

CNN 세션

제 1강

발표자: 5기 한영웅



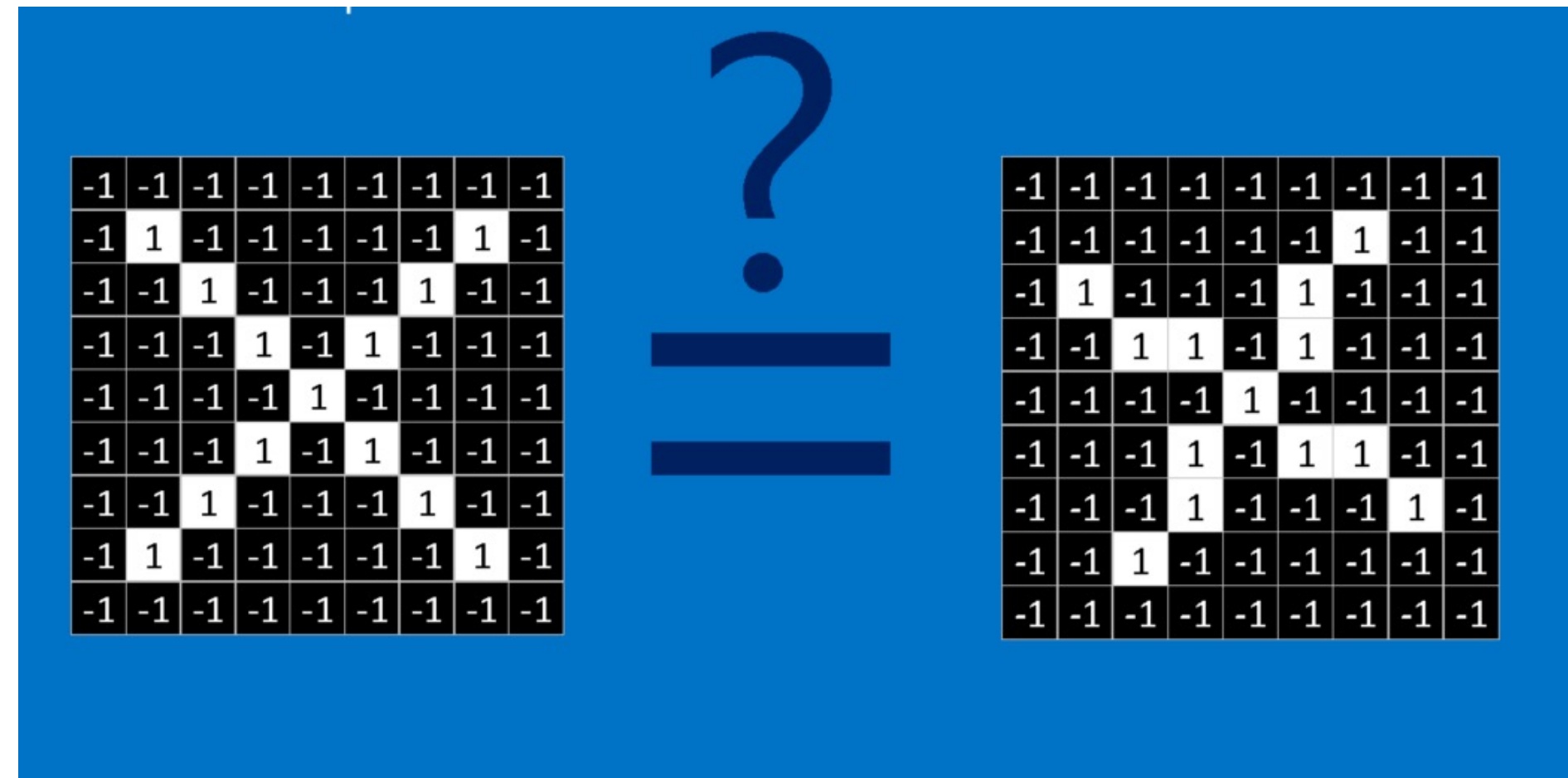
워밍업: 인간의 인지활동과 컴퓨터의 인지활동을 비교.

