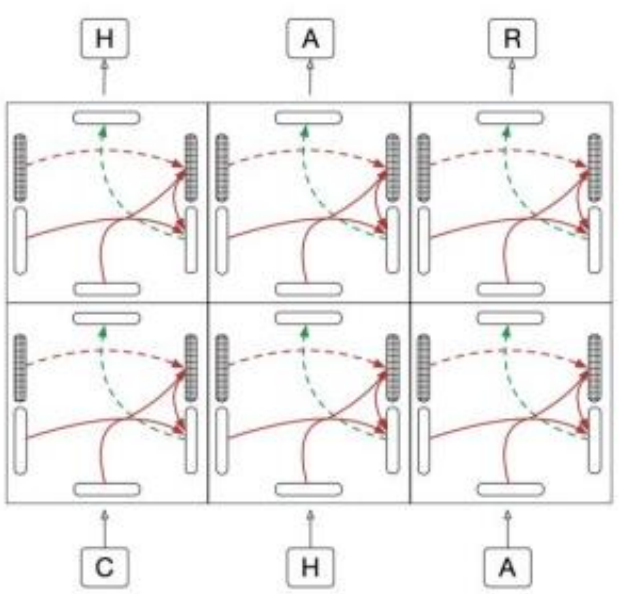
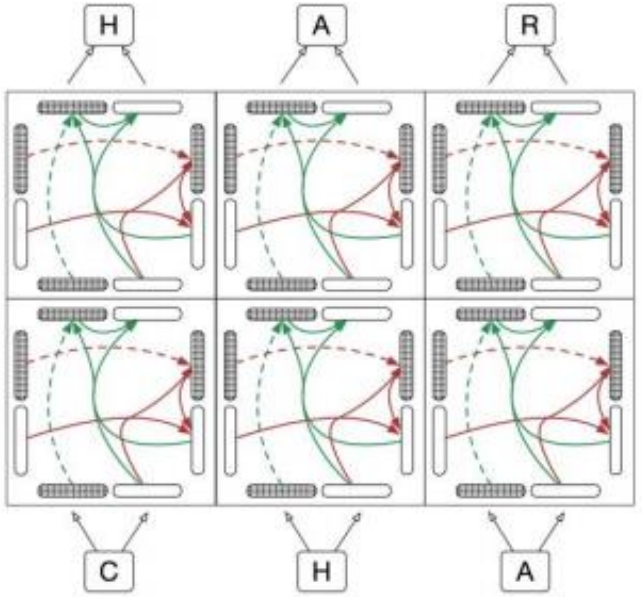
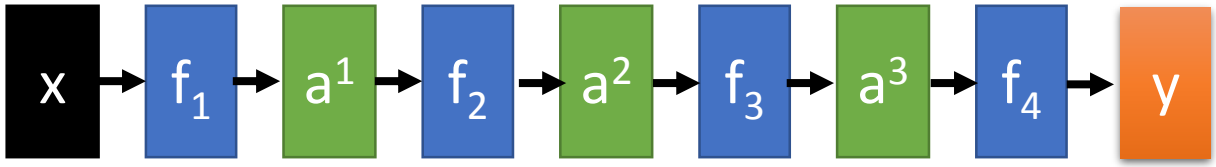
Highway Network & Grid LSTM





