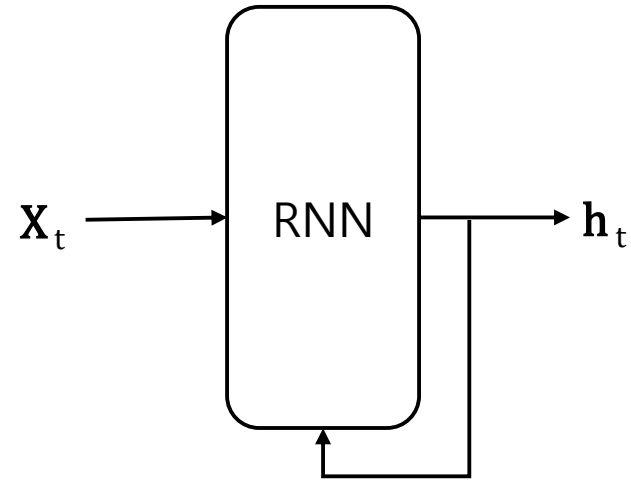
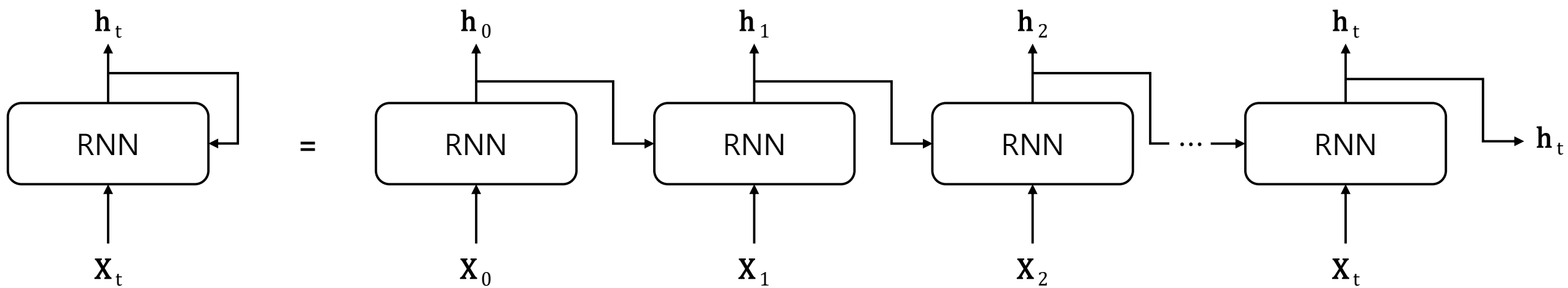


## Recurrent Neural Network (RNN)

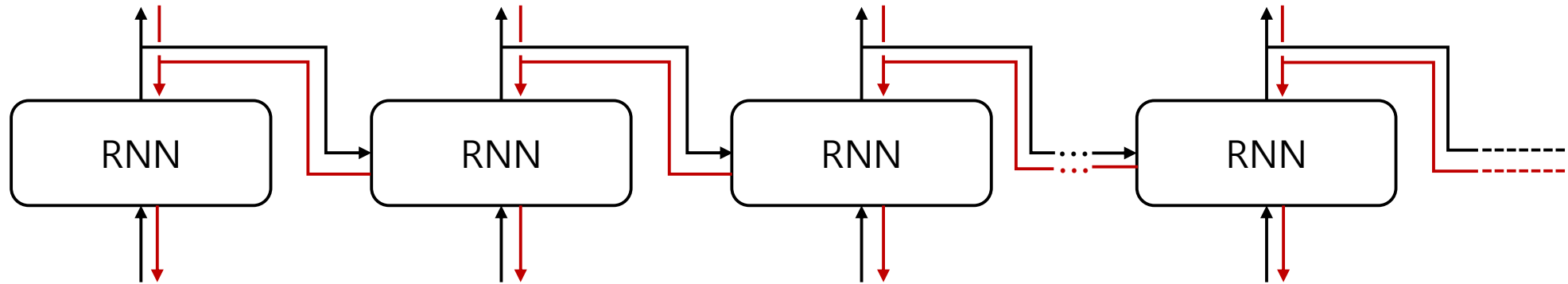


RNN 계층의 순환 구조 펼치기



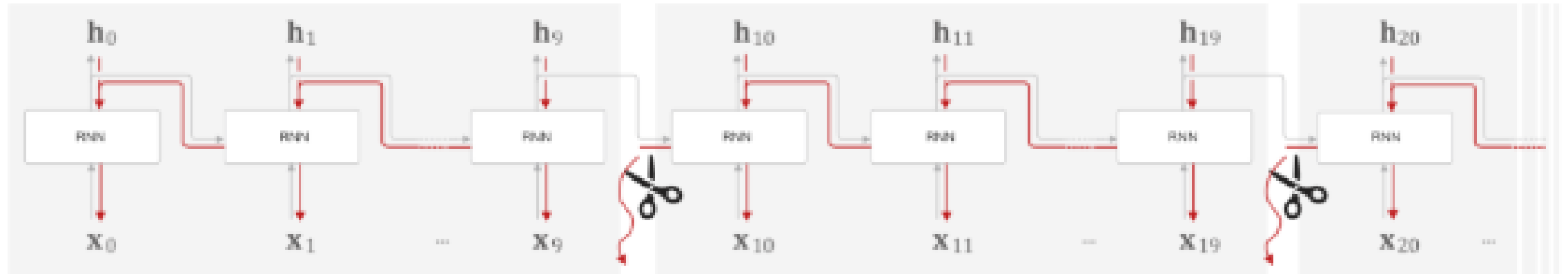
# Recurrent Neural Network (RNN)

## BPTT (Backpropagation Through Time)



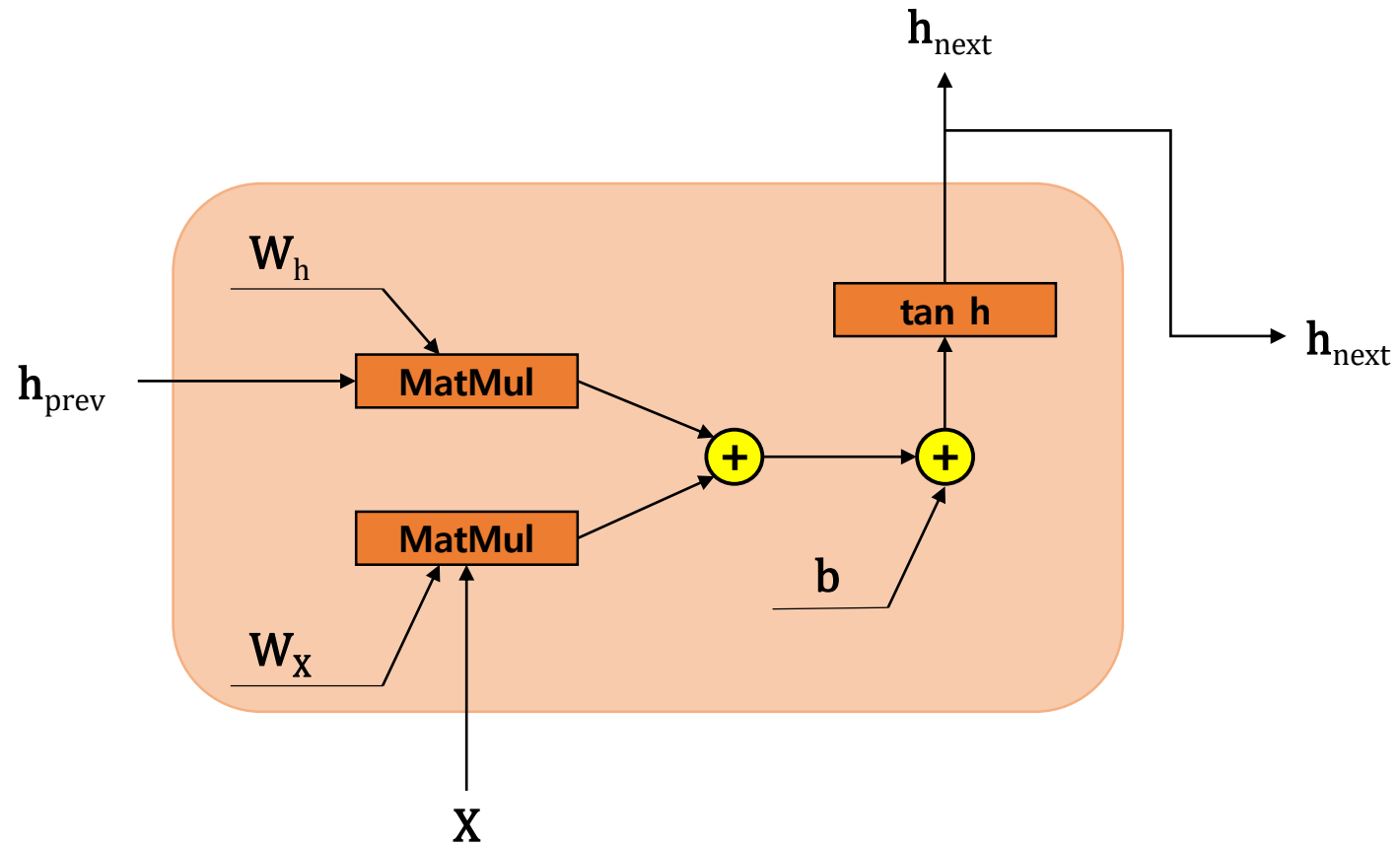
# Recurrent Neural Network (RNN)

## Truncated BPTT



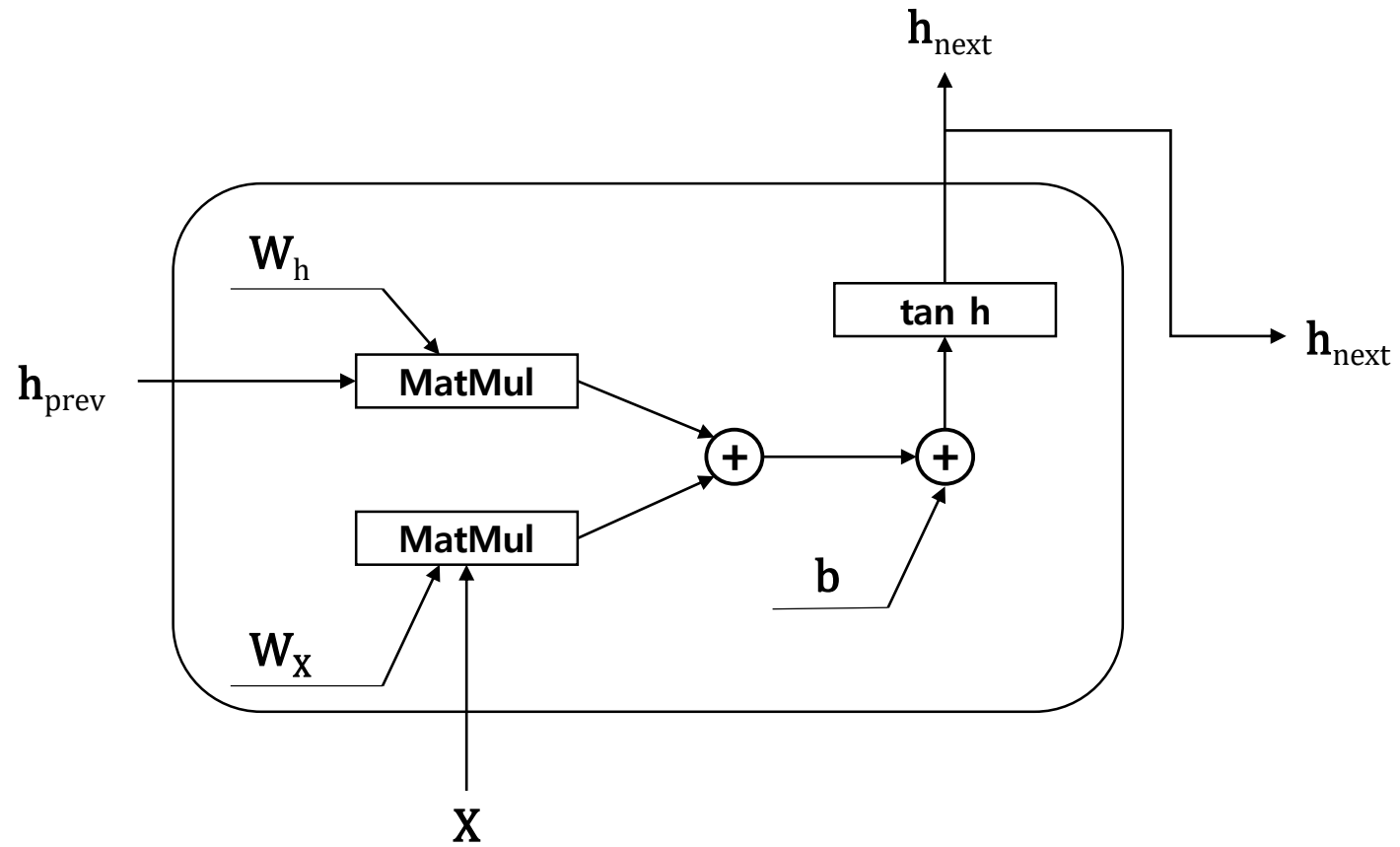
# Recurrent Neural Network (RNN)

forward (순전파)



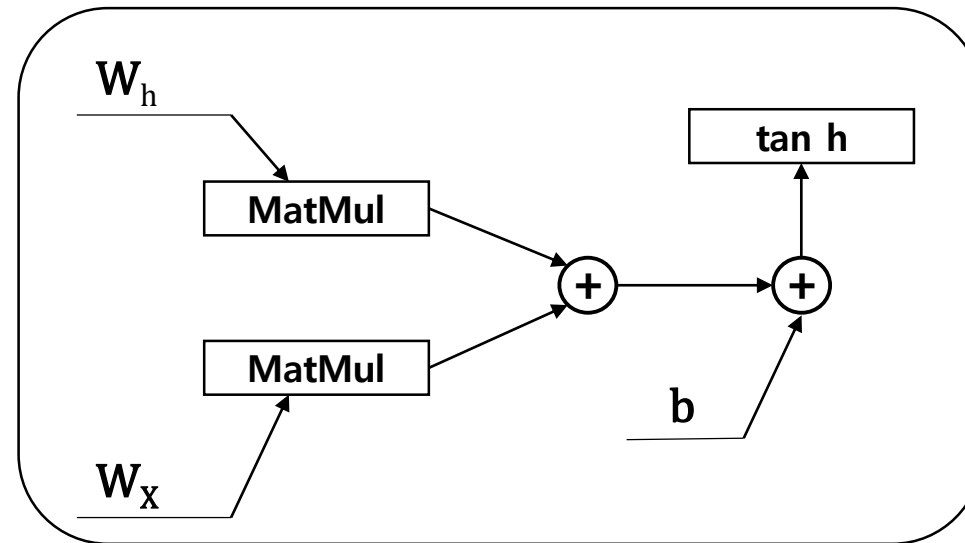
# Recurrent Neural Network (RNN)

forward (순전파)



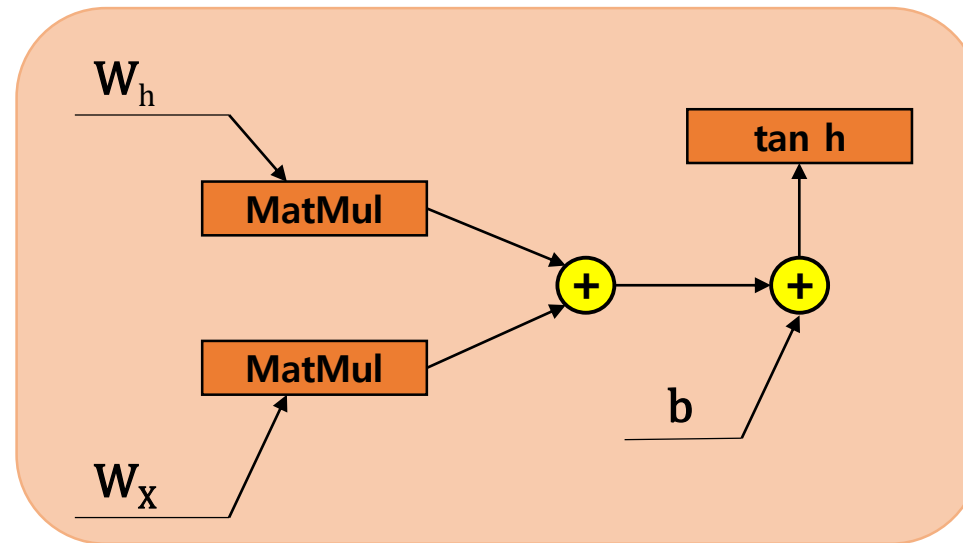
# Recurrent Neural Network (RNN)

forward (순전파)



