The paper presents a DRAW architecture to generate image iteratively.Image is encodes and decoded by RNN with write and read processes. The idea is to presents a more natural way of generating image step by step by attend a specific part of image. A lot o tricks are introduced in the paper.

For each time step t, the encoder receives the input image and hidden vector from previous decoder. The read operation concatenates the input image and error image. The hidden vector is used to parameters a distribution over latent vector z. Diagonal Gaussian distribution is used to preserve the differentiability for the convenience of SGD. As the input for decoder takes the information from input data and previous latent sample. The loss function need to sum the individual loss of input data and a sequence of latent distribution. The hidden vector produced by decoder contains the attention parameters forming a number of Gaussian filters. The Gaussians filters will extract a small patch from input image. The attention can shift with different weights in the filter. The output of write operation produces a canvas matrix which will be update through timestamps.