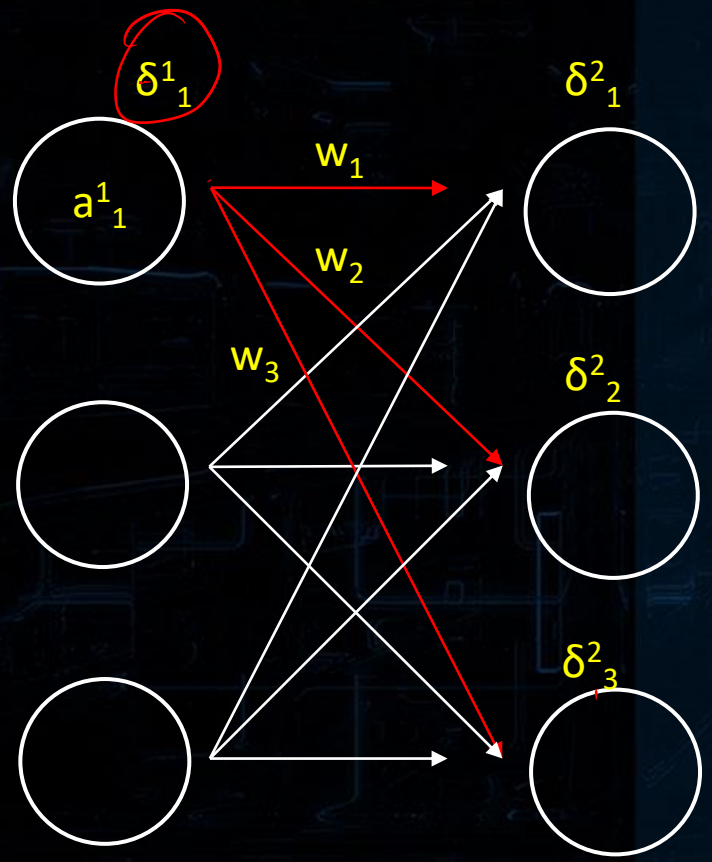
CNN神经网络

Steven Tang

卷积神经网络反向传播推导



$$\delta^1_1=(w_1*\delta^2_1+w_2*\delta^2_2+w_3*\delta^2_3)*f'(a^1_1)$$

卷积层前向过程回顾

a_{11}	a_{12}	a_{13}
a_{21}	a_{22}	a_{23}
a_{31}	a_{32}	a_{33}

$\bullet$

w_{11}	w_{12}
w_{21}	w_{22}

$\Rightarrow$

δ_{11} o_{11}	δ_{12} o_{12}
δ_{21} o_{21}	δ_{22} o_{22}

$$o_{11}=a_{11}*w_{11}+a_{12}*w_{12}+a_{21}*w_{21}+a_{22}*w_{22}$$

$$o_{12}=a_{12}*w_{11}+a_{13}*w_{12}+a_{22}*w_{21}+a_{23}*w_{22}$$

$$o_{21}=a_{21}*w_{11}+a_{22}*w_{12}+a_{31}*w_{21}+a_{32}*w_{22}$$

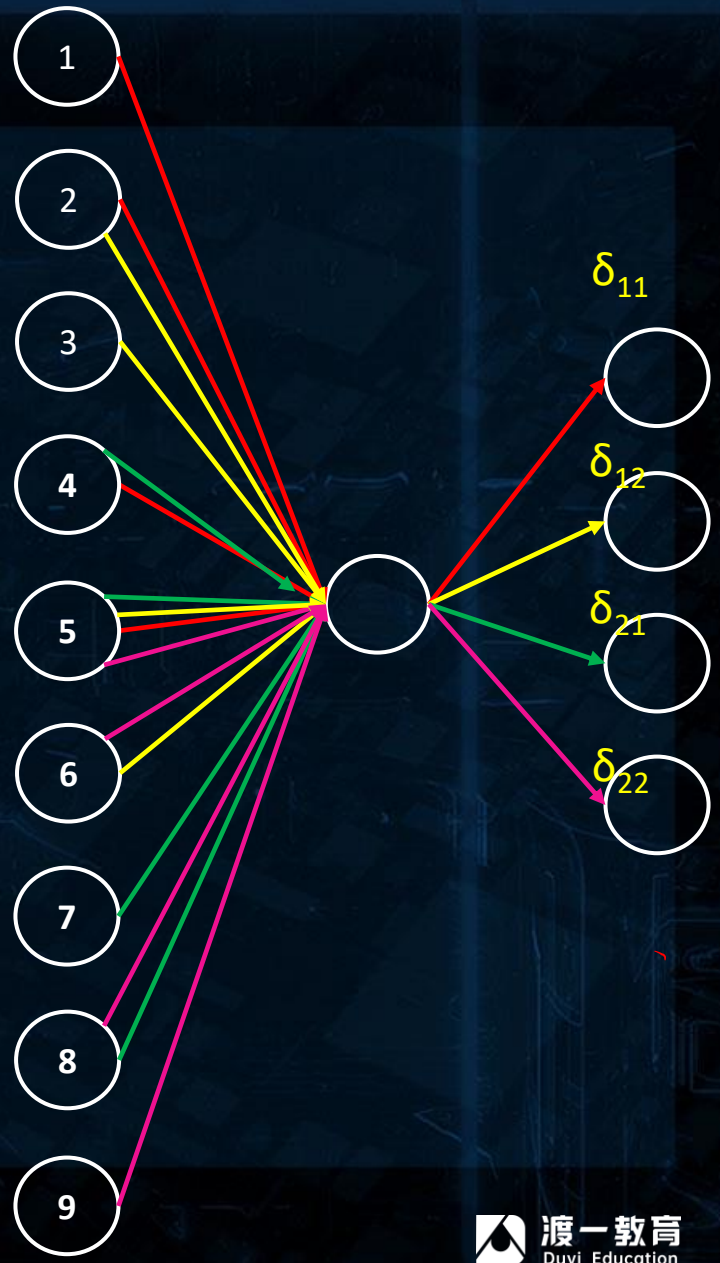
$$o_{22}=a_{22}*w_{11}+a_{23}*w_{12}+a_{32}*w_{21}+a_{33}*w_{22}$$

