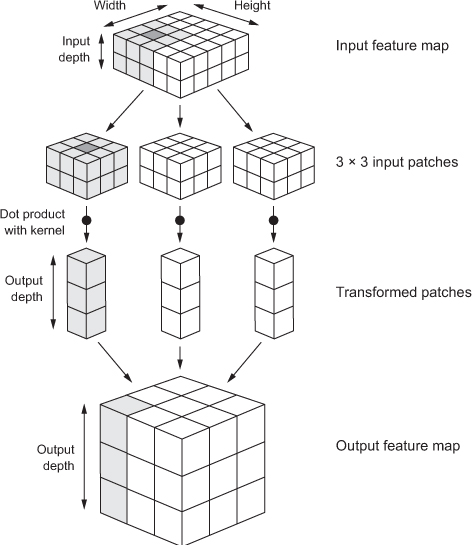
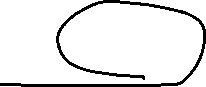
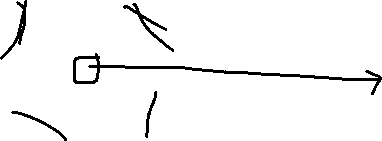
Building Blocks of CNN



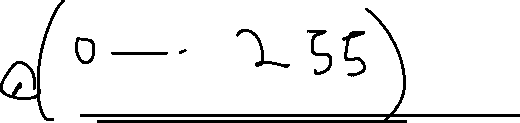
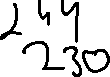
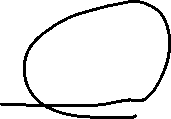




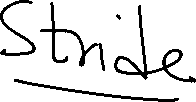


Layers in CNN :

* **The convolutional layer**
* **The pooling layer**
* **The fully connected layer**



**Convolutional Layer**

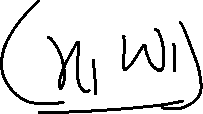


1. Filter or Kernel or filter layer



* 1. Body structure, face, legs, tail

Convolution operation that will extract all the important features from the image that characterize the horse



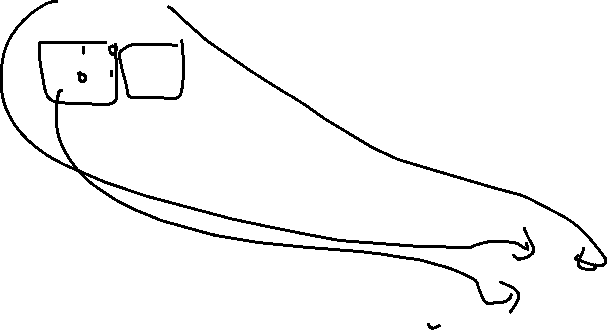
|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 3 | 2 | 1 |
| 2 | 9 | 1 | 1 |
| 1 | 3 | 2 | 3 |
| 5 | 6 | 1 | 2 |

|  |  |
| --- | --- |
| 1 | 0 |
| 0 | 1 |

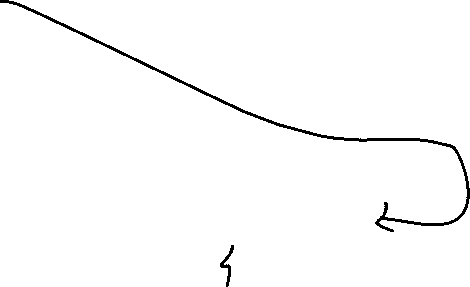


\*

Filter Matrix =F  
Input Matrix= I



Convolution Operation



|  |  |  |  |
| --- | --- | --- | --- |
| 1 1 | 30 | 2 | 1 |
| 20 | 91 | 1 | 1 |
| 1 | 3 | 2 | 3 |
| 5 | 6 | 1 | 2 |

|  |  |
| --- | --- |
| 10 | 3 |
| 7 | 4 |



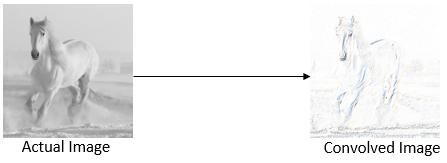
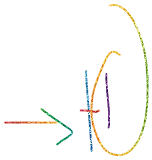
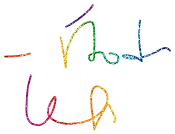
O1- 1X1+3X0+2X0+9X1=10