목차

1. CNN이란 무엇인가?
2. Filter 이해하기
3. 합성곱 이해하기
4. 패딩, 풀링 등 이해하기
5. 연습용 모델 학습

1. CNN이란 무엇인가?

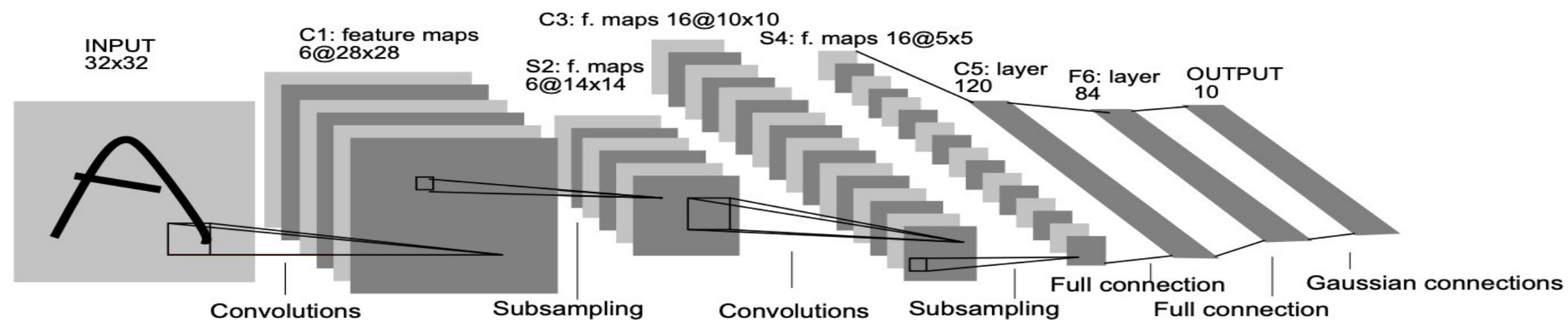
CNN은 Convolutional neural network의 줄임말로 합성곱을 이용한 인공 신경망입니다.



눈은 어떻게 두 사진이 X임을 쉽게 이해할까?

