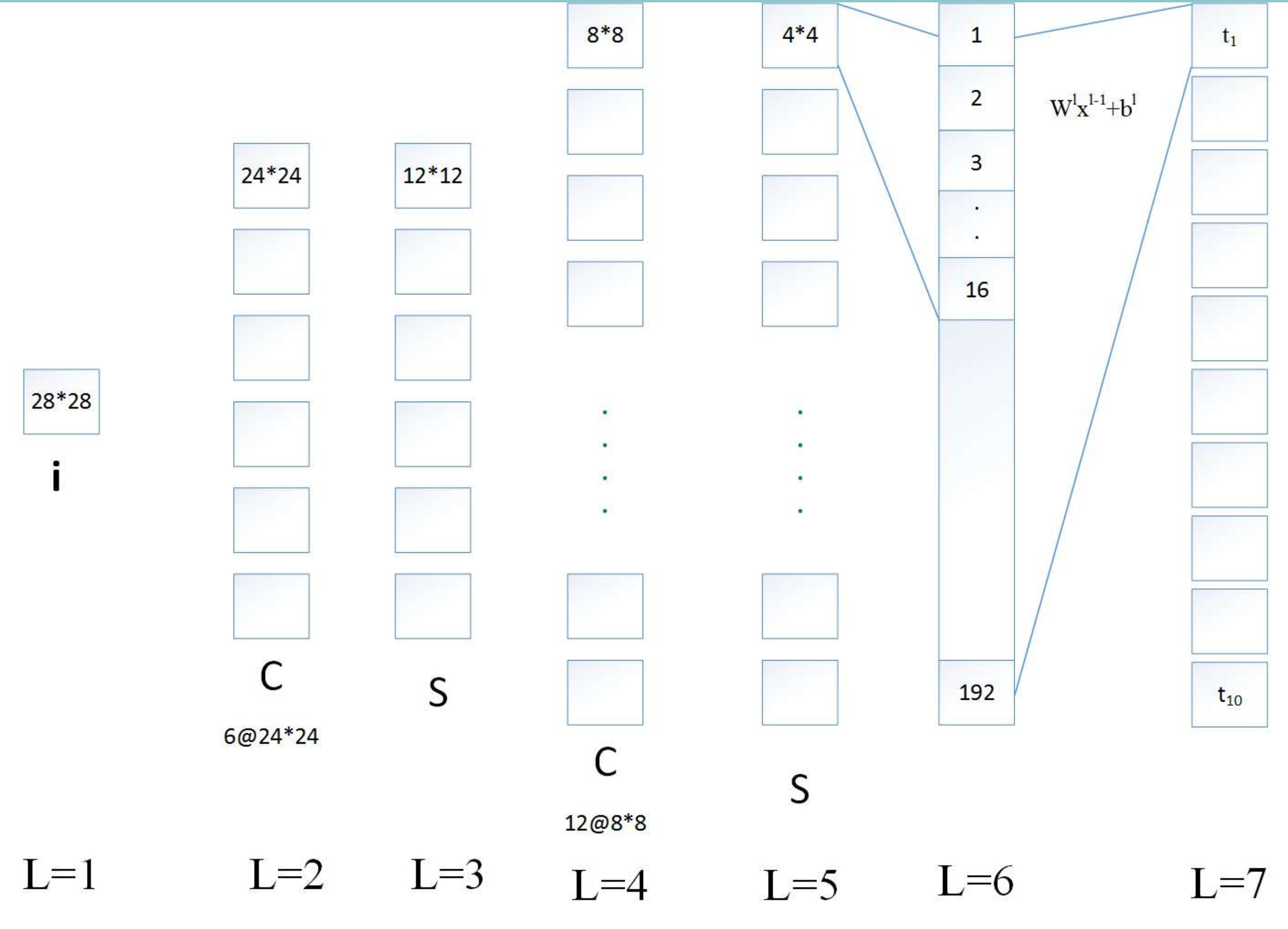
Lenet :


28*28
i
24*24
C
6@24*24
12*12
S
8*8
C
12@8*8
4*4
S
1
2
3
16
192
$W^l x^{l-1} + b^l$
t_1
t_{10}
L=1
L=2
L=3
L=4
L=5
L=6
L=7

