

사실 전자공학을 전공한 사람이라면 신호처리 수업을 들으면서 많이 들어봤겠지만, 그때의 Convolution은 Linear Time Invariant System상에서 이전값과 현재값을 연산하기 위해 주로 사용하던 연산이다. 의미는 그런식으로 담겨져 있었는데, Convolution Neural Network 상에서는 그런 큰 의미보다는 그냥 이미지내에서 feature를 뽑기 위한 용도로 연산을 활용한다.

일반적으로 다음의 케이스에 대해서 Convolution을 취한다.

0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Input Image

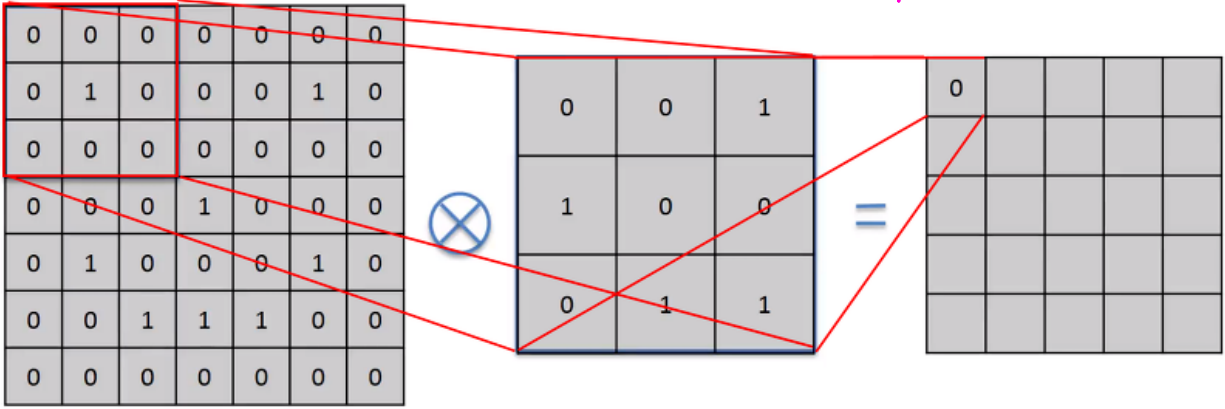
0	0	1
1	0	0
0	1	1

Feature Detector (특징 추출기)

(= Filter, kernel)

위와 같이 어떤 입력 이미지가 들어왔을 때 여기서 Feature를 감지할 Detector가 하나 존재한다. 이 Detector를 다르게는 Kernel이나 Filter라고 표현할 수 있는데, 말그대로 이 Detector가 짝 Input Image의 모든 영역을 훑으면서 특정 Feature를 뽑게 되는 것이다. 이를 바탕으로 해당 입력에 대한 Feature Map을 그리게 된다.

↑ "Feature Detector가 Feature Map을 그린다."



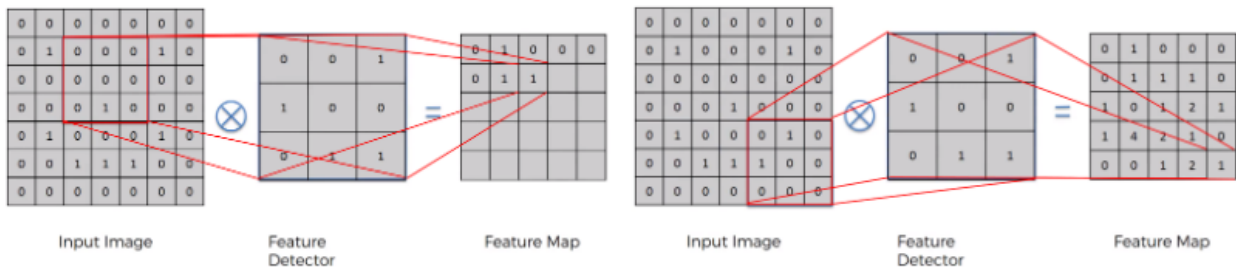
Input Image

Feature Detector (=filter)

Feature Map

• filter가 적용된 결과물.  
• 특징이 추출되어진 결과물.