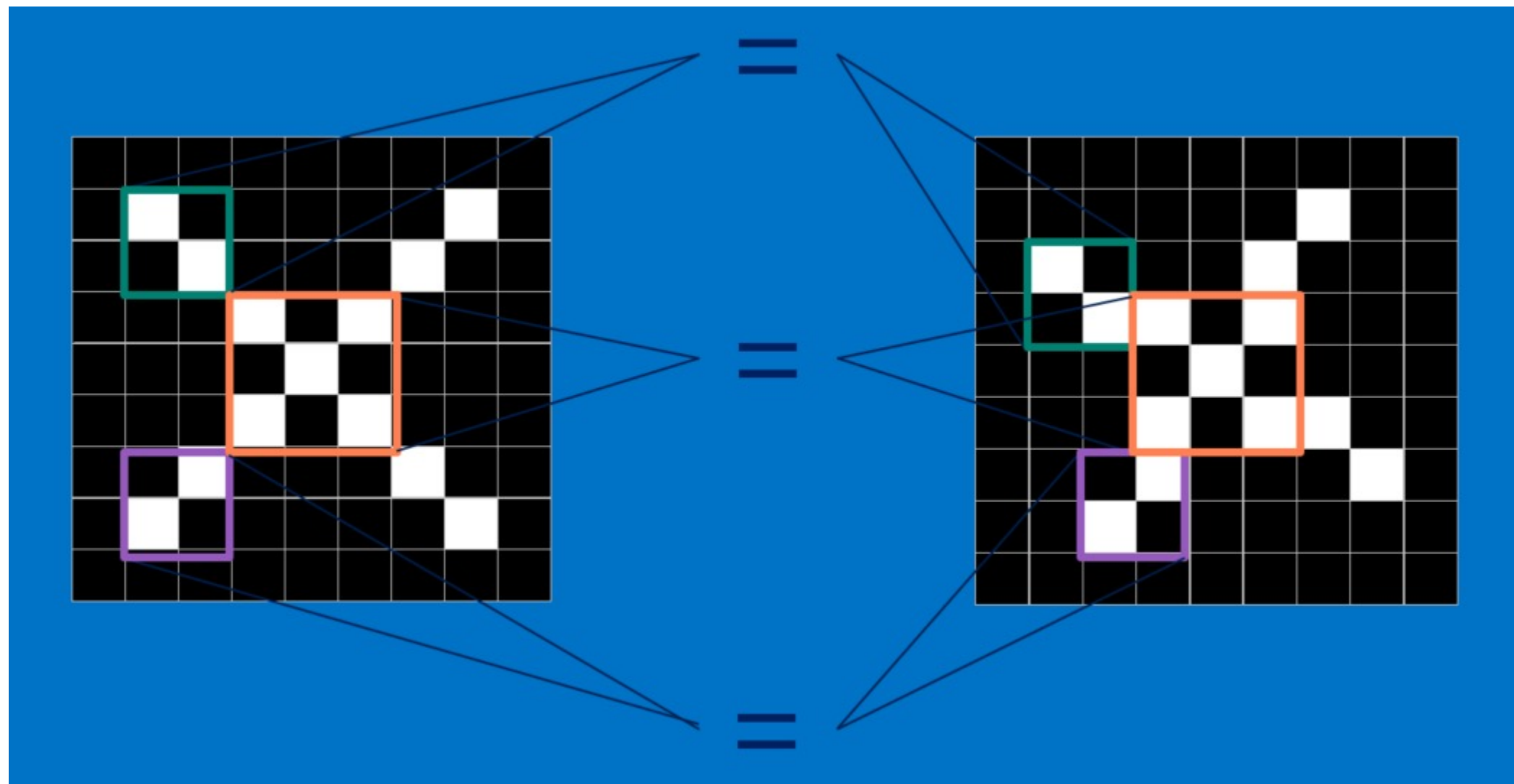
우리 눈에는 국소적인 영역을 보고 단순한 패턴에 자극을 받는 단순 세포와 넓은 영역을 보고 복잡한 패턴에 자극을 받는 복잡세포가 계층을 이루고 있습니다.

이것을 인공신경망으로 만드는 것이 CNN의 기본 아이디어입니다.



2. Filter 이해하기

Filter(혹은 kernel)은 CNN에서 ' 학습될 가중치 ' 입니다.