# Recurrent Neural Network (RNN)

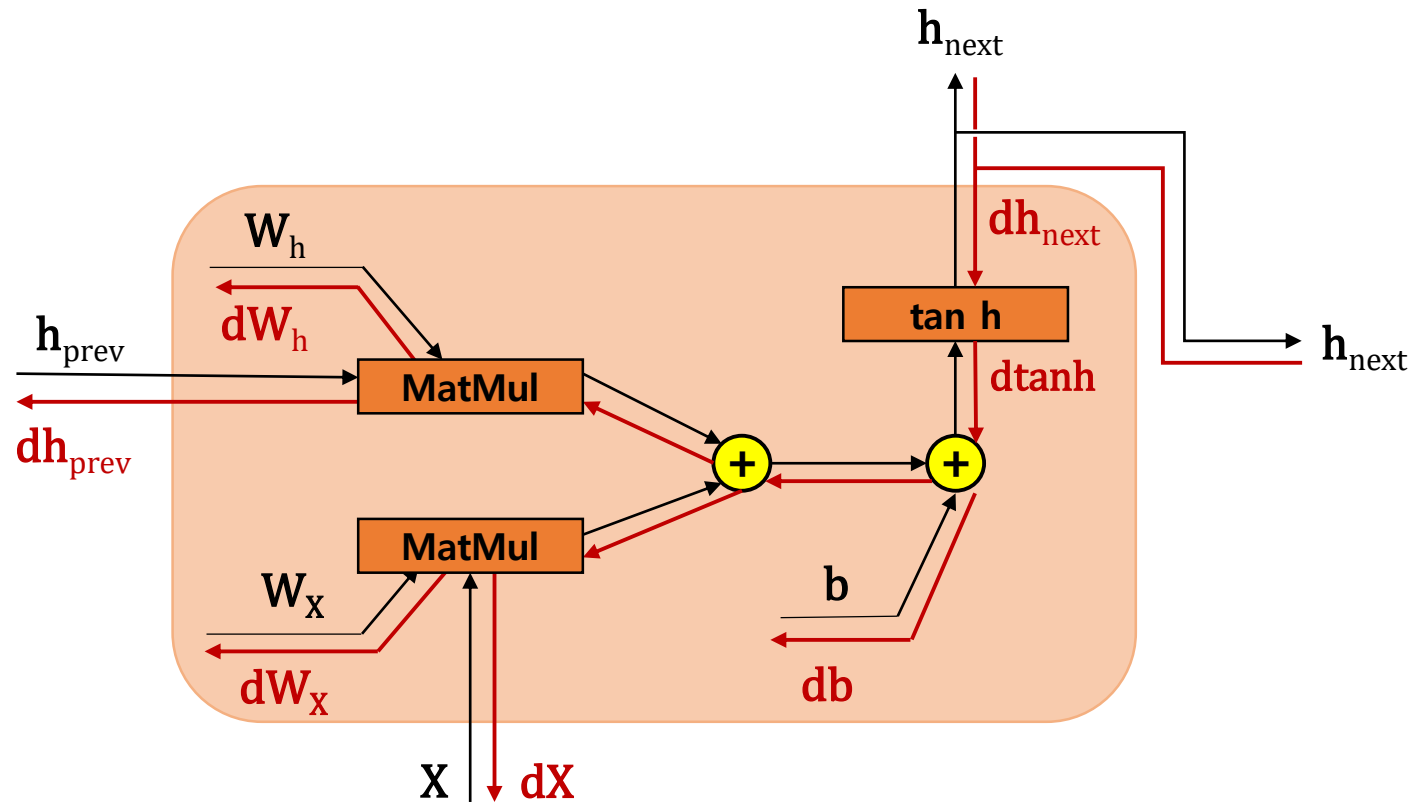
forward (순전파)





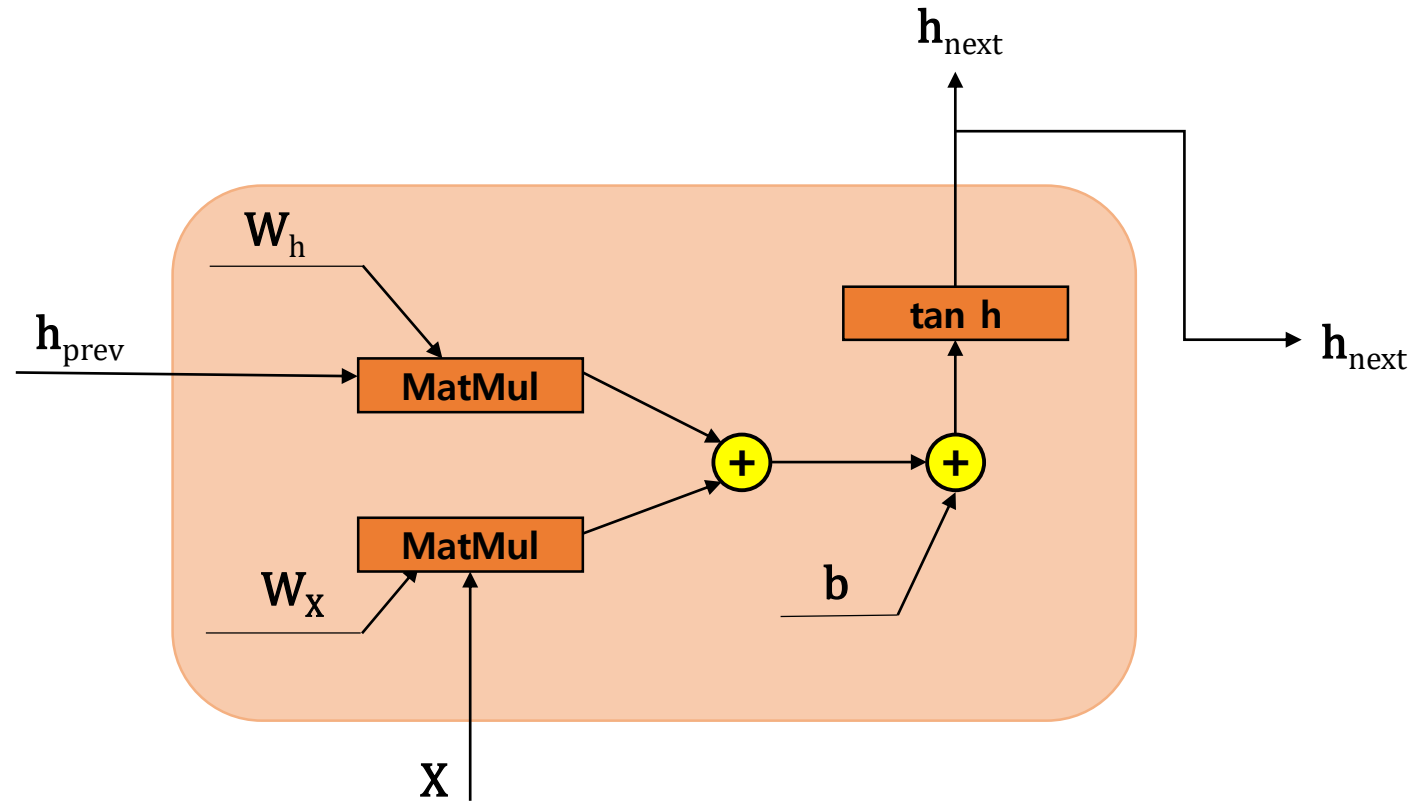
# Recurrent Neural Network (RNN)

backward (역전파)



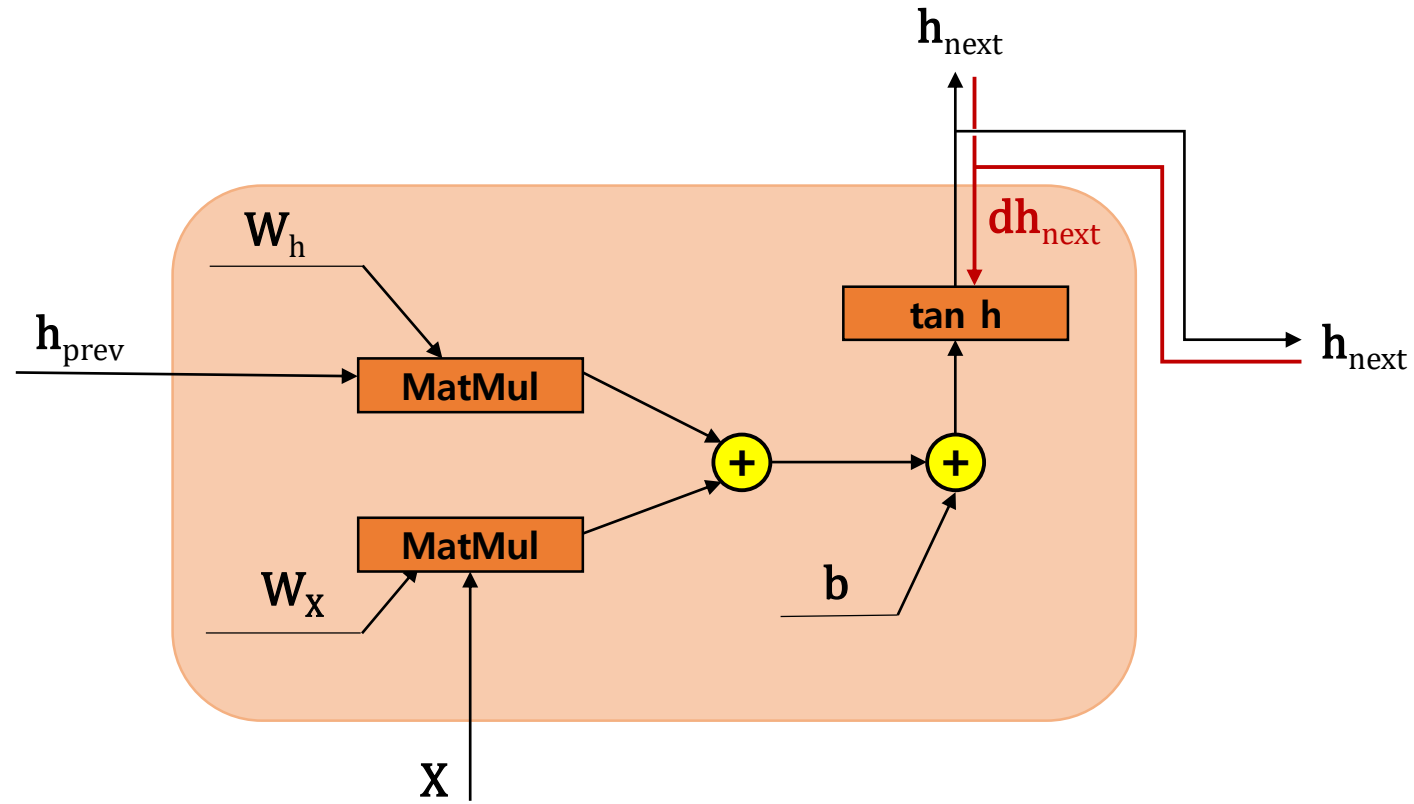
# Recurrent Neural Network (RNN)

backward (역전파) - 시작



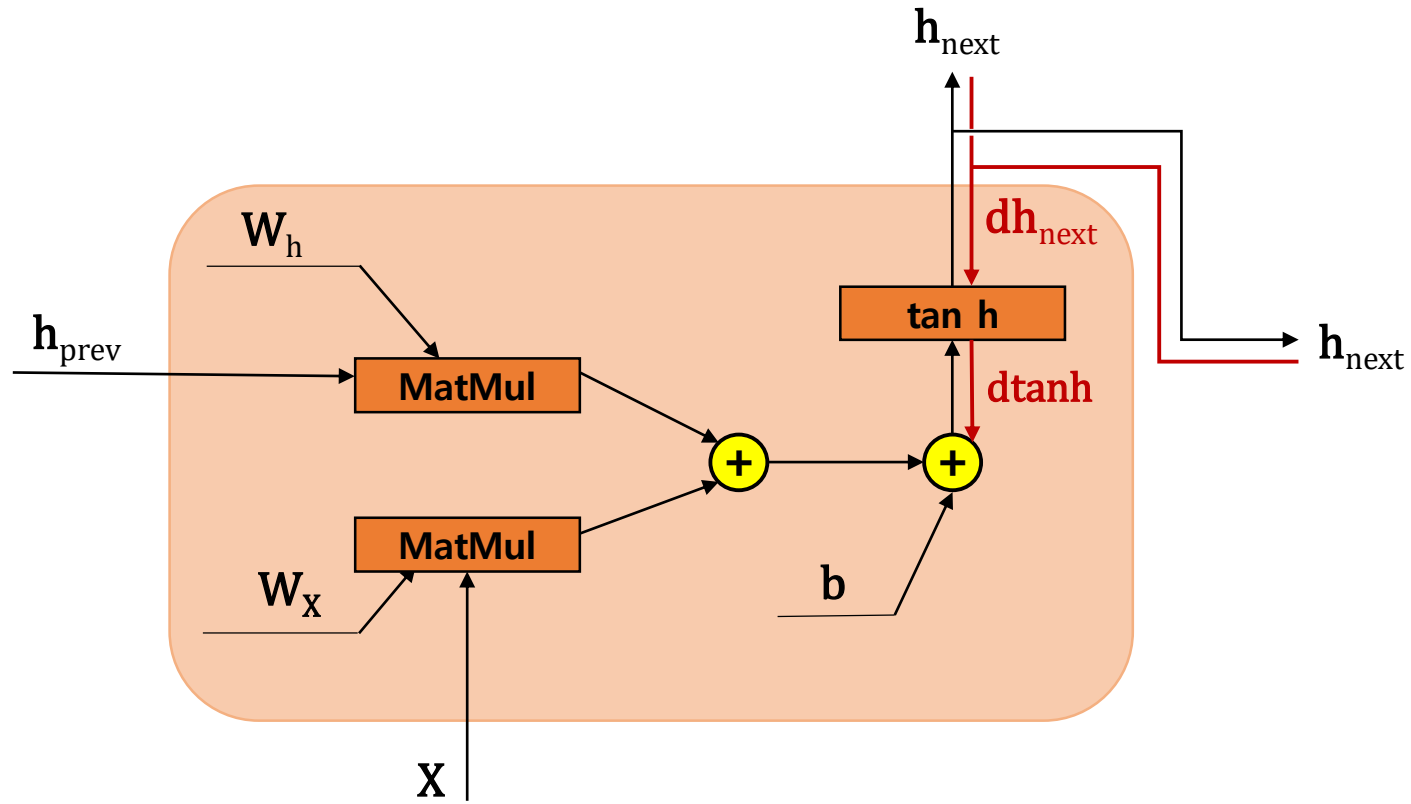
# Recurrent Neural Network (RNN)

