

CNN

Convolution Neural Network 의 구조와 이해

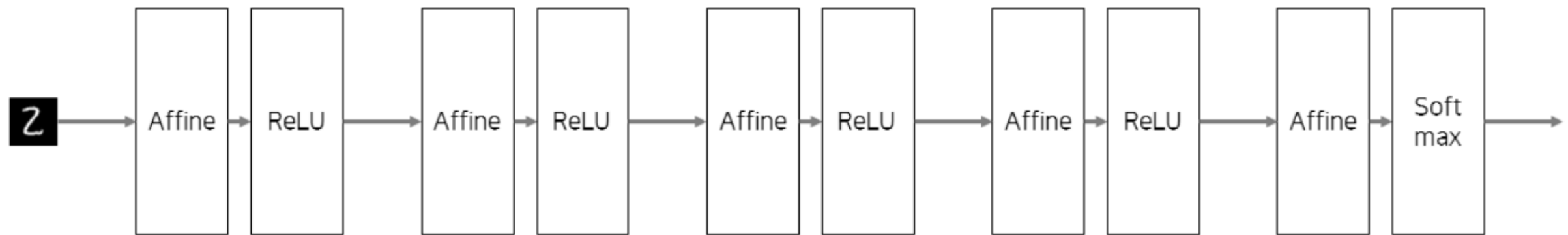
DAY 2

JARAM DeepLearning Study

박상원

기존 신경망의 구조

Affine 과 ReLU 로 구성된 인접한 뉴런끼리 연결되어있는 구조



완전 연결 신경망(Fully Connected Neural Network)

각각의 뉴런들이 인접하는 계층의 모든 뉴런과 연결되어 있다

위의 그림과 같이 Affine 계층 뒤에 활성화 함수를 갖는 ReLU 계층이 이어지는 조합이 여러개 있고, 마지막에 Softmax 계층을 사용하여 최종 결과(확률)를 출력하는 방식이다.

완전 연결 신경망의 문제점

이미지 처리에 적절하지 않은 문제점

입력 데이터를 1차원 데이터로 평탄화한 뒤 모든 입력 데이터를
동등한 뉴런으로 취급하여 처리하기에 데이터의 형상이 무시된다.

이미지의 경우 세로, 가로, 채널(색상)로 구성된 3차원 데이터를 1차원 데이터로 평탄화하기에
공간적으로 가까운 픽셀의 값이 비슷하다던가, RGB 의 각 채널이 서로 밀접하게 관계되어 있다는 등의
3차원 속에서 의미를 갖는 본질적인 패턴을 찾을 수 없게 된다.

데이터의 형상 뿐만 아니라 완전 연결 계층은 모든 뉴런을 연결하기 때문에 너무 많은 연결이 생기게 된다.
이는 CNN 을 통해 층을 부분적으로 연결함으로써 이 문제를 해결할 수 있다.

CNN 이란

데이터의 형상을 유지할 수 있는 Convolution Neural Network

con·vo·lute

미국식 [kánvəlù:t]   | 영국식 [kón-]  

