

# 딥러닝 시작하기 (CNN 입문)

# 목차

- ▶ 01 기본 이해
- ▶ 02 초기 신경망 완전 연결층(FCL)의 한계
- ▶ 03 합성곱 신경망(CNN)
- ▶ 04 합성곱 신경망(CNN) - 특징맵 생성과정
- ▶ 05 합성곱 신경망 - 스트라이드(stride)
- ▶ 06 완전 연결층(FCL)과 합성곱 층의 연산의 차이
- ▶ 07 합성곱 신경망
- ▶ 08 풀링 연산(Pooling)

# 01 기본 이해

## ▶ CNN이란 무엇일까요?

- Convolutional Neural Network의 약자
- 인공 신경망의 한 종류
- CNN은 주로 이미지 인식, 물체 탐지, 음성 인식 등의 작업에 활용된다.
- CNN의 핵심 아이디어는 입력 데이터에서 지역적 연관성(local connectivity)와 공간 계층 구조(spatial hierarchy)를 잘 포착하는 것이다.

# 01 기본 이해

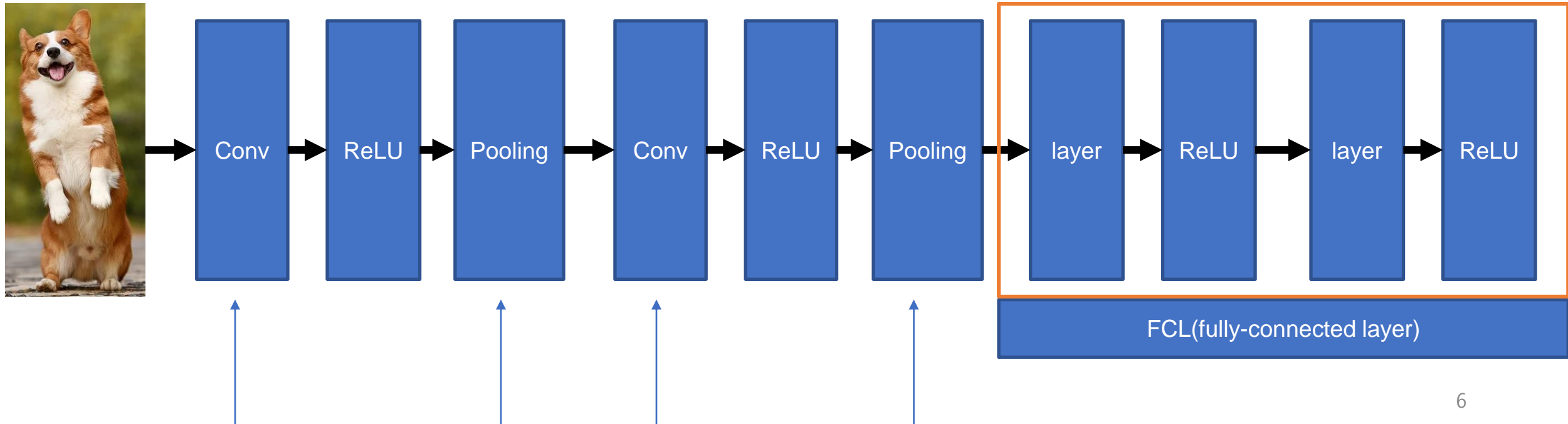
## ▶ CNN이란 무엇일까요?

- CNN은 입력 데이터에 대해 필터(filter)연산을 수행하여 특징을 추출한다. 이 과정에서 **합성곱(convolution)연산이 사용**된다.
- 추출된 특징 맵(feature map)은 다음 층으로 전달되어 더 높은 수준의 특징을 학습하게 된다.
- CNN은 이미지나 음성 인식 외에도 자연어 처리, 신호 처리 등의 다양한 분야에서 **활용**되고 있음.

# 01 기본 이해 CNN

- LeCun에 의해 처음 소개됨

(1) <https://dl.acm.org/doi/10.1162/neco.1989.1.4.541>



# 01 기본 이해

## ▶ CNN이 활용되는 모델

- **ResNet**(Residual Network) : 2015년 MS 개발. 이미지 분류(ResNeXt)
- **Inception**(GoogleNet) : 구글에서 개발된 모델.
- **DenseNet**(Dense Convolutional Network)
- **U-Net** : 생물의학 영상 분할(segmentation) 문제를 해결하기 위해 개발한 CNN 기반 모델. 인코더-디코더 구조를 가진다.

