

# 딥러닝 시작하기 (CNN 입문)

# 목차

- ▶ 01 기본 이해
- ▶ 02 CNN의 등장 배경
- ▶ 03 CNN 활용 분야
- ▶ 04 합성곱 계층(Convolutional Layer)
- ▶ 05 Pooling Layer(풀링 연산)
- ▶ 06 합성곱 연산층과 완전연결층(FCL)의 차이
- ▶ 07 합성곱 신경망 요약
- ▶ 08 사전 훈련 네트워크
- ▶ 09 과대 적합을 줄이는 방법

# 01 기본 이해

## ▶ CNN이란 무엇일까요?

- Convolutional Neural Network의 약자
- 딥러닝 모델의 한 종류 이미지 인식, 물체 탐지, 음성 인식 등의 작업에 활용된다.
- CNN은 합성곱 층, 풀링 층, 완전 연결 층으로 이루어져 있다.
- CNN의 핵심 아이디어는 입력 데이터에서 지역적 연관성(local connectivity)와 공간 계층 구조(spatial hierarchy)를 잘 포착하는 것이다.

# 01 기본 이해

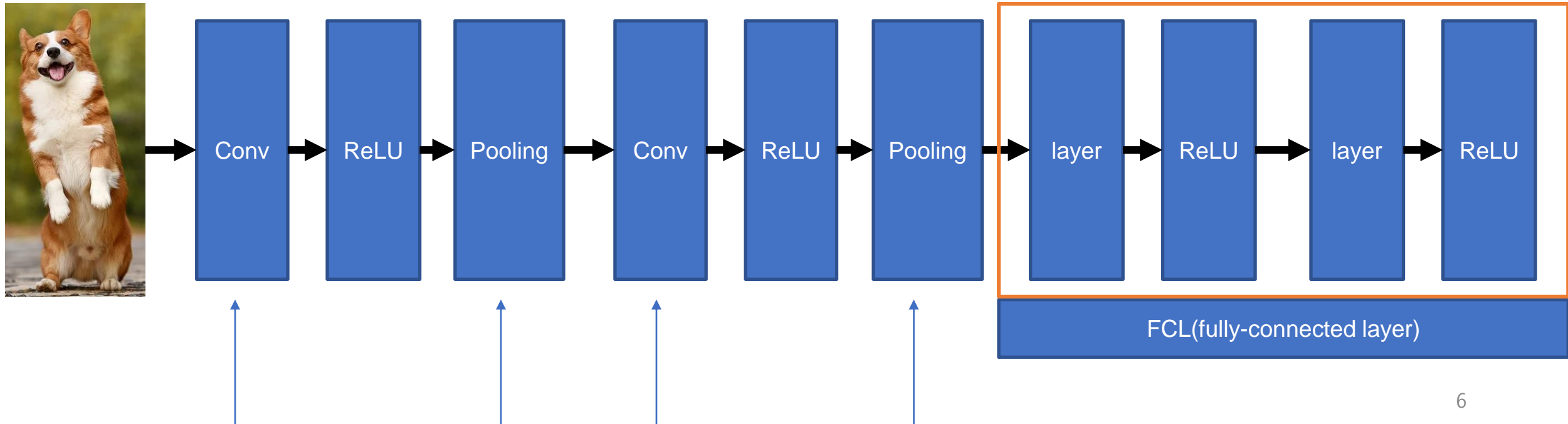
## ▶ CNN이란 무엇일까요?

- ▶ CNN은 입력 데이터에 대해 필터(filter)연산을 수행하여 특징을 추출한다. 이 과정에서 합성곱(convolution)연산이 사용된다.
- ▶ 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 활용되고 있음.
- ▶ 이미지를 인식하기 위한 패턴을 찾는데 특히 유용
  - => 이미지의 공간 정보를 유지한 상태로 학습이 가능한 모델(CNN)

# 01 기본 이해

- LeCun에 의해 처음 소개됨

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



# 01 기본 이해 CNN

## ▶ 장점

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

# 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 - 용어

## ▶ 스트라이드, stride

- 필터를 적용하는 간격의 크기 stride라 함.

