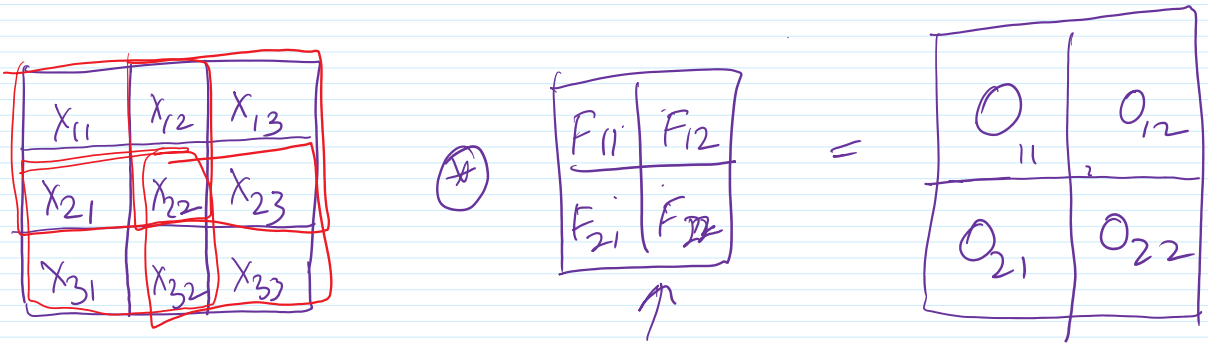


$$O_{11} = X_{11}F_{11} + X_{12}F_{12} + X_{21}F_{21} + F_{22}F_{22}$$

$O_{12} =$ ———————

$O_{21} =$ ———————

$O_{22} =$ ———————

$\frac{\partial O}{\partial X}, \frac{\partial O}{\partial F} \rightarrow \partial$

** $\frac{\partial L}{\partial F}$ is convolution between input X and Loss Gradient from the next layer $\frac{\partial L}{\partial O}$

** $\frac{\partial L}{\partial X}$ is convolution between 180° rotated F and Loss Gradient from the next layer $\frac{\partial L}{\partial O}$

Backpropagation in Pooling Layers :—

⇒

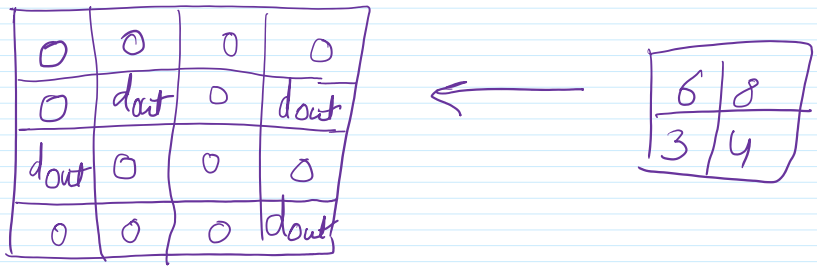
1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

max pooling (2×2) filter stride=2 →

6	8
3	4

⇒ the gradient will be propagated thrayen this layer.

local gradients



⇒ Average pooling :- Gradient is divided by the area of pooling block ($k \times k$)