출판사별  **동아출판** YBM 교학사 슈프림 영영사전

형용사 | 타동사-자동사 | 명사

형용사

1. (식물·패류) 한쪽으로 감긴, 소용돌이꼴의, 회선상(回旋狀)의
2. 둘둘 말린

타동사-자동사

감아 넣다, 둘둘 말다[감다]: 뒤엎히다

Convolution 의 동사형인 Convolute 를 보면

둘둘 말린, 회선상의 라는 의미이다. Convolution 은

의미 그대로 입력 특징 맵(입력 데이터)을 필터(가중치)에 대해

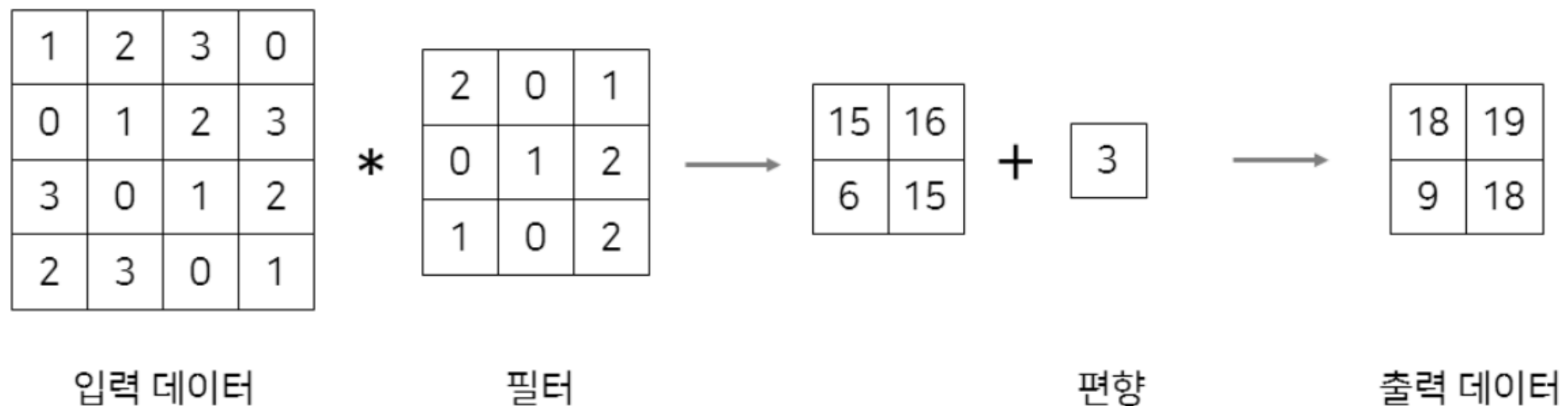
돌려가며 연산하여 출력 특징 맵(출력 데이터)을 만드는 것이다.

참조 : [CNN convolution](#)

Feature map을 Activation map 이라고도 한다.

합성곱 연산

Fused Multiply-Add, FMA



CNN의 합성곱 연산은 필터의 윈도우를 일정 간격으로 이동해가며 입력 데이터의 대응하는 원소끼리 곱한 후 총합을 구해 해당 장소에 저장하는 데 이를 단일 곱셈-누산(FMA)이라고 한다.

사실 CNN은 합성곱이 아니라 교차상관을 사용하지만 딥러닝에서는 잘 구분하지 않는다고 한다.