## ▶ 특징맵(feature map, activation map)

- 필터를 적용해서 얻어진 결과를 말함.



## 02 CNN의 등장 배경 - FCL의 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 CNN - 활용 분야

### ▶ Classification

(A) 분류는 주어진 이미지가 어떤 클래스에 속하는지를 결정하는 작업.

### ▶ Localization

(A) 이미지 내에서 특정 객체의 위치를 식별하는 작업. CNN은 객체의 경계 상자를 예측하여 이미지의 객체 위치를 표시.

### ▶ Object Detection

(A) 이미지 내에서 여러 객체를 동시에 탐지. 각 객체의 클래스와 위치를 식별하는 작업.

(예) 자율주행차에서 보행자, 신호등, 도로 표지판 등을 탐지

## 03 CNN - 활용 분야

### ▶ Object segmentation

(A) 이미지의 각 픽셀을 특정 클래스에 할당하는 작업으로 객체의 경계를 정확하게 식별.

(예) 의료 이미지에서 종양의 경계를 식별, 자율 주행차에서 도로와 보행자를 구분하는 작업

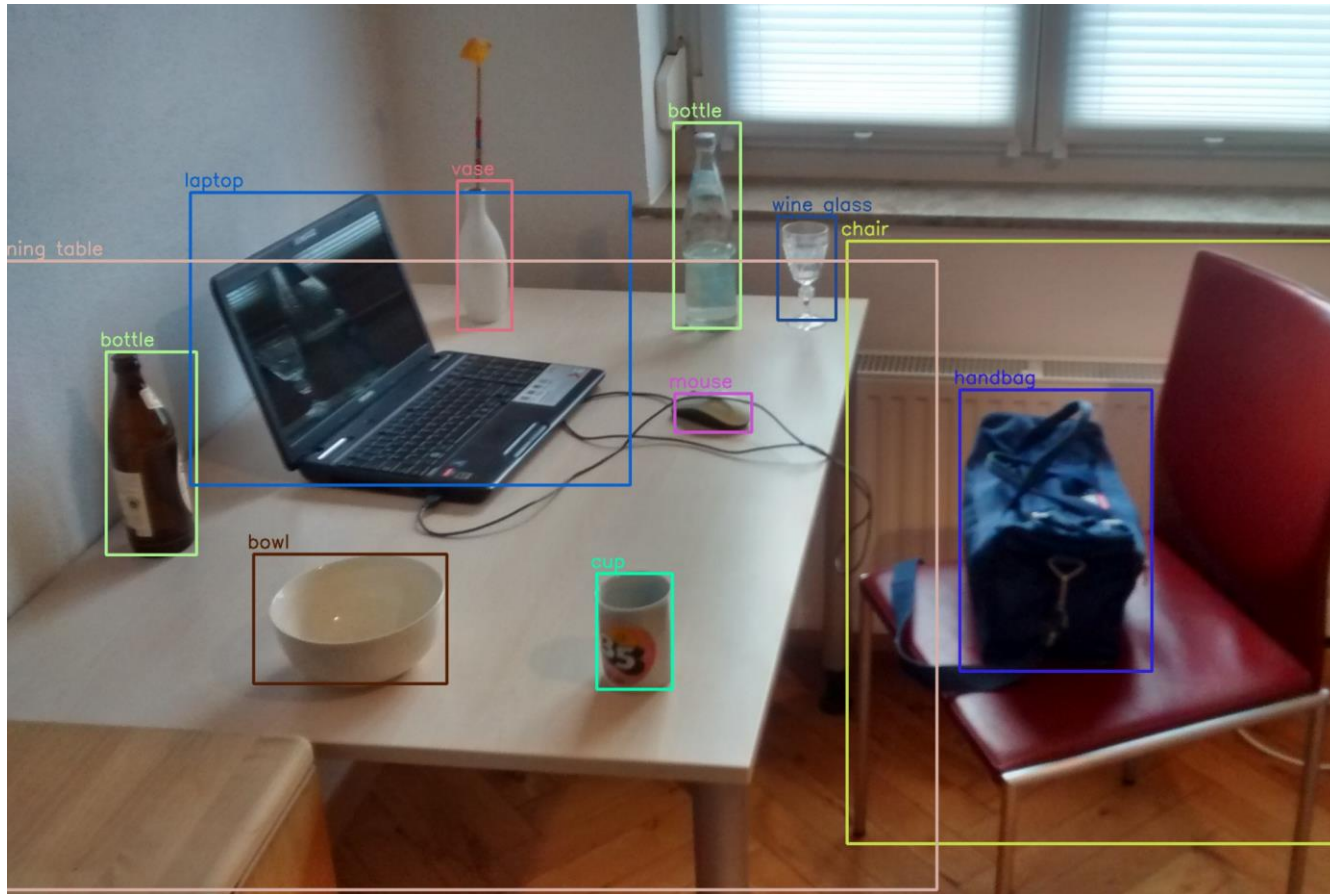
### ▶ 생성 AI

(A) 생성 AI는 새로운 데이터를 생성하는 모델을 의미. CNN은 이미지 생성 및 변환 작업에 중요한 역할.