O2= 2X1+1X0+1X0+1X1=3

O3= 1X1+3X0+5X0+6X1=7

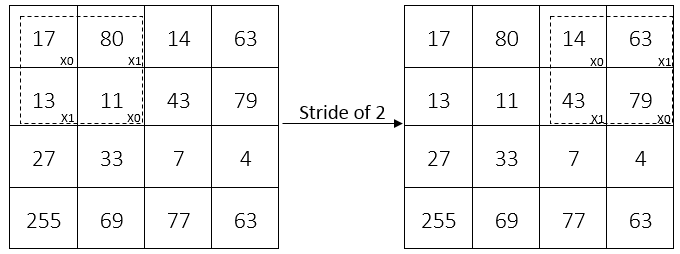
O4= 2X1+3X0+1X0+2X1= Output(Convoluted Image or feature map)

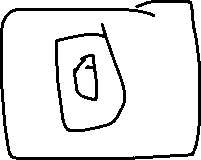
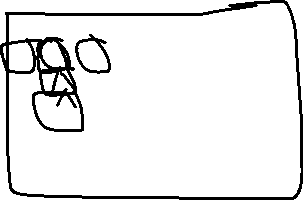




**Stride**

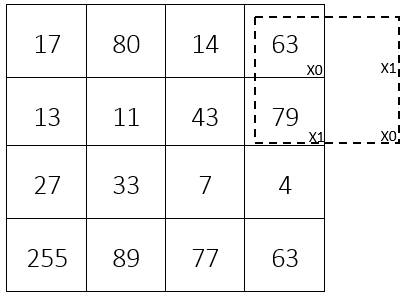
The number of pixels we slide over the input matrix by the filter matrix is called a **stride**





**Padding**

we move our filter matrix by two pixels, it reaches the border and the filter matrix does not fit the input matrix. That is, some part of our filter matrix is outside the input matrix





**Problem Being Solved by padding**- Shrinking Out put and Throwing away lot of information from the edges

**Same padding** or **zero padding**: pad the input matrix with zeros so that the filter can fit the input matrix

Padding with 0’s

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | X | X | X | X | X |
| X |  |  |  |  | X |
| X |  |  |  |  | X |
| X |  |  |  |  | X |
| X |  |  |  |  | X |
| X | X | X | X | X | X |



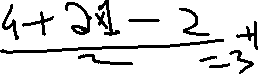
|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |



