Generative model as data source.

method:

input dataset $\{x_i\}_{i=1}^{N}$

Encoder: $f : x \to e$

Generator: $G : z \to x$

Conditional variant $G_y : z, y \to x$ → class label

contrastive framework:

$$\mathcal{L}_{NCE} = -E \left[\log \frac{e^{f(x_a)^T f(x_p)}}{e^{f(x_a)^T f(x_p)} + \sum_{k=1}^{K} e^{f(x_a)^T f(x_n^k)}} \right]$$
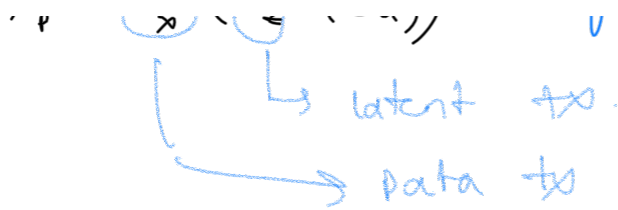
Pair Sampling

During contrastive latent view + pixel to.

$z_a, z_n^k \sim P_z$ ← sample generator

$x_a \sim T_x(G(z_a))$ ← generate anchor

$x_n \sim T_x(T_z(z_a))$ ← generate +ve

$$\downarrow \text{ latent } t_0.$$
$$\longrightarrow \text{Data } t_0$$

$$\hat{x}_n^k \sim T_x\left(a\left(T_z\left(\hat{\varepsilon}_n^k\right)\right)\right) \leftarrow \text{gen} \quad \text{-ve}$$

Properties of T_z

(i) **Gaussian latent view**
Truncated gaussian offset

$$
\begin{bmatrix}
W_{gauss} \sim \mathcal{N}^+(\mu, \sigma, \mathcal{t}) \quad // \text{ truncated Normal} \\
\qquad\qquad\qquad\qquad\qquad\qquad \text{Distribution} \\
\\
T_z(z) = z + W_{Gauss}
\end{bmatrix}
$$

(ii) **steered latent view:**

$$T_z(z) = z + W_{steer}$$

$$W_{steer} = \text{argmin}_{\substack{z, \alpha \\ w}} \; E \left[\| G(z + \alpha w) - T(G(z), \alpha) \| \right]$$

$$\left[\text{Aims to match target pixel space } t_0 \right]$$