Right:

t_k=[0.01 0.02 0.01 0.03 0.01 0.03 0.05 0.02 0.04 0.98] k=1:10

y_k=[0 0 0 0 0 0 0 0 0 1]

Wrong:

t=[0.01 0.02 0.78 0.03 0.01 0.03 0.05 0.02 0.04 0.02]

y=[0 0 0 0 0 0 0 0 0 1]

y_k=[1 0 0 0 0 0 0 0 0 0] （汽车）

y_k=[0 1 0 0 0 0 0 0 0 0] （飞机）

$$E^N = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{c}(t_k^n - y_k^n)^2$$

C:类别个数（10）
N:训练样本个数

用l表示第几层,不考虑样本数（每份样本都是一样的），cnn中每一层：

$$E = \frac{1}{2}\sum_{k=1}^{c}(t_k^l - y_k^l)^2$$

理想：误差E=0

$$t^l = f\left(u^l\right) = sigmod(u^l)$$

$$u^l = W^l x^{l-1} + b^l$$

（我们要求的是变量W和b）

反向传播（**Backpropagation Pass**， BP）

$$W = W - \alpha \frac{\partial}{\partial W} J(W, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(W, b)$$

$$J(W, b) = E = \frac{1}{2} \sum_{k=1}^{c} (t_k^l - y_k^l)^2$$

$$f(u_1^{l-1})$$

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial u} \frac{\partial u}{\partial b} = \frac{\partial E}{\partial u} = \delta \qquad (\frac{\partial u}{\partial b} = 1)$$