CNN 와 시각피질

사람의 이미지 인식 방식과 CNN

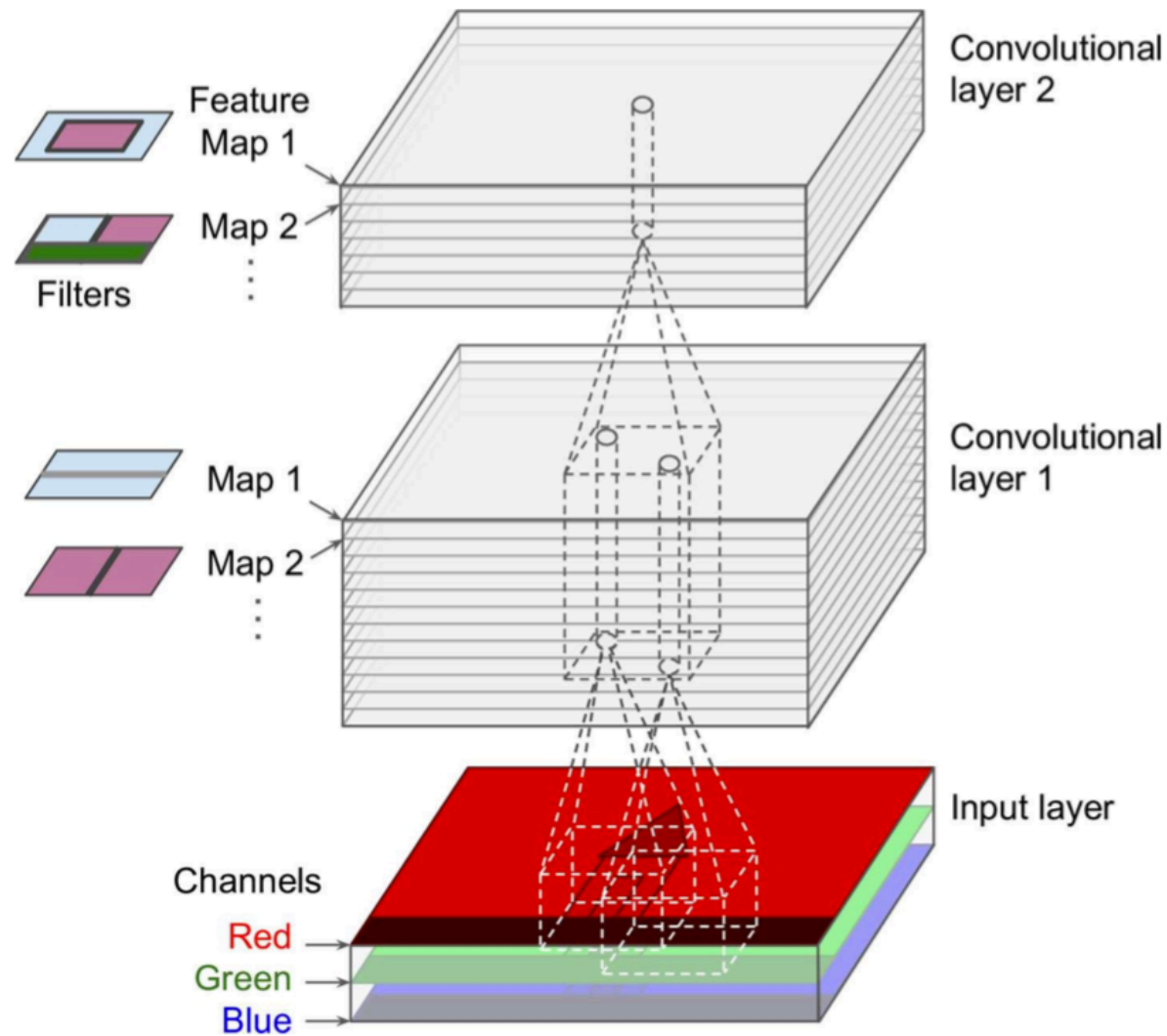


Figure 13-6. Convolution layers with multiple feature maps, and images with three channels

CNN 은 사람이 이미지를 처리할 때 시각 피질 안의 많은 뉴런들이 일부 범위 안에 있는 시각 자극에만 반응하고, 이런 뉴런들이 서로 합쳐지며 전체 시야를 감싸게 된다는 연구에서 영감을 받은 신인식기가 점진적으로 진화해가며 만들어진 것.

