
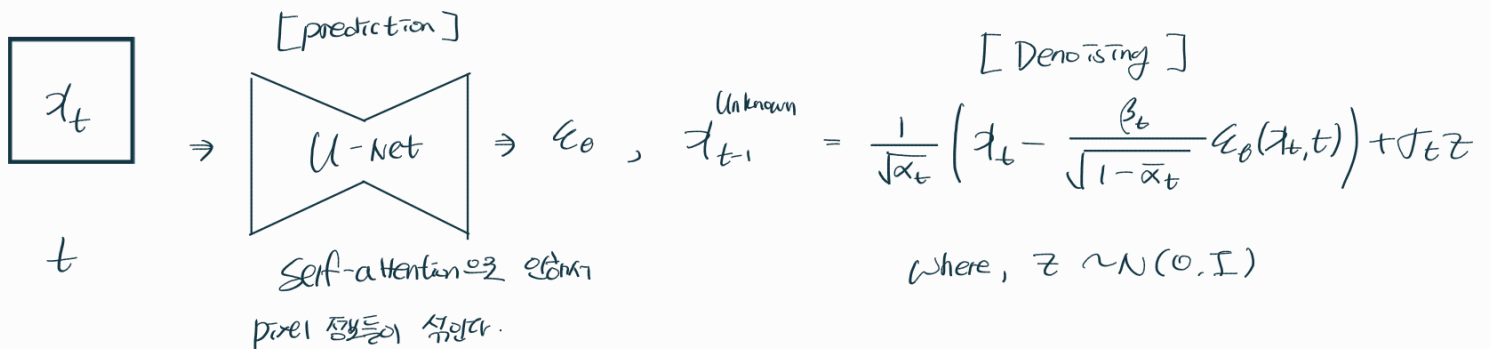
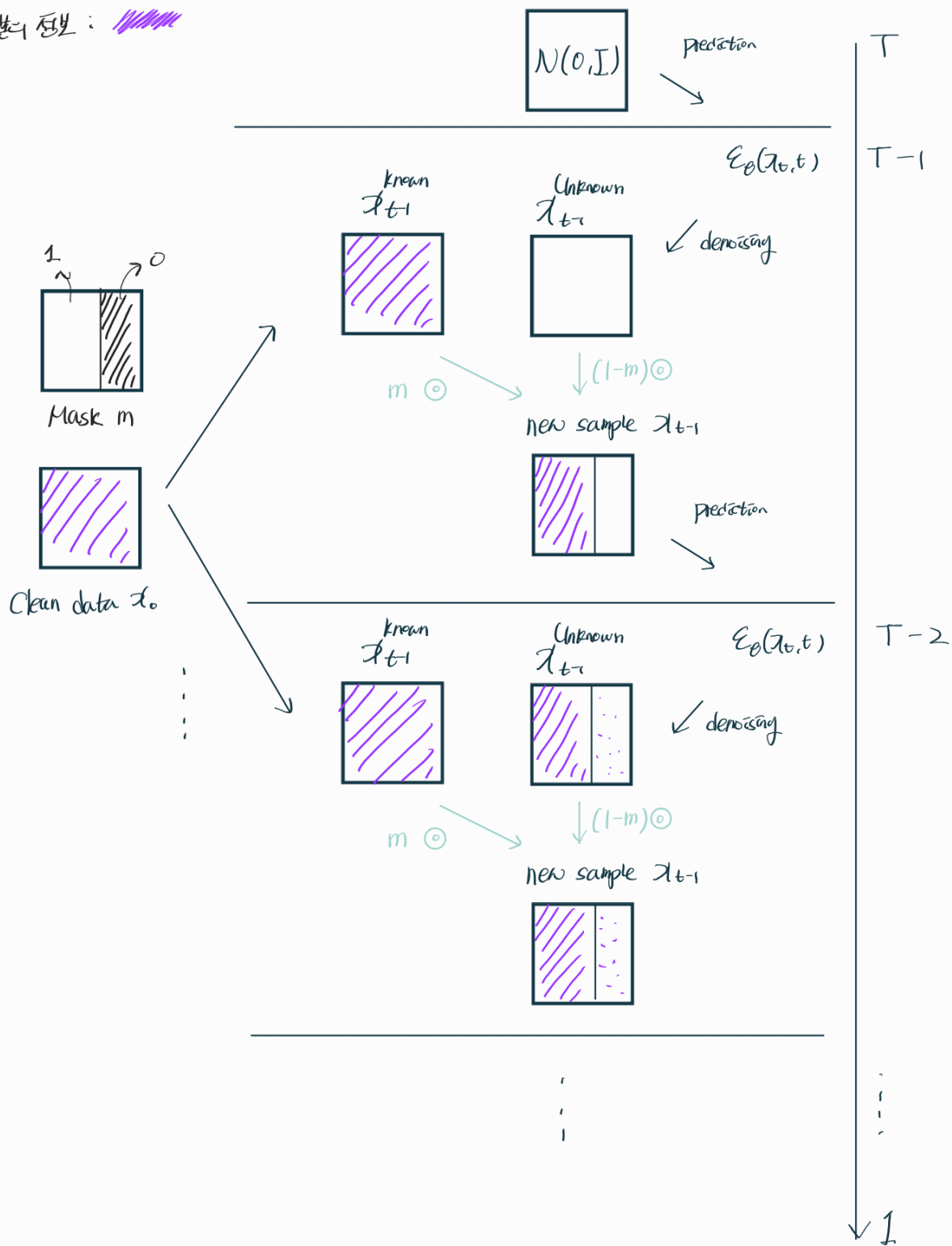