# 01 기본 이해

## ▶ CNN이 활용되는 모델

- YOLO(You Only Look Once) : 실시간 물체 탐지를 위한 단일 신경망 모델. 많이 사용됨
- Mask R-CNN : 물체 탐지 뿐만 아닌 인스턴스 분할(instance segmentation)도 수행
- U-Net : 생물의학 영상 분할(segmentation) 문제를 해결하기 위해 개발한 CNN 기반 모델. 인코더-디코더 구조를 가진다.

# 01 기본 이해 - 용어

## ▶ 채널, Channel

- 이미지 픽셀 하나하나를 실수, 컬러 사진은 천연색을 표현하기 위해 각 픽셀을 RGB 3개의 실수로 표현한 3차원 데이터. 컬러 이미지는 3개의 채널로 구성.

## ▶ 필터(Filter)

- 필터는 이미지의 특징을 찾아내기 위한 공용 파라미터. Filter를 Kernel이라고도 한다.
- CNN에서 학습의 대상은 필터 파라미터가 된다.



# 01 기본 이해

## ▶ 스트라이드, stride

- 필터를 적용하는 간격의 크기 stride라 함.

## ▶ 특징맵(feature map, activation map)

- 필터를 적용해서 얻어진 결과를 말함.

## 02 초기의 신경망 완전 연결층(FCL)의 한계

- ▶ Fully Connected Layer 만으로 구성된 인공 신경망의 입력 데이터는 1차원 형태로 한정.
- ▶ 3차원 사진 데이터를 1차원으로 평면화 시켜야 한다. 이때 공간 정보 손실 발생.
  - 한 장의 사진은 3차원 데이터, 배치 모드 사용되는 여러장 사진은 4차원 데이터
- ▶ Fully Connected Layer 파라미터 수가 많음.

## 03 합성곱 신경망(Convolutional Neural Network, CNN)

- ▶ LeCun에 의해 처음 소개됨

(1) <https://dl.acm.org/doi/10.1162/neco.1989.1.4.541>

- ▶ 이미지를 인식하기 위한 패턴을 찾는데 특히 유용

=> 이미지의 **공간 정보를 유지한 상태로 학습**이 가능한 모델(CNN)

## 03 CNN - 활용 분야

### ▶ Classification

(A) 이미지 안의 객체(Object)의 종류를 구분하는 행위.

### ▶ Localization

(A) 모델이 주어진 이미지 안의 Object가 이미지 안의 어느 위치에 있는지 위치 정보를 출력.

### ▶ Object Detection

(A) 보편적으로 Classification과 Localization이 동시에 수행되는 것을 의미.