참조 : CNN 이미지 인식

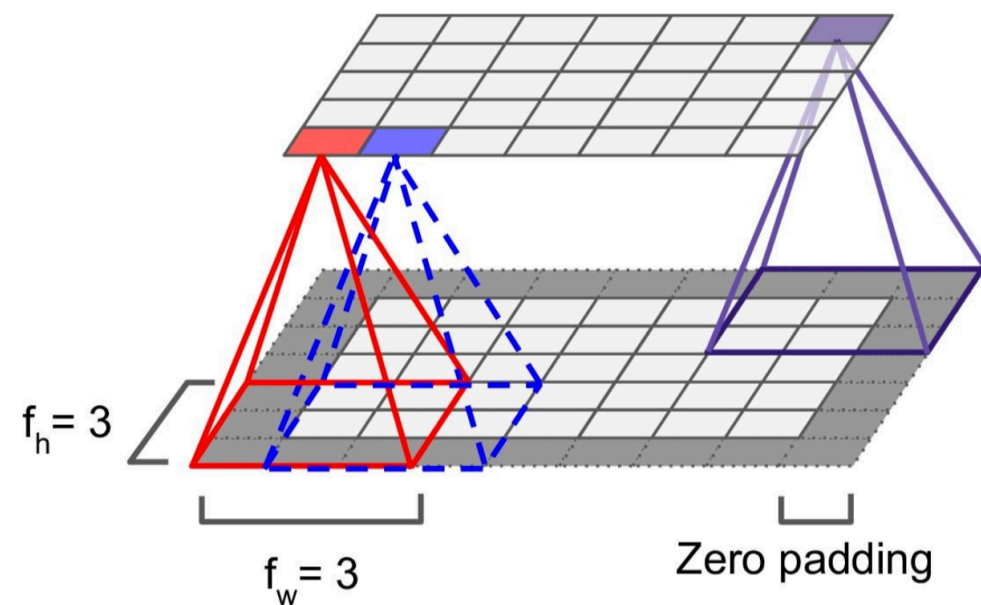
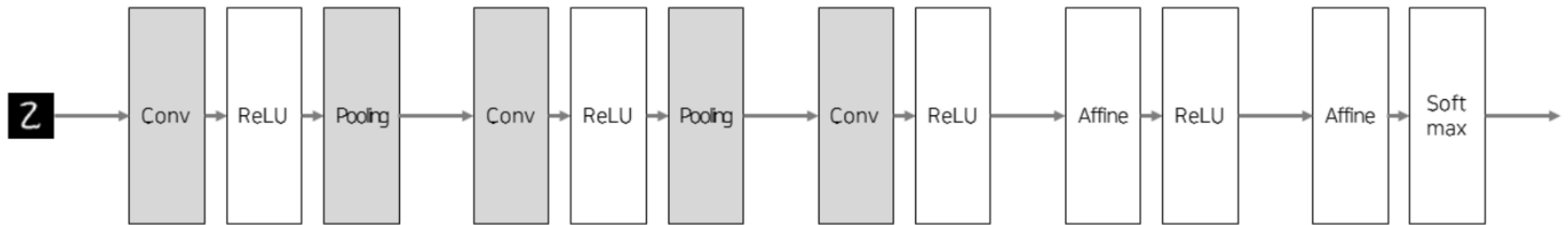


Figure 13-3. Connections between layers and zero padding

그림 출처 : Hands-On Machine Learning with Scikit-Learn & TensorFlow

CNN의 구조

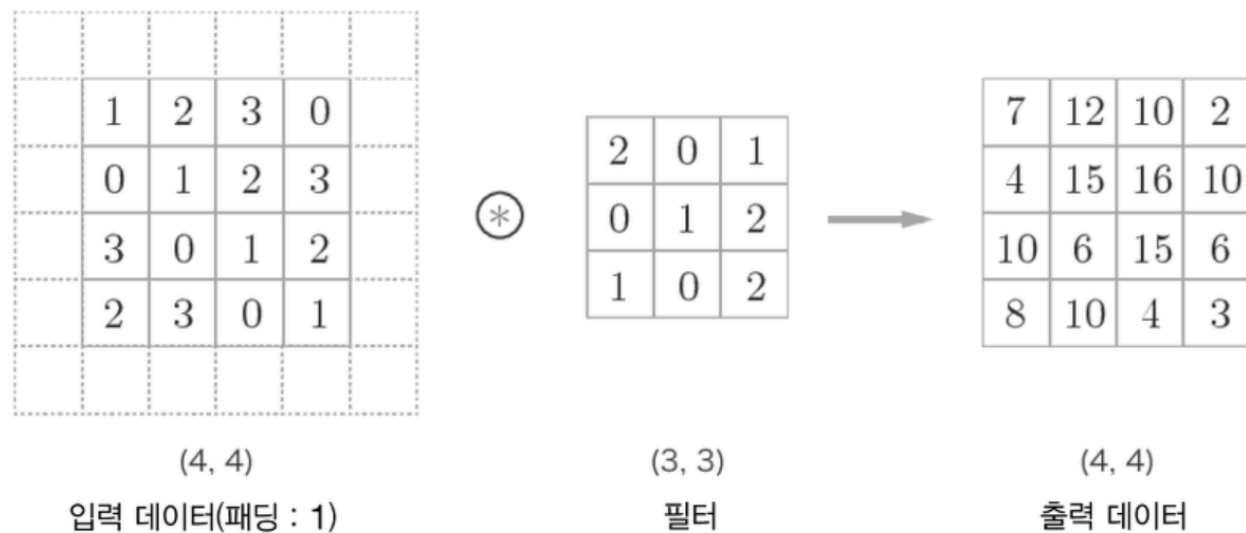
Convolution 계층과 Pooling 계층이 추가된 CNN



출력에 가까운 층에서는 지금까지의 Affine-ReLU 조합을 사용할 수 있고,
또 마지막 출력 계층에서는 Affine-Softmax 조합을 그대로 사용한다.