backward (역전파) - (1)  $dh_{next}$



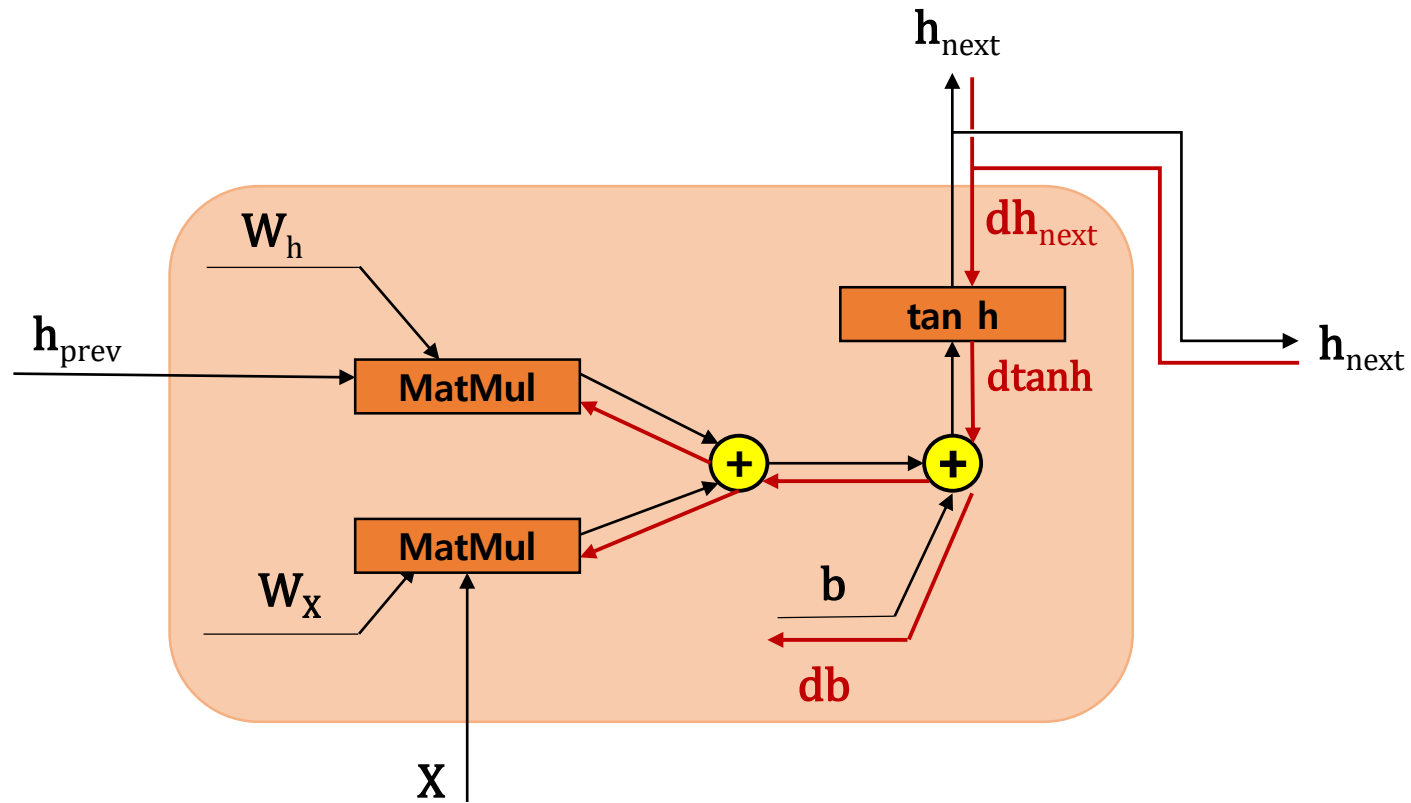
# Recurrent Neural Network (RNN)

backward (역전파) - (2) dtanh



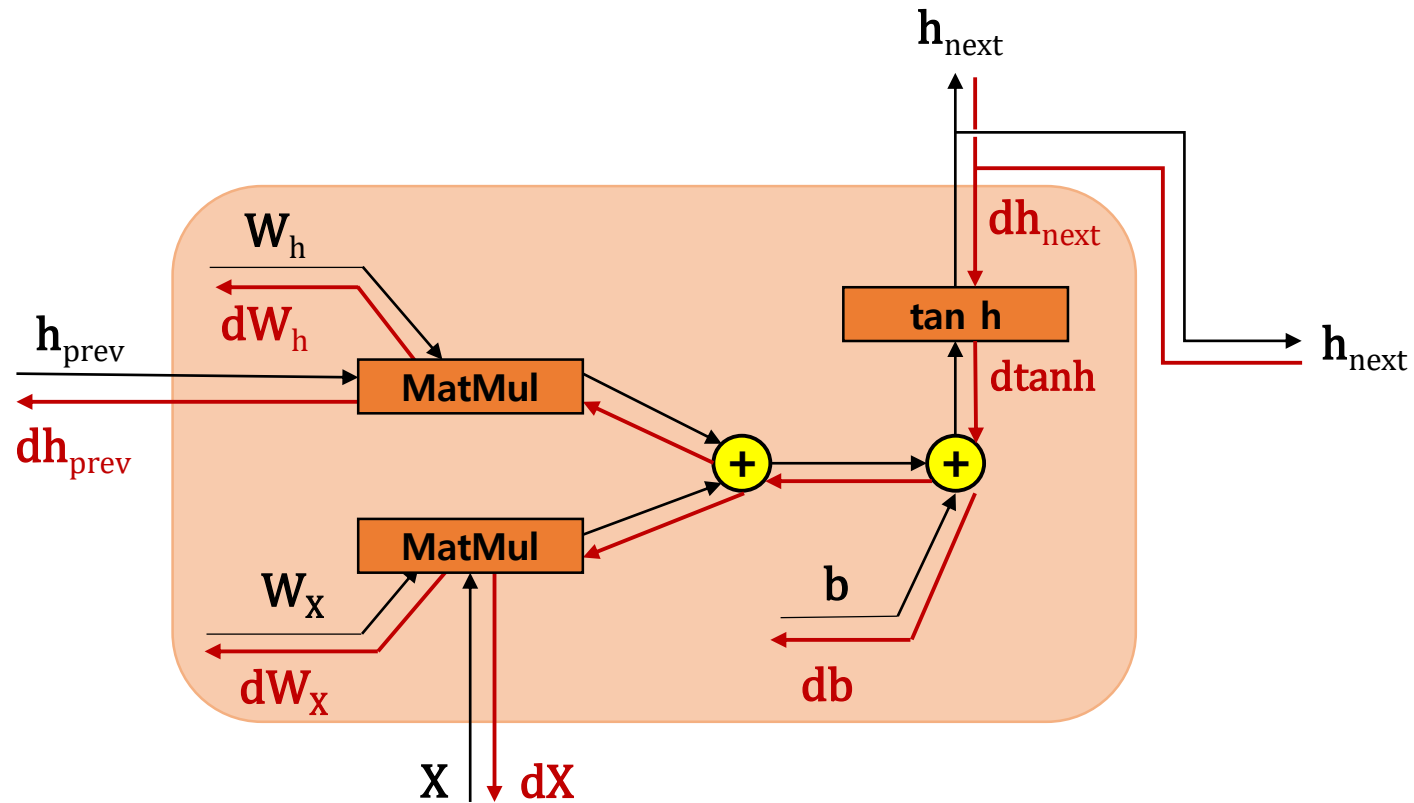
# Recurrent Neural Network (RNN)

backward (역전파) - (3) 덧셈 노드



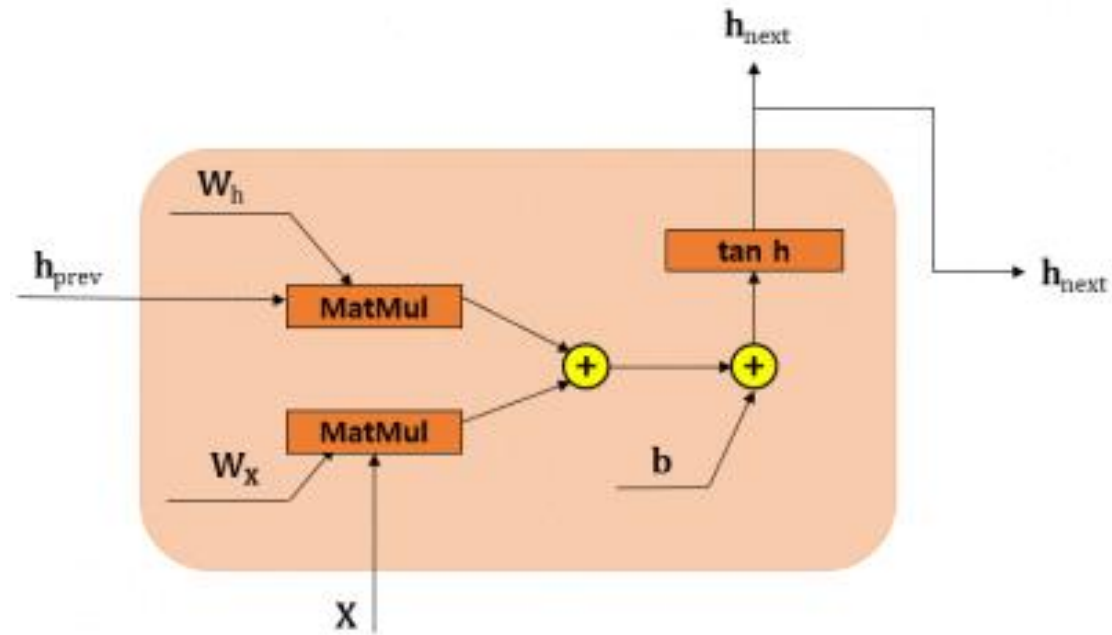
# Recurrent Neural Network (RNN)

