

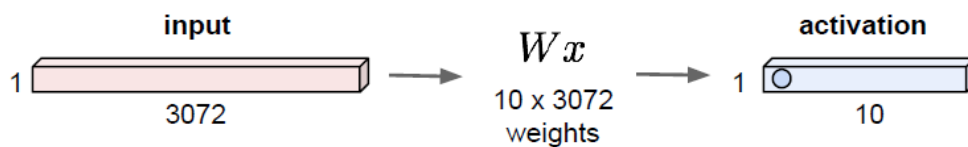
● Convolution Layer

Neural Network 는 선형 함수와 비선형 함수의 합으로 계속 이어져 있다.

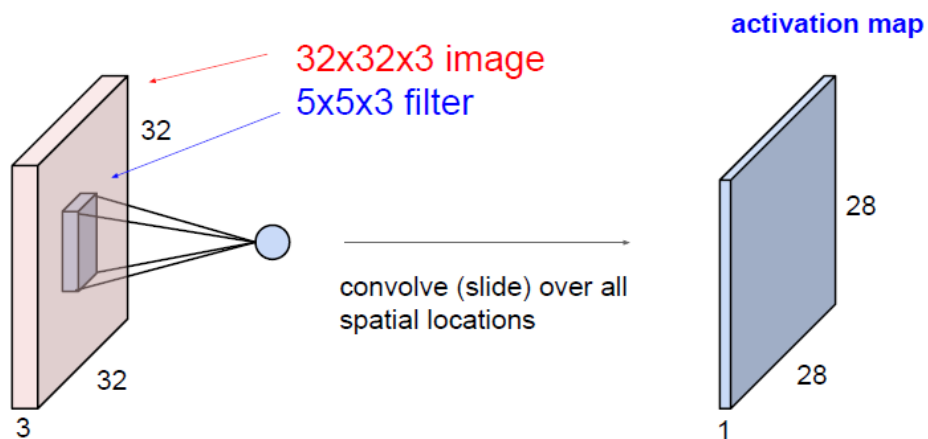
Convolution Neural Network 의 큰 특징은 **공간적 구조를 유지**하는 layer 를 가지고 있다는 점.

여기서 Convolution 이라는 개념은 좀 더 rough 하게 생각하면, 우리가 신호처리에서 배운 개념과는 다르게 적용된다.

Fully connected layer 의 같은 경우, (32,32,3)의 이미지가 입력으로 들어오면, (3072,1)의 형태로 만들어준 다음, 계산을 수행한다.

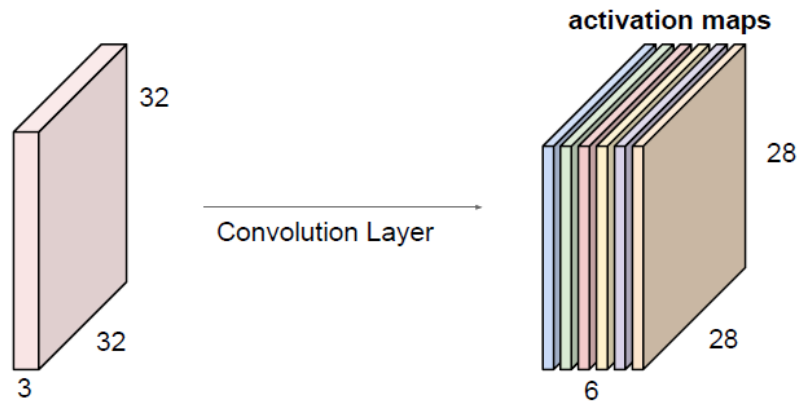


Convolution layer 의 경우 Filter 를 하나 만들고, 그 Filter 가 움직이면서 Output 을 만들어낸다. 필터를 w 라고 했을 때, Output 의 한 점은 $wTx + bwTx + b$ 로 내적을 시킨 값을 넣어준다.



