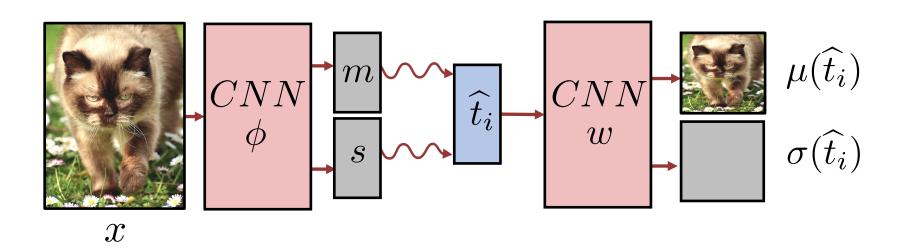
Scaling up Variational EM

maximize
$$\sum_{i} \mathbb{E}_{q_i} \log \frac{p(x_i \mid t_i, w) p(t_i)}{q_i(t_i)}$$
subject to
$$q_i(t_i) = \mathcal{N}(m(x_i, \phi), \operatorname{diag}(s^2(x_i, \phi)))$$



$$\widehat{t}_i \sim \mathcal{N}(m(x_i, \phi), \operatorname{diag}(s^2(x_i, \phi)))$$