δ称为这一层偏置b的灵敏度或残差

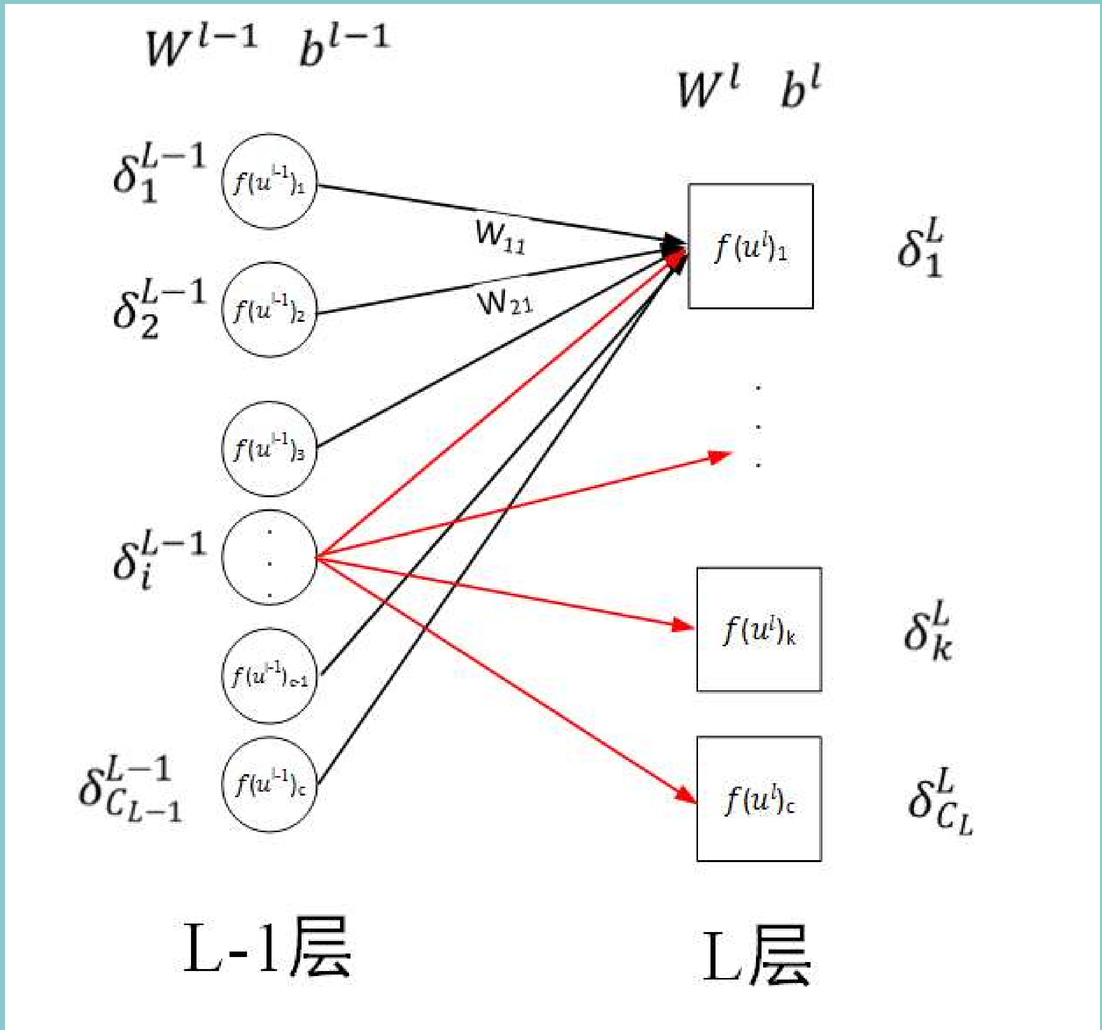
求最后一层的偏置:　$t^l = f(u^l) = sigmod(u^l)$

$$\delta_k^l = \frac{\partial E}{\partial u_k^l} = \frac{\partial}{\partial u_k^l}\{\frac{1}{2}[f(u_1^l) - y_1^l]^2 + \frac{1}{2}[f(u_2^l) - y_2^l]^2 + \cdots + \frac{1}{2}[f(u_k^l) - y_k^l]^2 + \cdots + \frac{1}{2}[f(u_2^l) - y_2^l]^2\}$$

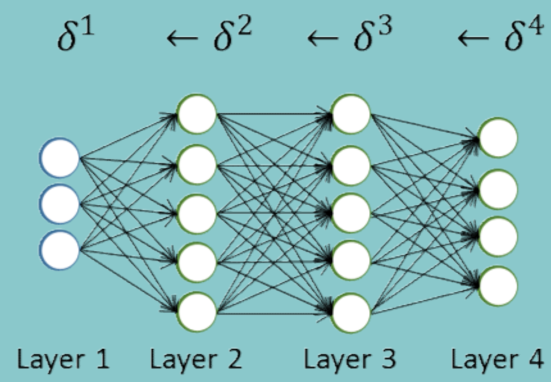
$$= (f(u_k^l) - y_k^l)(f(u_k^l) - y_k^l)'$$

$$= (f(u_k^l) - y_k^l)f'(u_k^l)$$

$$f'(u^l) = \left(\frac{1}{1+e^{-u^l}}\right)' = \frac{1}{1+e^{-u^l}}\left(1 - \frac{1}{1+e^{-u^l}}\right) = f(u^l)(1 - f(u^l))$$


W^{l-1} b^{l-1}
W^{l} b^{l}
δ_1^{L-1}
δ_2^{L-1}
δ_i^{L-1}
$\delta_{C_{L-1}}^{L-1}$
$f(u^{l-1})_1$
$f(u^{l-1})_2$
$f(u^{l-1})_3$
$f(u^{l-1})_{c-1}$
$f(u^{l-1})_c$
W_{11}
W_{21}
$f(u^{l})_1$
$f(u^{l})_k$
$f(u^{l})_c$
δ_1^{L}
δ_k^{L}
$\delta_{C_L}^{L}$
L-1层
L层

