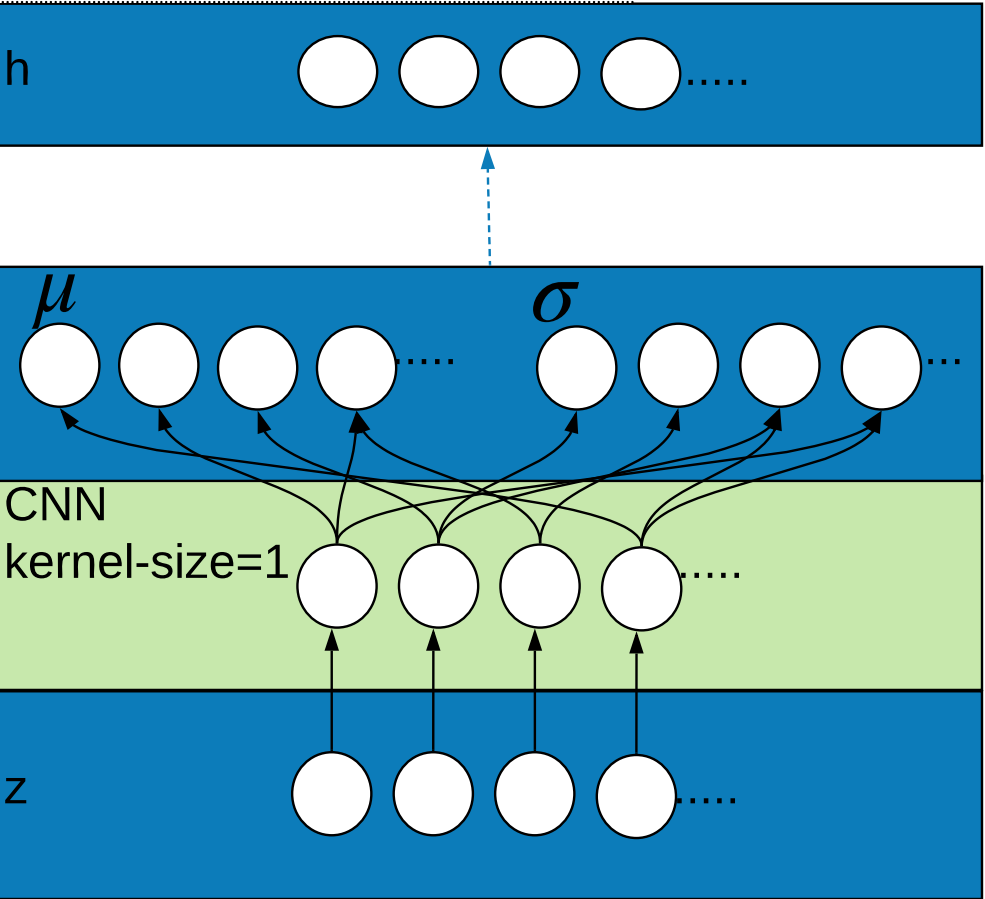
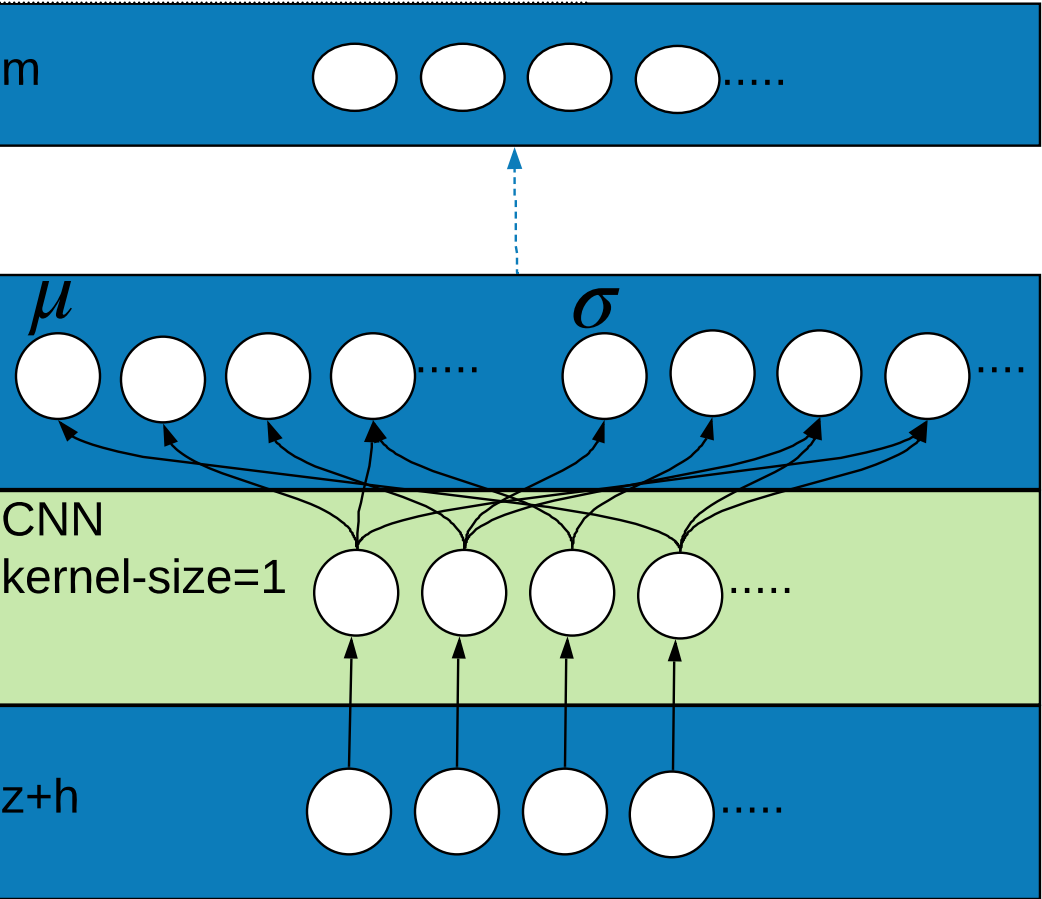