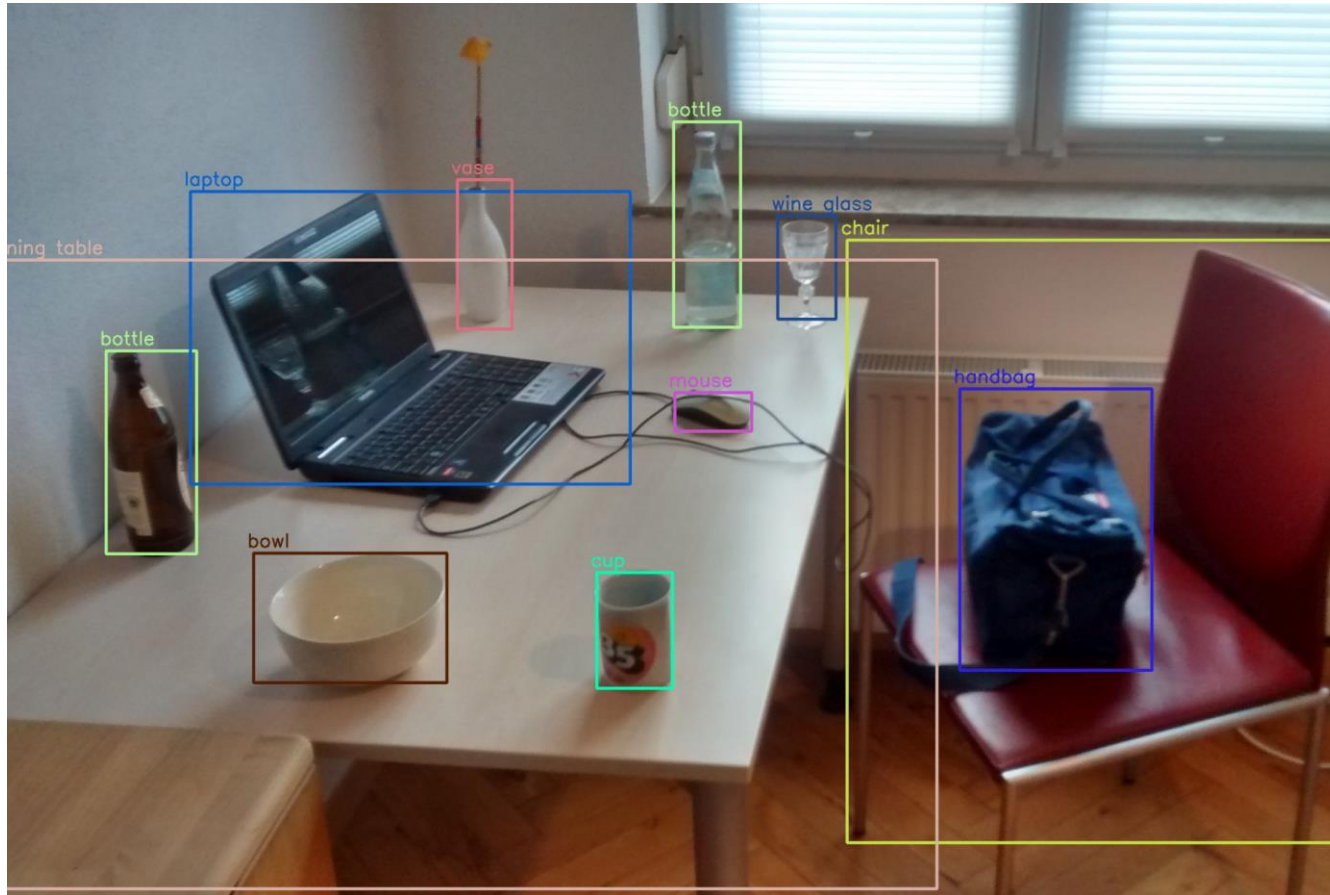
### ▶ Object segmentation

(A) Object detection을 통해 검출된 object 형성을 따라 object의 영역을 표시.