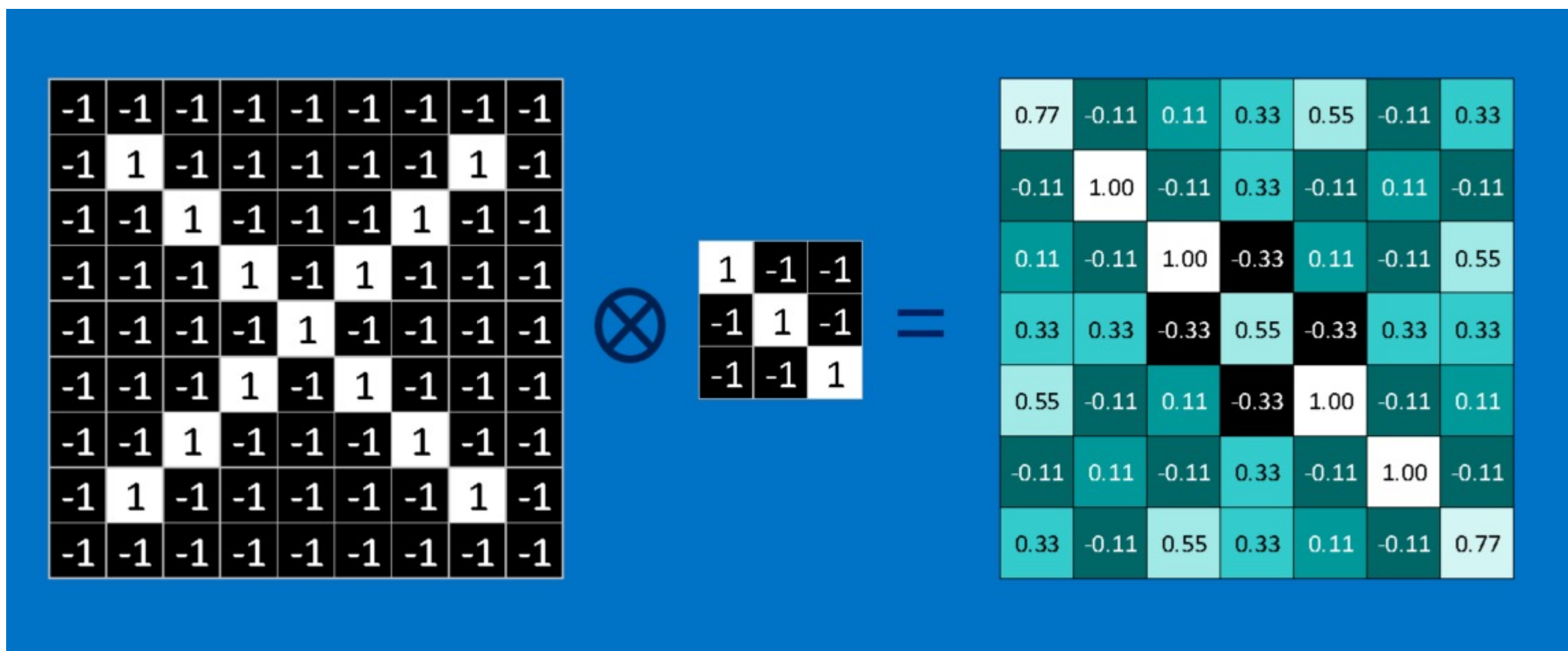


Input image의 'feature (특징)'들을 filter를 통하여 추출합니다.

앞에서 언급했듯이 추출되는 특징에는 단순한 특징과 복잡한 특징 모두가 포함됩니다.

3. 합성곱 이해하기

합성곱 신경망에서는 하나의 함수가 다른 함수와 **얼마나 일치하는가** 라는 질문에 답을 하는 데에 사용됩니다.



3. 합성곱 이해하기

Stride(스트라이드): 합성곱연산과정에 Filter가 이동하는 단위

Activation map (활성화 지도) , Feature map(특성 지도):
Filter 하나와 input 이미지가 합성곱 연산되어 도출된 결과

Stride=2인 경우

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

×

w0	w1
w2	w3

=

중간 복습!

- 1. 이미지 전체를 한 번에 학습하는 것이 아니라 특정 크기만큼 차례대로 훑는다.**
- 2. 1의 '특정 크기'라는 건 filter의 크기를 의미한다.
Kernel_size라고도 말한다.**
- 3. Cnn 모델에서 filter가 ' 학습되는 가중치' 이다.**
- 4. 이미지의 특정 부분과 filter가 가지고 있는 특징이 얼마나 일치하는가를
계산하기 위해 합성곱을 사용한다.**
- 5. Stride는 합성곱 연산을 위해 filter가 이동하는 단위이다.**