1	2	3
4	5	6
7	8	9

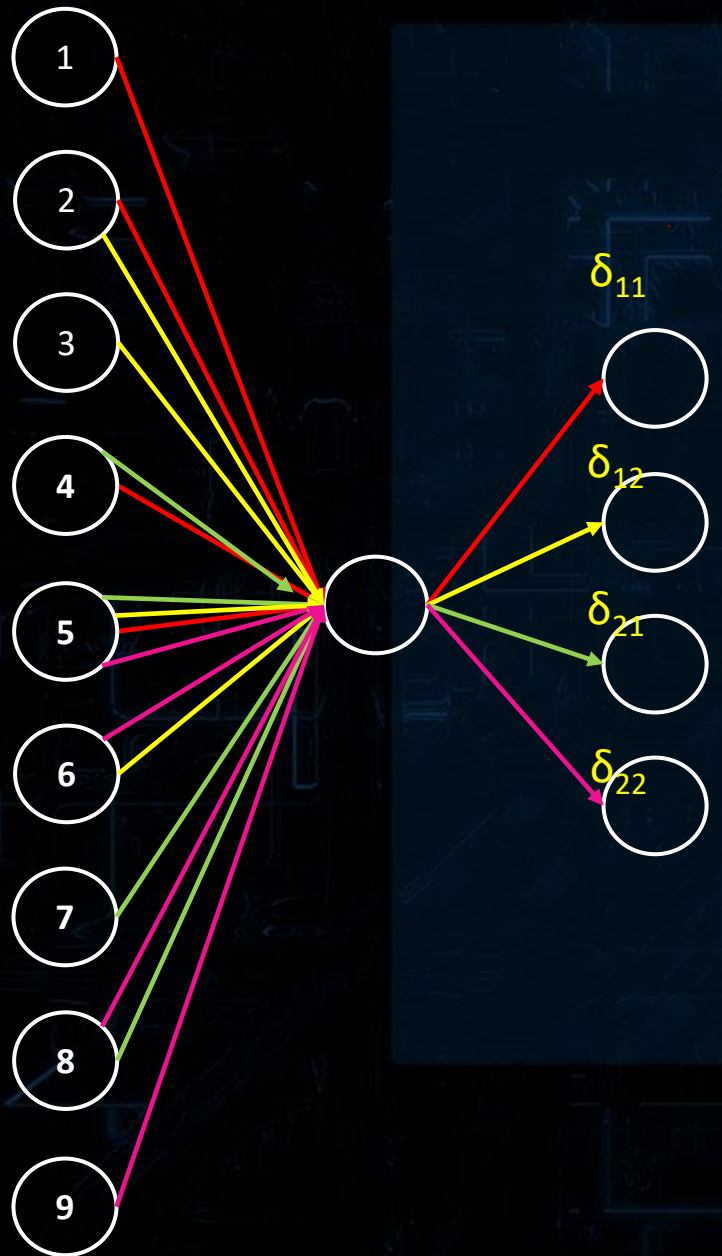
1	2	3
4	5	6
7	8	9

1	2	3
4	5	6
7	8	9

1	2	3
4	5	6
7	8	9



卷积层反向传播过程



$$o_{11}=a_{11}*w_{11}+a_{12}*w_{12}+a_{21}*w_{21}+a_{22}*w_{22}$$

$$o_{12}=a_{12}*w_{11}+a_{13}*w_{12}+a_{22}*w_{21}+a_{23}*w_{22}$$

$$o_{21}=a_{21}*w_{11}+a_{22}*w_{12}+a_{31}*w_{21}+a_{32}*w_{22}$$

$$o_{22}=a_{22}*w_{11}+a_{23}*w_{12}+a_{32}*w_{21}+a_{33}*w_{22}$$

$$\nabla a_{11}=\delta_{11}w_{11}$$

$$\nabla a_{12}=\delta_{11}w_{12}+\delta_{12}w_{11}$$

$$\nabla a_{13}=\delta_{12}w_{12}$$

$$\nabla a_{21}=\delta_{11}w_{21}+\delta_{21}w_{11}$$