(예) 스타일 전이, 이미지 복원, 새로운 이미지 생성 GAN 모델

## 03 CNN - 활용 분야

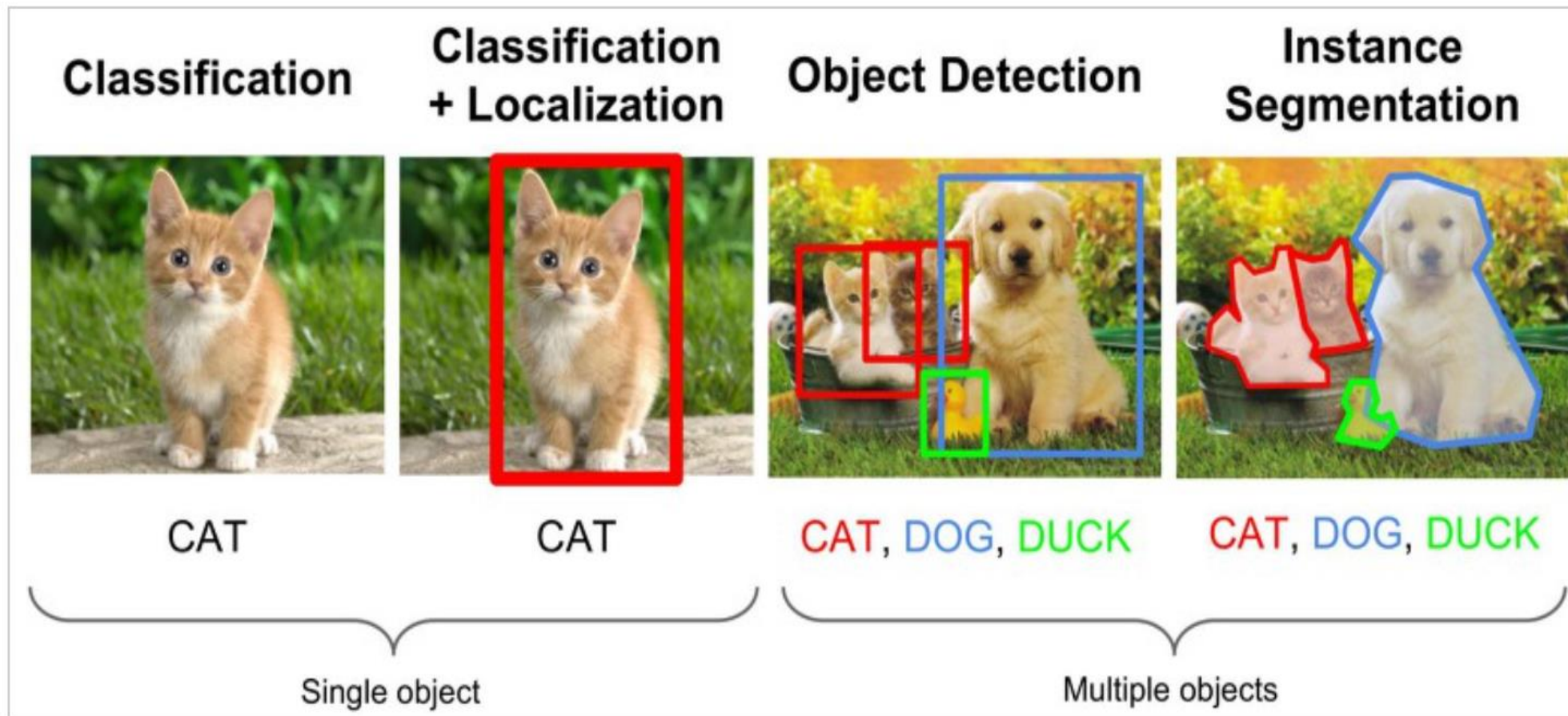
### ▶ 어느 분야에 사용되는가? Object Detection



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

## 03 CNN - 활용 분야

### ▶ 어느 분야에 사용되는가? Segmentation



## 03 CNN - 활용 분야

### ▶ CNN이 활용되는 모델

- Inception(GoogLeNet) : 2014년 구글에서 개발된 모델. 여러 크기의 필터를 동시 적용하여 다양한 스케일의 특징을 추출.
- VGGNet : 깊은 네트워크 구조를 가진 모델로, 3x3 필터를 사용하여 더 깊은 층을 쌓아 특징을 추출. VGG16, VGG19 (2014년)
- ResNet(Residual Network) : 2015년 MS 개발. 잔차 학습 개념을 도입하여 매우 깊은 네트워크를 효과적으로 학습할 수 있게 만듦. 이미지 분류(ResNeXt)
- DenseNet(Dense Convolutional Network) : 2017년

## 03 CNN - 활용 분야

### ▶ CNN이 활용되는 모델

- YOLO(You Only Look Once) : 실시간 물체 탐지를 위한 단일 신경망 모델. 많이 사용됨(2016년)
- Mask R-CNN : 물체 탐지 뿐만 아닌 인스턴스 분할(instance segmentation)도 수행(2017년)
- EfficientNet : EfficientNet은 모델 크기와 성능 간의 균형을 최적화하기 위해 "Compound Scaling" 방법을 사용합니다. 이 방법은 네트워크의 깊이, 너비, 해상도를 동시에 조정하여 효율성을 극대화. 2019년도
- Vision Transformers (ViT) : 이미지 처리에 Transformer 아키텍처를 적용한 모델로, 2020년에 발표
- 이미지 생성 AI인 GAN에도 동일하게 CNN이 사용될 경우도 많다.