(B) 이미지의 각 pixel을 classification해서 결과를 도출. 전경과 배경을 구분하는 용도

# 03 합성곱 신경망(Convolutional Neural Network, CNN)

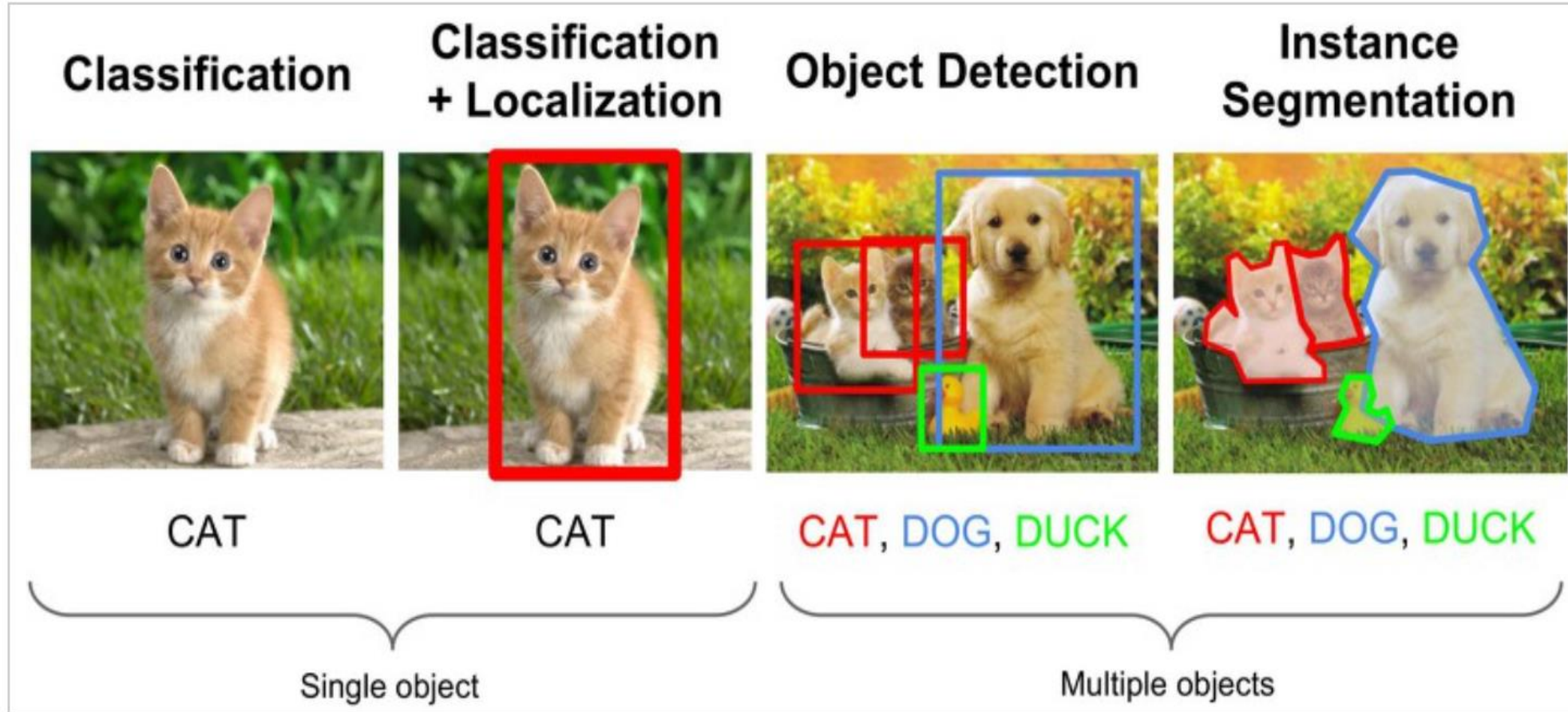
## ▶ 어느 분야에 사용되는가? Object Detection



80개의 공통 객체를 탐지할 수 있는 YOLOv3 모델. 딥 뉴럴 네트워크로 감지된 물체

# 03 합성곱 신경망(Convolutional Neural Network, CNN)

## ▶ 어느 분야에 사용되는가? Segmentation



## 03 합성곱 신경망(Convolutional Neural Network, CNN)

### ▶ Image Segmentation

(A) 이미지의 영역을 분할하는 것.