|  |  |
| --- | --- |
| 1 | 0 |
| 0 | 1 |

Filter Output

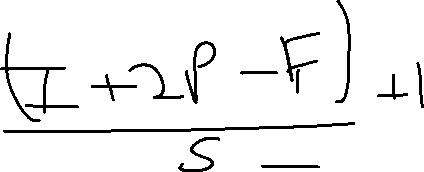
6 X 6 -Input after paddling



Output matrix= (I+2P-F/S)+1= (4+2-2)/2 +1= 3

I = Input matrix

P= Padding



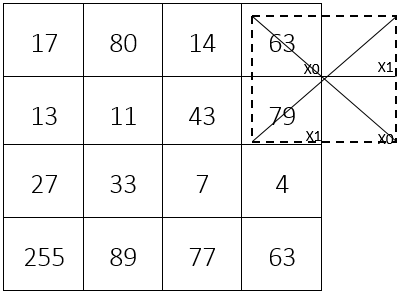
F= Filter

S= Stride

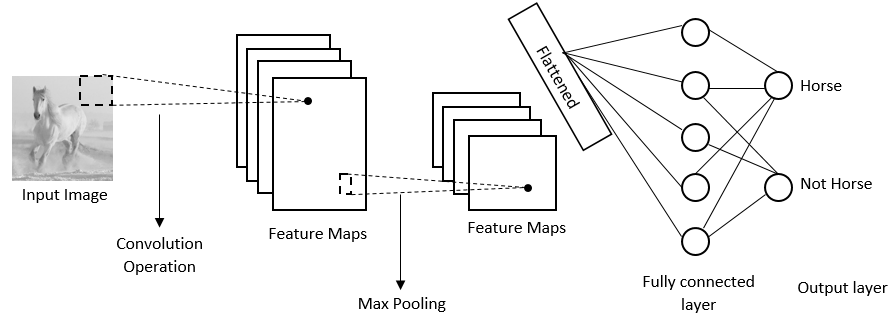
O= (( I + 2P -F)/S) +1



Valid padding- padding them with zeros, we can also simply discard the region of the input matrix where the filter doesn't fit in









Pooling Layer

For example, to recognize a horse from the image, we need to extract and keep only the features of the horse; we can simply discard unwanted features, such as the background of the image and more.

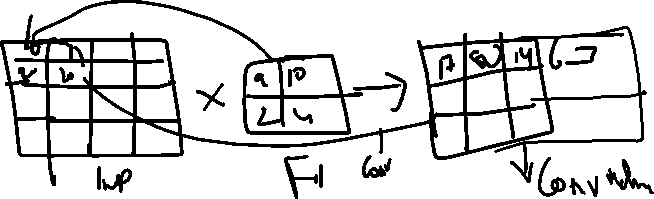


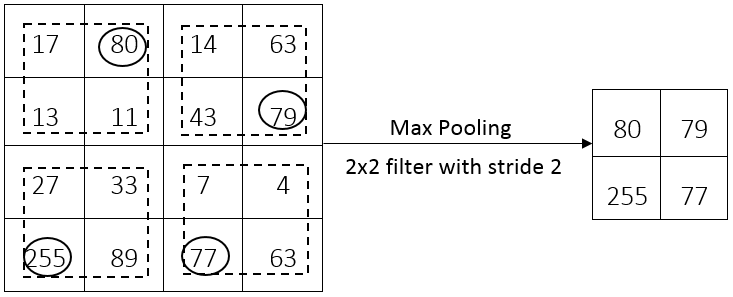
pooling operation is also called a **downsampling** or **subsampling** operation

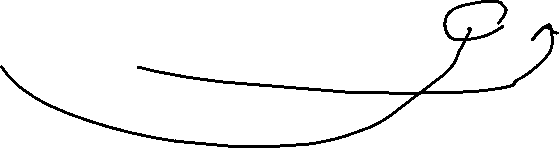
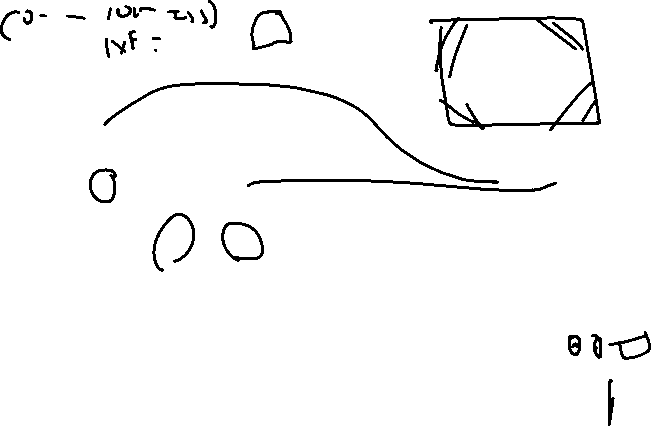
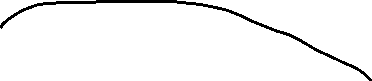
*The pooling operation will not change the depth of the feature maps; it will only affect the height and width.*



*Max pooling-*max pooling, we slide over the filter on the input matrix and simply take the maximum value from the filter window







*Max pooling is one of the most commonly used pooling operations.*

***Average pooling***, we take the average value of the input matrix within the filter window,

***sum pooling,*** we sum all the values of the input matrix within the filter window