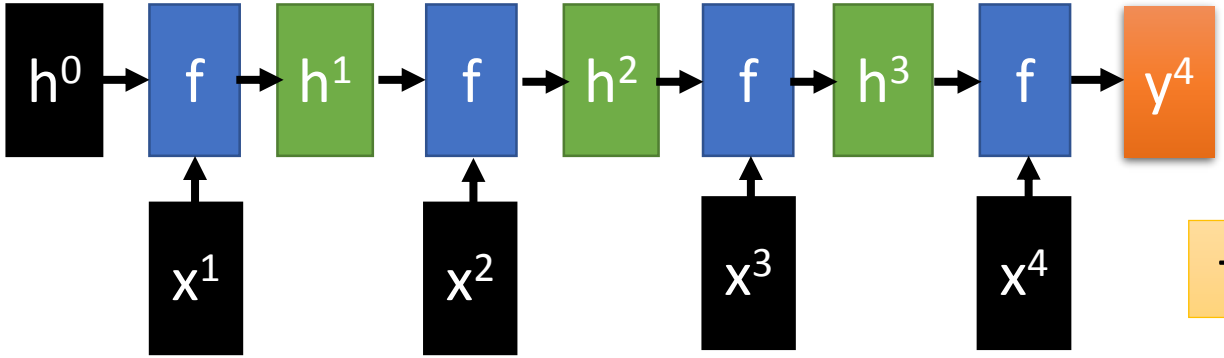
Feedforward v.s. Recurrent

1. Feedforward network does not have input at each step
2. Feedforward network has different parameters for each layer



$$a^t = f_l(a^{t-1}) = \sigma(W^t a^{t-1} + b^t)$$

t is layer



t is time step

$$h^t = f(h^{t-1}, x^t) = \sigma\left(W^h h^{t-1} + W^i x^t + b^i\right)$$

Applying gated structure in feedforward network

GRU → Highway Network

把RNN的network豎直從另一個方向看

No input x^t at each step

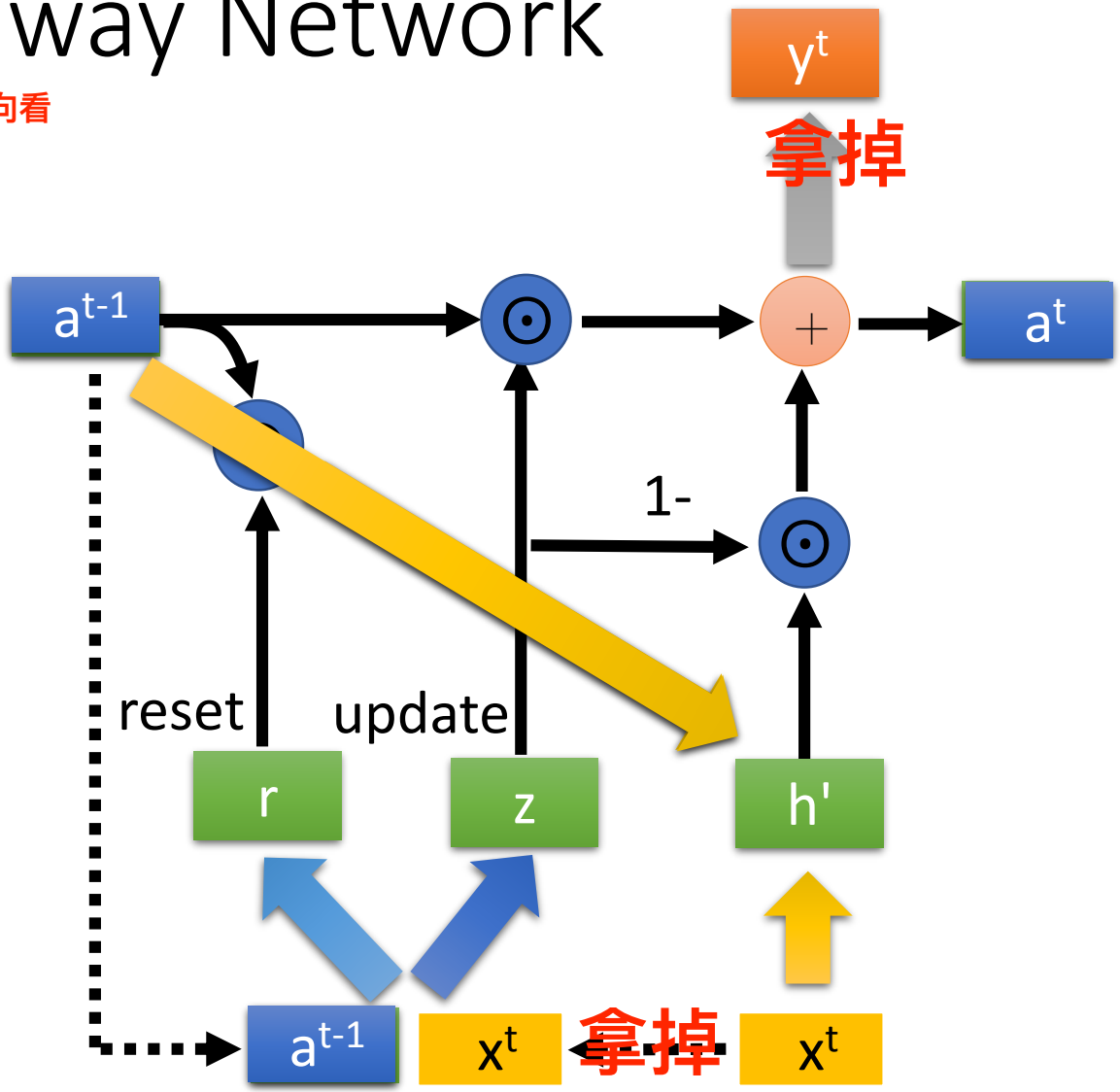
No output y^t at each step

a^{t-1} is the output of the (t-1)-th layer

a^t is the output of the t-th layer

No reset gate

因為每一層layer一定要有上一層的input
reset gate永遠是處於開啟的狀態



at–1前一個hidden layer的output

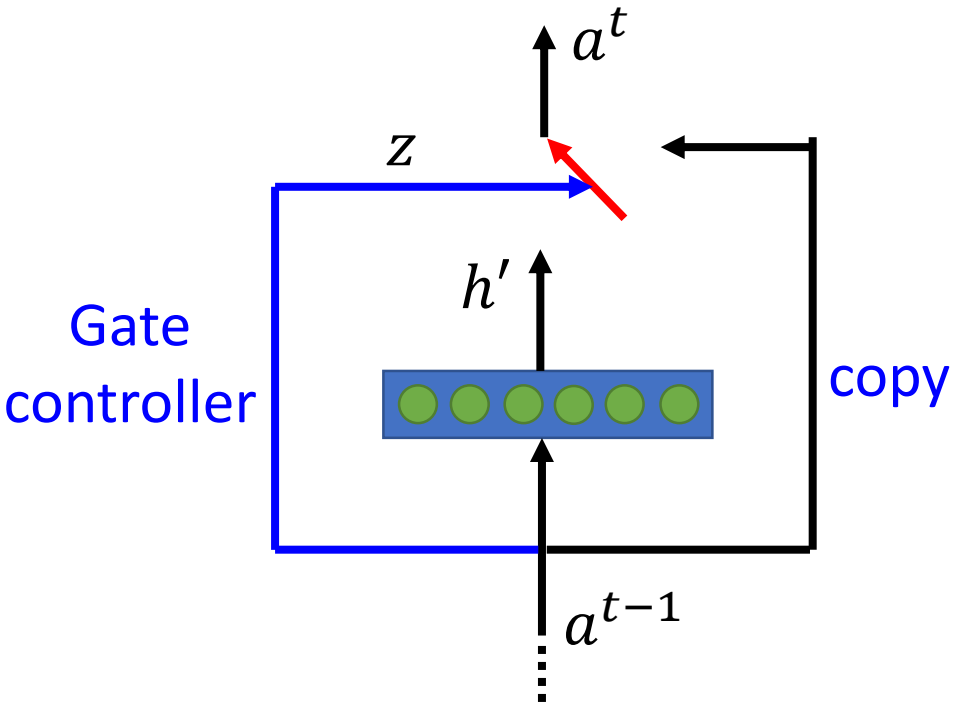
為了避免gradient vanishing問題，因為LSTM/GRU的設計就是為了解決這個問題