패딩

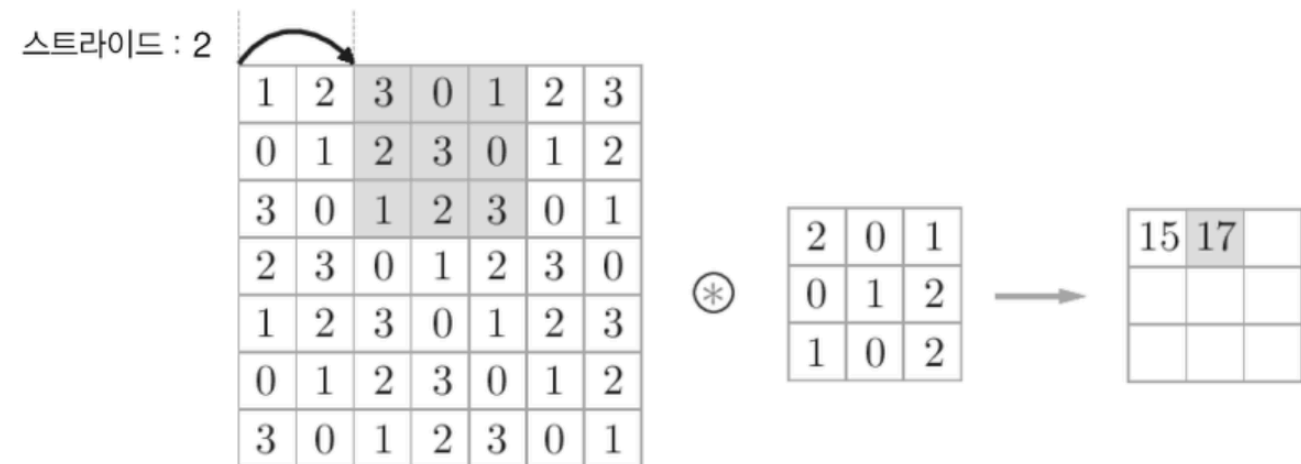
출력 특징맵의 크기를 키우기 위해 사용하는 Padding



입력 데이터의 주변을 특정 값으로 채우는 것
주로 0을 채워 넣는데 이를 제로 패딩이라고 한다.

스트라이드

출력 특징맵의 크기를 줄이기 위해 사용하는 Stride



필터를 이동할 때의 간격을 조정하는 것

출력 크기를 조정하는 이유

합성곱 연산을 계속 적용하기 위해서

(4,4) 입력 데이터에 (3,3) 필터를 적용한다고 가정할 경우 출력은 (2,2)가 되어, 입력보다 2만큼 줄어든다.

합성곱 연산을 거칠 때마다 크기가 작아지면 어느 시점에서는 출력 크기가 1이 되어 합성곱 연산을 적용할 수가 없게 된다. 이는 합성곱 연산을 몇 번이나 되풀이하는 심층 신경망에서는 문제가 될 수 있다.

그렇기에 우리는 출력 데이터의 공간적 크기를 조정하여 다음 계층에 전달하는 것이다.

출력 관계 수식

필터와 스트라이드 값에 따른 출력 데이터 크기 계산

$$OH = \frac{H + 2P - FH}{S} + 1$$

$$OW = \frac{W + 2P - FW}{S} + 1$$

OH Output Height

OW Output Width

H Height

W Width

P Padding

S Stride

FH Filter Height

FW Filter Width

결과 크기가 정수로 나누어 떨어지는 값이어야 한다는 점에 주의해야 한다.

(딥러닝 프레임워크에서 딱 나누어떨어지지 않을 경우 반올림하는 등의 방법으로 에러가 나지 않도록 구현할 수 있다.)