$$u^l = W^l f(u^{l-1}) + b^l$$


$\delta^1 \quad \leftarrow \delta^2 \quad \leftarrow \delta^3 \quad \leftarrow \delta^4$
Layer 1
Layer 2
Layer 3
Layer 4

L-1层残差：

$$\delta_i^{l-1} = \frac{\partial E}{\partial u_i^{l-1}} = \frac{\partial}{\partial u_i^{l-1}} \frac{1}{2} \sum_{k=1}^{c_l} (f(u_k^l) - y_k^l)^2$$

$$= \sum_{k=1}^{c_l} (f(u_k^l) - y_k^l) \frac{\partial}{\partial u_i^{l-1}} (f(u_k^l) - y_k^l)$$

$$= \sum_{k=1}^{c_l} (f(u_k^l) - y_k^l) \frac{\partial}{\partial u_i^{l-1}} f(u_k^l) \frac{\partial u_k^l}{\partial u_k^l}$$

$$= \sum_{k=1}^{c_l} (f(u_k^l) - y_k^l)\, f'(u_k^l) \frac{\partial u_k^l}{\partial u_i^{l-1}}$$

$$= \sum_{k=1}^{c_l} \delta_k^l \frac{\partial u_k^l}{\partial u_i^{l-1}}$$

$$u^l = W^l f(u^{l-1}) + b^l$$

$$= \sum_{k=1}^{c_l} \delta_k^l \frac{\partial}{\partial u_i^{l-1}} [\sum_{j=1}^{c_{l-1}} W_{jk}^l\, f(u_j^{l-1}) + b_k^l]$$