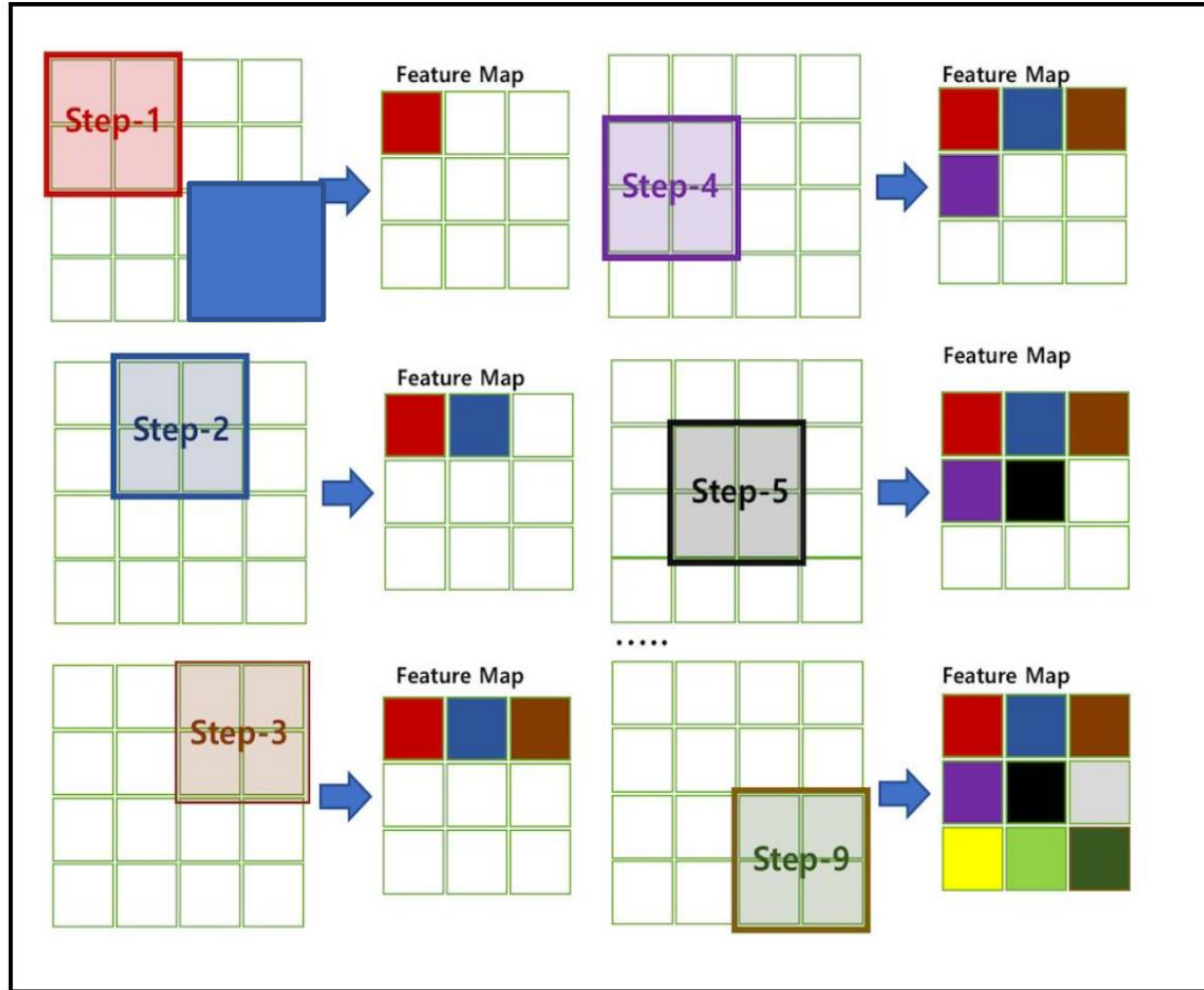
## 03 합성곱 신경망(Convolutional Neural Network, CNN)

### ▶ 장점

- 이미지의 **공간 정보 유지**
- 필터를 공유 파라미터로 사용하여, 일반 인공 신경망과 비교하여 **학습 파라미터가 작다.**



## 04 합성곱 신경망- 특징맵(feature map) 생성과정



- ▶ 윈도우가 이미지 위를 이동하며 특징 맵을 만들어 낸다.
- ▶ 컨볼루션 층에서는 필터와 이미지 합성곱으로 특징맵을 만든다.
- ▶ 풀링 층에서는 정해진 조건을 만족하는 값을 생성해 낸다.

