

## 1.1 动机

这篇文章使用了深度生成模型来进行图像的一次性泛化学习。它建立在反馈（**feedback**）和注意（**attention**）的原理上。**Attention** 模型结合了注意机制来学习图像。在这篇文章中，实现了空间变换（**ST**）的注意过程，其对形状和大小是不变的。它使用参数  $\lambda$  处理输入图像  $x$  以生成输出。

## 1.2 原理

Latent variables:  $z_t \sim \mathcal{N}(z_t | \mathbf{0}, \mathbf{I}), t = 1, \dots, T$

Context:  $v_t = f_v(h_{t-1}, x'; \theta_v)$

Hidden State:  $h_t = f_h(h_{t-1}, z_t, v_t; \theta_h)$

Hidden Canvas:  $c_t = f_c(c_{t-1}, h_t; \theta_c)$

Observation:  $x \sim p(x | f_o(c_T; \theta_o))$

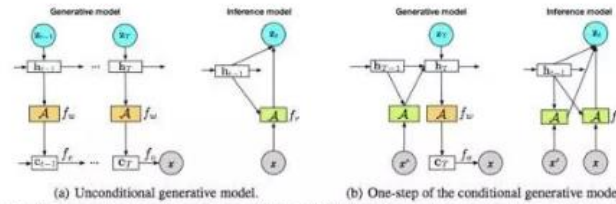


Figure 2. Stochastic computational graph showing conditional probabilities and computational steps for sequential generative models.  $\mathcal{A}$  represents an attentional mechanism that uses function  $f_w$  for writings and function  $f_r$  for reading.

在每个步骤  $t = 1 \dots T$ ,  $K$  维隐含变量  $z_t$  都是独立进行采样的。可以使用注意机制来合并外部上下文或侧面信息片段,该模型是现有模型(如 **DRAW**、复合 **VAE** 和 **AIR**)的泛化。然而,它引入了隐藏的 **canvas** 步骤,其允许在隐藏空间中构造预图像。另一个需要注意的是,它允许对模型进行采样,而不将画布的结果反馈到隐藏状态,这会使其更高效。

生成新样本或新类型。这篇文章使用上述生成模型对给定小样本进行了演示。它类似于上面论文中描述的能力。然而,模型不是通过学习单个图像(或少量图像)来产生样本,而是生成单个类型的新样本,该单个类型并不是训练集的一部分。

此论文使用深度生成过程来模仿人类的一次性泛化能力。它表明,当提供未见过的样本时,该过程可以产生新样本或新类型。相比于使用对象自然发生的因果结构的 **Lake** 的 **BPL** 模型,这种方法适用于许多场景。然而,还是有一些的限制,这还需要未来进一步研究。它仍然需要合理的数据量来避免过度拟合。虽然它适用于任何图像(通过不使用专用结构的方式),但它不能将结构解析为更小的部分。