

$$\mathcal{L}_{pixel\_loss} = w(\sigma) \|x - x'\|^2$$

# 训练过程



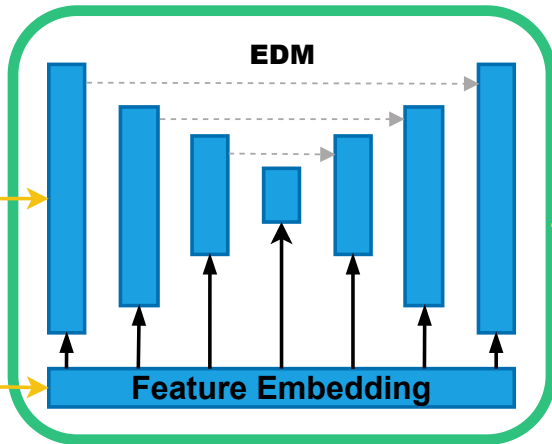
原始图像  $x$



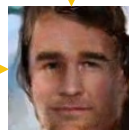
人脸识别模型

$$t = F(x)$$

$t$



$$x' = f_{\theta}(x_0 + \sigma\epsilon, y_0, \sigma)$$



重建图像  $x'$



$$t' = F(x')$$



人脸识别模型

$$\mathcal{L}_{\text{feat}}(\theta) = \max \left( 0, m + \frac{\langle t', t_{\text{neg}} \rangle}{\|t'\|_2 \|t_{\text{neg}}\|_2} - \frac{\langle t', t \rangle}{\|t'\|_2 \|t\|_2} \right)$$

$$\mathcal{L}_{\text{div}}(\theta) = \sum \left( 1 - \frac{\langle t'_i, t'_j \rangle}{\|t'_i\|_2 \|t'_j\|_2} \right)$$