## 04 합성곱 계층(Convolutional Layer)

▶ 합성곱 연산은 핵심적인 2개의 파라미터로 정의

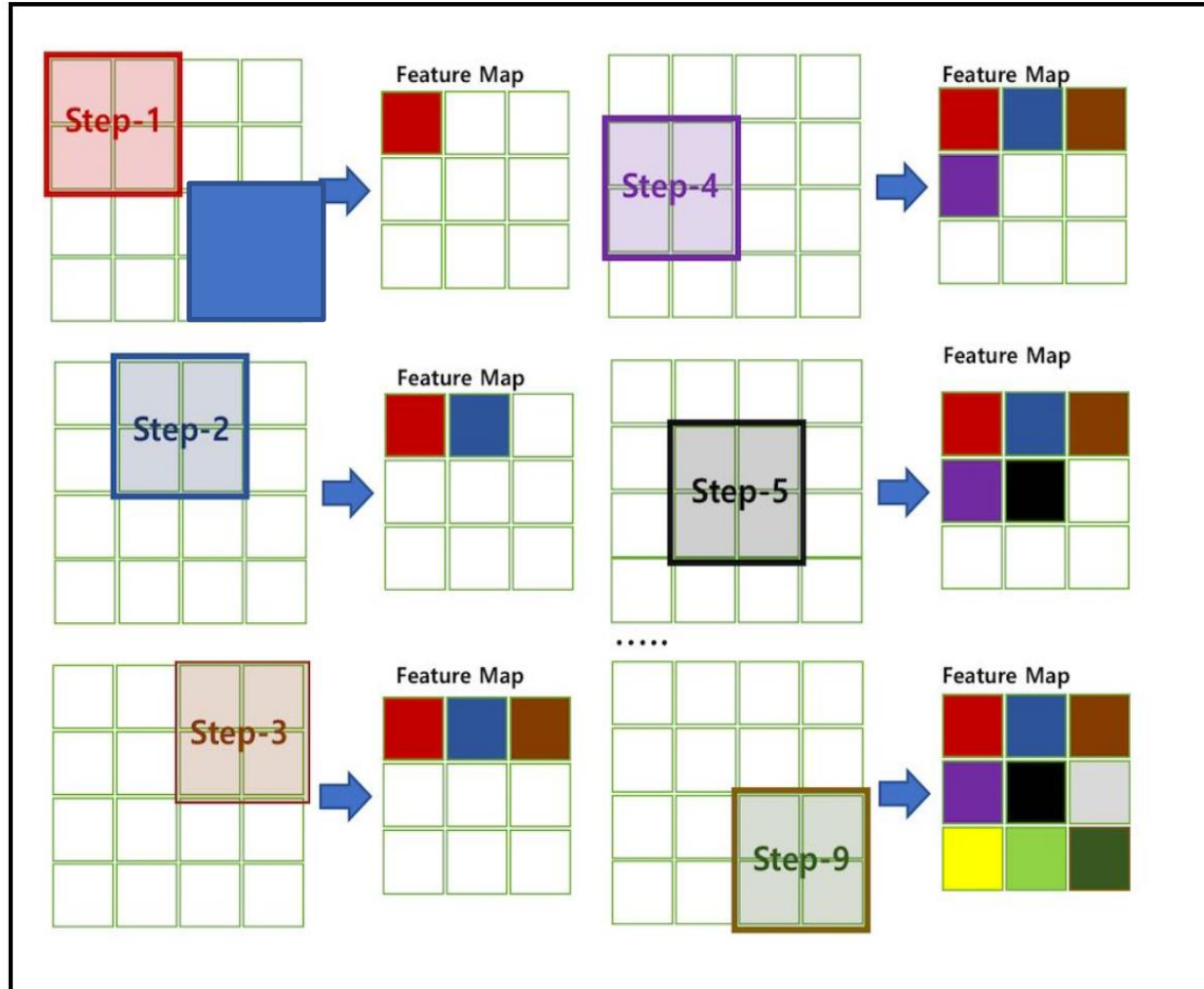
(1) 입력으로부터 하나의 합성곱 연산을 할 필터 :  $3 \times 3$ ,  $5 \times 5$

(2) 합성곱으로 계산할 필터의 수

▶ 합성곱 연산은 필터의 수만큼 특성 맵을 만들어냅니다.



