

# Training Normalizing Flows

Density evaluation and training is based on change of variable formula

$$p_{\mathbf{x}}(\mathbf{x} = \text{[Image of a man]}) = p_z(T_{\theta}(\mathbf{x}) = \text{[Image of noise]}) \left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\mathbf{x}} \right|$$

↓

Impossible to calculate!

↓

Easy to calculate!  
=> log likelihood  
related to  $\|f_{\theta}(x)\|_2$

↓

All of Normalizing Flow literature is trying to build  $f_{\theta}$  such that this is tractable to calculate.

# Implementing Normalizing Flows

In general, a determinant calculation is  $\mathcal{O}(d^3)$

$$\left| \det \frac{\partial T_{\theta}(\mathbf{x})}{\mathbf{x}} \right| \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} :$$