取j=i时才有值

$$= \sum_{k=1}^{c_l} \delta_k^l\, W_{ik}^l f'(u_i^{l-1})$$

$$\delta^{l-1} = (W^l)^T \delta^l .* f'(u^{l-1})$$

L层：

$$b = b - \alpha \frac{\partial}{\partial b} J(W, b)$$

$$= b - \alpha * \delta^{l}$$

$$W = W - \alpha \frac{\partial}{\partial W} J(W, b)$$

$$\frac{\partial E}{\partial W} = \frac{\partial E}{\partial u} \frac{\partial u}{\partial W} = \delta^{l} * f(u^{l-1})$$

$$\frac{\partial E}{\partial W_{ik}} = \frac{\partial E}{\partial u_k^l} \frac{\partial u_k^l}{\partial W_{ik}} = \delta_k^l * f(u_i^{l-1})$$

$$W = W - \alpha \frac{\partial}{\partial W} J(W, b) = W - \alpha * \delta_k^l * f(u_i^{l-1})$$

L=7

$$\delta_k^l = \left(f\left(u_k^l\right) - y_k^l\right) f'(u_k^l)$$

L=6

$$\delta^{l-1} = \left(W^l\right)^T \delta^l .* f'\left(u^{l-1}\right)$$

L=5 S

Reshape 12@4*4

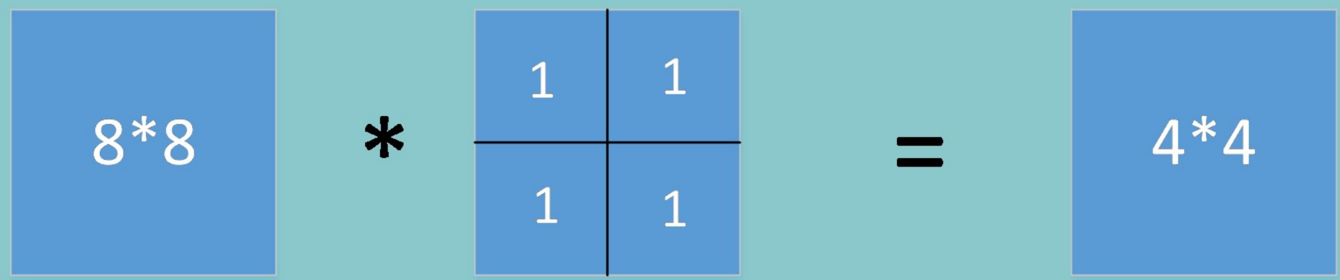
L=4 C

δ =

0.3517	0.9172	0.3804	0.5308
0.8308	0.2858	0.5678	0.7792
0.5853	0.7572	0.0759	0.9340
0.5497	0.7537	0.0540	0.1299