# 05 합성곱 신경망 - 스트라이드(stride)

필터 가중치는 공유

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

1	0
0	1

필터

1	0
1	1

1	0
0	0

8	19	12	14
18	20	22	24
28	30	32	34
38	40	42	44

8	19	12	14
18	20	22	24
28	30	32	34
38	40	42	44

8	19	12	14
18	20	22	24
28	30	32	34
38	40	42	44



풀링

20	24
40	44

20	24
40	44

20	24
40	44

2 x 2
-------

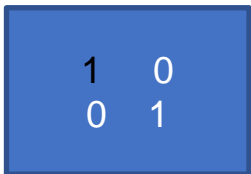


FCL

파라미터

## 05 합성곱 신경망 - 스트라이드(stride)

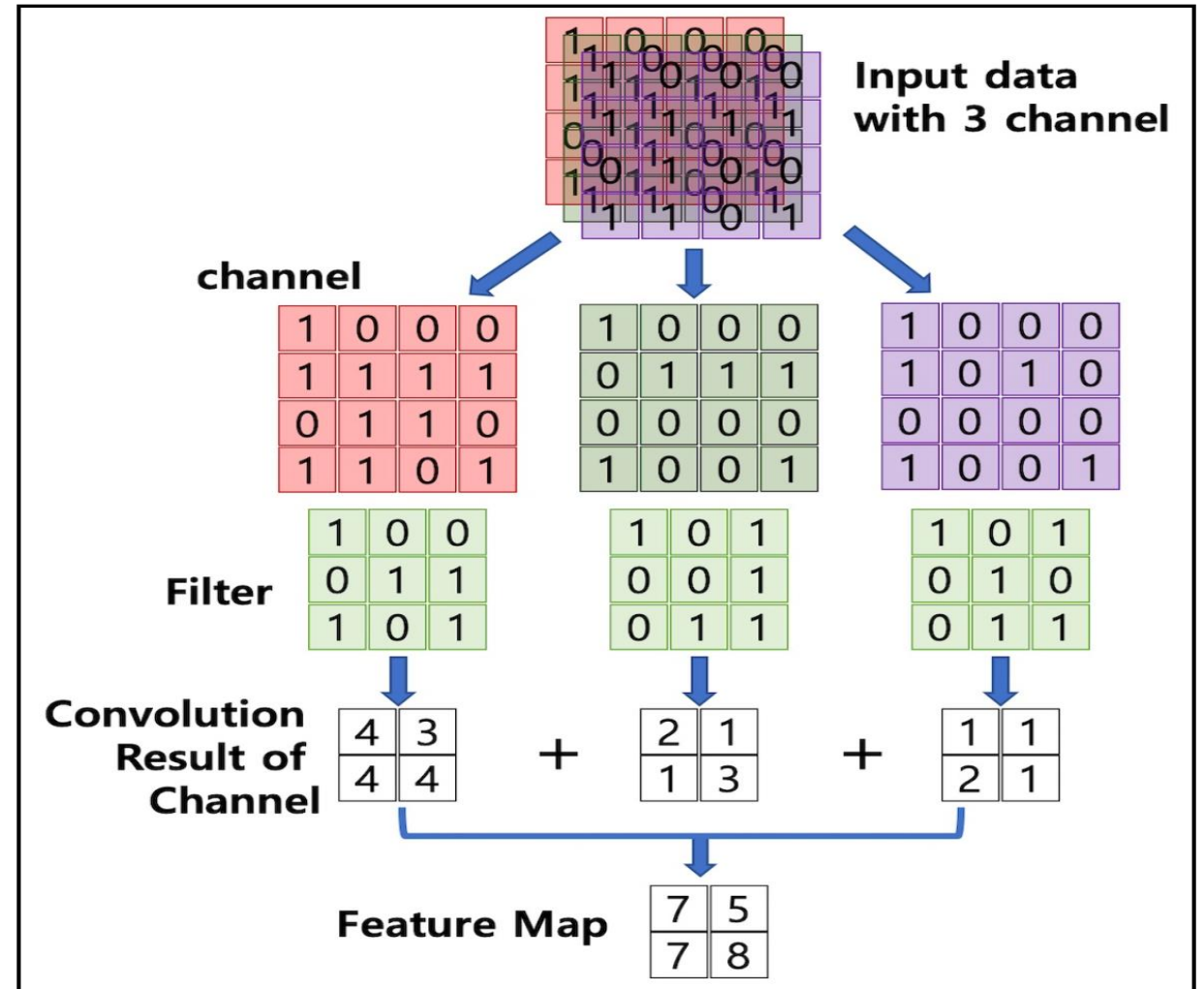
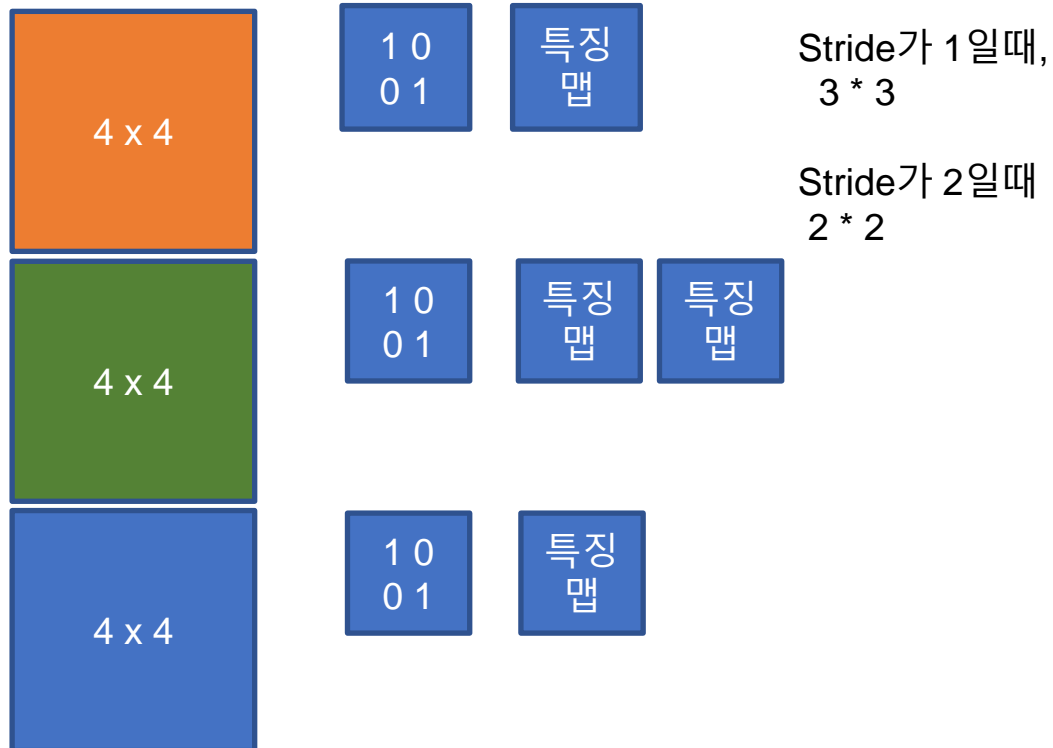
- ▶ 이것들이 반복적으로 이루어져 하나의 필터가 하나의 특성맵을 만들어냅니다.
- ▶ 이때 윈도우가 슬라이딩하는 이동의 크기를 **스트라이드**라고 합니다.
- ▶ 스트라이드는 기본적으로 1입니다. 만약 2를 사용하면 특성 맵의 높이와 너비가 2의 배수로 **다운샘플링 된 것을 의미**합니다.