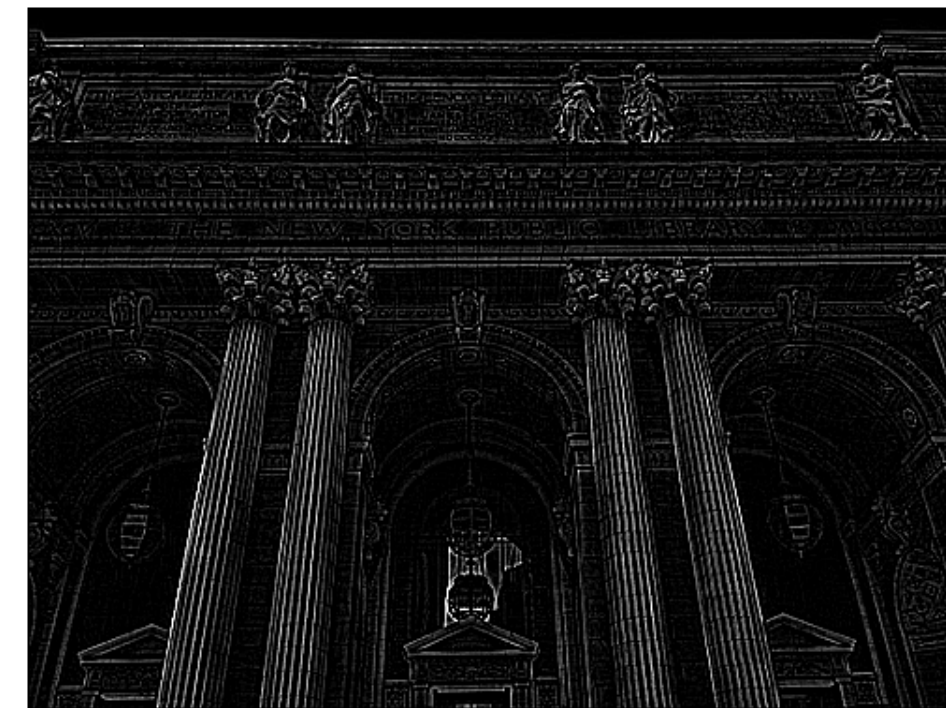
Filter에 선택에 따른 결과물



→
엠보스 필터 사용



→
경계선 필터 사용



이러한 역할을 하는 필터가 우리가 지정한 갯수만큼 학습이 됩니다. 이 필터들을 통해 우리는 그림의 다양한 특징들을 뽑아내 다음 은닉층으로 보냅니다. 필터를 통해 뽑아낸 특성들이 중첩됨에 따라 모델은 더 복잡하고 다양한 형태를 구분할 수 있게 됩니다.

<https://setosa.io/ev/image-kernels/>

4-1. 패딩 이해하기

학습을 진행할수록 feature map의 크기가 작아지는 현상이 발생!

Input_size가 매우 크지 않은 이상, 신경망을 깊게 만들 수 없다.

이 현상을 방지하기 위해 패딩을 사용.

0	0	0	0	0	0
0	1	2	3	0	0
0	0	1	2	3	0
0	3	0	1	2	0
0	2	3	0	1	0
0	0	0	0	0	0

⊛

2	0	1
0	1	2
1	0	2



7	12	10	2
4	15	16	10
10	6	15	6
8	10	4	3

padding x의 경우

padding =1의 경우

padding =2의 경우

<https://excelsior-cjh.tistory.com/79>

패딩 값을 0으로 줄 수도 있고 (제로 패딩) 이미지의
경계선 값을 복사하거나 임의의 값을 패딩 값에 넣을
수도 있습니다.

매우 중요한 수식! 꼭 외우세요!

Feature map의 크기: O
(복습: feature map은 filter 하나를
지난 후에 생긴 결과물!)

이미지의 크기: I
필터의 크기: K
Stride: S
패딩의 크기: P

직접 쓰기!

4-2. 풀링 이해하기

이미지 데이터의 크기가 큰 경우 굳이 패딩이 필요 없다. 오히려 층이 깊어질수록 feature map의 수가 많아져서 연산해야 하는 수의 크기가 너무나도 커진다.