이때 Input Image의 pixel과 Feature Detector의 각 kernel이 곱한 결과를 합하게 되고, 그 결과가 Feature Map의 하나를 채우게 된다. 이런 과정을 Feature Detector가 입력 이미지를 모두 훑을 때까지 계속하게 된다.



Input Image

Feature Detector

Feature Map

Input Image

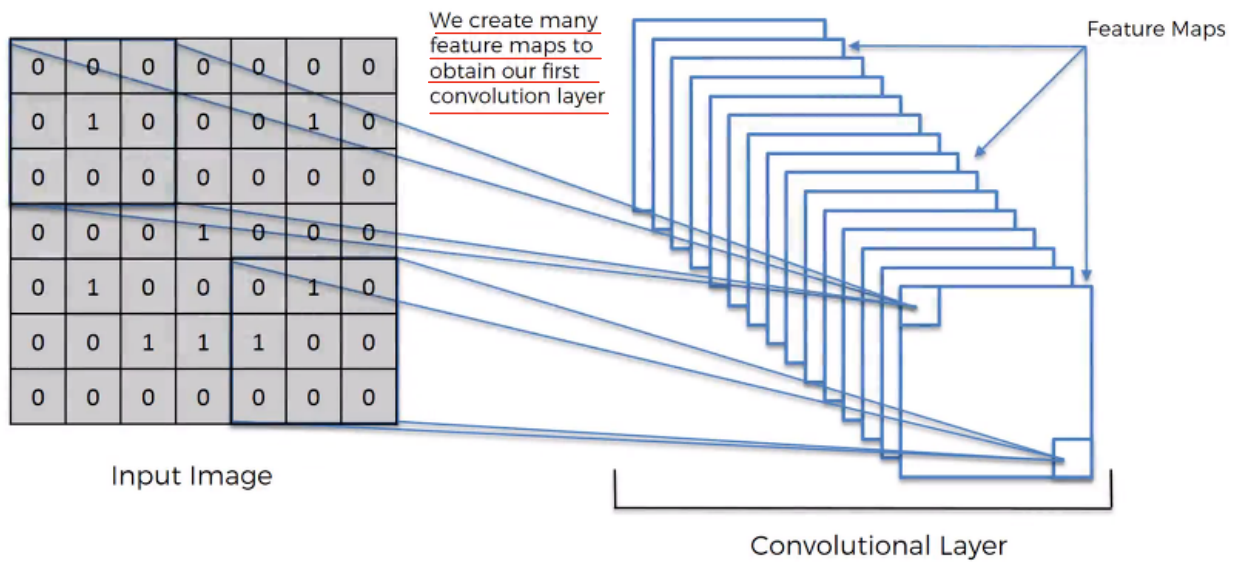
Feature Detector

Feature Map

이렇게 Feature Map을 만들고, 원래 입력 이미지와 크기를 비교해보면 원래의 7x7 이미지가 5x5 이미지로 작아진 것을 알 수 있다. 다시 말해 pixel 별로 비교해도 49회 걸리던 일이 25회만 비교하면 되는 것이다. 이로 인해 이미지 비교시 속도라는 이점을 얻을 수 있는게 Convolution연산의 특징이다. 물론 속도가 빠른 만큼 원래 정보가 소실된다는 점도 있긴 하다. 우리의 목적 (이 모든 pixel을 비교하는 게 아닌, 이미지의 feature, 특징을 뽑아내는데 주안점을 둔만큼 이런 과정이 필요하다. 다시 말해 우리가 얻고자 하는 것은 원래의 입력 이미지와 정확히 똑같은 이미지를 분류하는 게, 아니라 Feature Map을 통해 Feature를 뽑는 것, 예를 들어 4라는 값이 Feature Map상에 있고, 그 옆에 2가 위치해 있다던지의 Pattern을 얻어내는 것이다.)

