1. Compute one convolution operation followed by max pooling operation (2×2) for the below data with stride = 1.

SOLUTION:

[Convolution is a mathematical operation to merge two sets of information].

[Here, the input data is merged with kernal filter].

INPUT IMAGE:

$$\begin{bmatrix} 1 & 2 & 3 & 2 & 2 & 1 & 0 \\ 3 & 2 & 1 & 4 & 1 & 1 & 1 \\ 2 & 1 & 3 & 2 & 1 & 0 & 1 \\ 5 & 3 & 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 & 6 & 0 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Kernal filter:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

Stride value = 1.

Input size = 7 x 7

Kernal size : 3 x 3

stride : 1

Output size:

$$\text{Output height} = \frac{\text{input height} - \text{Kernal height}}{\text{Stride}} + 1$$

$$= \frac{7-3}{1} + 1 = 5$$

$$\text{Output width} = \frac{\text{Input width} - \text{Kernal width}}{\text{Stride}} + 1$$

$$= \frac{7-3}{1} + 1 = 5$$

Total number of receptive fields

= Output height x output width

= 5 x 5

= 25.

First receptive field:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 4 & 9 \\ 9 & 4 & 1 \\ 4 & 1 & 9 \end{bmatrix}$$

$1 + 4 + 9 + 9 + 4 + 1 + 4 + 1 + 9$

$= 42.$

Second receptive field:

$$\begin{bmatrix} 2 & 3 & 2 \\ 2 & 1 & 4 \\ 1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 6 & 6 \\ 6 & 2 & 4 \\ 2 & 3 & 6 \end{bmatrix}$$

Sum = 37

Third receptive field:

$$\begin{bmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ 3 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 4 & 6 \\ 3 & 8 & 1 \\ 6 & 2 & 3 \end{bmatrix}$$

Sum = 36.

~~OUTPUT~~ By further calculations, we can arrive at OUTPUT FEATURE MAP:

$$\begin{bmatrix} 42 & 37 & 35 & 29 & 13 \\ 37 & 36 & 29 & 13 & 8 \\ 34 & 30 & 27 & 20 & 9 \\ 29 & 23 & 18 & 15 & 8 \\ 21 & 16 & 13 & 10 & 6 \end{bmatrix}$$

MAX POOLING:

Max pooling with 2x2 window and stride 2 on 5x5 feature map.

- **First max-pooling operation:**

$$\begin{bmatrix} 42 & 37 \\ 37 & 36 \end{bmatrix}$$

Max. value = 42.

- **Second max pooling:**

$$\begin{bmatrix} 36 & 29 \\ 29 & 13 \end{bmatrix}$$

Max. value : 36

- **Third max pooling:**

$$\begin{bmatrix} 34 & 30 \\ 29 & 23 \end{bmatrix}$$

Max. value : 34

- **Fourth max pooling:**

$$\begin{bmatrix} 18 & 15 \\ 16 & 13 \end{bmatrix} \quad \begin{bmatrix} 27 & 20 \\ 18 & 15 \end{bmatrix}$$

Max. value : 27

Final pooled feature Map:

$$\begin{bmatrix} 42 & 36 \\ 34 & 27 \end{bmatrix}$$