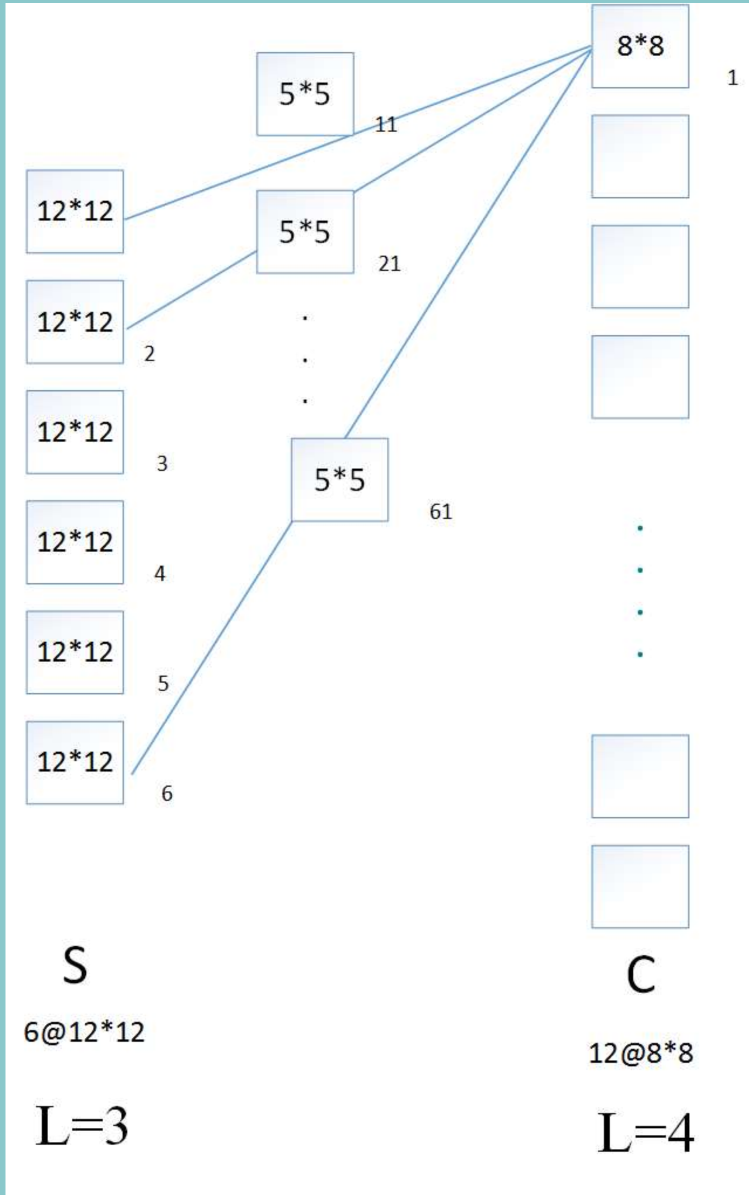
ans =

0.3517	0.3517	0.9172	0.9172	0.3804	0.3804	0.5308	0.5308
0.3517	0.3517	0.9172	0.9172	0.3804	0.3804	0.5308	0.5308
0.8308	0.8308	0.2858	0.2858	0.5678	0.5678	0.7792	0.7792
0.8308	0.8308	0.2858	0.2858	0.5678	0.5678	0.7792	0.7792
0.5853	0.5853	0.7572	0.7572	0.0759	0.0759	0.9340	0.9340
0.5853	0.5853	0.7572	0.7572	0.0759	0.0759	0.9340	0.9340
0.5497	0.5497	0.7537	0.7537	0.0540	0.0540	0.1299	0.1299
0.5497	0.5497	0.7537	0.7537	0.0540	0.0540	0.1299	0.1299