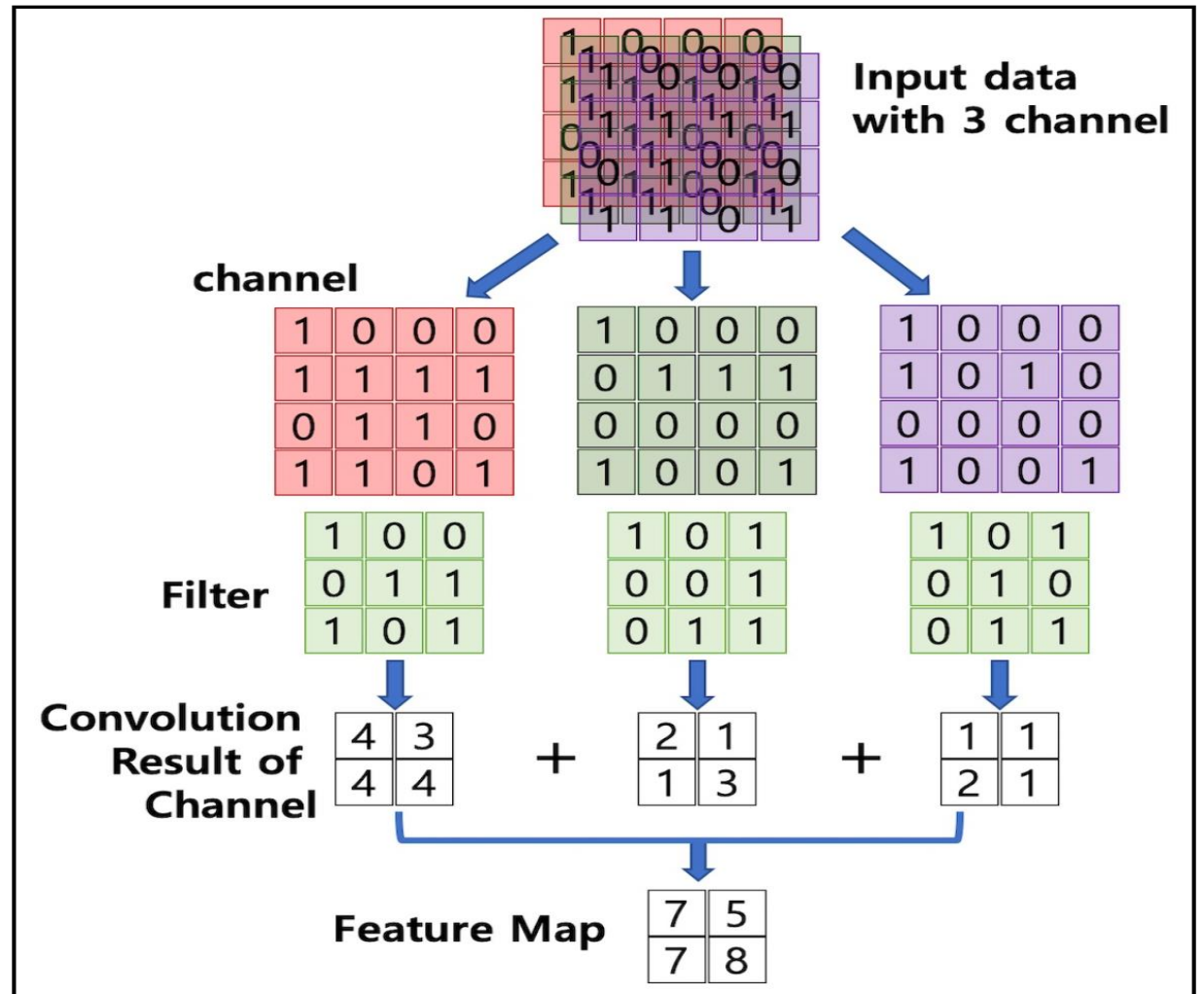
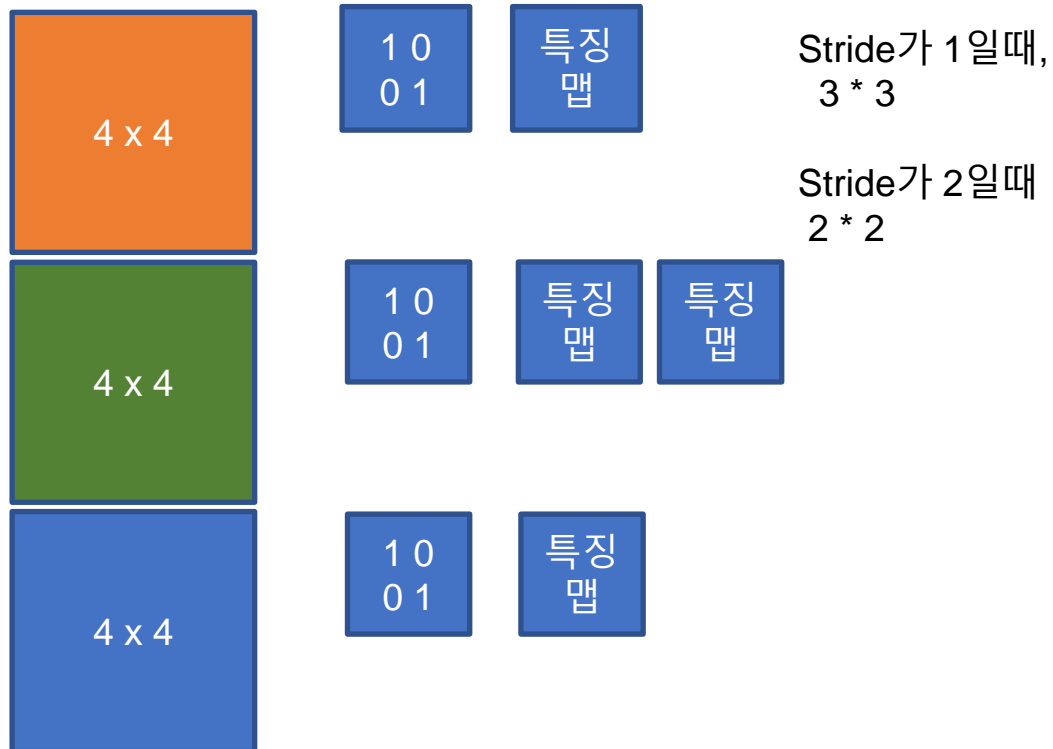
# 04 합성곱 계층(Convolutional Layer) - 특징맵 생성과정