Highway Network

可以拿來train network with very deep layer

$$h' = \sigma(W a^{t-1})$$
$$z = \sigma(W' a^{t-1})$$
$$a^t = z \odot a^{t-1} + (1 - z) \odot h'$$

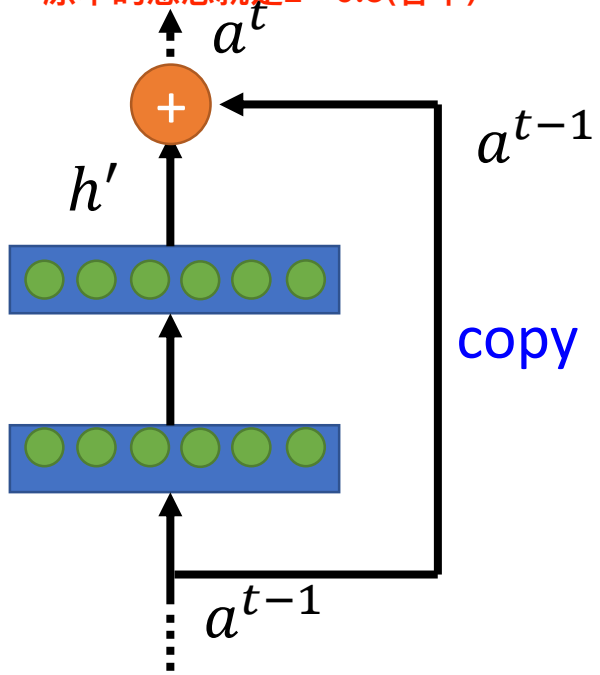
- **Highway Network**



Training Very Deep Networks
https://arxiv.org/pdf/1507.06228v2.pdf

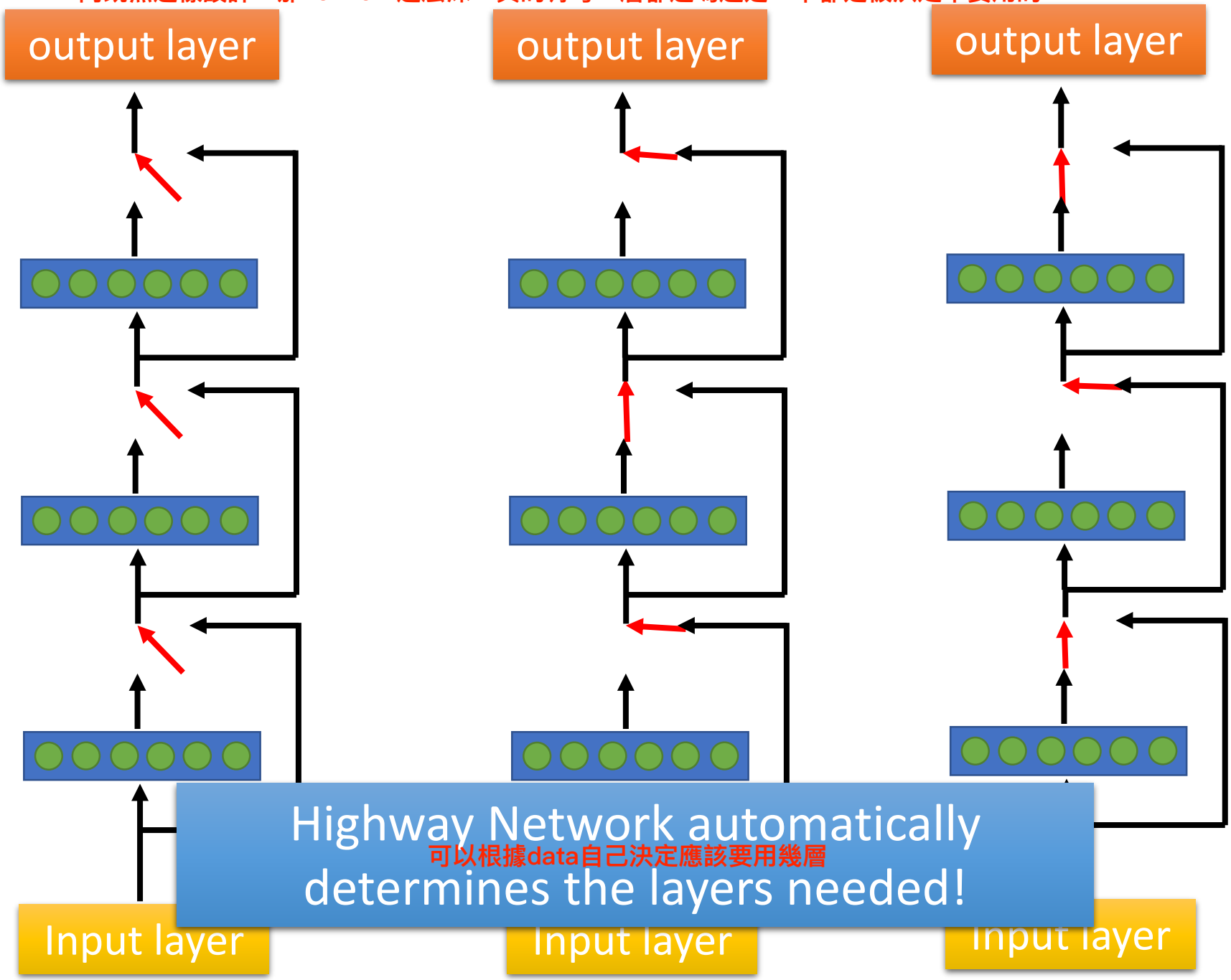
- **Residual Network**

原本的意思就是z= 0.5(各半)