풀링 계층

Feature Map 을 임의로 줄이는 Pooling 계층



Convolution 계층을 통해서 특징이 어느정도 추출되었다면 특징 맵을 줄이는 작업을 한다.

굳이 모든 특징을 가지고 판단하지 않아도 괜찮은 결과를 얻을 수 있기 때문이다.

그래서 추출된 특징 맵을 인위로 줄이는 작업을 하는데,

이 작업을 Subsampling 또는 Pooling 이라고 한다.

최대 풀링, 평균 풀링, L2 풀링 등이 있고 최대 풀링이 주로 쓰인다.

풀링 계층 특징

Feature Map 을 임의로 줄이는 Pooling 계층

1. 학습해야 할 매개 변수가 없다

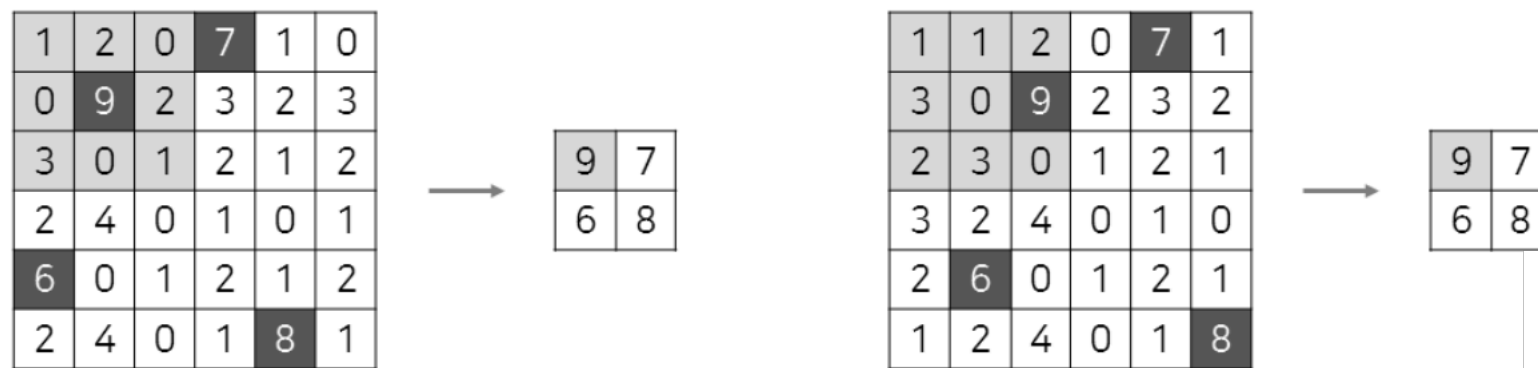
대상 영역에서 최댓값이나 평균을 취하는 명확한 처리이므로 학습할 것이 없다.

2. 채널 수가 변하지 않는다

채널마다 독립적으로 계산하기 때문에 입력 데이터의 채널 수와 출력 데이터의 채널 수가 같다.

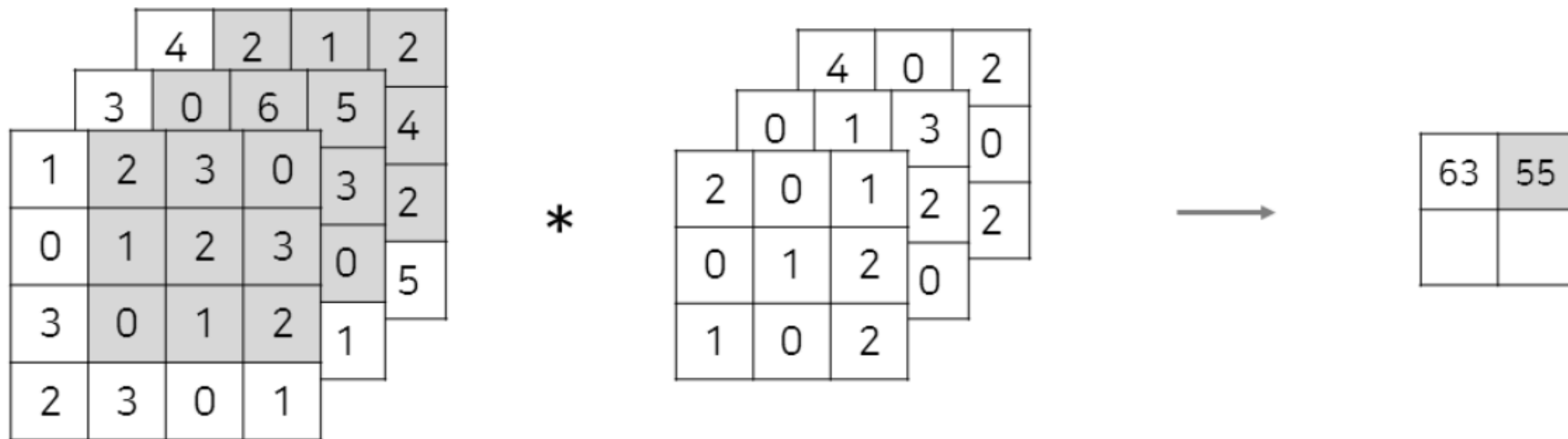
3. 입력의 변화에 영향을 적게 받는다 (강건하다)

