

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_{θ} such that this is tractable to calculate.

Training Normalizing Flows

Maximum likelihood training to find parameters θ that make our training samples likely under our parameterized model.

$$\max_{\theta} \mathbb{E}_{\mathbf{x} \sim p_x} p_{\theta}(\mathbf{x})$$