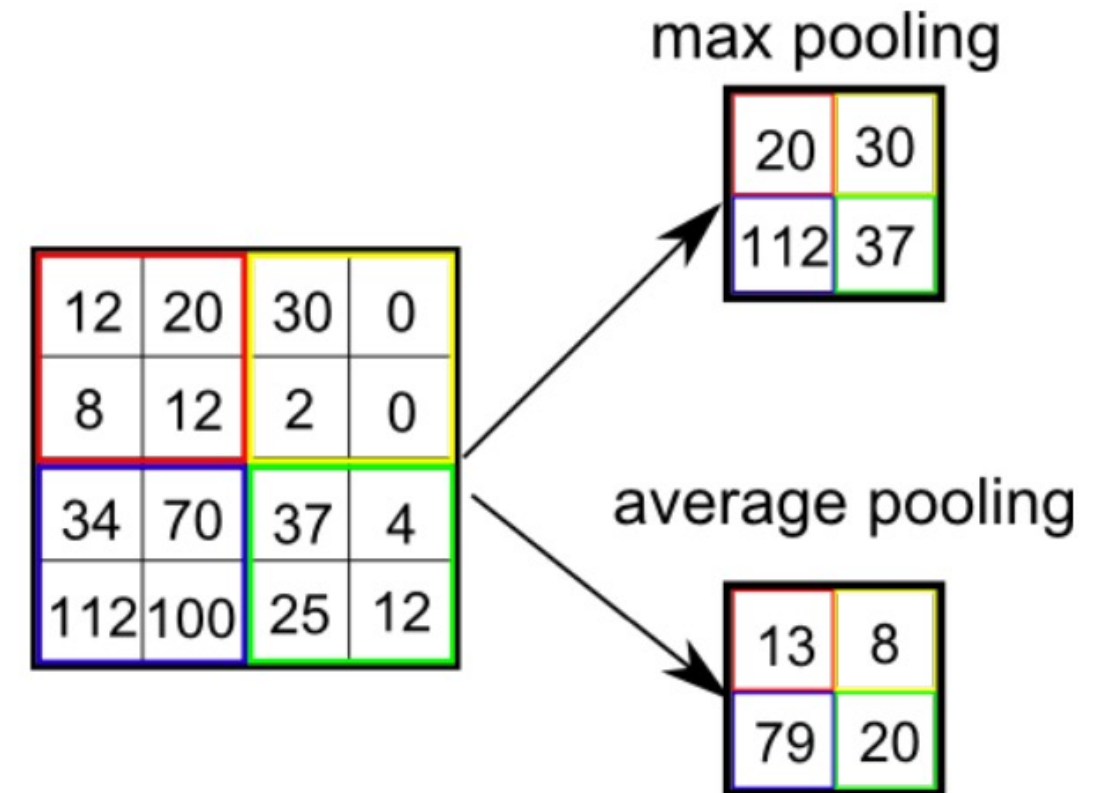
상황에 따라서 높은 화질이 필요하지 않을 수도 있고, 넓게 봐야 파악할 수 있는 특징도 있을 수 있다.

Max Pooling: 일정 크기의 구간 내에서 가장 큰 값을 전달하고 다른 정보는 버리는 것. 자극의 관점에서 봤을 때 가장 강한 자극만 남기고 나머지는 무시하는 것.

Average Pooling: 일정 크기의 구간 내의 값들의 평균을 전달하는 방법으로 자극의 관점에서 봤을 때 평균적 자극을 전달하는 것.

