d\bar z}_{KL(q\|p)}$$

$$\text{Const} = \underset{q \to \max}{L(q)} + \underset{q \to \min}{KL(q\|p)}$$

$$L(q) = \int q(\bar z) \log \frac{p(\bar x, \bar z)}{q(\bar z)} d\bar z = \int q(\bar z) \log \frac{p(\bar z)\, p(\bar x|\bar z)}{q(\bar z)} d\bar z =$$

$$= \underbrace{\int q(\bar z) \log p(\bar x|\bar z) d\bar z}_{\int q(\bar z) \log \mathcal{N}(\bar x \mid f(\bar z), cI) d\bar z} - \underbrace{\int q(\bar z) \log \frac{q(\bar z)}{p(\bar z)} d\bar z}_{KL(q(\bar z) \,\|\, p(\bar z))}$$

$$\text{Const} - \frac{1}{2c} \int q(\bar z) \cdot \|\bar x - f(\bar z)\|^2 d\bar z$$

$$\mathcal{N}(\bar z|\bar 0, I)$$

$$\mathcal{L}(q) = -\left(\underbrace{\frac{1}{2c}\cdot \mathbb{E}_{q(\bar z)}\left[\|\bar x - f(\bar z)\|^2\right]}_{\text{Reconstruction loss}} + \underbrace{KL\left(q(\bar z)\,\|\,p(\bar z)\right)}_{\text{Regularizer}}\right)$$

$$q(\bar z) = \mathcal{N}(\bar z \mid \bar\mu_x, \bar\sigma_x) \qquad \begin{pmatrix} \sigma_{x_1} & 0 \\ 0 & \sigma_{x_d} \end{pmatrix} \qquad p(\bar z) = \mathcal{N}(\bar z \mid \bar 0, I)$$

$$KL(q\|p) = KL\left(\prod_i q_i(z_i) \,\Big\|\, \prod_i p_i(z_i)\right) =$$

$$= \int \prod_i q_i \cdot \log \frac{\prod_i q_i}{\prod_i p_i}\, d\bar z = \int \prod_i q_i(z_i)\left(\log\frac{q_1}{p_1} + \dots + \log\frac{q_d}{p_d}\right) d\bar z =$$

$$= \sum_{j=1}^{d} \left(\underbrace{\int \prod_{i\neq j} q_i(z_i)\, d\bar z_{-j}}_{=1}\right)\cdot\left(\int q_j(z_j)\cdot \log\frac{q_j}{p_j}\, dz_j\right) =$$

$$= \sum_{j=1}^{d} KL\left(q_j(z_j)\,\|\,p_j(z_j)\right)$$

$$KL\left(\mathcal{N}(z_j \mid \mu_{x_j}, \sigma^2_{x_j})\,\|\,\mathcal{N}(z_j \mid 0, 1)\right) =$$

$$= \int \frac{1}{\sqrt{2\pi\sigma^2_{x_j}}}\, e^{-\frac{1}{2\sigma^2_{x_j}}(z_j-\mu_{x_j})^2} \log\frac{\mathcal{N}(-)}{\mathcal{N}(-)}\, dz_j =$$

$$= \mathbb{E}_{\mathcal{N}(z_j \mid \mu_{x_j}, \sigma^2_{x_j})}\left[-\frac{1}{2}\log\sigma^2_{x_j} - \frac{1}{2\sigma^2_{x_j}}(z_j-\mu_{x_j})^2 + \frac{1}{2}z_j^2\right] =$$

$$\mathbb{E} z_j^2 = \mu^2_{x_j} + \sigma^2_{x_j} \qquad \mathbb{E} z = \mu_{x_j}$$

