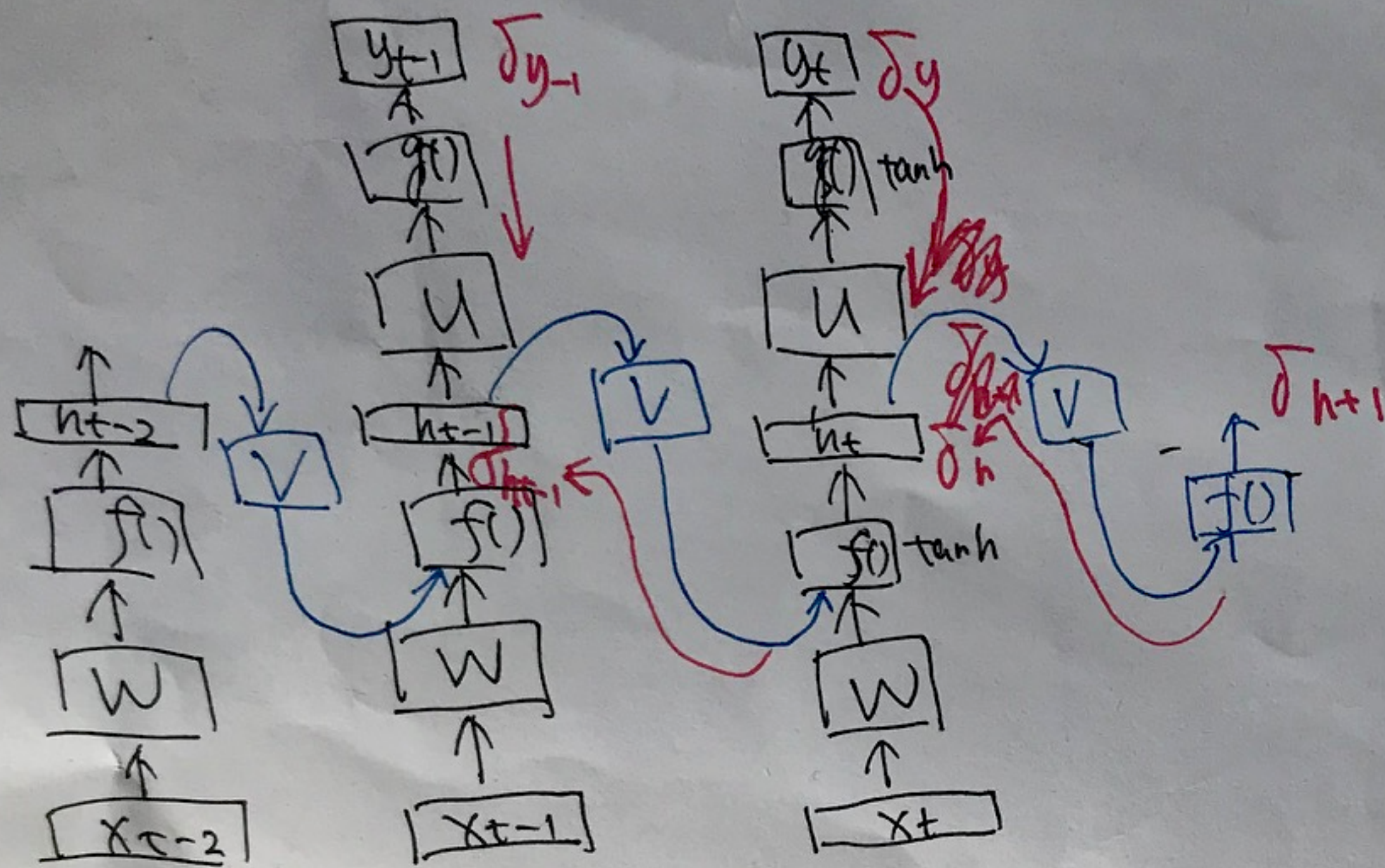


RNN



$$h_t = f(Wx_t + Vh_{t-1} + b_h)$$

$$y_t = g(Uh_t + b_y)$$

$$\frac{\partial L}{\partial U} = \underbrace{\frac{\partial L}{\partial a_y}}_{\text{error term}} \cdot \frac{\partial a_y}{\partial z_y} \cdot \frac{\partial z_y}{\partial U} \stackrel{= h_t}{=} L' g' h_t = \underbrace{\delta y}_{\delta y}$$

$$z_y = Uh_t + b_y$$

$$a_y = g(z)$$

(same as FFNN)

$$\frac{\partial L}{\partial W} = \underbrace{\frac{\partial L}{\partial a_h}}_{\text{error term}} \cdot \frac{\partial a_h}{\partial z_h} \cdot \frac{\partial z_h}{\partial W} \stackrel{= x_t}{=} \delta_h x_t$$

$$z_h = Wx_t + Vh_{t-1} + b_h$$

$$a_h = f(z)$$

$$\frac{\partial L}{\partial V} = \underbrace{\frac{\partial L}{\partial a_h}}_{\text{error term}} \cdot \frac{\partial a_h}{\partial z_h} \cdot \frac{\partial z_h}{\partial V} \stackrel{= h_{t-1}}{=} \delta_h h_{t-1}$$

$$\delta_h = g' u \delta y + f' v \delta_{h+1}$$