입력 데이터가 조금 변해도 풀링의 결과는 잘 변하지 않는다.



3차원 데이터의 합성곱 연산

Channel 의 수를 같게 유지한다!

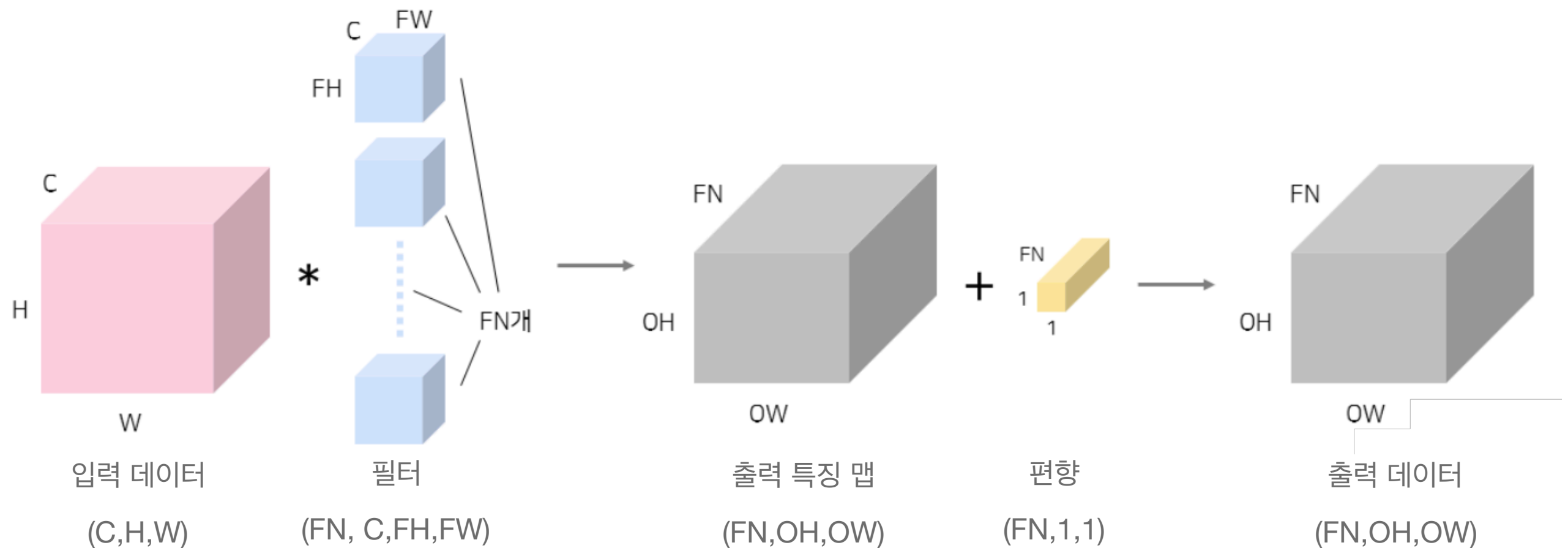


이미지는 세로, 가로에 채널까지 고려한 3차원 데이터이다.