Predictor

likelihood of h : $p(h \mid z, x)$

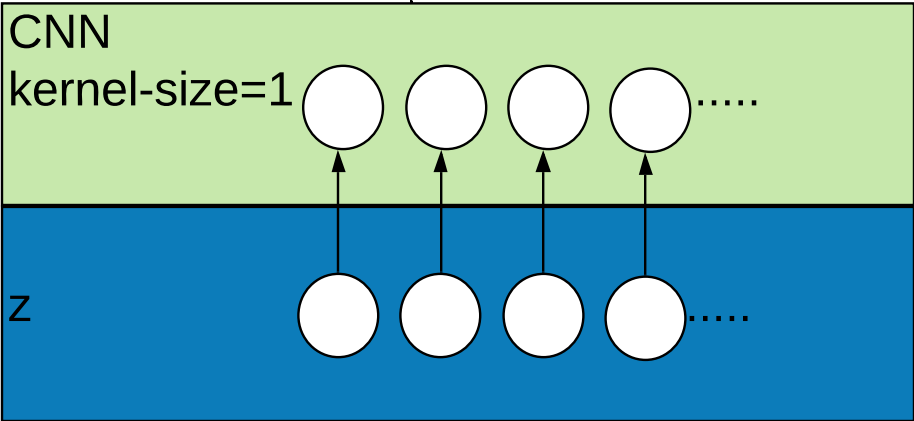
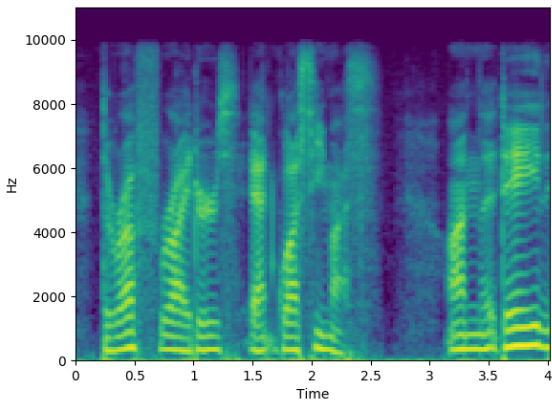