backward (역전파) - (4) 곱셈 노드

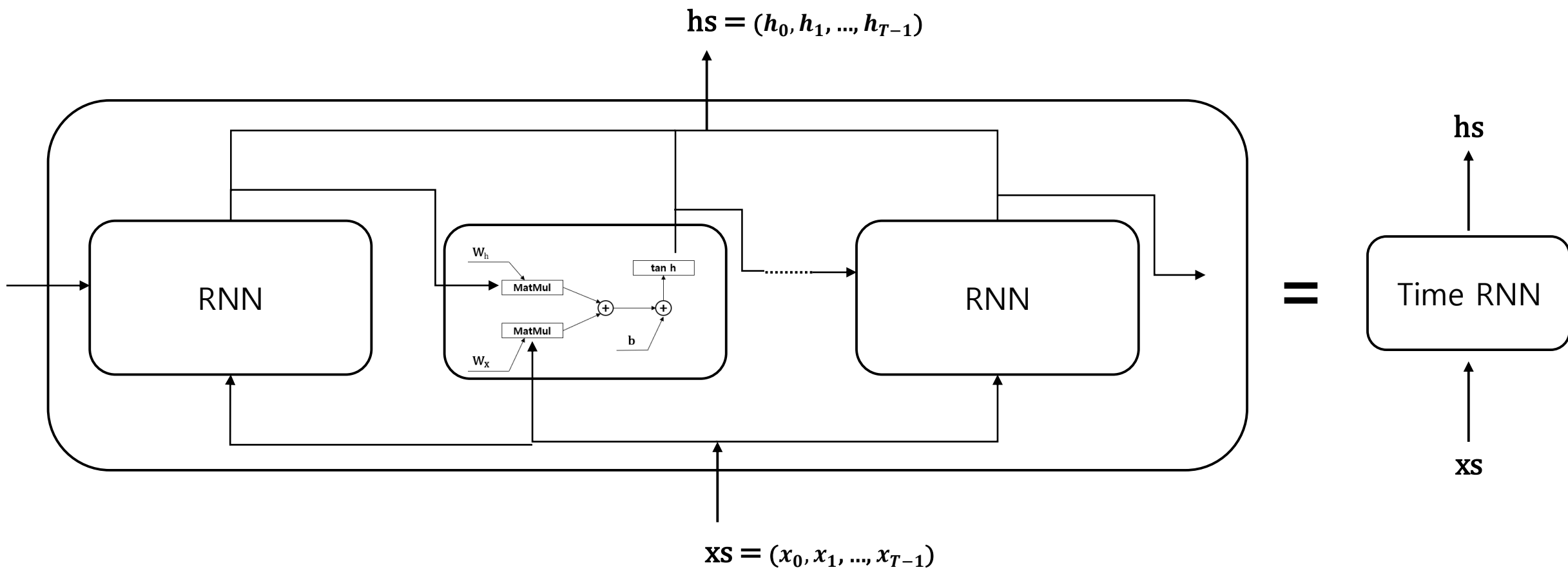


# Recurrent Neural Network (RNN)

backward (역전파) – gif

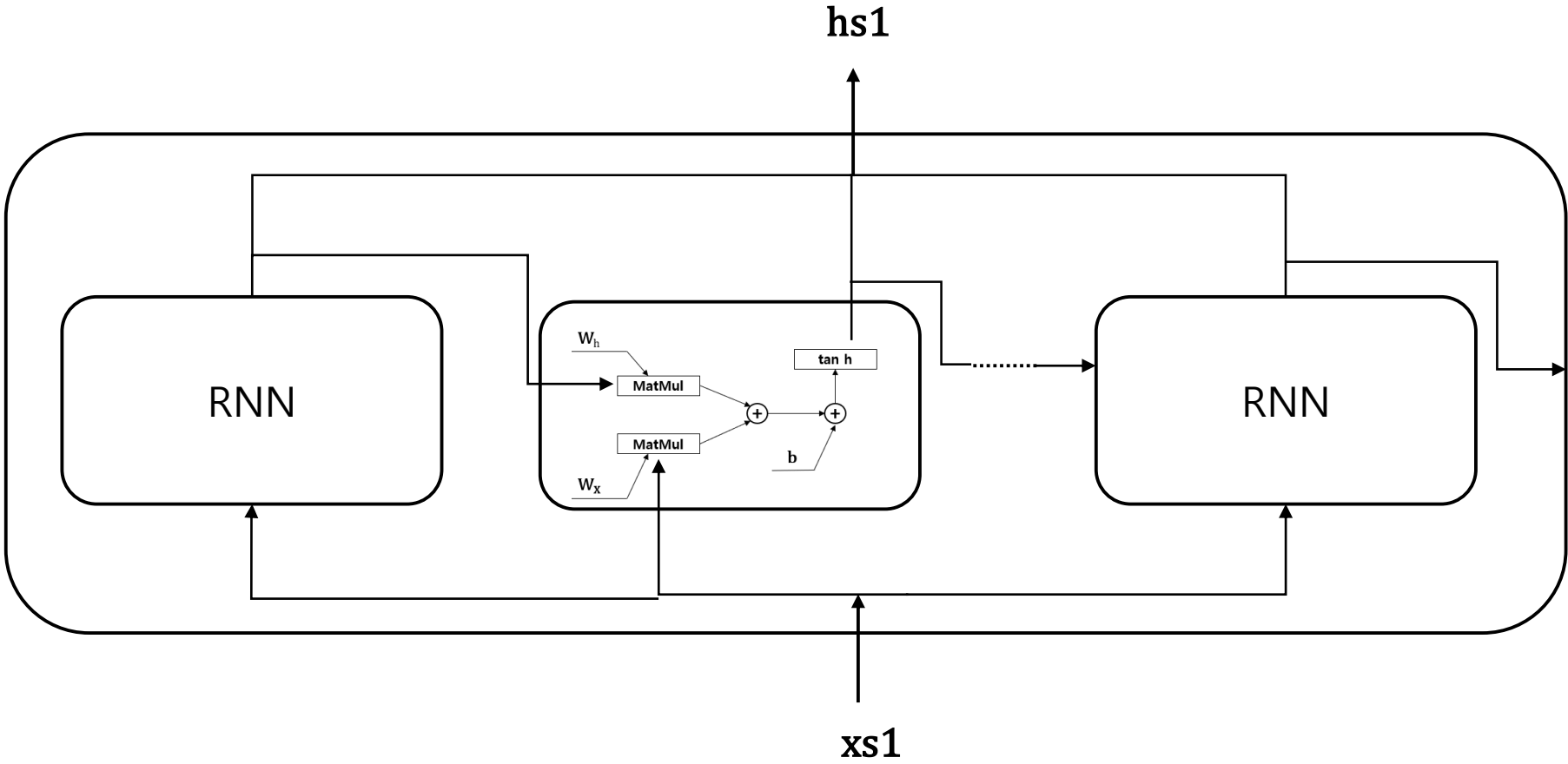


# Time RNN 계층과 RNN 계층

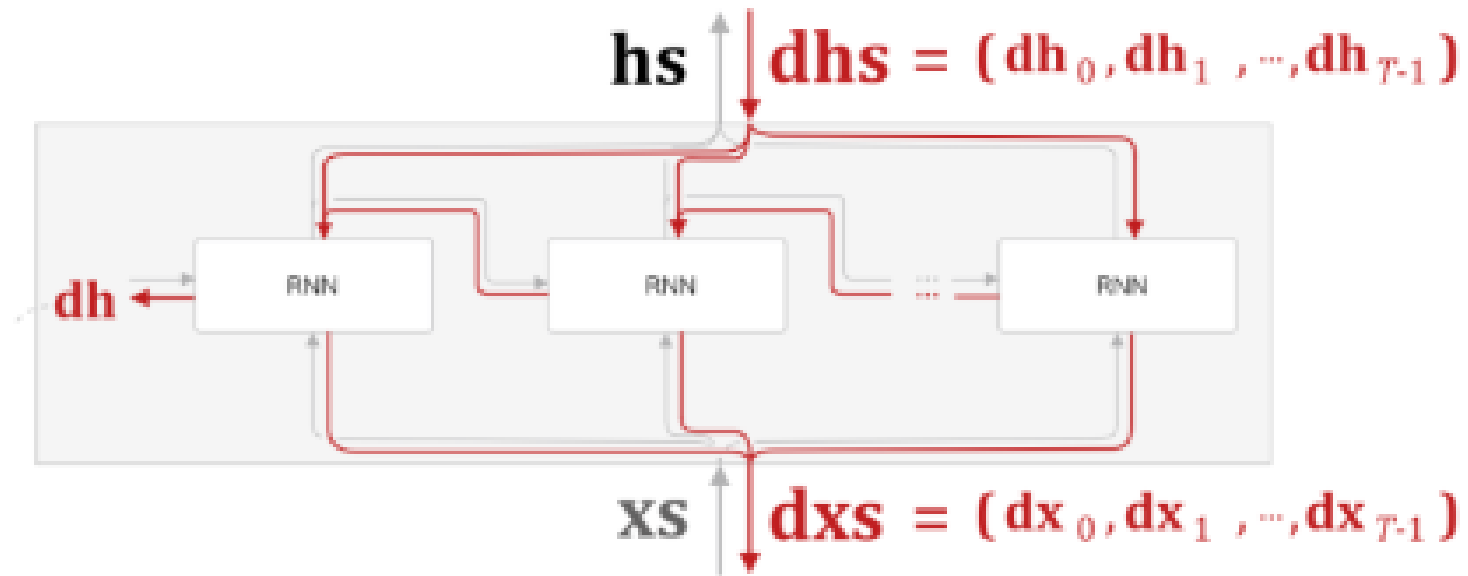




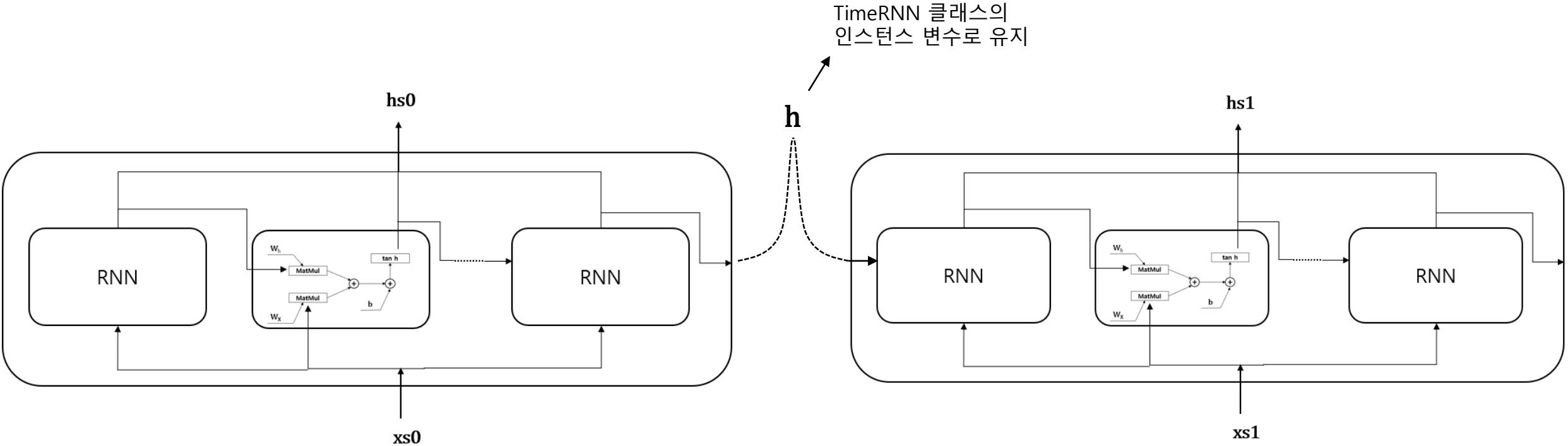
Time RNN 계층과 RNN 계층



## Time RNN 계층의 역전파

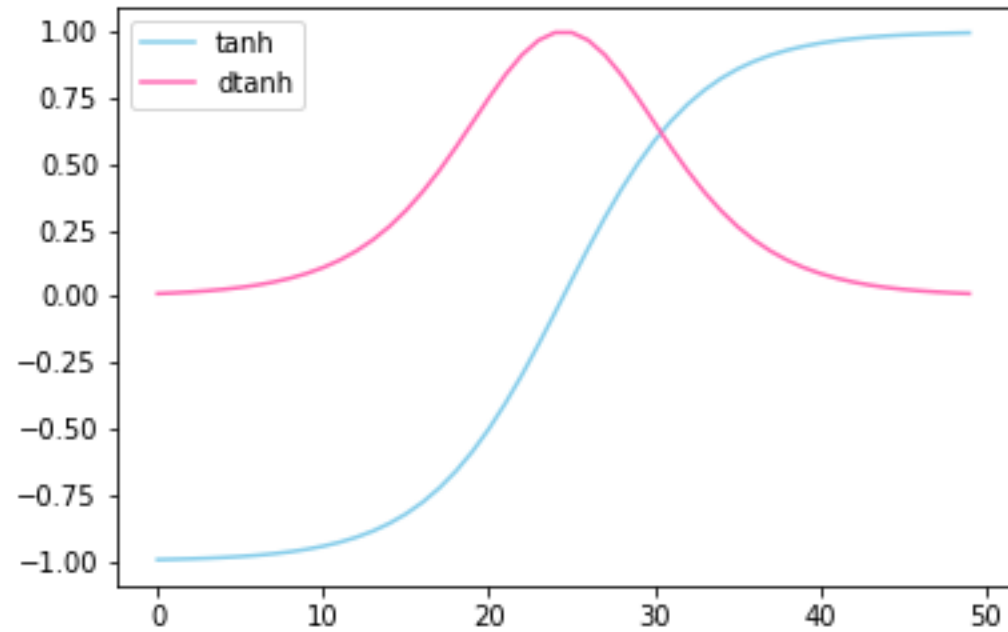


# Time RNN 계층과 RNN 계층

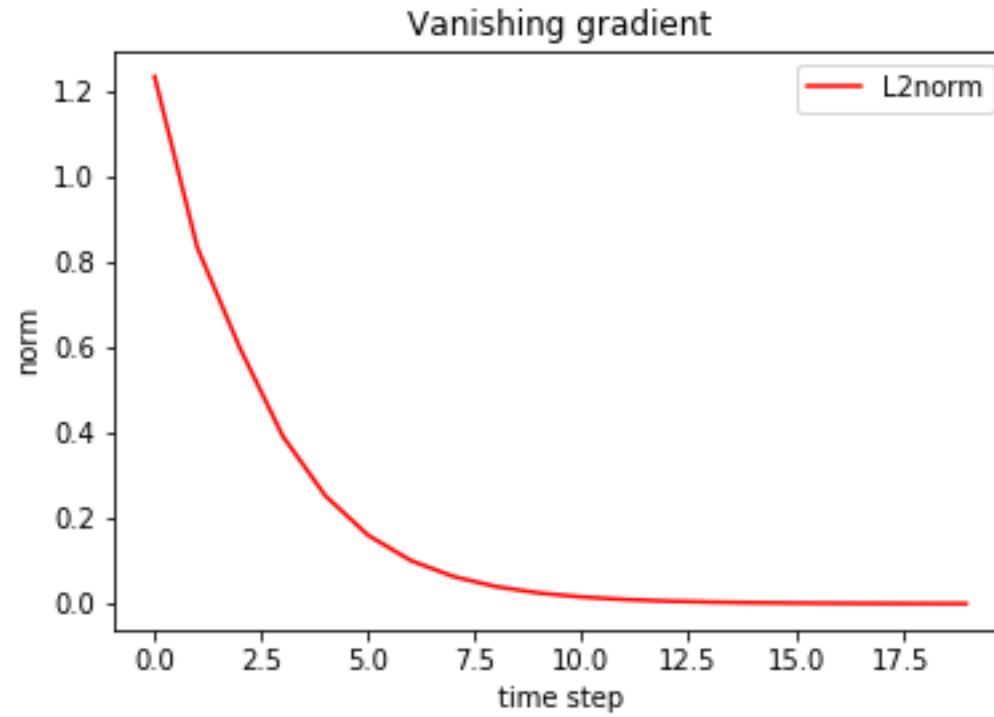


# Hyperbolic tangent

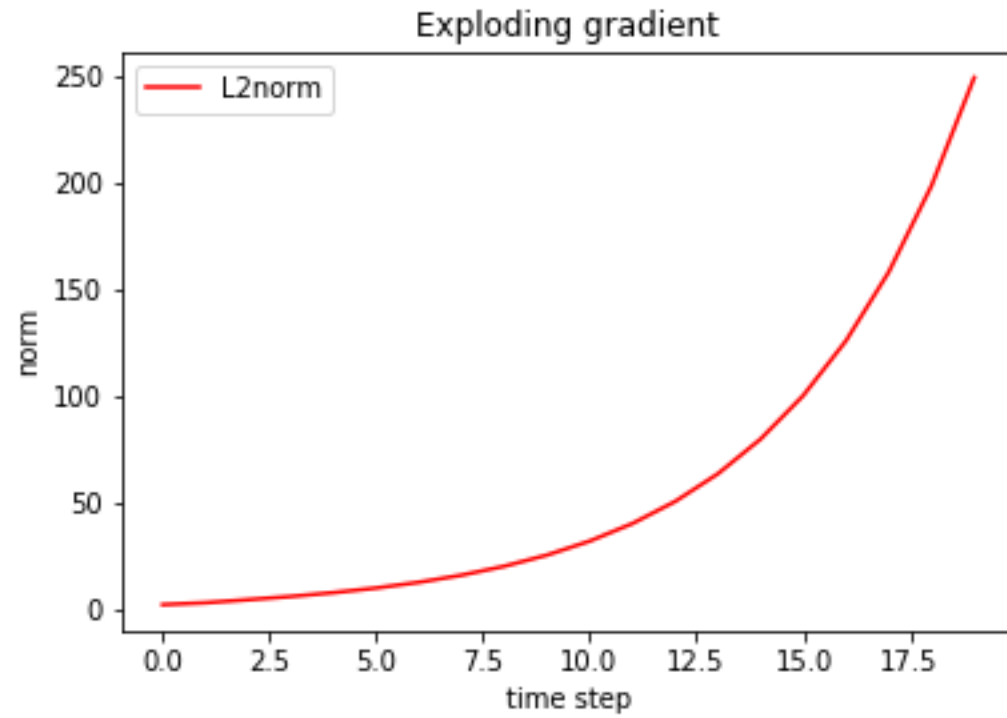