Deep Residual Learning for Image Recognition
http://arxiv.org/abs/1512.03385

阿既然這樣設計，那network這麼深，真的有每一層都選嗎還是一半都是被決定不要用的？

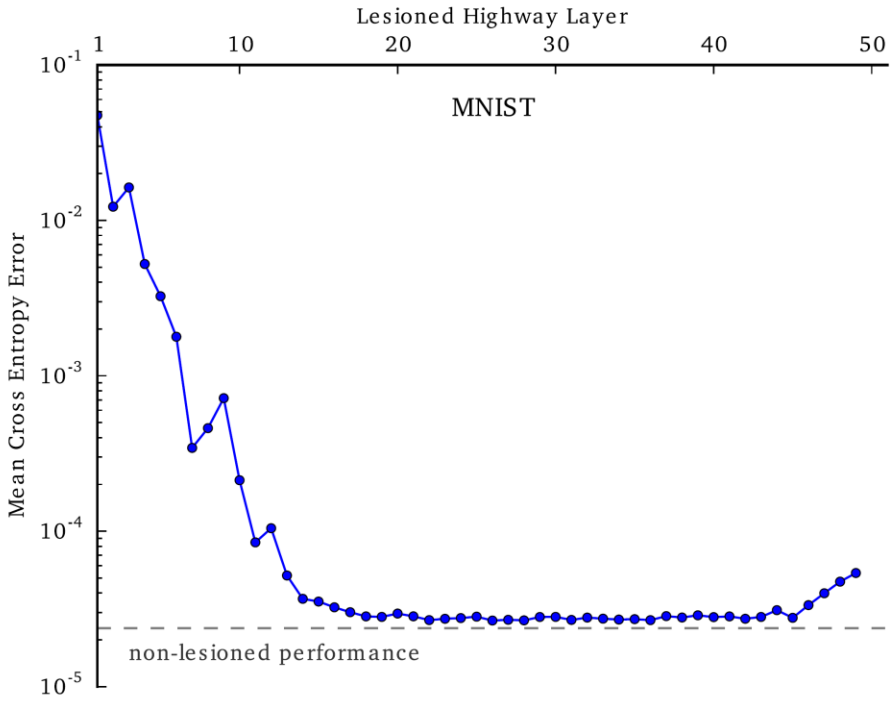


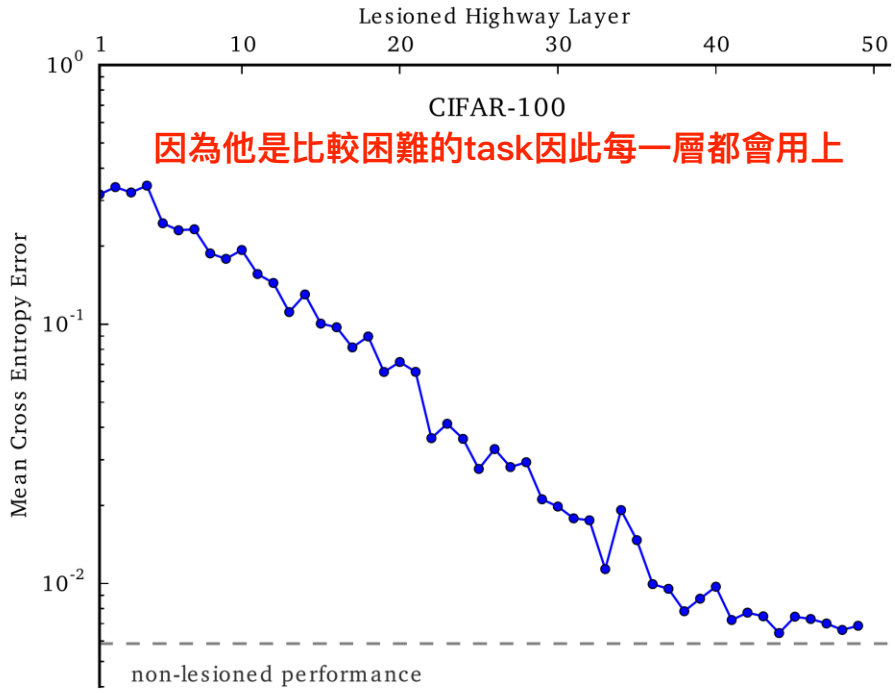
Highway Network automatically determines the layers needed!