7	9	11
15	17	19
23	25	27

# 05 합성곱 신경망 연산

- ▶ 컨볼루션 계층
- ▶ 필터는 채널 수 만큼의 깊이를 갖는다.
- ▶ 각 채널의 가중치는 채널마다 다르다.



<http://taewan.kim/post/cnn/> 그림 참조

## 06 완전연결층(FCL)과 합성곱 연산층의 차이

▶ 완전 연결층(FCL)과 합성곱층 사이의 차이는 다음과 같습니다.

(1) 완전 연결층의 Dense 층은 입력 특성 공간에 있는 **모든 픽셀에 대한 패턴학습**

(2) 합성곱층은 **지역 패턴을** 학습.

## 07 합성곱 신경망

- ▶ 합성곱 연산은 핵심적인 2개의 파라미터로 정의
  - (1) 입력으로부터 하나의 합성곱 연산을 할 필터 :  $3 \times 3$ ,  $5 \times 5$
  - (2) 합성곱으로 계산할 필터의 수
- ▶ 합성곱 연산은 필터의 수만큼 특성 맵을 만들어냅니다.

## 08 풀링(Pooling) 연산

입력 특성 맵에서 윈도우에서 조건에 맞는 하나의 값을 추출한다.

### (1) 최대 풀링 연산(Max Pooling)

해당 윈도우에서 가장 최대의 값을 추출한다. 2x2라면 해당 영역안에서 가장 높은 값을 취한다.

### (2) 평균 풀링 연산(Average Pooling)

해당 윈도우에서 평균값을 추출한다.

### (3) 최소 풀링 연산(Min Pooling)

해당 윈도우에서 가장 최소의 값을 추출한다. 2x2라면 해당 영역안에서 가장 낮은 값을 취한다.

## 08 풀링(Pooling) 연산

- ▶ 컴퓨터 비전, 이미지 인식 분야에서는 주로 Max-Pooling을 사용한다.