채널마다 행렬을 가지며, 주의할 점은 입력 데이터의 채널 수와 필터의 채널 수를 같게 해야한다는 것이다.

3차원 데이터의 합성곱 연산

블록으로 생각하면 이해하기 쉽다!

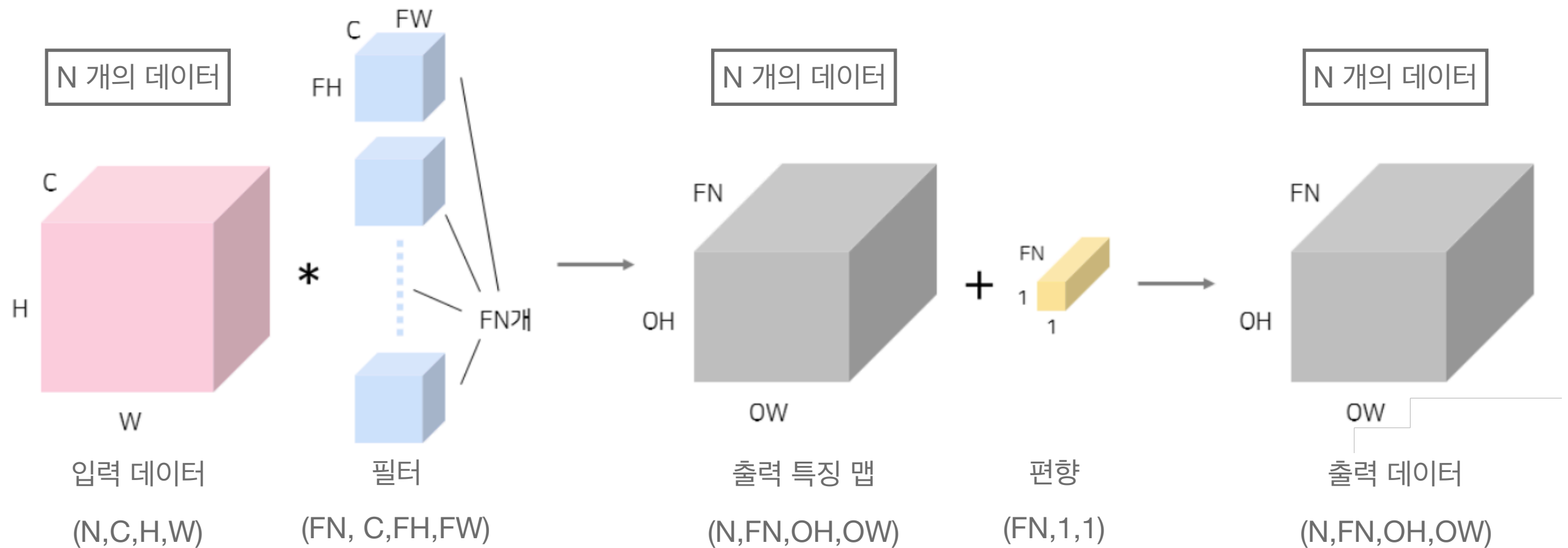


원래 출력 특징 맵은 1개 이지만 필터를 FN 개로 설정하여 출력 특징 맵의 수가 FN 개로 나오게 하였다.

(C : 채널 수, H : 높이(세로), W : 폭(가로))

배치 처리

처리 시간을 줄여주는 Batch Processing



데이터를 N 개로 묶어 4차원 데이터로 저장하여 처리하는 방식으로
한번에 N 회분의 합성곱 연산을 수행한다.

Q & A

자료 출처

밑바닥부터 시작하는 딥러닝

핸즈온 머신러닝

투빅스 6&7기 Deep Learning Seminar

A beginner's guide to Understanding CNN - 이승은님

조대협 의 블로그

cs231n

The background of the slide features a dark, textured surface, likely a wooden table. On the left side, there is an open notebook with lined pages, one of which is yellowed. On the right side, there is a cup of dark coffee. A white rectangular frame is centered on the slide, containing the text.

Thank you!

DAY 2

JARAM DeepLearning Study

박상원