## $\tanh$ & $\text{d}\tanh$



## Vanishing gradient

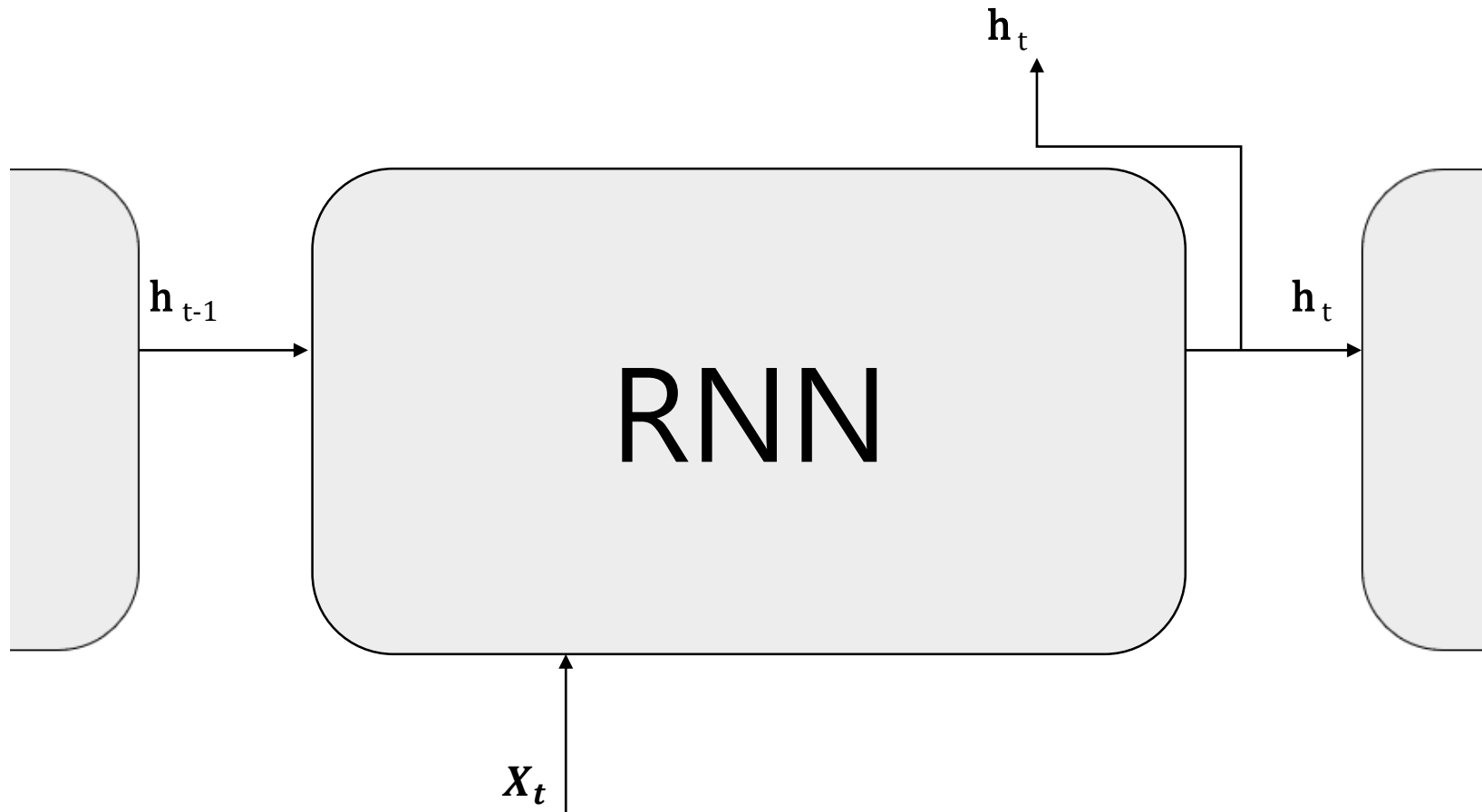


## Exploding gradient

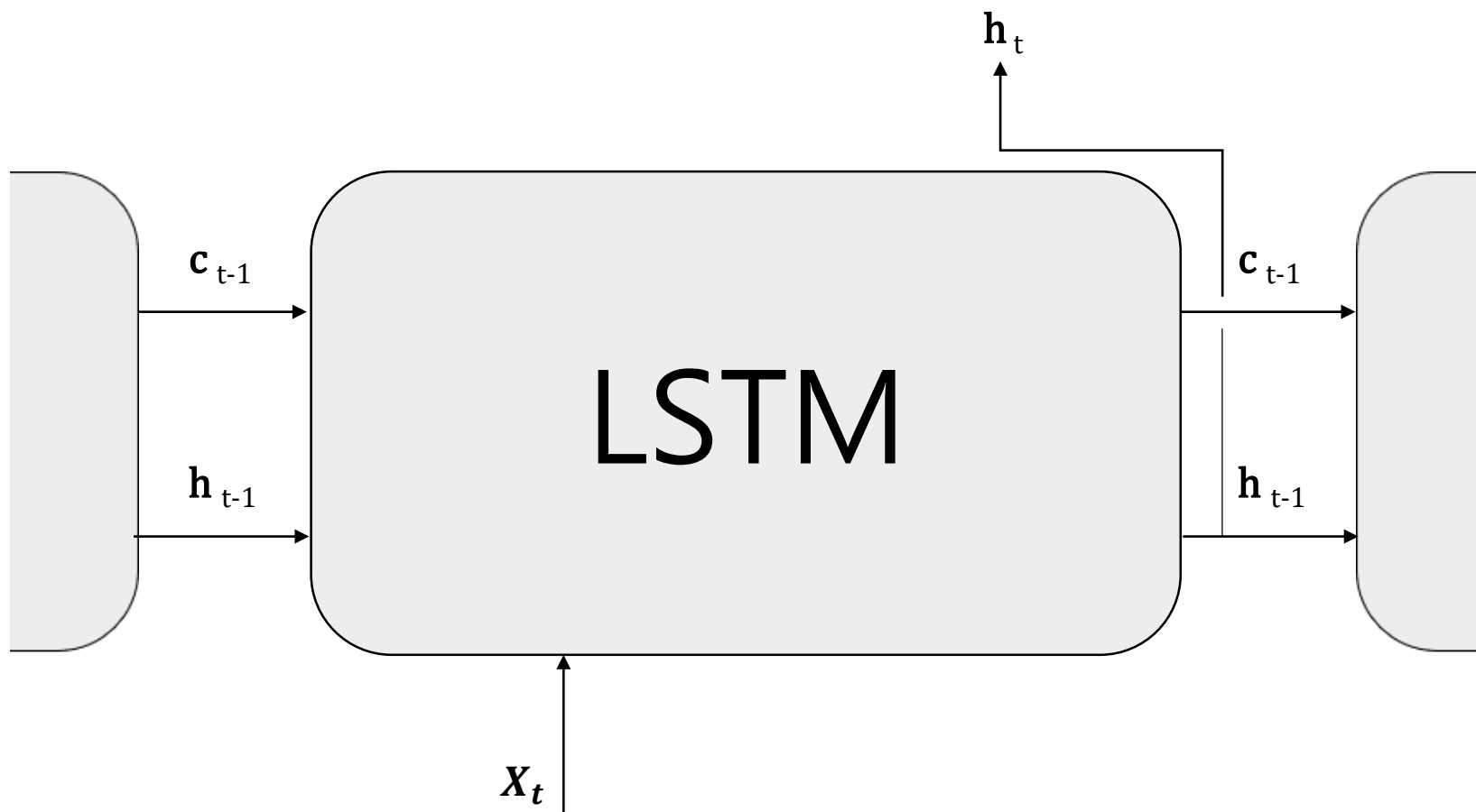


# Recurrent Neural Network (RNN)

## Interface



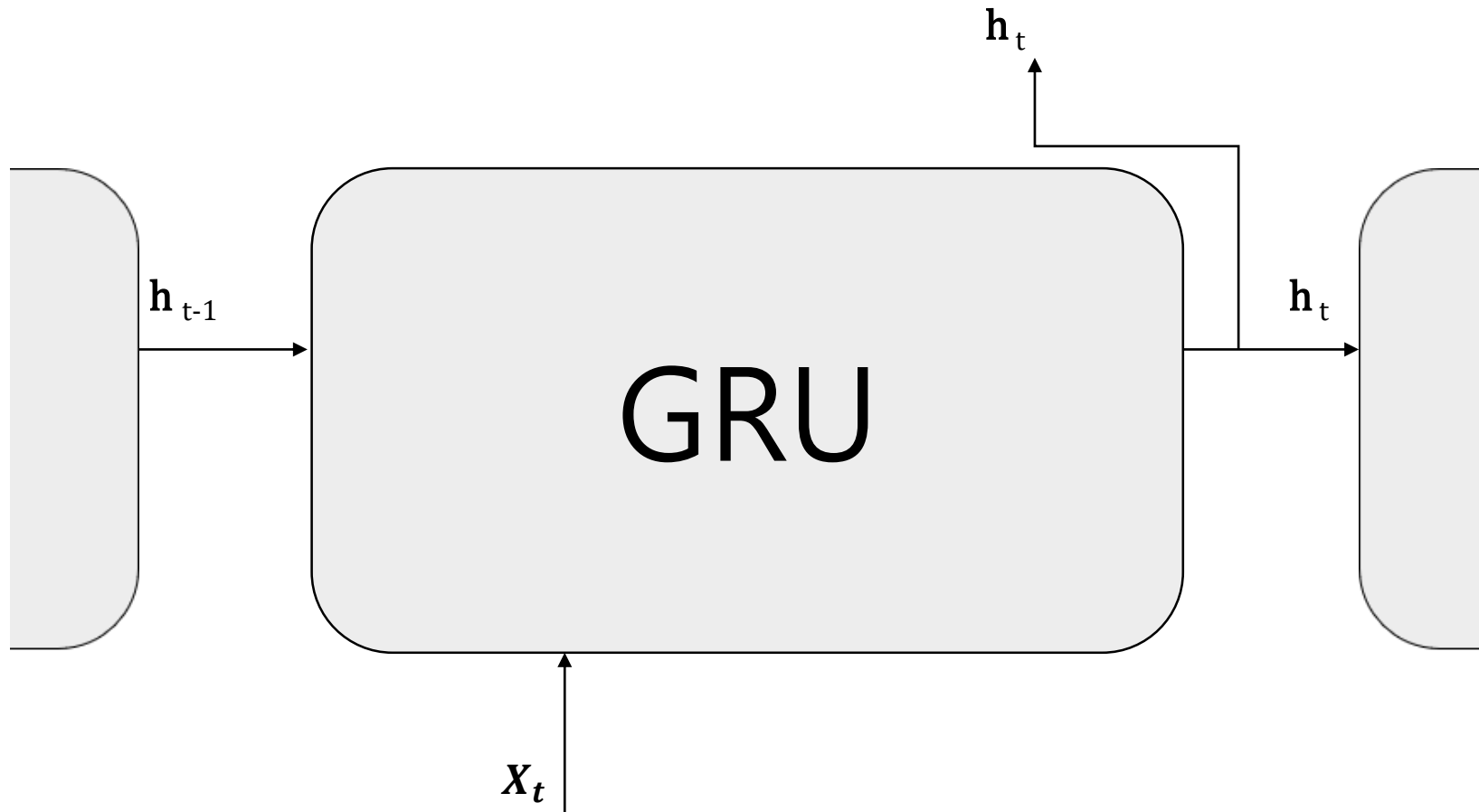
# Long Shor Term Memory (LSTM) Interface



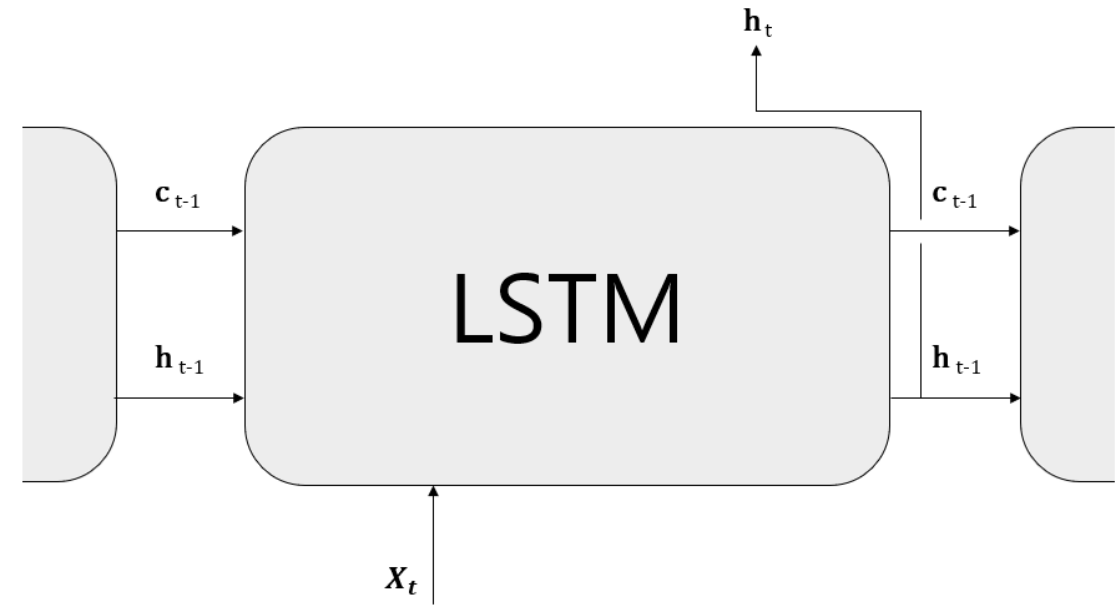
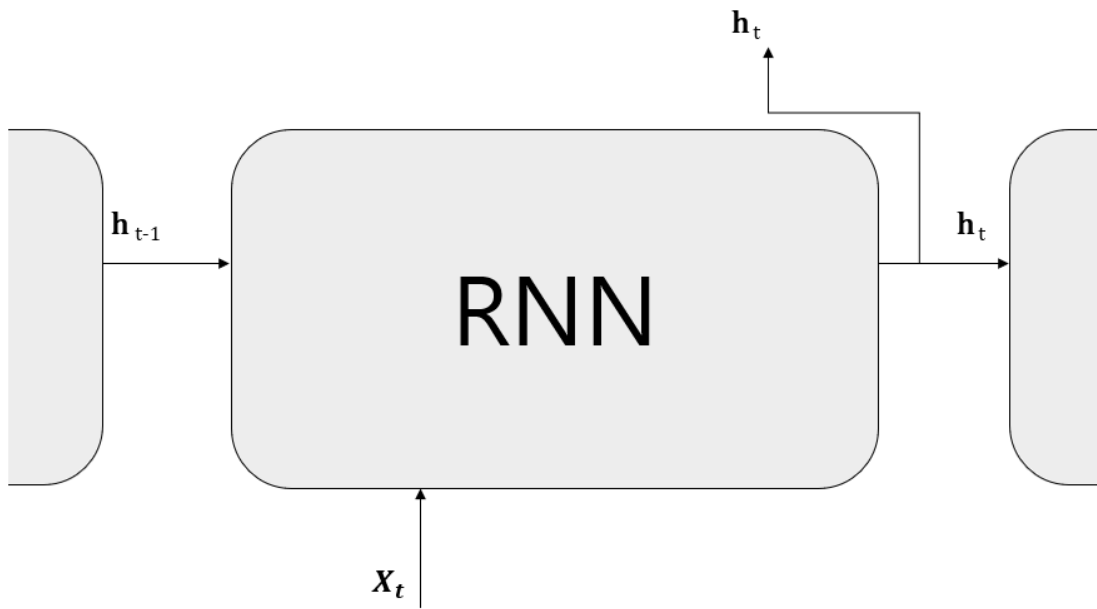


# Gate Recurrent Unit (GRU)

## Interface

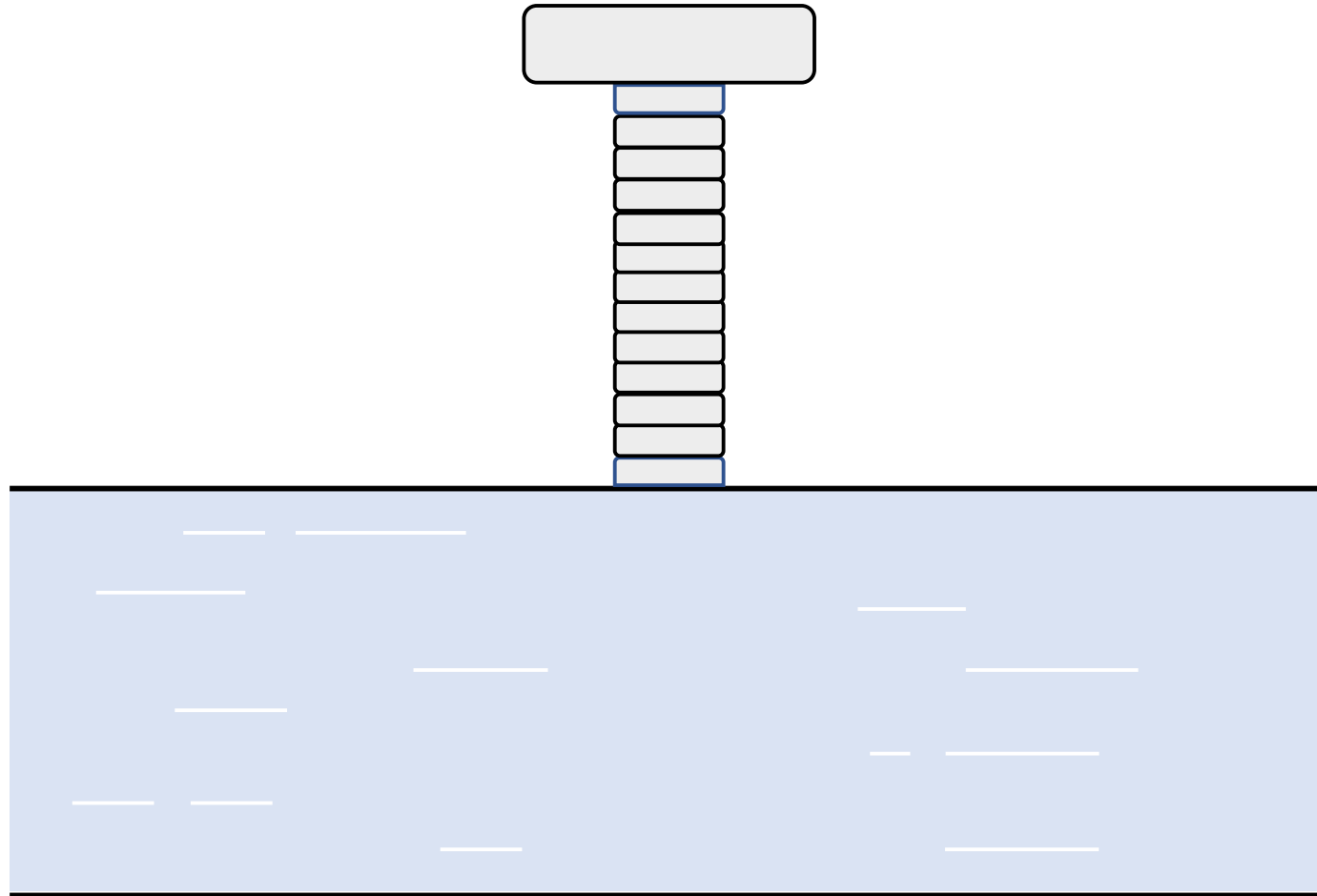


## RNN과 LSTM Interface 비교



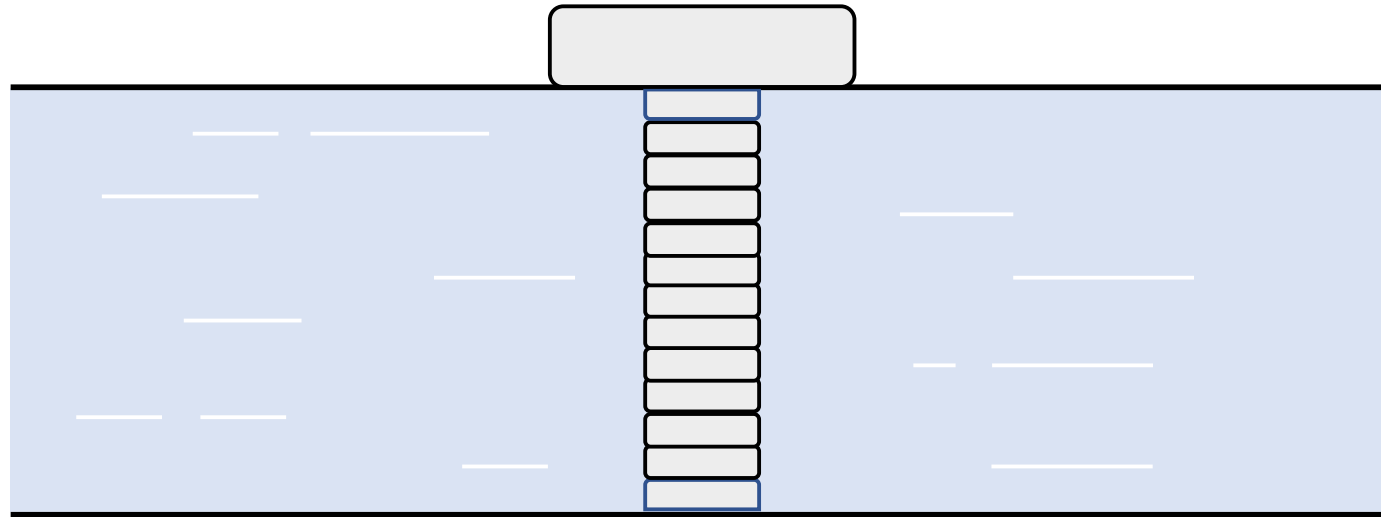
# Long Short Term Memory (LSTM)