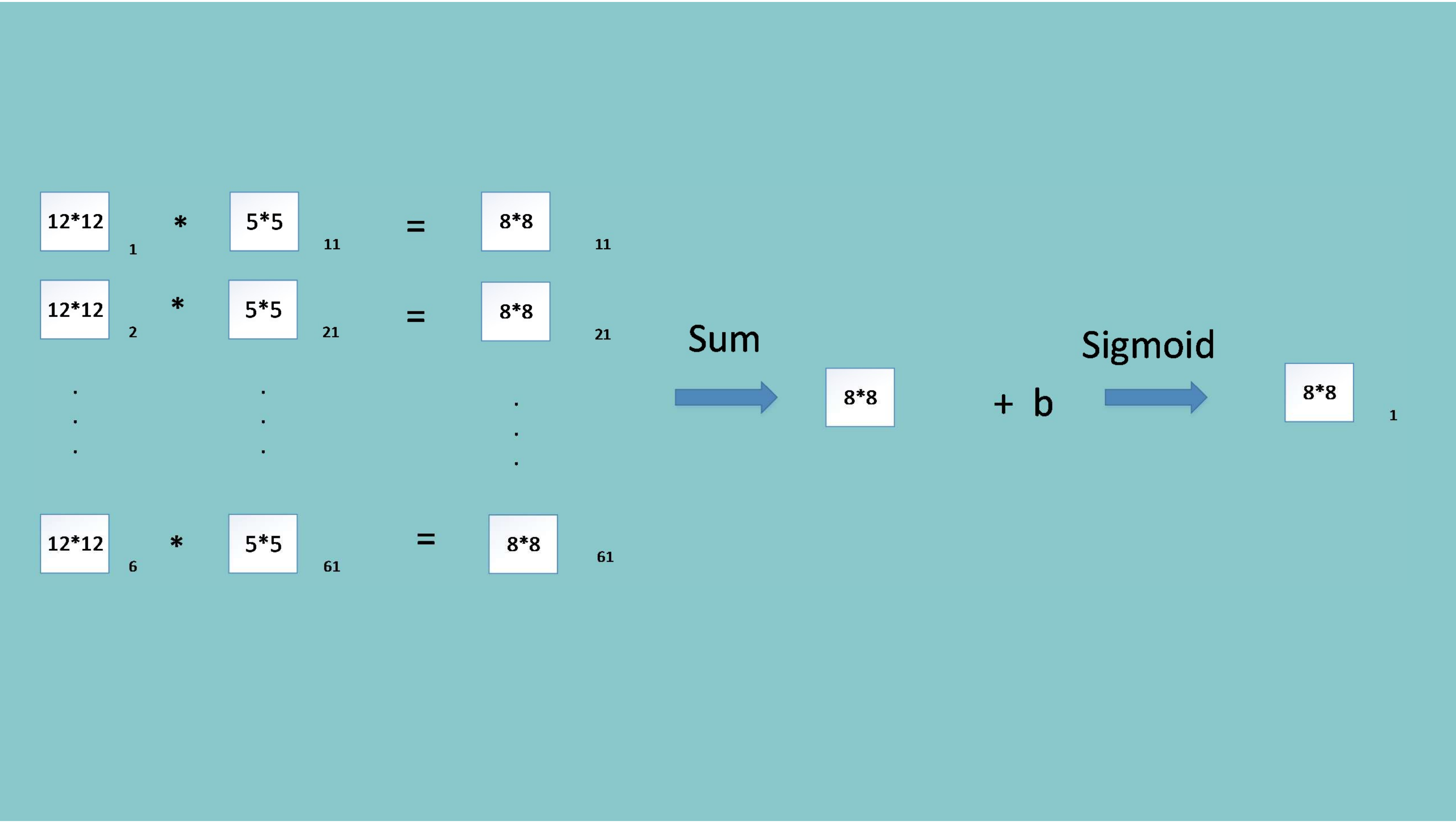


$$\delta^{l-1} = (W^{l})^{T}\delta^{l} .* f'(u^{l-1})$$

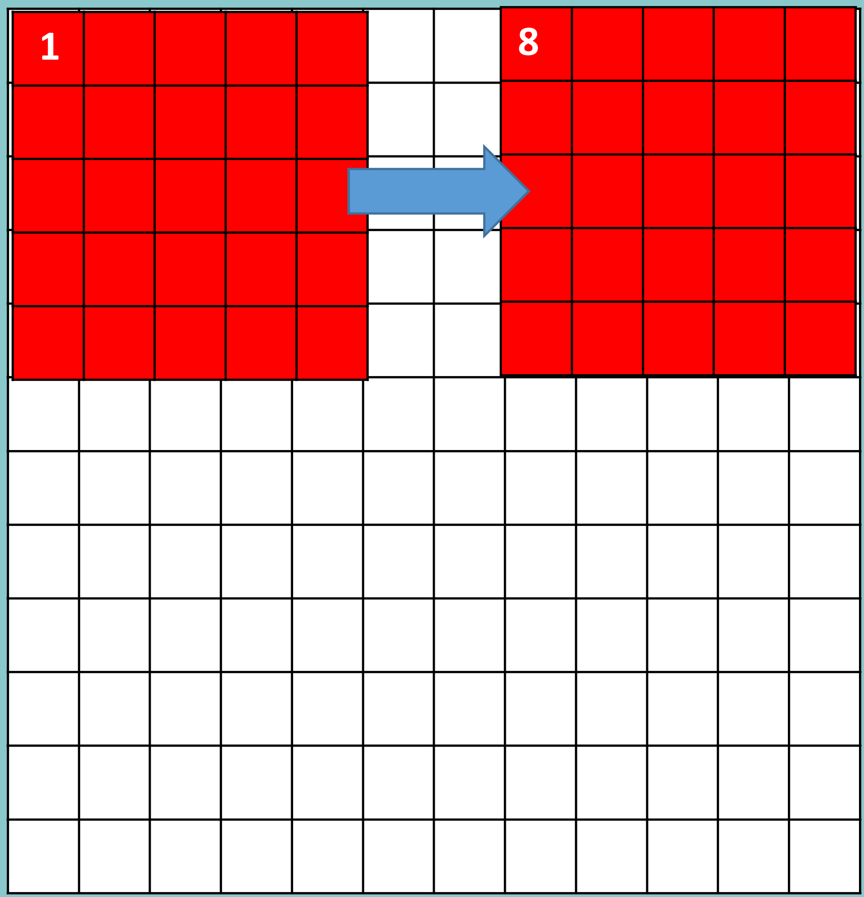
L=3 S



一共6*12个卷积核



12*12
1
*
5*5
11
=
8*8
11
12*12
2
*
5*5
21
=
8*8
21
12*12
6
*
5*5
61
=
8*8
61
Sum
8*8
+ b
Sigmoid
8*8
1