可以根據data自己決定應該要用幾層

Highway Network

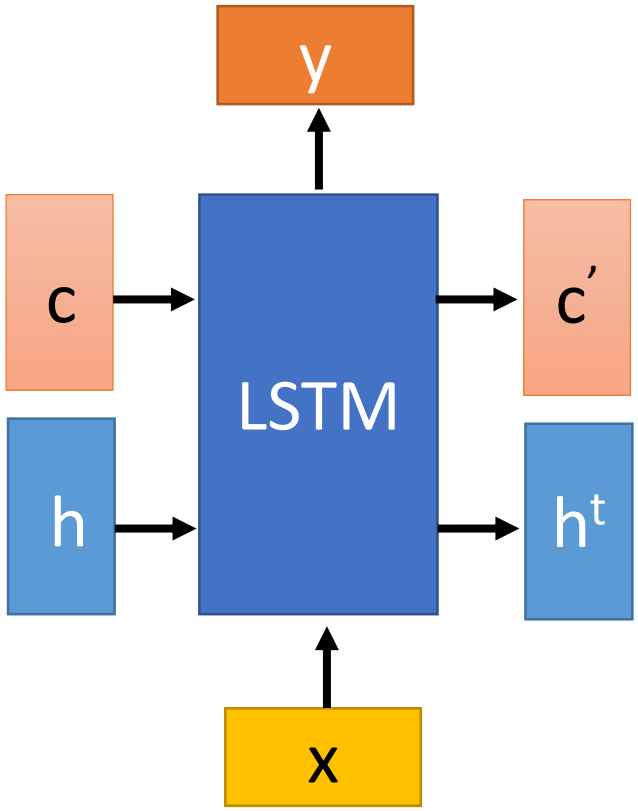
分類問題：把第一層到第五十層分別拔掉



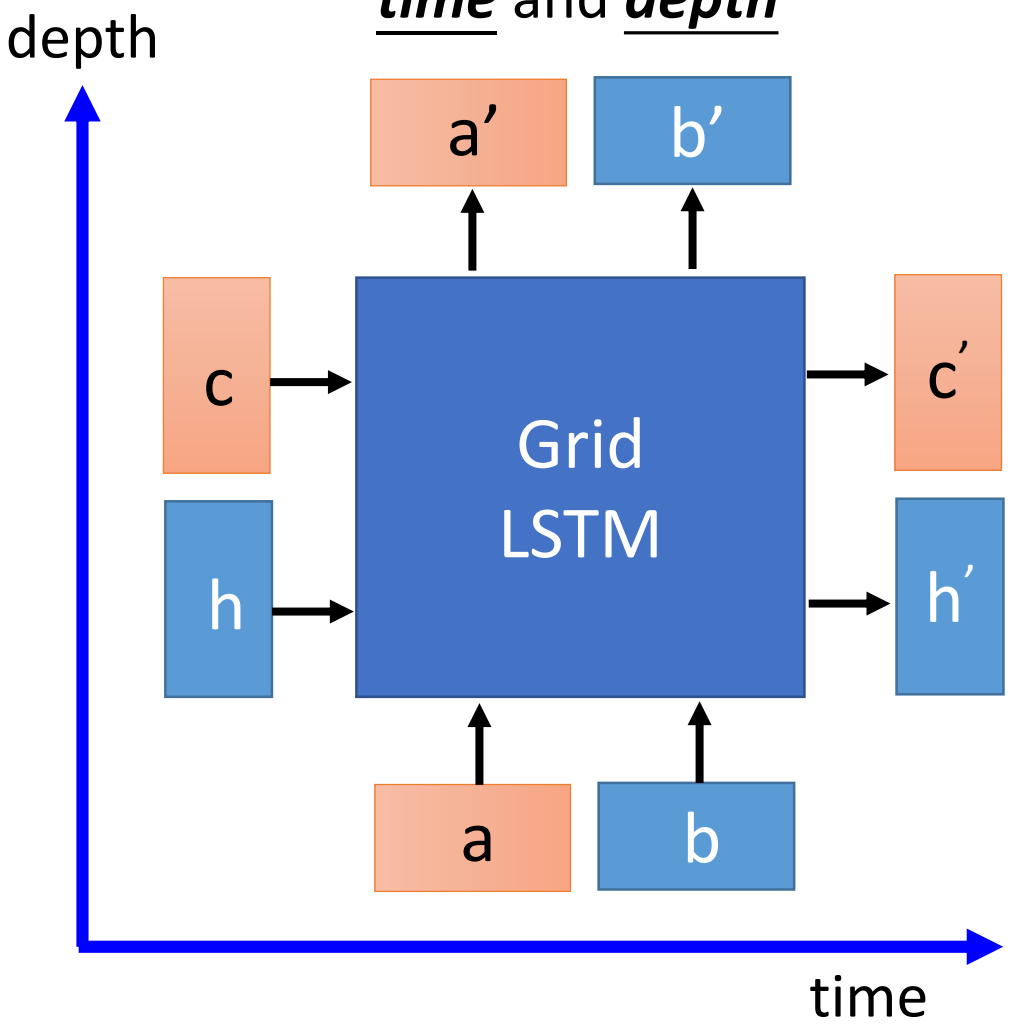


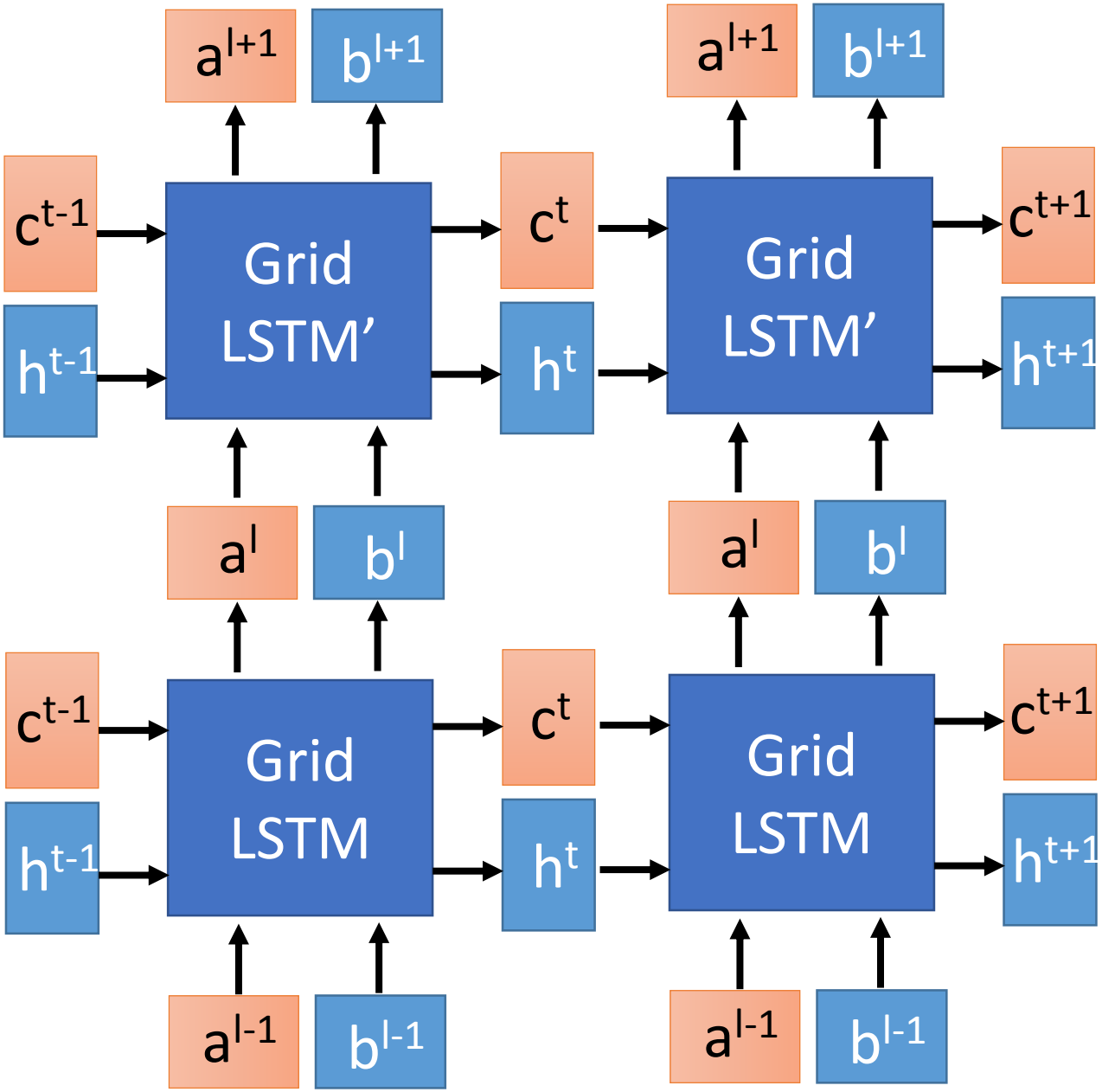
Grid LSTM

