(a) Padding = 1 is used, stride = 1 is used to get a 3x5 feature map.

∴ Padded image:

0	0	0	0	0	0	0
0	1	0	2	3	1	0
0	3	2	0	7	0	0
0	0	6	1	1	4	0
0	0	0	0	0	0	0

Sobel Kernel

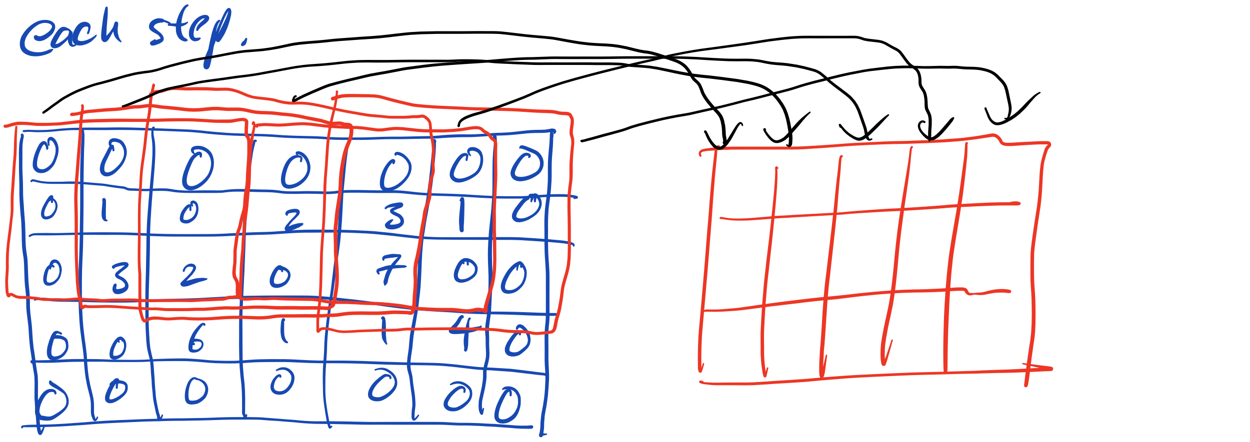
-1	0	1
-2	0	2
-1	0	1

To apply the kernel, we first need to flip up and flip left.

∴ Resultant kernel:

1	0	-1
2	0	-2
1	0	-1

The kernel is applied as a sliding window, moving 1 square each step.



The kernel and portion of the image undergoes an element-wise multiplication and is summed to get the value of the top left cell. As the window

Shifts right, the resultant value of the convolution is the value of the next cell on the right. This also occurs as the window shifts down, resulting in the value to be the one belonging to the cell below.

E.g.

0	0	0	0	0	0	0
0	1	0	2	3	1	0
0	3	2	0	7	0	0
0	0	6	1	1	4	0
0	0	0	0	0	0	0

1	0	-1
2	0	-2
1	0	-1

-2				

$0\times1 + 0\times0 + 0\times-1 + 0\times2 + 1\times0 + 0\times-2 + 0\times1 + 3\times0 + 2\times-1 = -2$

0 0 0 0 0 0 0 0 -2

0	0	0	0	0	0	0
0	1	0	2	3	1	0
0	3	2	0	7	0	0
0	0	6	1	1	4	0
0	0	0	0	0	0	0

1	0	-1
2	0	-2
1	0	-1

-2	1			

$0\times1 + 0\times0 + 0\times-1 + 1\times2 + 0\times0 + 2\times-2 + 3\times1 + 2\times0 + 0\times-1 = 1$

0 0 0 2 0 -4 3 0 0

0	0	0	0	0	0	0
0	1	0	2	3	1	0
0	3	2	0	7	0	0
0	0	6	1	1	4	0
0	0	0	0	0	0	0

1	0	-1
2	0	-2
1	0	-1

-2	1			
-10				

$0\times1 + 1\times0 + 0\times-1 + 0\times2 + 3\times0 + 2\times-2 + 0\times1 + 0\times0 + 6\times-1 = -10$

0 0 0 0 0 -4 0 0 -6

0	0	0	0	0	0	0
0	1	0	2	3	1	0
0	3	2	0	7	0	0
0	0	6	1	1	4	0
0	0	0	0	0	0	0

1	0	-1
2	0	-2
1	0	-1

-2	1	-11	2	13
-10	4	-8	-2	18
-14	1	5	-6	9

After applying the kernel across the padded this is the resultant feature map.

-2	1	-11	2	13
-10	4	-8	-2	18
-14	1	5	-6	9

(b) Max Pooling. It chooses the highest activation in it's kernel, reducing sensitivity to small pixel shifts.

(c) With the formula,

$$W_2 = (W_1 - F_w + 2P_w)/S_w + 1$$

and $W_2 = W_1 = 5$, $S_w = 1$, $F_w = 5$

$$W_2 - 1 = 2P_w$$
$$4 = 2P_w$$

$P_w = 2$

Similarly for height,

$P_h = 2$

$\therefore$ Padding of 2.

1d) Width:

$504 = (512 - F_w + 0)/1 + 1$

$503 = 512 - F_w$

$F_w = 9$

Height:

$504 = (512 - F_H + 0)/1 + 1$

$503 = 512 - F_H$

$F_H = 9$

Kernel dimension: $(9) \times (9)$

1e) $W_2 = (504 - 2 + 0)/2 + 1$

$= 252$

$H_2 = (504 - 2 + 0)/2 + 1$

$= 252$

Spatial dimension: $(252) \times (252)$

1f) $W_2 = (252 - 3 + 0)/1 + 1$

$= 250$

$H_2 = (252 - 3 + 0)/1 + 1$

$= 250$

Spatial Dimension: $(250) \times (250)$

1g) Filter of 5×5, padding $= 2$ and stride $= 1$

↳ Same padding, dimensions stay same.

MaxPool2D and Flatten has no trainable parameters. MaxPool2D halves the dimension given a kernel size of 2.

Layer 1: $(5\times5+1)\times 32 = 832$ parameters.

$\text{Dimensions}_{out} = [16, 16, 32]$

Layer 2: $(5\times5+1)\times 64 = 1664$ parameters.

$\text{Dimensions}_{out} = [8, 8, 64]$

Layer 3: $(5\times5+1)\times 128 = 3328$ parameters.

$\text{Dimension}_{out} = [4, 4, 128]$

After flattening,

Dimensions = $[4 \times 4 \times 128,] = [2048,]$

Layer 4: $(2048+1) \times 64 = 131,136$ params

Layer 5: $(64+1) \times 10 = 650$ parameters.

total number of parameters $= 832 + 1664 + 5328$

$+ 131136 + 650$

$= 137,610$ parameters