하나의 Filter 는 항상 입력의 채널의 수만큼 확장이 되어야 한다!

For example, if we had 6 5x5 filters, we'll get 6 separate activation maps:



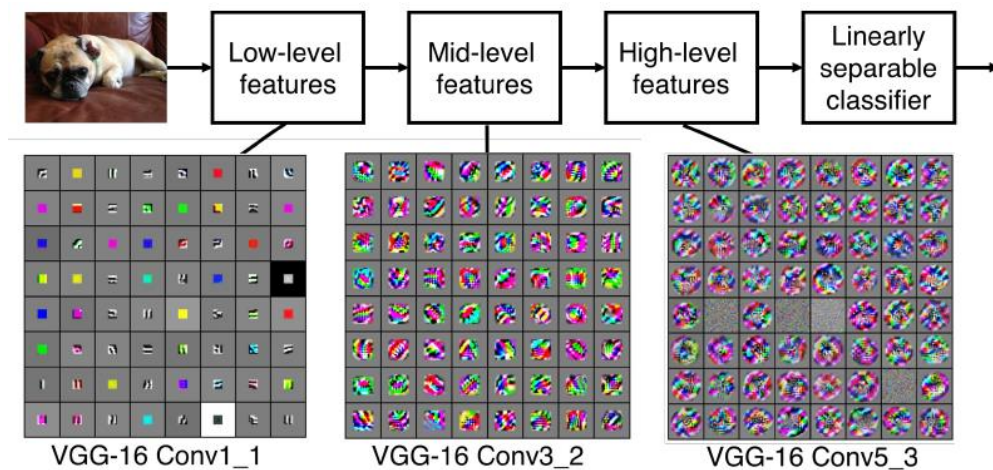
또한, Filter 의 수로 Output 의 activation map 의 수를 결정할 수 있다

다른 Filter 들로 같은 위치에서 다른 특징들을 뽑아 낼 수 있다

Convolution layer 를 여러 번 수행하면 다음과 같이 나타낼 수 있다.

입력을 이미지로 넣었을 때 각 layer 마다 어떤 feature 들이 뽑아질 수 있는지 다음의 예시를 참고하면 직관적으로 이해할 수 있다.

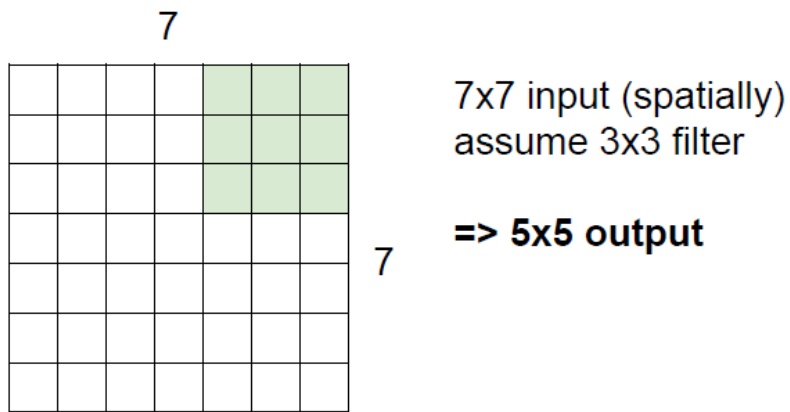
low-level 특징점부터 high-level 특징점까지 다양한 특징을 추출할 수 있다.