在時間跟layer兩個方向都設計gate

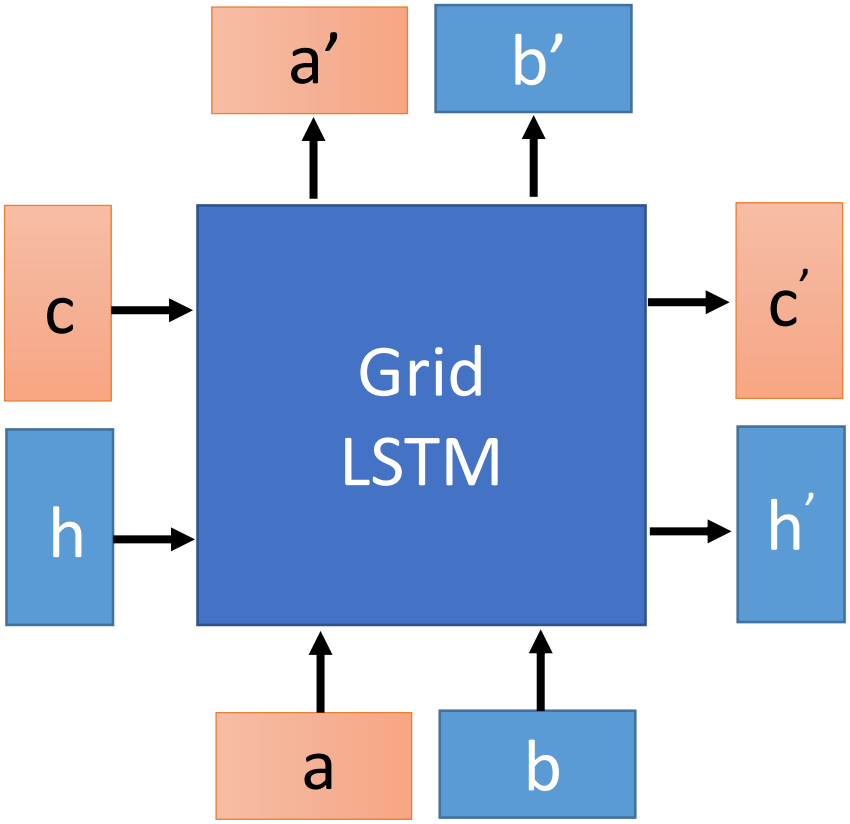


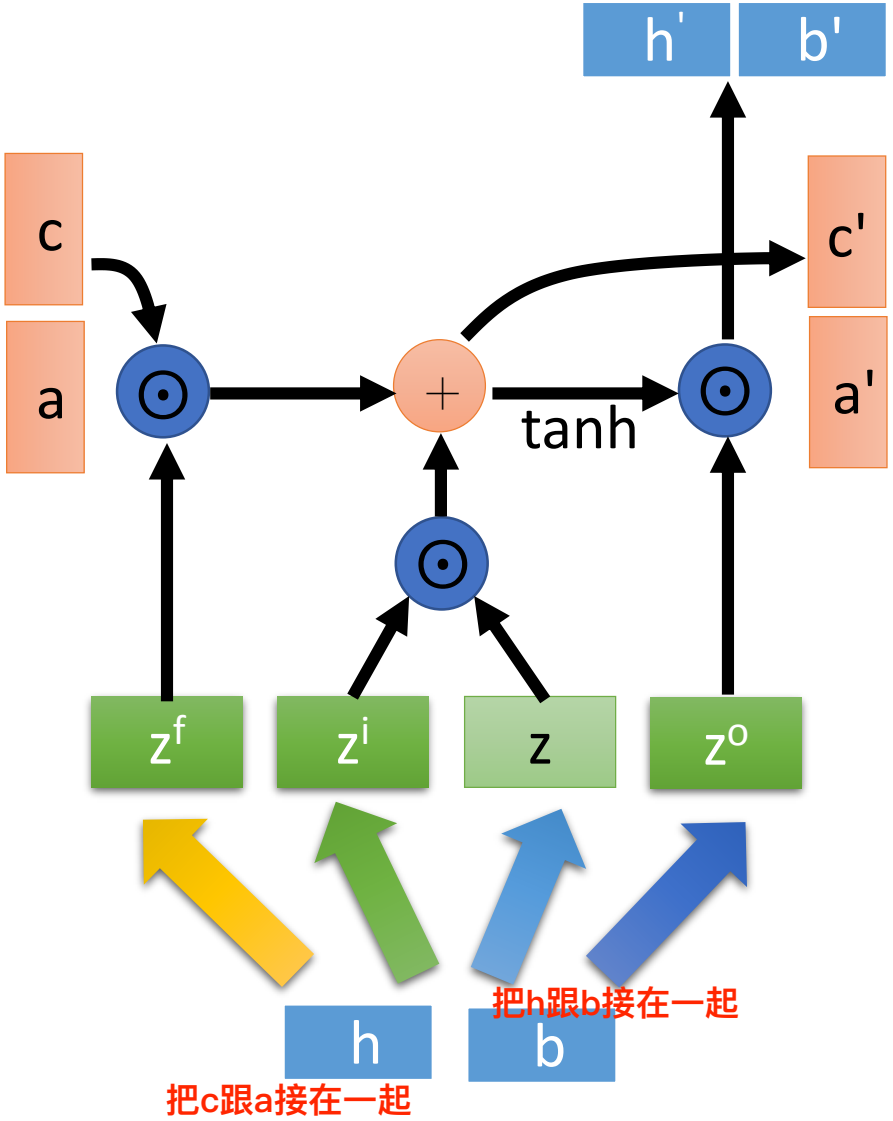
Memory for both ***time*** and ***depth***



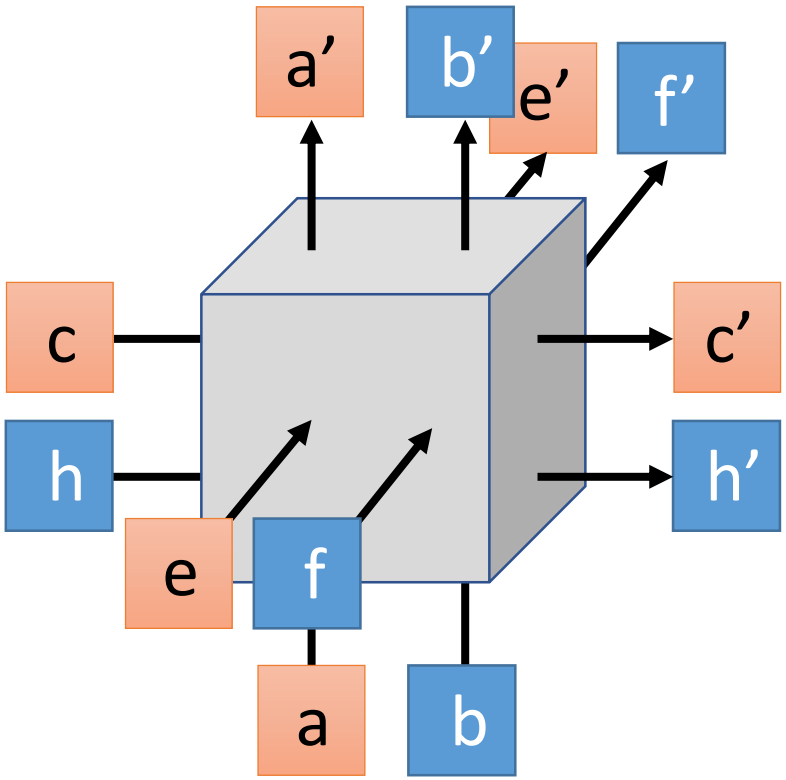

a^{l+1}
b^{l+1}
a^{l+1}
b^{l+1}
c^{t-1}
Grid LSTM'
c^t
Grid LSTM'
c^{t+1}
h^{t-1}
h^t
h^{t+1}
a^l
b^l
a^l
b^l
c^{t-1}
Grid LSTM
c^t
Grid LSTM
c^{t+1}
h^{t-1}
h^t
h^{t+1}
a^{l-1}
b^{l-1}
a^{l-1}
b^{l-1}