$$\nabla a_{22}=\delta_{11}w_{22}+\delta_{12}w_{21}+\delta_{21}w_{12}+\delta_{22}w_{11}$$

$$\nabla a_{23}=\delta_{12}w_{22}+\delta_{22}w_{12}$$

$$\nabla a_{31}=\delta_{21}w_{21}$$

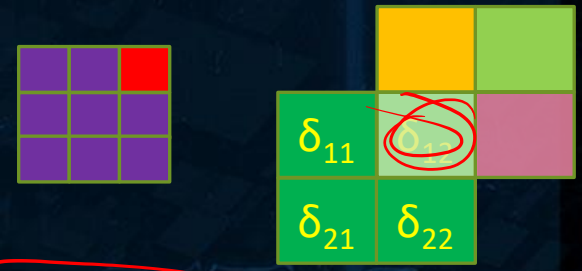
$$\nabla a_{32}=\delta_{21}w_{22}+\delta_{22}w_{21}$$

$$\nabla a_{33}=\delta_{22}w_{22}$$

卷积层反向传播过程

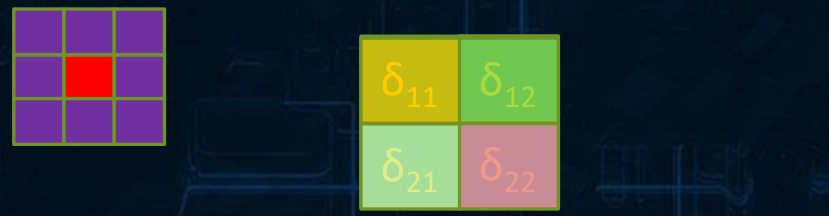
$$\nabla a_{11} = \delta_{11} w_{11}$$

$$\nabla a_{12} = \delta_{11} w_{12} + \delta_{12} w_{11}$$

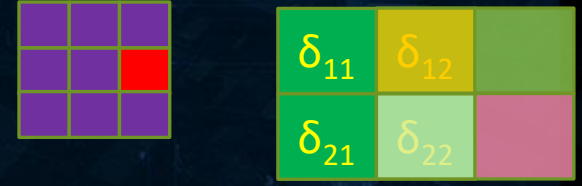


$$\nabla a_{13} = \delta_{12} w_{12}$$

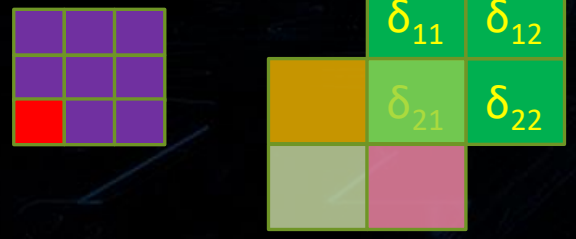
$$\nabla a_{21} = \delta_{11} w_{21} + \delta_{21} w_{11}$$