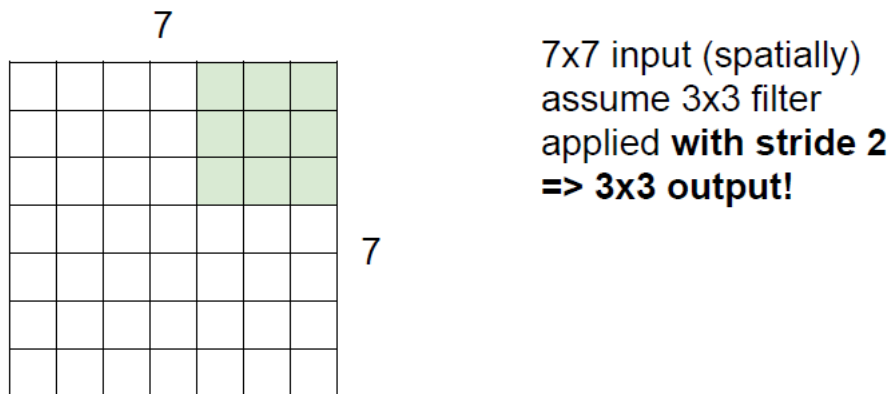
- 간단한 용어 정리

Stride : Filter 가 입력 Image 를 sliding 할 때, 몇 칸씩 띄어가면서 계산을 할 것인가 알려주는 것.

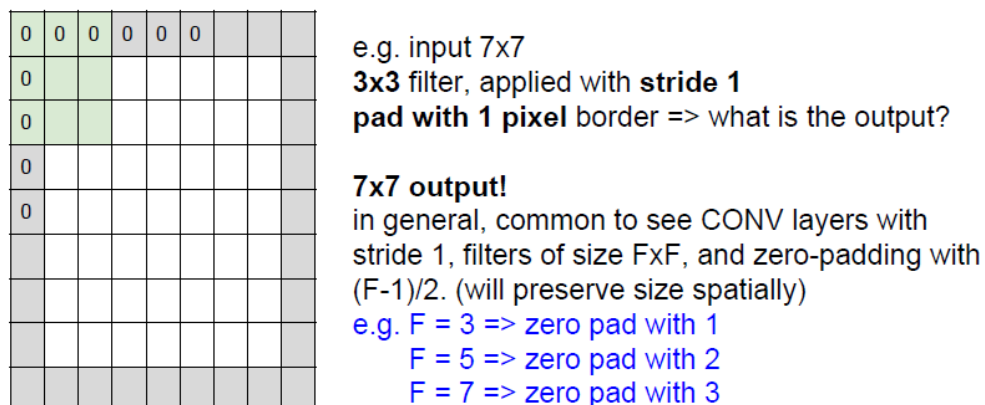
예를 들면, Stride = 1 일 때



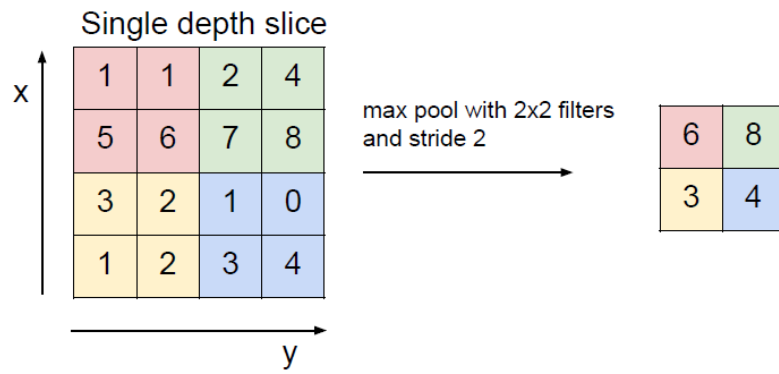
Stride = 2 일 때



Padding : convolution 과정을 거치면 결과 activation map 의 가로,세로가 줄어들게 되는데, Input Size 를 유지하고 싶어서 Image 가장자리에 값을 채워주는 것. 여기서는 zero padding 을 예로 들었다.



Pooling : 강제로 downsampling 을 하고 싶을 때, 사용하는 기법. depth 에는 영향을 주지 않는다. 파라미터의 수를 줄어들게 하는 효과가 있다. 주로 Maxpooling 이라는 기법을 많이 사용한다.



아래처럼 Convolution layer 를 여러개 사용하고 Fully connected layer 를 마지막에 사용하면 CNN 딥러닝 모델이 만들어 진다.