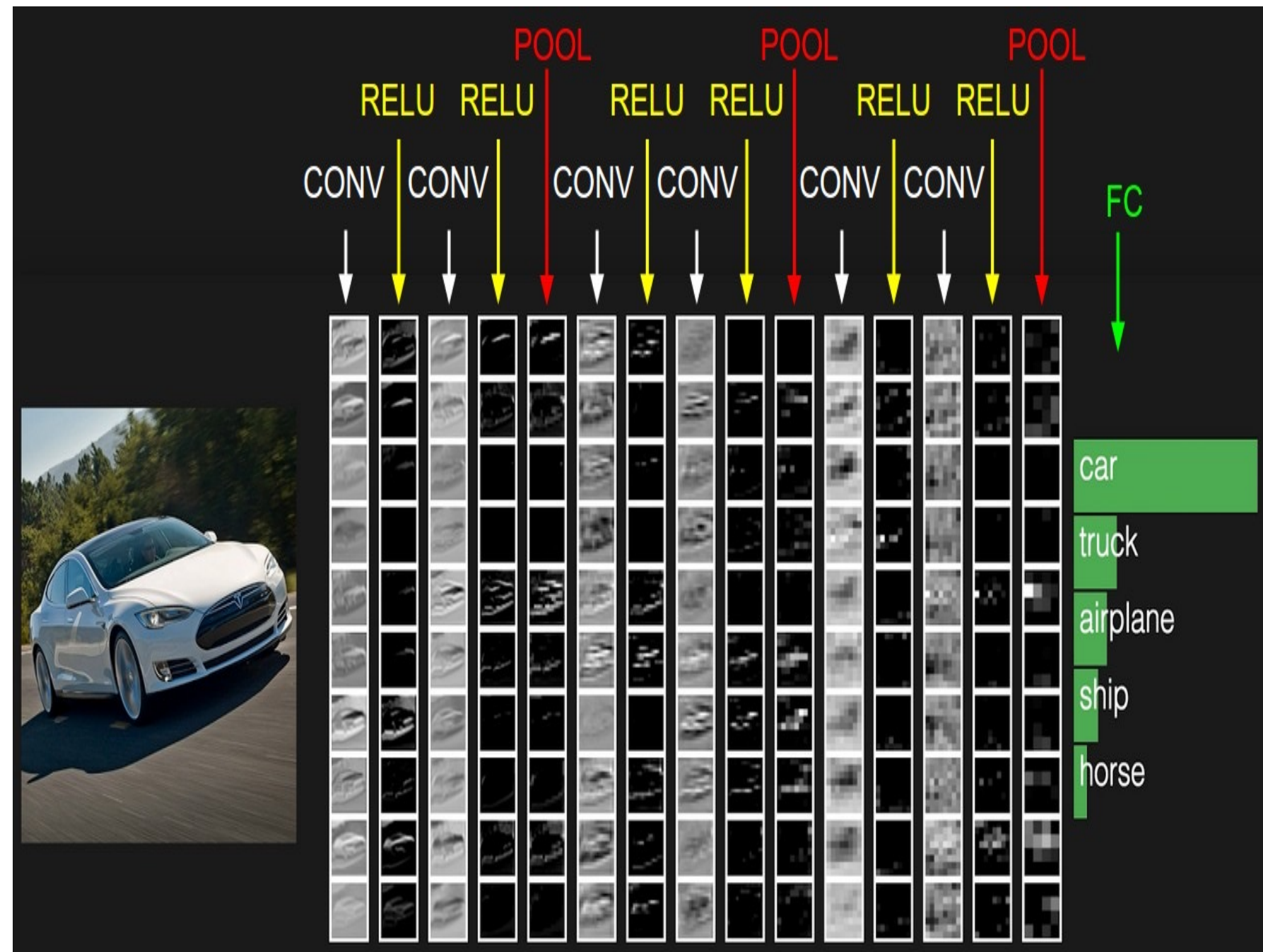
(합성곱연산의 관점에서는 해당 필터의 모양과 평균적으로 얼마나 일치하는지를 측정하는 것)

→ **Pooling**



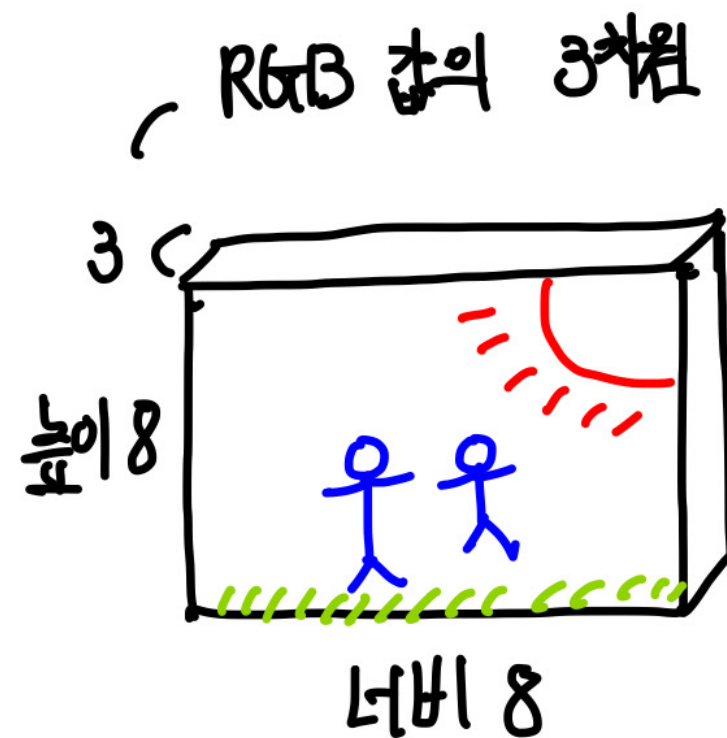
합성곱 신경망의 일반적인 구조

1. 입력값에 대해 몇 번의 합성곱 연산을 활성화 함수와 함께 적용한다.
2. 풀링으로 전체 크기를 줄여준다.
3. 1, 2를 반복한다.
4. 어느 정도 특성을 다 뽑으면 뽑은 특성들을 입력으로 받는 인공 신경망을 이용하여 각 클래스별 확률을 추출하거나, 특정 수치를 추출한다.