비유하자면 게이트는 물의 흐름을 제어한다



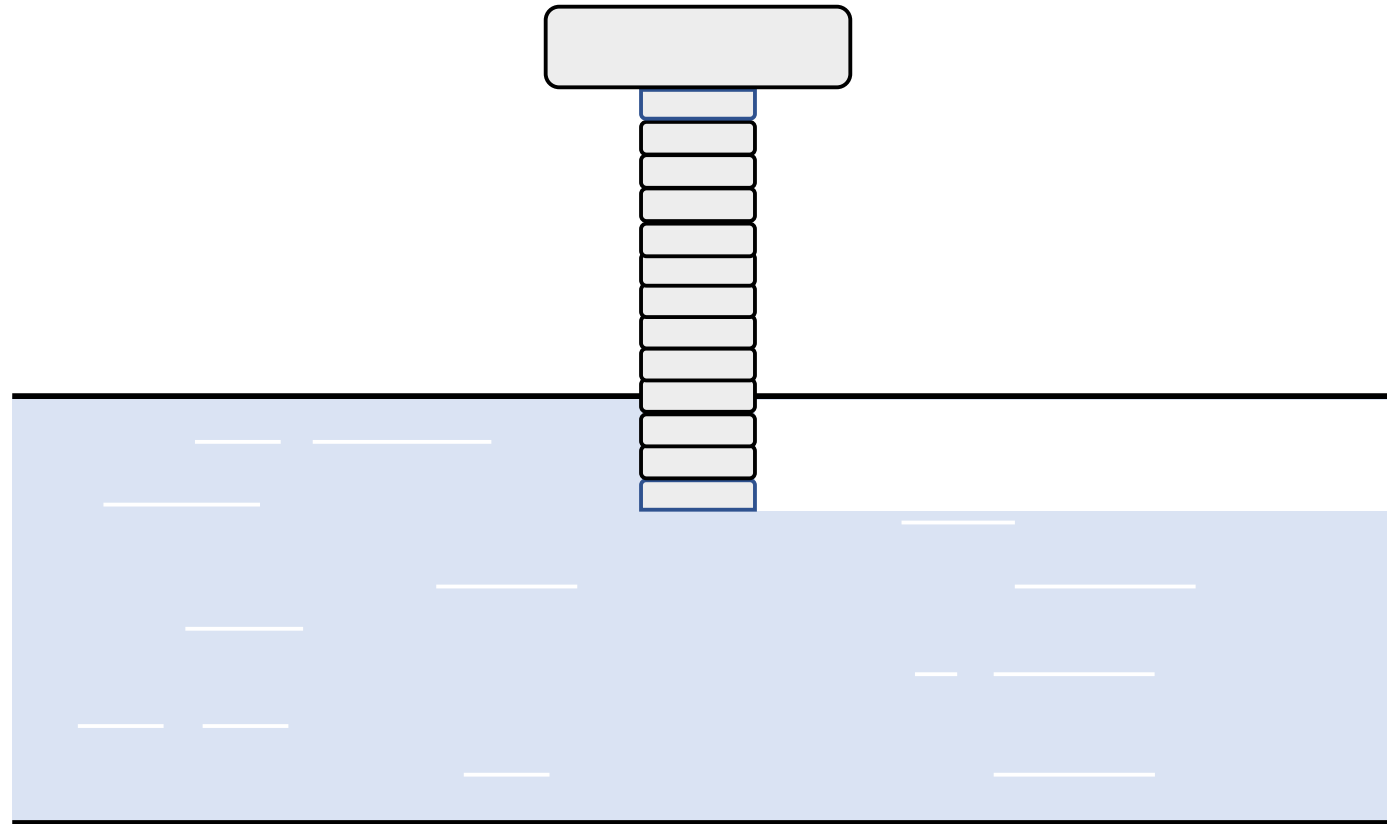
# Long Shor Term Memory (LSTM)

비유하자면 게이트는 물의 흐름을 제어한다



## Long Short Term Memory (LSTM)

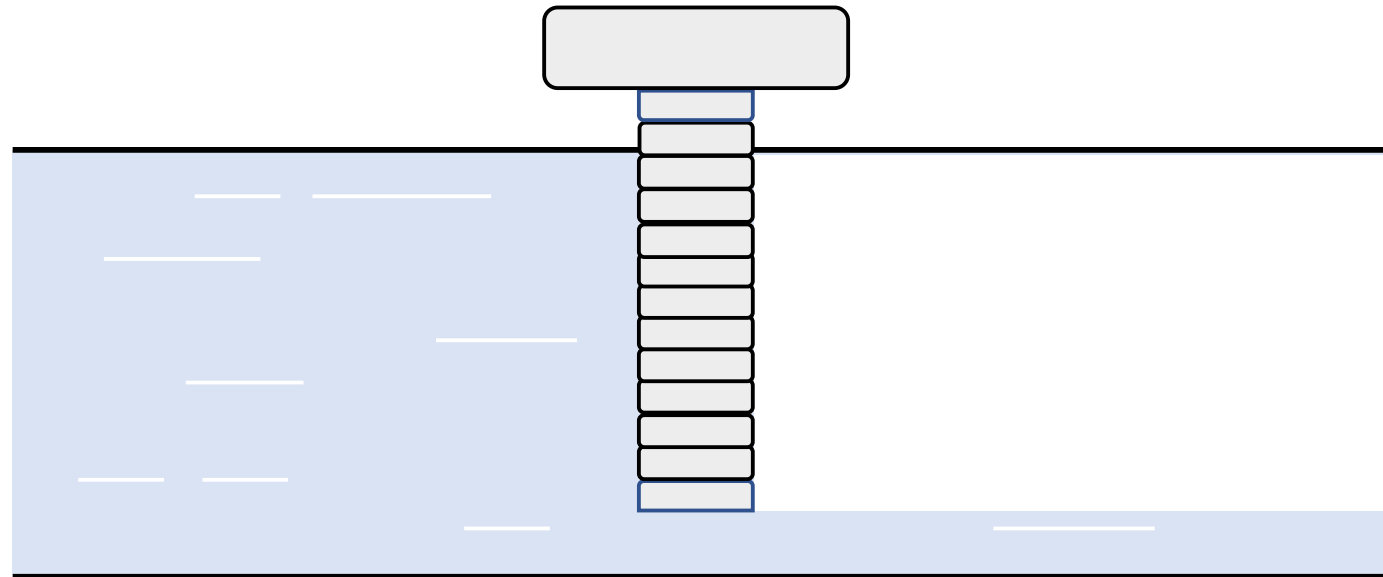
물이 흐르는 양을 0.0 ~ 1.0 범위에서 제어한다.



0.7 (70%)

## Long Short Term Memory (LSTM)

물이 흐르는 양을 0.0 ~ 1.0 범위에서 제어한다.



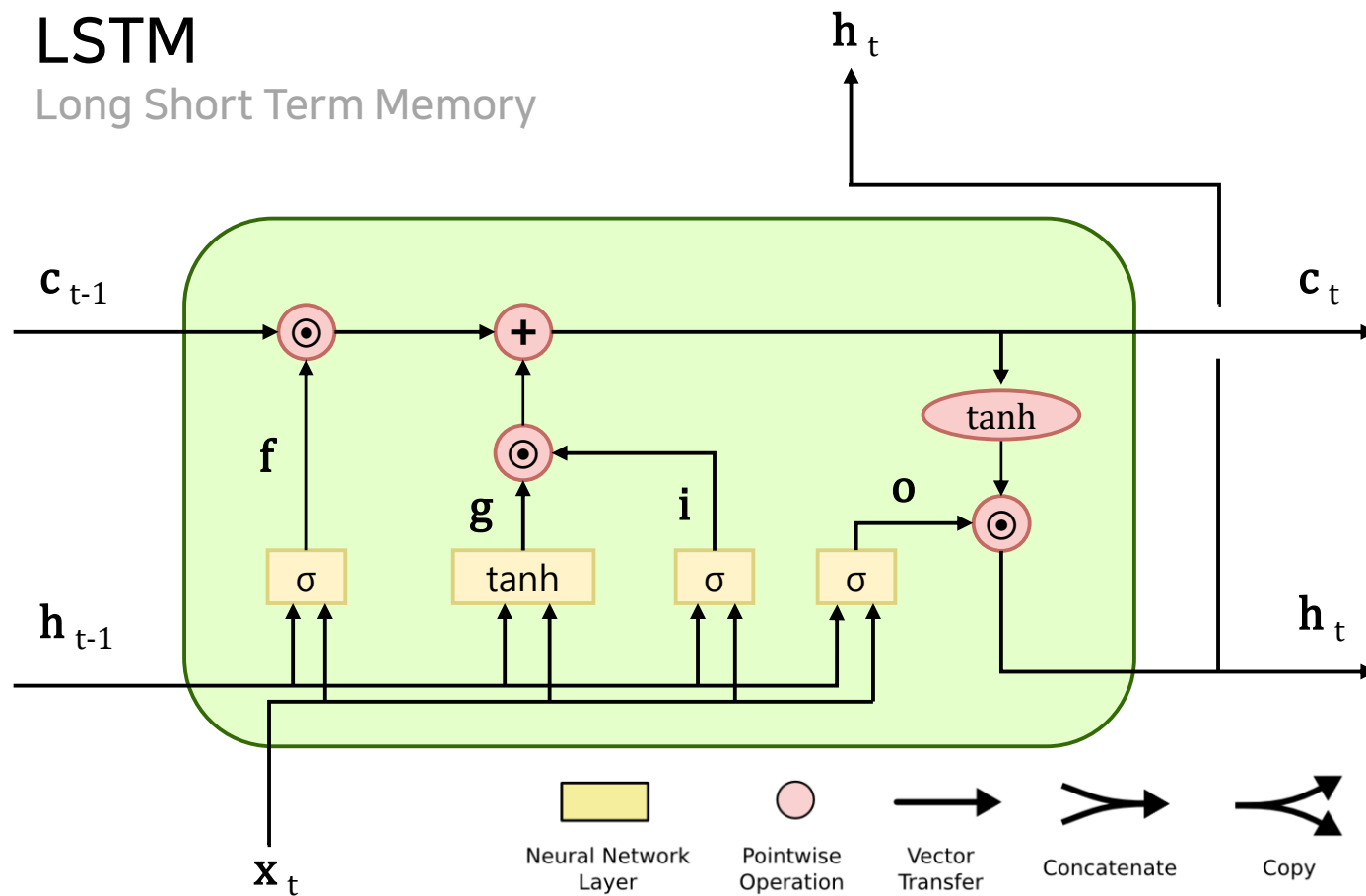
0.2 (20%)

# Long Short Term Memory (LSTM)

## LSTM의 계산 그래프

### LSTM

Long Short Term Memory



$$f = \sigma(x_t W_x^{(f)} + h_{t-1} W_h^{(f)} + b^{(f)})$$

$$g = \tanh(x_t W_x^{(g)} + h_{t-1} W_h^{(g)} + b^{(g)})$$

$$i = \sigma(x_t W_x^{(i)} + h_{t-1} W_h^{(i)} + b^{(i)})$$

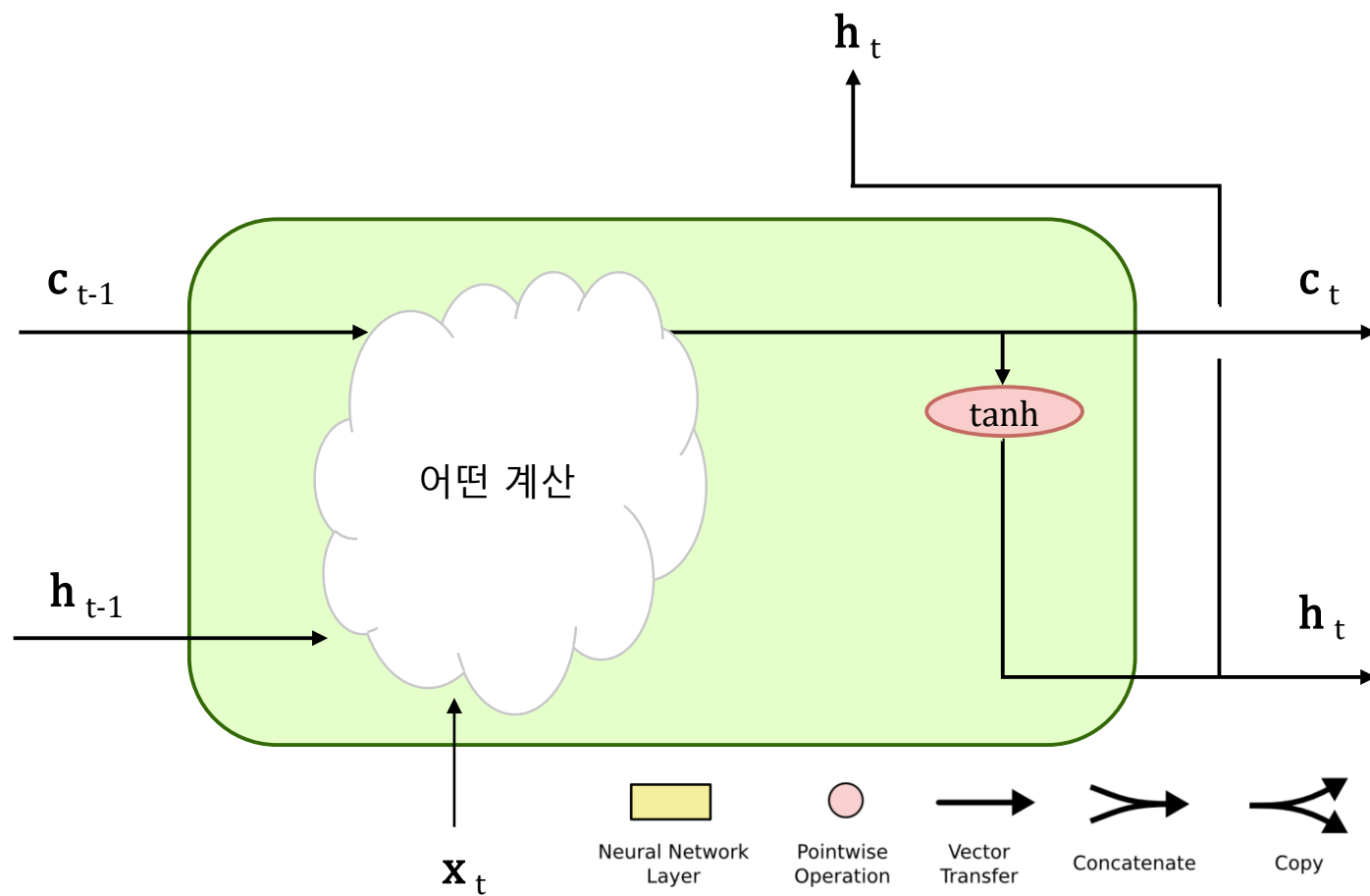
$$o = \sigma(x_t W_x^{(o)} + h_{t-1} W_h^{(o)} + b^{(o)})$$