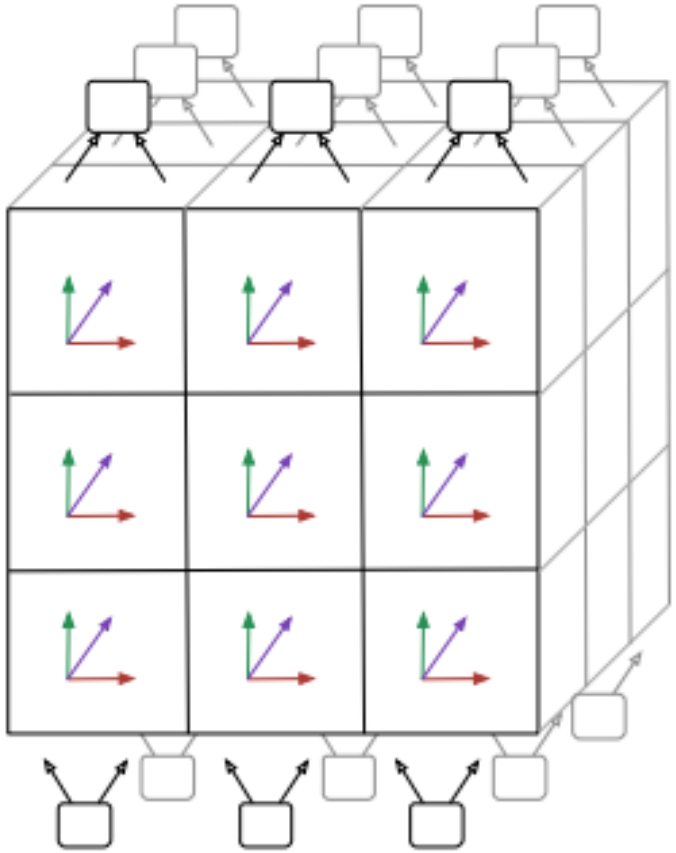
Grid LSTM


a'
b'
c
h
Grid
LSTM
c'
h'
a
b


h'
b'
c
a
c'
a'
tanh
z^f
z^i
z
z^o
h
b
把h跟b接在一起
把c跟a接在一起

3D Grid LSTM

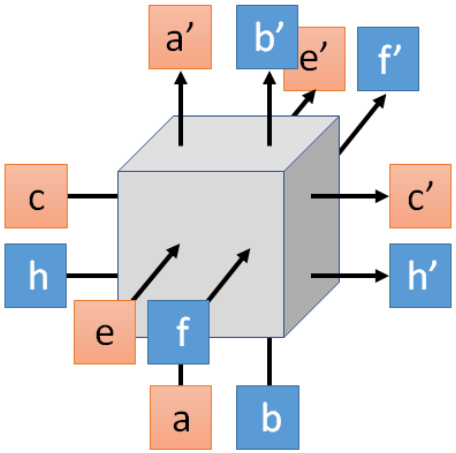


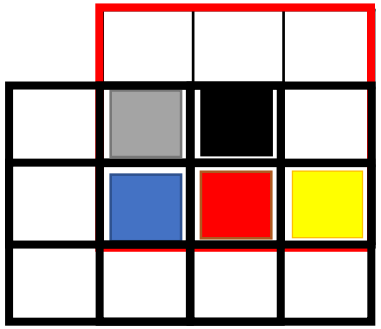


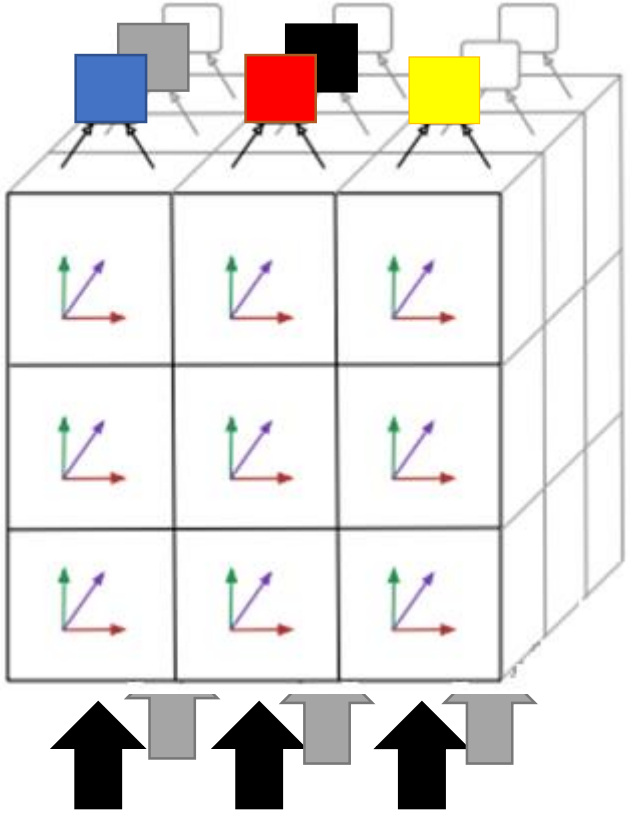
3D Grid LSTM

pixelRNN

- Images are composed of pixels







3 x 3 images