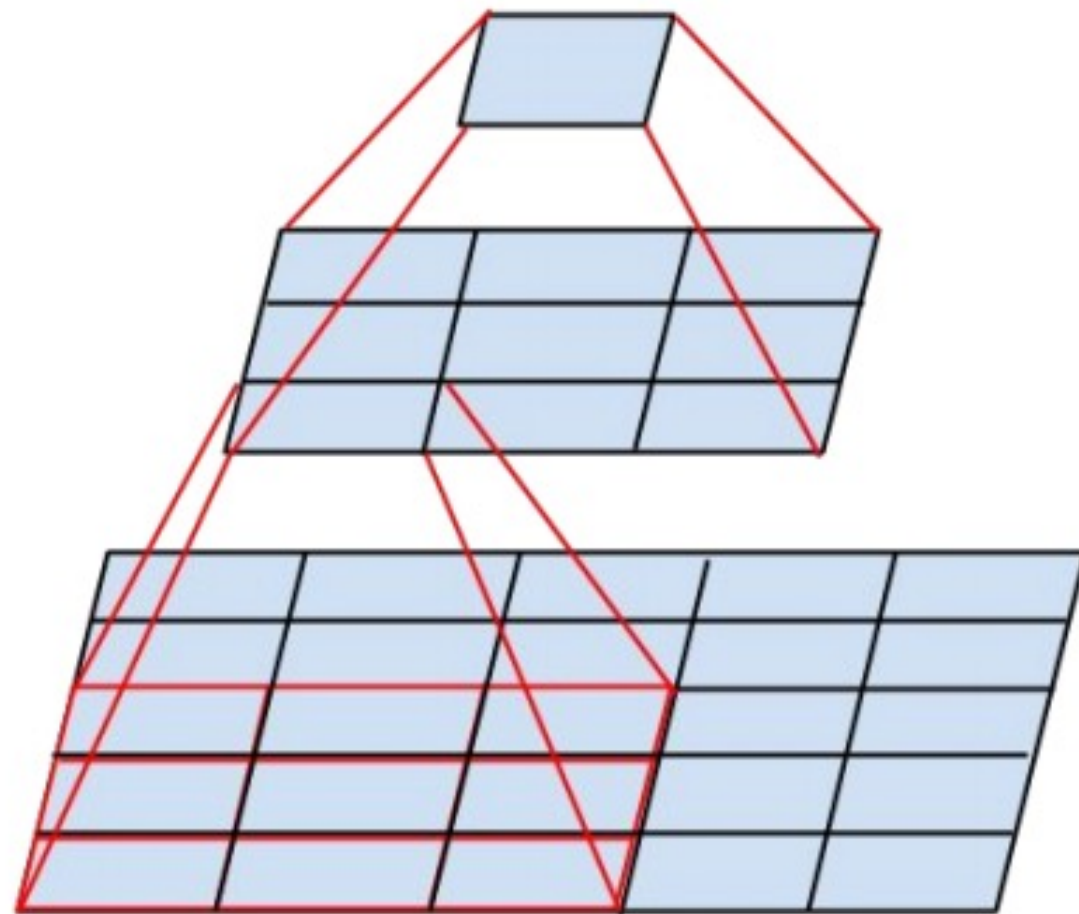
CNN 더욱 이해하기

[색 이미지로 복습]



합성곱
⇒
3x3 필터
슬라이드 1
패딩 1
채널수 16

**그런데 왜 깊은 층에서 나온 feature
map이 얇은 층에서 나온 feature
map보다 복잡한 이미지의 특성을 가지고
있나요?**



— ۲۳ —