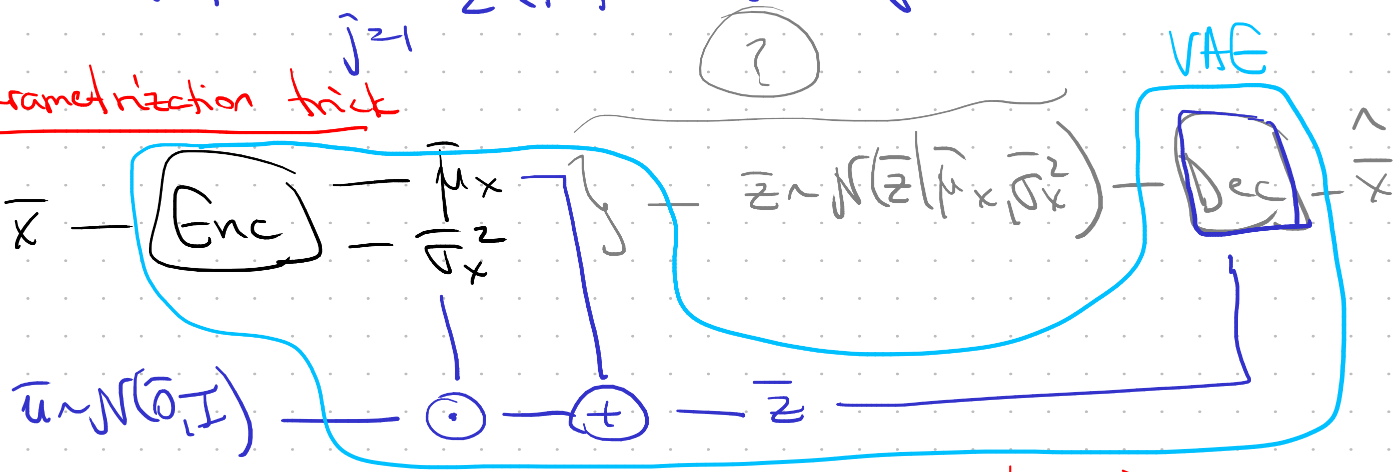
$$= -\frac{1}{2}\log\sigma^2_{x_j} - \frac{\mu^2_{x_j}}{2\sigma^2_{x_j}} + \mathbb{E}\left[\frac{1}{2}\left(1 - \frac{1}{\sigma^2_{x_j}}\right) z_j^2 + \frac{\mu_{x_j}}{\sigma^2_{x_j}}\, z_j\right] =$$

$$= -\frac{1}{2}\log\sigma_{x_j}^2 - \cancel{\frac{\mu_{x_j}^2}{2\sigma_{x_j}^2}} + \frac{1}{2}\left(1-\frac{1}{\sigma_{x_j}^2}\right)\left(\mu_{x_j}^2+\sigma_{x_j}^2\right) + \cancel{\frac{\mu_{x_j}^2}{\sigma_{x_j}^2}}$$

$$= -\frac{1}{2}\log\sigma_{x_j}^2 + \frac{1}{2}\mu_{x_j}^2 + \frac{1}{2}\sigma_{x_j}^2 - \frac{1}{2}$$

$$KL(q\|p) = \sum_{j=1}^{d} \frac{1}{2}\left(\mu_{x_j}^2 + \sigma_{x_j}^2 - \log\sigma_{x_j}^2 - 1\right)$$

Reparametrization trick

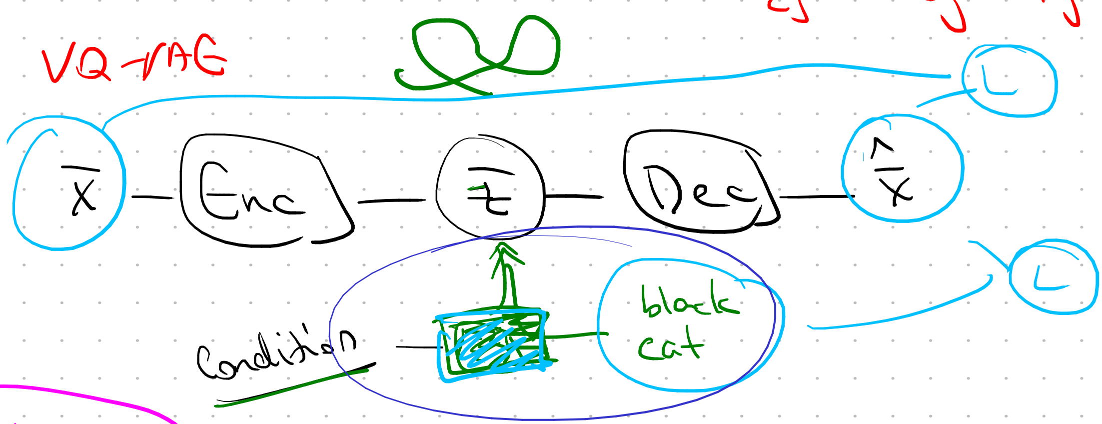


$$z_j \sim N(z_j|\mu_{x_j}, \sigma_{x_j}^2)$$

$$u_j \sim N(u_j|0,1)$$

$$z_j = u_j \cdot \sigma_{x_j} + \mu_{x_j}$$

VQ-VAE



NeRF