$$c_t = f \odot c_{t-1} + g \odot i$$

$$h_t = o \odot \tanh(c_t)$$

# Long Short Term Memory (LSTM)

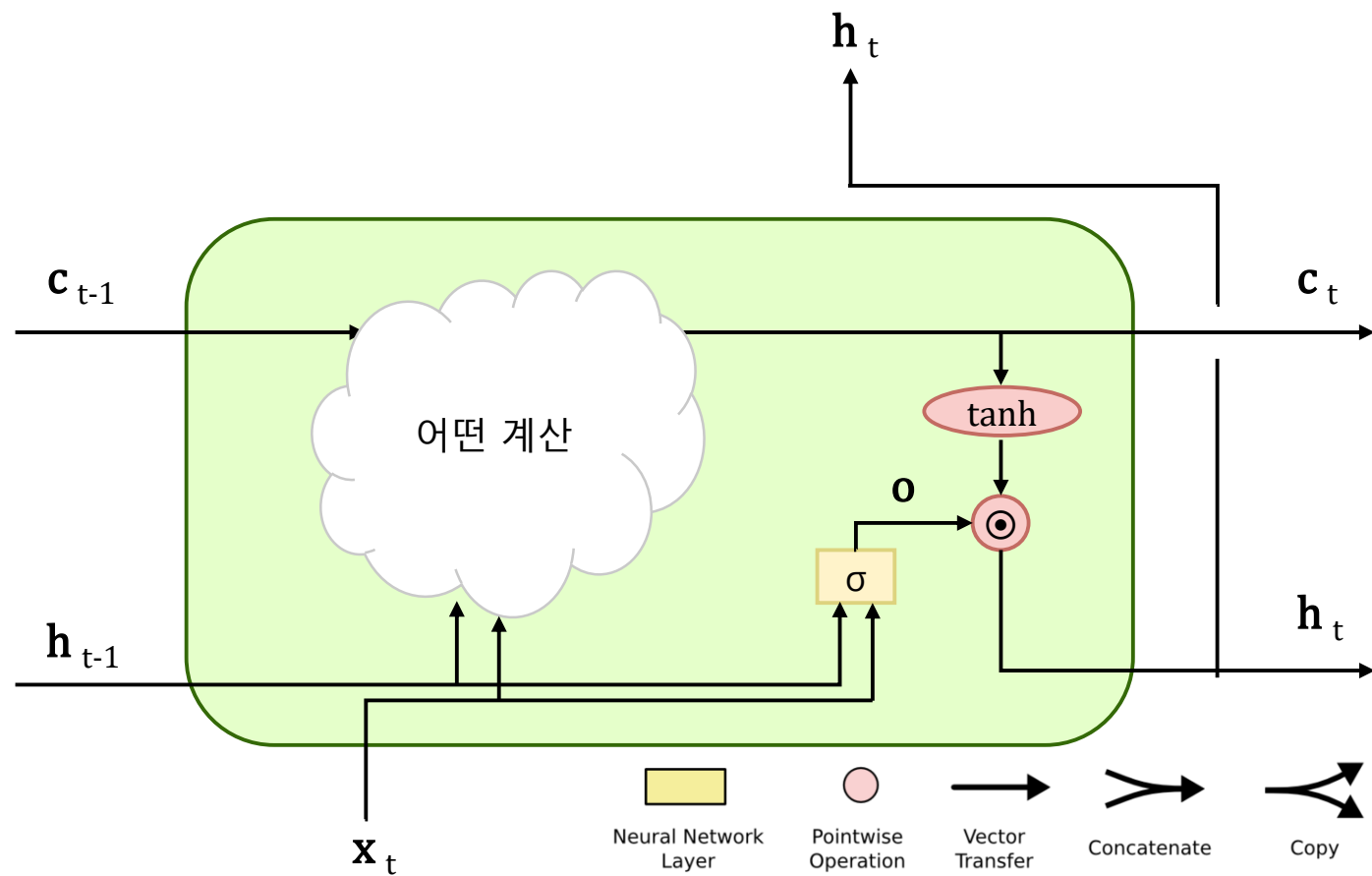
기억 셀  $c_t$ 를 바탕으로 은닉상태  $h_t$ 를 계산하는 LSTM 계층





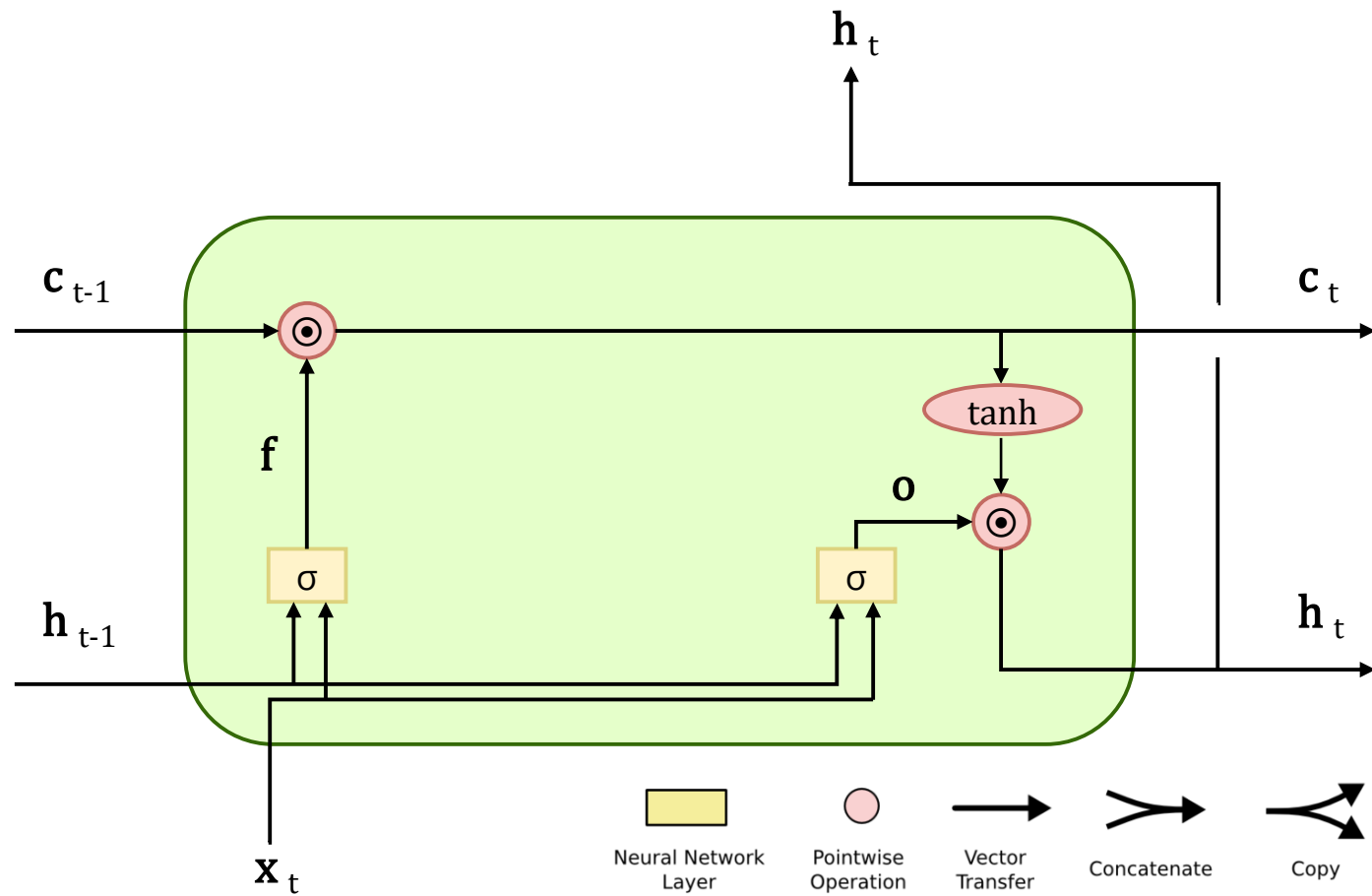
# Long Short Term Memory (LSTM)

output 게이트 추가 (o gate)



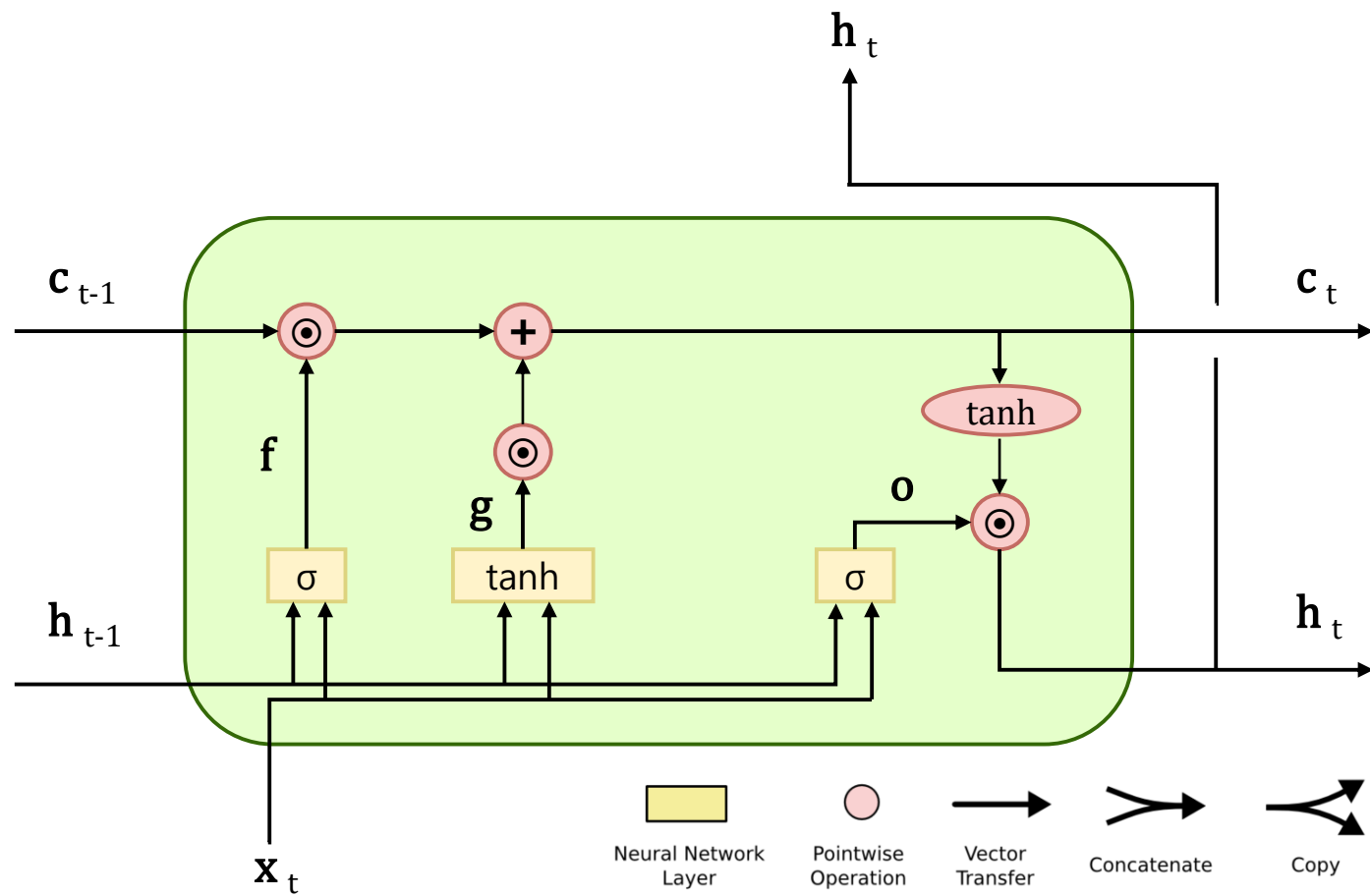
# Long Short Term Memory (LSTM)

forget 게이트 추가 (f gate)



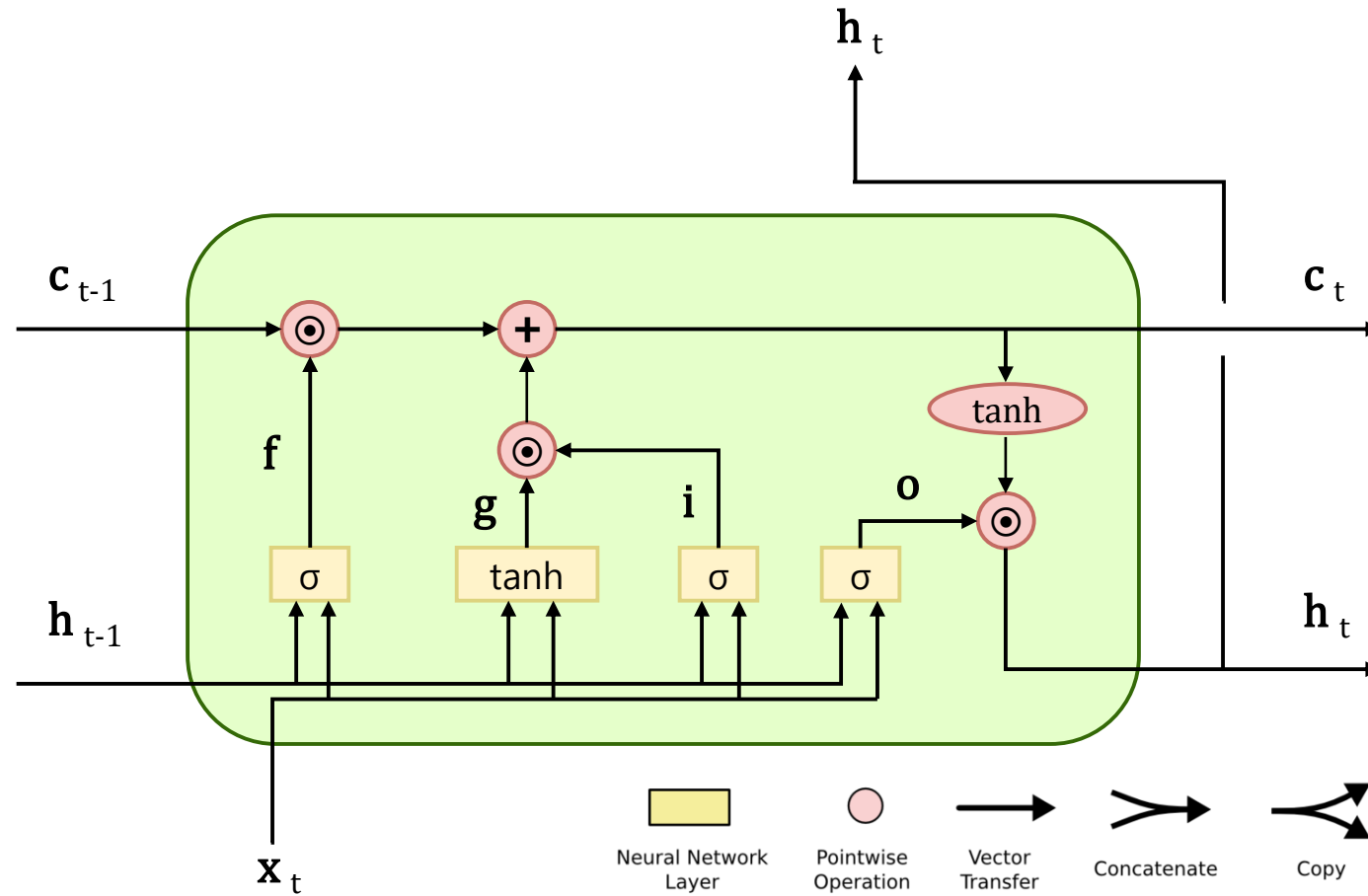
# Long Short Term Memory (LSTM)

새로운 기억 셀에 필요한 정보를 추가 (g gate)