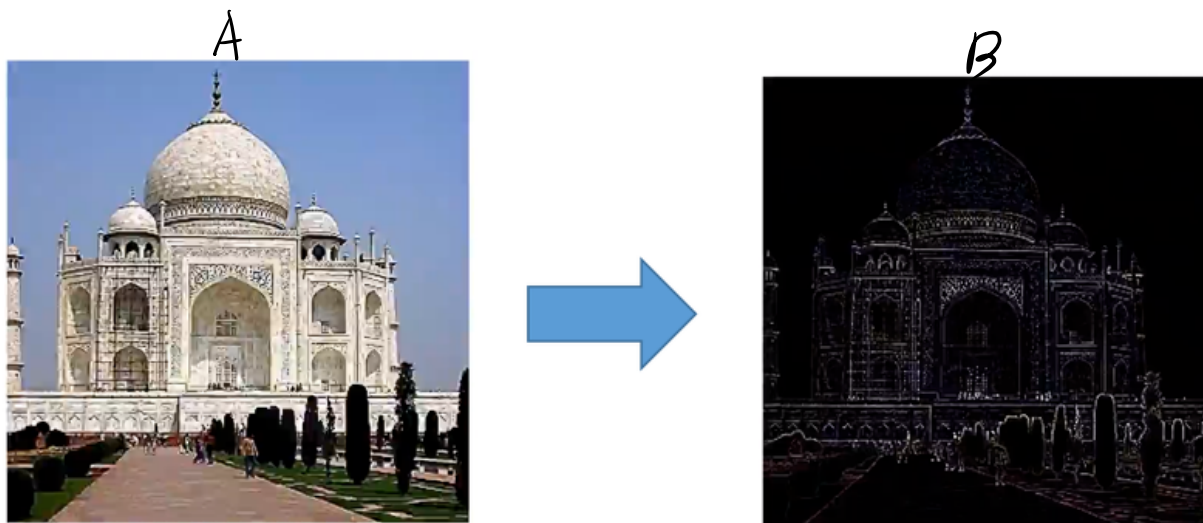
※

그리고 우리가 Feature Detector를 위의 예시와 같이 저렇게 정의된 것만 쓸 수 있는 것이 아니다. 다양한 Feature Detector를 사용할 수 있고, 각각 이에 대한 Feature Map도 뽑아낼 수 있게 된다.



결국 이렇게 뽑아낸 Feature Map들을 잘 살펴보면 얻어낸 Feature 중에서도 가장 중요한 Feature를 뽑아낼 수 있을 것이고, 이를 위해서 학습하는 역할을 CNN중 Convolutional Layer에서 수행한다.

사실 이런식으로 Filter를 적용하는 것은 굳이 이런 Deep Learning 영역이 아니어도 그 쓰임새가 있었다. 보통 이미지 보정을 할때도 sharpen을 준다면 Blur 효과를 주는 것도 원 이미지에 특정 Filter를 적용해서 효과를 누릴 수 있다. 이렇게 본래 Filter라고 불리던 것이 CNN에서 Feature Detector로써 사용되는 것이다. 결국 그런 이미지 보정을 하는 것도 다른 시점에서 보면 그 그림의 "특징"을 부각시켜 보이게 하는 것이다.



위의 이미지같은 경우도 영상처리에서 Sobel filter라는 것을 적용하게 되면 오른쪽과 같은 그림이 나오게 되는데, 이로써 이미지 내에서 특정 물체의 윤곽이라는 특징을 얻어낼 수 있다.

간단히 요약하면, 기존에 나와있던 Convolution 연산을 활용해서 입력 이미지의 특징을 찾는 과정을 수행하는 것이 CNN내에서 Convolution Layer가 수행하는 역할이 되겠다.