likelihood of m : $p(m \mid z, h, x)$

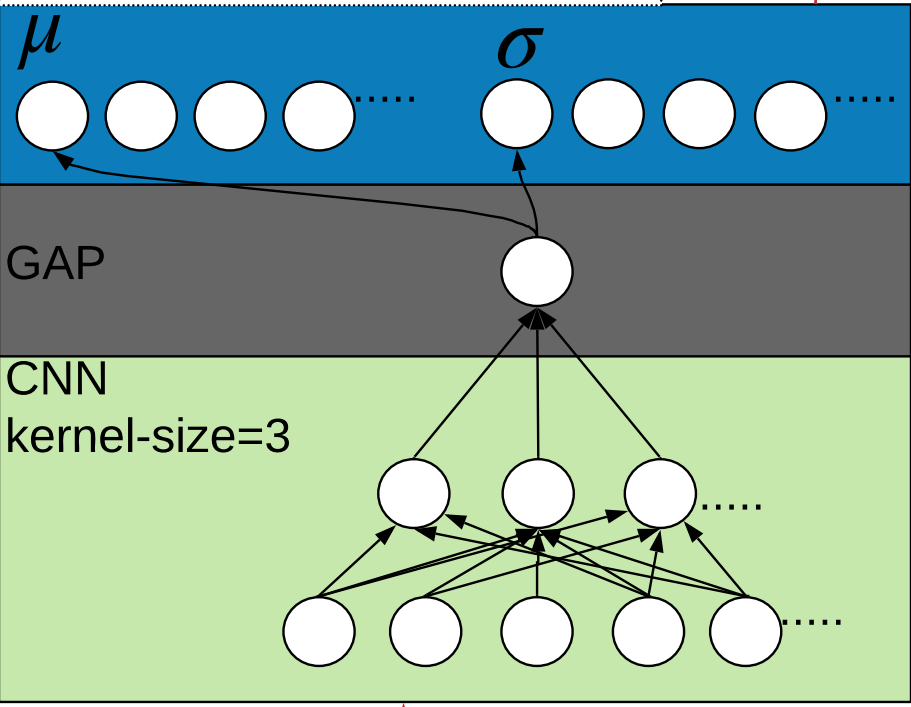


Decoder

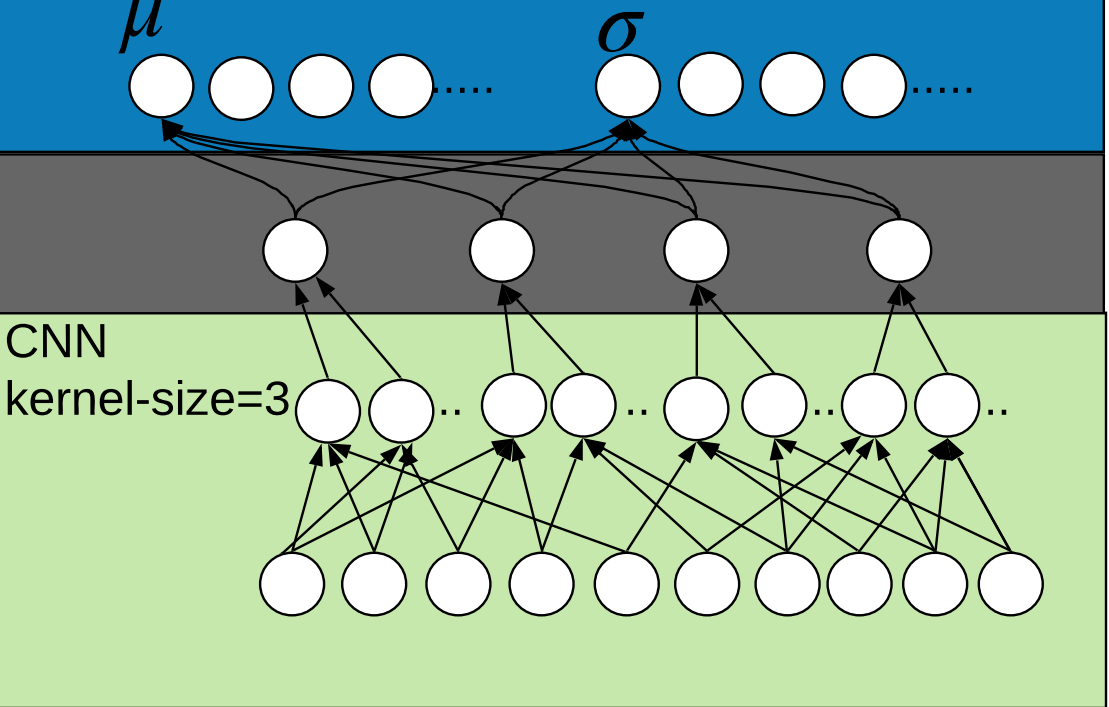
Reconstruction $p(x \mid z)$



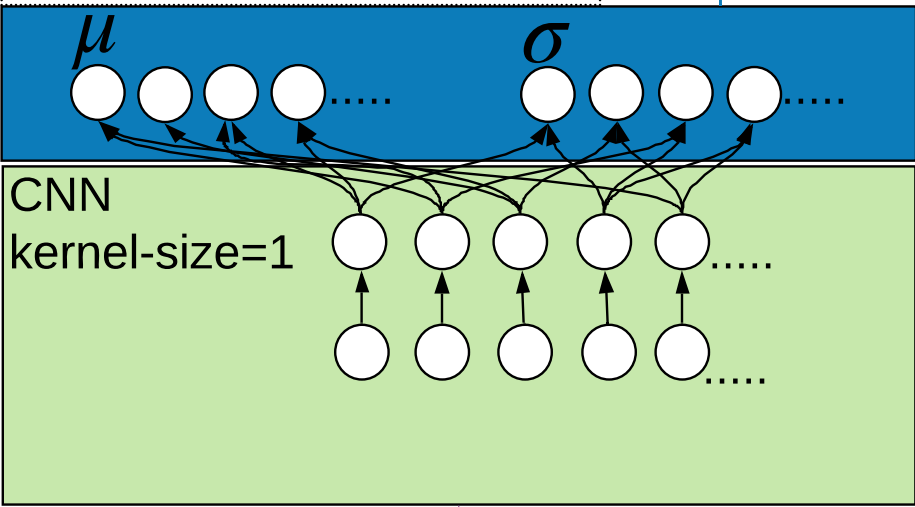
global time-scale $p(h \mid x)$



medium time-scale $p(m \mid h, x)$



local time-scale $p(z \mid h, m, x)$



Encoder

input

input+h

input+m

input

input