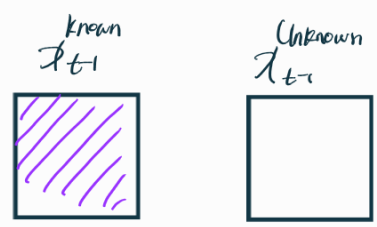
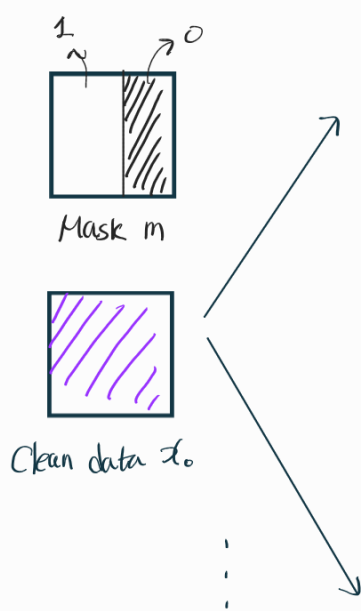
음성 정보: 



입력 정보: 



prediction



$\epsilon_\theta(x_t, t)$

denoising

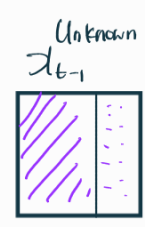
$m \odot$ $\downarrow (1-m) \odot$
new sample x_{t-1}



Add Noise

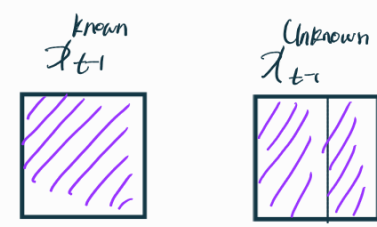


prediction



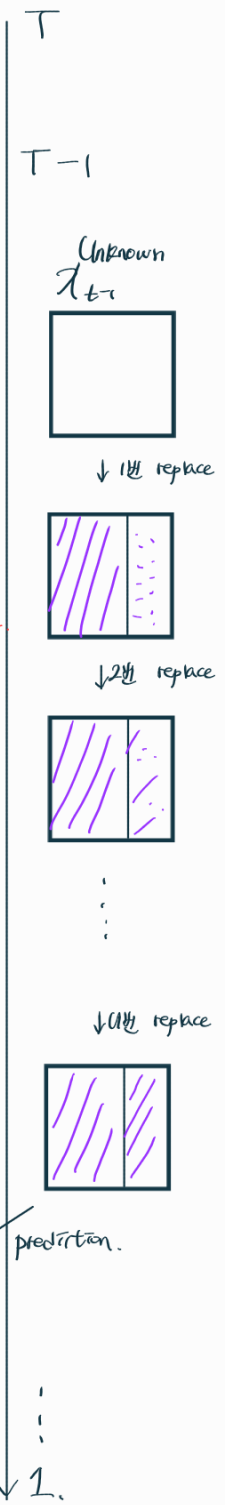
$\epsilon_\theta(x_t, t)$

denoising

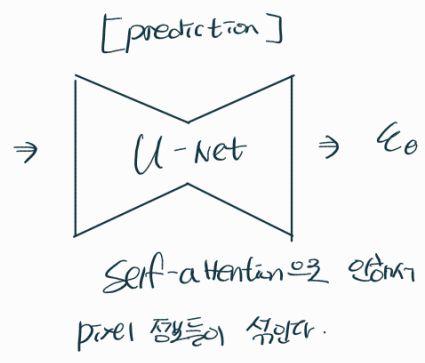


denoising

$\epsilon_\theta(x_t, t)$



t



[prediction] $\Rightarrow \epsilon_\theta$, $x_{t-1}^{\text{Unknown}} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1-\alpha_t}} \epsilon_\theta(x_t, t) \right) + \sqrt{\beta_t} z$

where, $z \sim N(0, I)$

[Add noise]

$x_t = \sqrt{1-\beta_t} x_{t-1} + \sqrt{\beta_t} \epsilon$