$$\nabla a_{22} = \delta_{11} w_{22} + \delta_{12} w_{21} + \delta_{21} w_{12} + \delta_{22} w_{11}$$



$$\nabla a_{23} = \delta_{12} w_{22} + \delta_{22} w_{12}$$



$$\nabla a_{31} = \delta_{21} w_{21}$$

$$\nabla a_{32} = \delta_{21} w_{22} + \delta_{22} w_{21}$$

$$\nabla a_{33} = \delta_{22} w_{22}$$

卷积层反向传播过程

0	0	0	0
0	δ_{11}	δ_{12}	0
0	δ_{21}	δ_{22}	0
0	0	0	0

$\bullet$

w_{22}	w_{21}
w_{12}	w_{11}

∇a_{11}	∇a_{12}	∇a_{13}
∇a_{21}	∇a_{22}	∇a_{23}
∇a_{31}	∇a_{32}	∇a_{33}

$$\delta_i^l = pad(\delta_i^{l+1}) * rot180(W^{l+1}) \bullet f'(a_i^l)$$

池化层前向过程回顾

1	1	2	4
2	4	7	8
3	2	1	0
1	1	2	4

maxpooling

4 δ_{11}	8 δ_{12}
3 δ_{21}	4 δ_{22}

反向传播

0	0	0	0
0	δ_{11}	0	δ_{12}
δ_{21}	0	0	0
0	0	0	δ_{22}

$$\delta_i^l = upsampling(\delta_i^{l+1}) \bullet f'(a_i^l)$$

$$f'(a_i^l) = \begin{cases} 1 & \text{if } a_i^l \text{ is max} \\ 0 & \text{else} \end{cases}$$

池化层前向过程回顾

$$\delta_i^l = upsampling(\delta_i^{l+1}) \bullet f'(a_i^l)$$

4 δ_{11}	8 δ_{12}
3 δ_{21}	4 δ_{22}

上采样

δ_{11}	δ_{11}	δ_{12}	δ_{12}
δ_{11}	δ_{11}	δ_{12}	δ_{12}
δ_{21}	δ_{21}	δ_{22}	δ_{22}
δ_{21}	δ_{21}	δ_{22}	δ_{22}

$\bullet$

0	0	0	0
0	1	0	1
1	0	0	0
0	0	0	1

反向传播

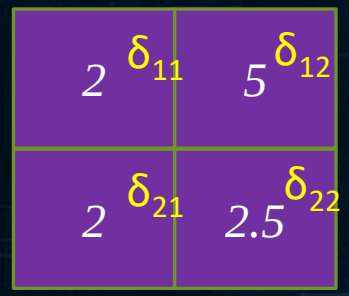
0	0	0	0
0	δ_{11}	0	δ_{12}
δ_{21}	0	0	0
0	0	0	δ_{22}

$$f'(a_i^l) = \begin{cases} 1 & \text{if } a_i^l \text{ is max} \\ 0 & \text{else} \end{cases}$$

池化层前向过程回顾

1	1	2	3
2	4	7	8
3	2	1	3
1	2	2	4

Average pooling



反向传播

$\delta_{11}/4$	$\delta_{11}/4$	$\delta_{12}/4$	$\delta_{12}/4$
$\delta_{11}/4$	$\delta_{11}/4$	$\delta_{12}/4$	$\delta_{12}/4$
$\delta_{21}/4$	$\delta_{21}/4$	$\delta_{22}/4$	$\delta_{22}/4$
$\delta_{21}/4$	$\delta_{21}/4$	$\delta_{22}/4$	$\delta_{22}/4$

$$\delta_i^l = upsampling(\delta_i^{l+1}) \bullet f'(a_i^l)$$

$$f'(a_i^l) = \frac{1}{n}$$