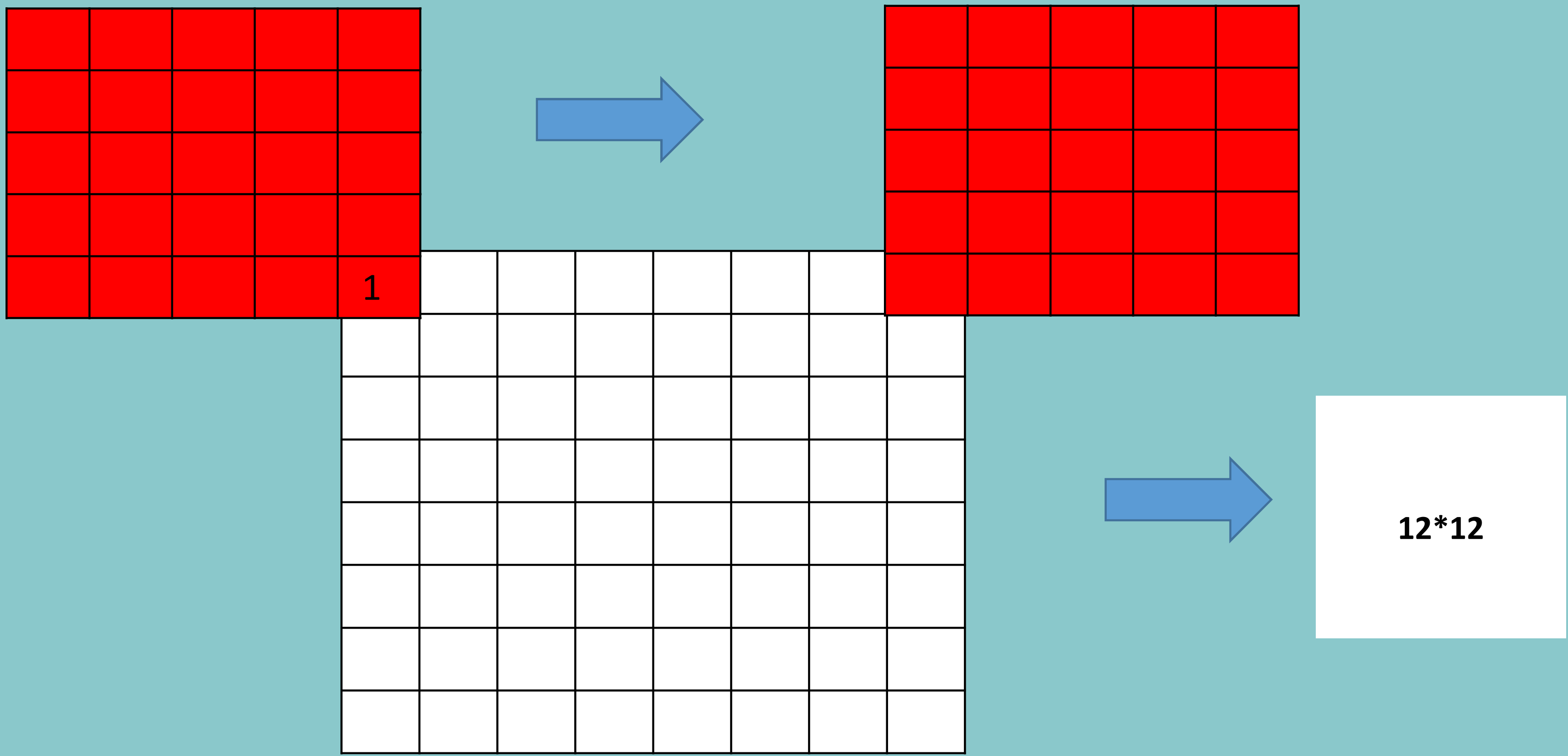


Image_size:L*L=[12, 12]

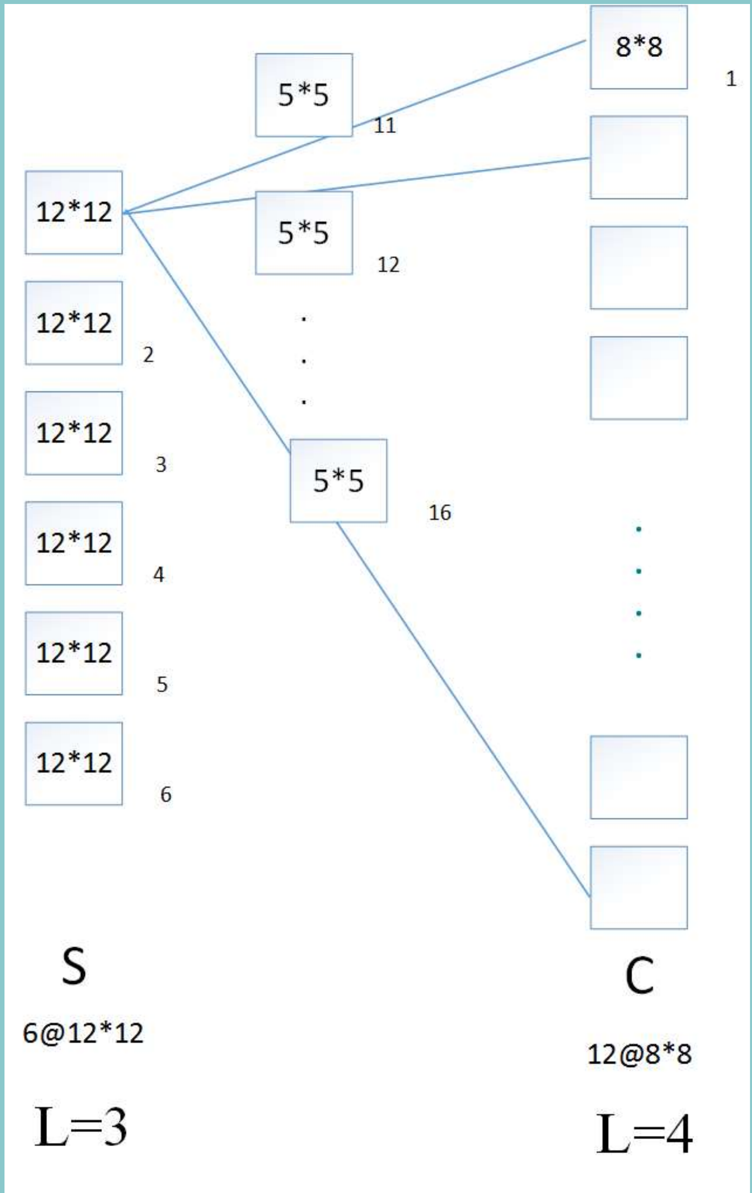
Weight_size:k*k=[5*5]

移动步长：stride:1

Output_image_size:
(L-k)/stride+1=(12-5)/1+1=8*8



12*12=convn(8*8,5*5,’full’)



12个12*12的残差相加得
到L3第一个12*12残差

L=2 C

类似L=4 扩展，公式。

L=1 I

无参数及残差

更新参数W、b

$$W = W - \alpha \frac{\partial}{\partial W} J(W, b)$$

$$b = b - \alpha \frac{\partial}{\partial b} J(W, b)$$

再输入一批训练图片得到输出t，继续训练