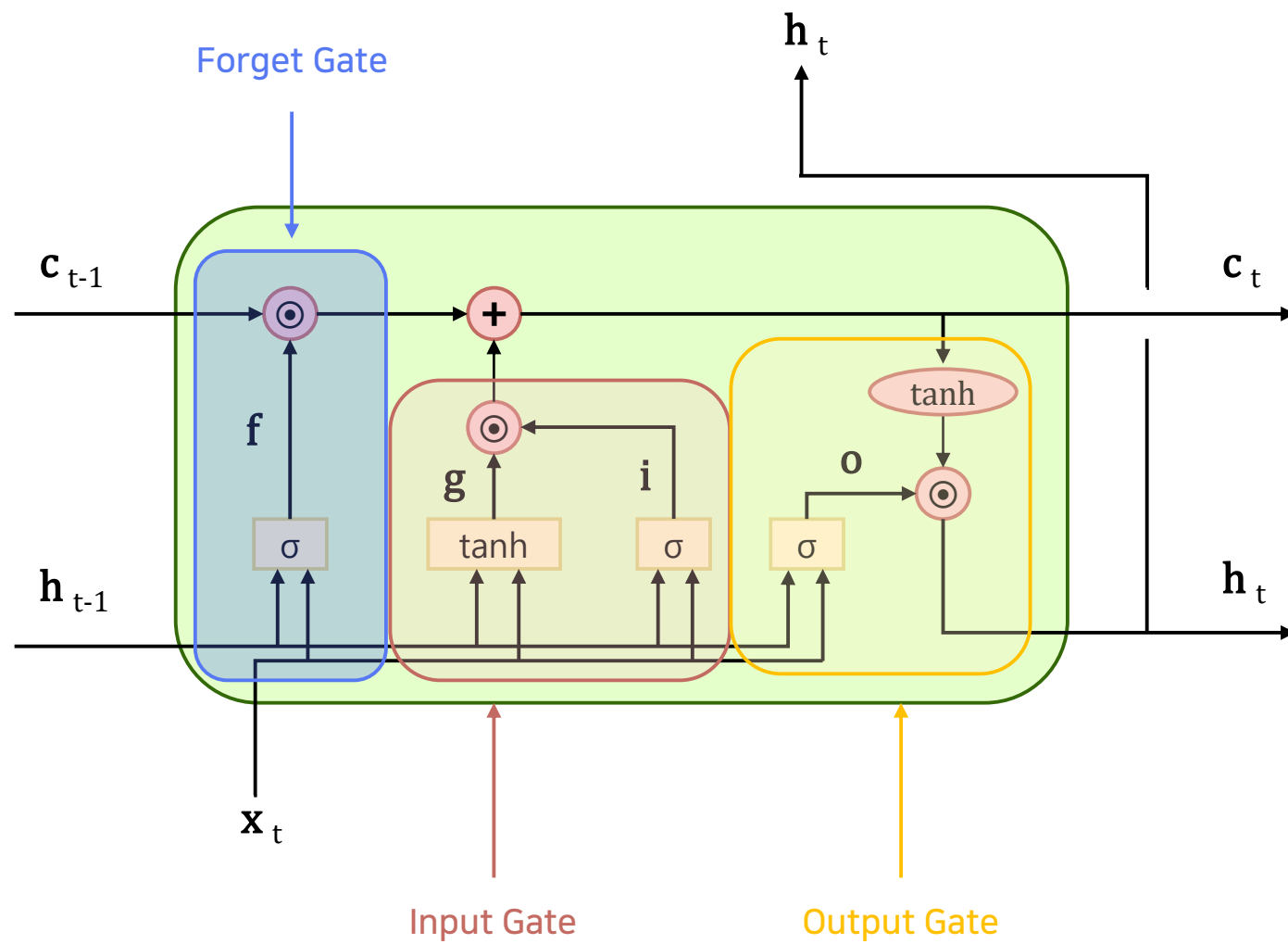


# Long Short Term Memory (LSTM)

## Input 게이트 추가 (i gate)

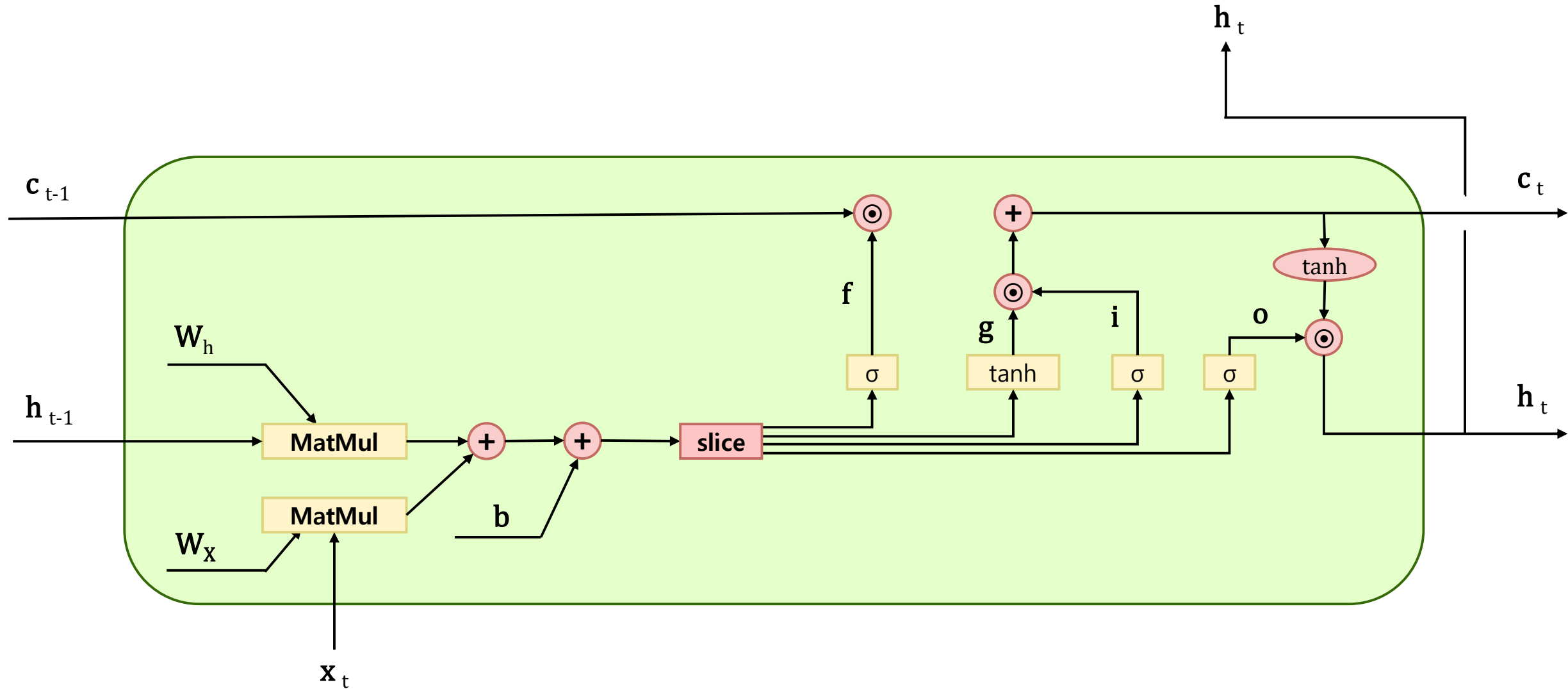


# Long Short Term Memory (LSTM)

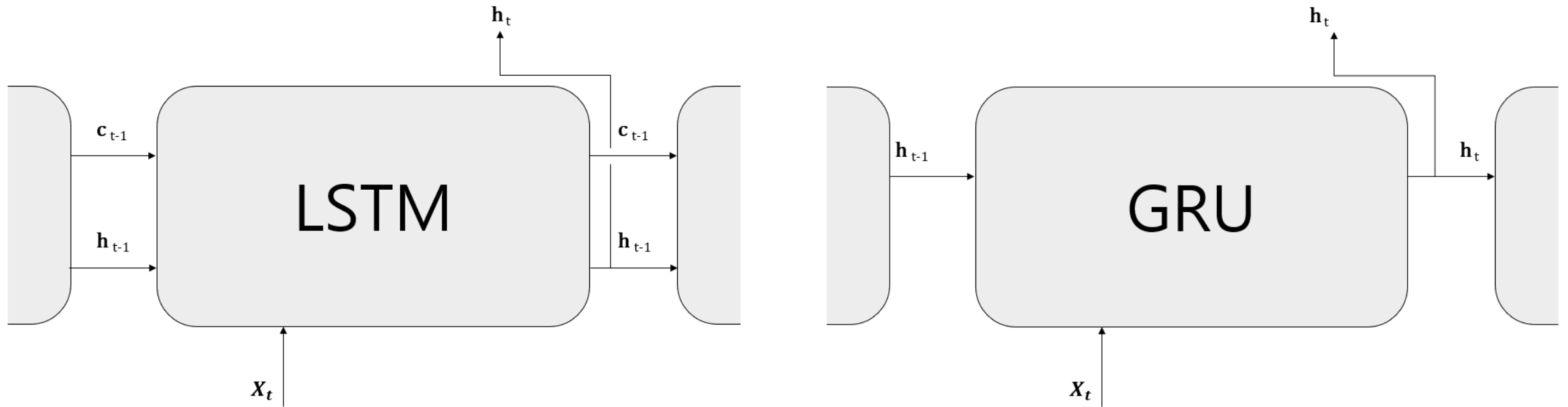


# Long Short Term Memory (LSTM)

4개분의 가중치를 모아 아핀 변환을 수행하는 LSTM의 계산그래프



## LSTM과 GRU Interface 비교

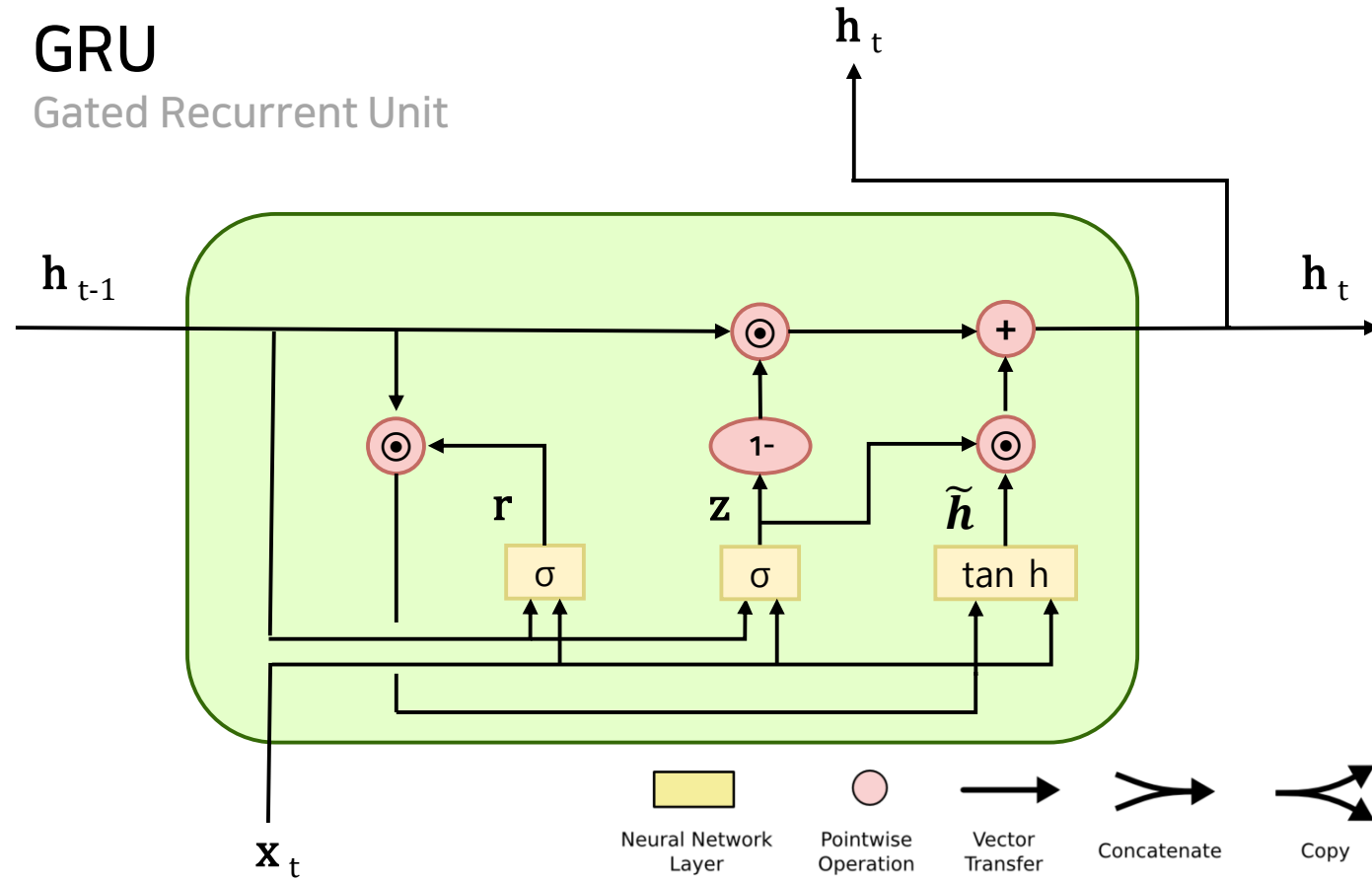


# Gated Recurrent Unit (GRU)

## GRU의 계산 그래프

GRU

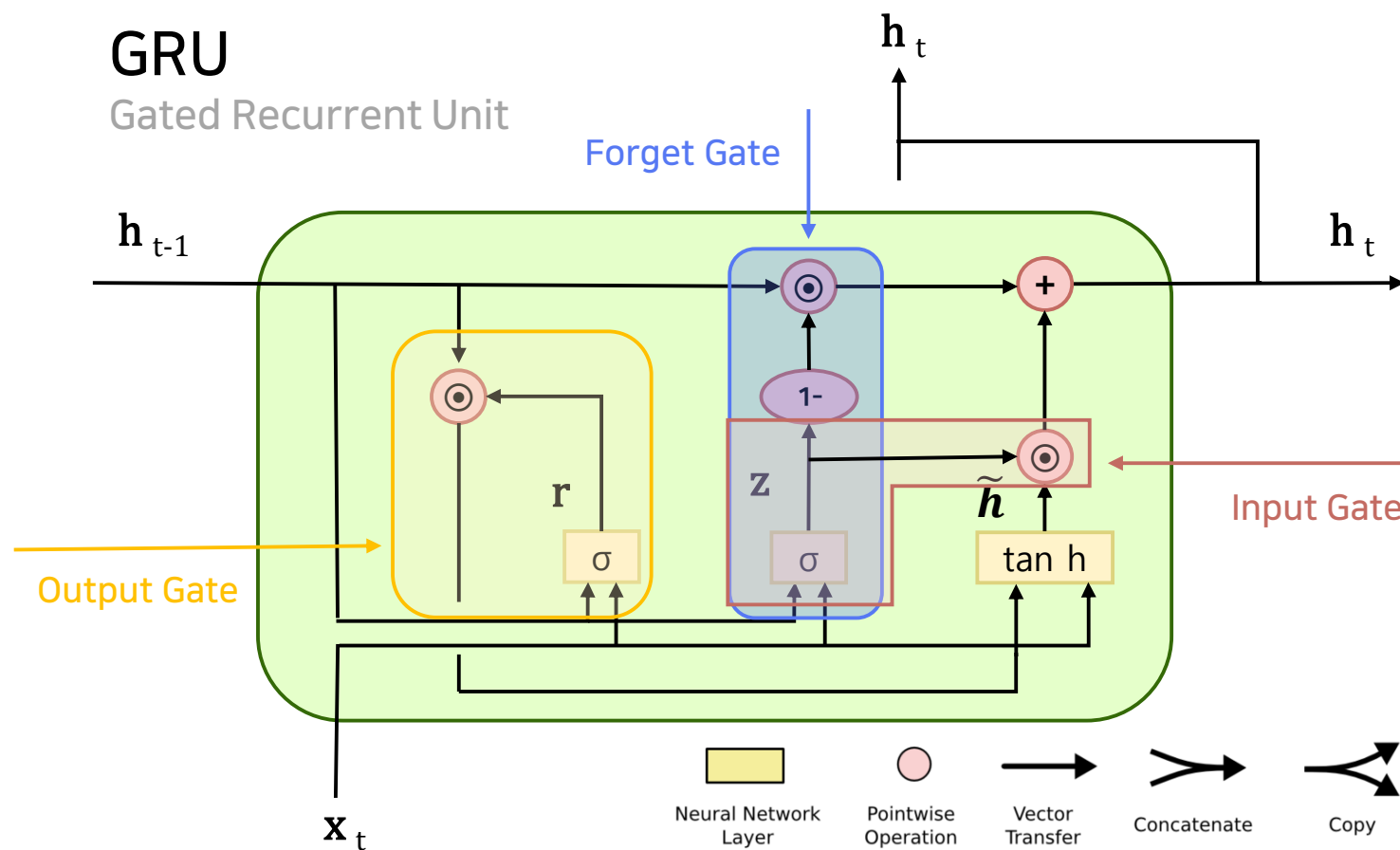
Gated Recurrent Unit





# Gated Recurrent Unit (GRU)

## GRU의 Forget Gate 와 Input Gate



# Gated Recurrent Unit (GRU)

GRU의 Forget Gate 와 Input Gate