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

# 04 합성곱 계층(Convolutional Layer) - Conv연산

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



<http://taewan.kim/post/cnn/> 그림 참조

## 04 합성곱 계층(Convolutional Layer) - 패딩

- ▶ 특성맵 이미지의 크기를 유지하기 위해 고안된 방법

- ▶ Convolution Filter를 통과하면 Input 이미지가 작아진다. 단, Padding을 이용하여 그대로 유지가 가능하다.

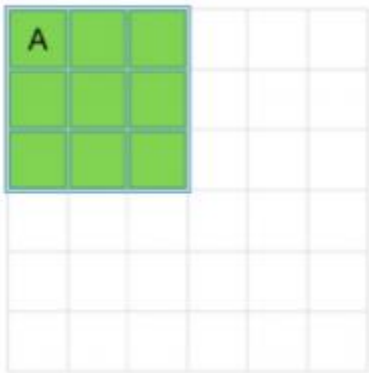
- ▶ Edge쪽 픽셀 정보를 잘 이용하기 위한 방법

(1) 컨볼루션 레이어를 적용할 때, 이미지 주변의 픽셀이 많이 사용되지 않아, 손실되는 경향이 있다.

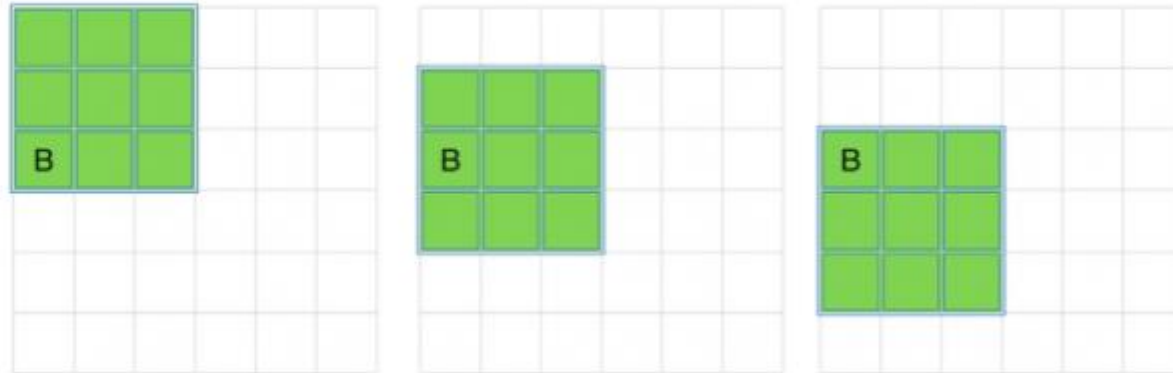
## 04 합성곱 계층(Convolutional Layer) - 패딩

- ▶ 컨볼루션 레이어를 적용할 때, 이미지 주변의 픽셀이 많이 사용되지 않아, 손실되는 경향이 있다.

Corner Pixel



Edge Pixel



## 04 합성곱 계층(Convolutional Layer) - 패딩

- ▶ 컨볼루션 레이어를 적용할 때, 이미지 주변의 픽셀이 많이 사용되지 않아, 손실되는 경향이 있다.



## 04 합성곱 계층(Convolutional Layer) - 패딩

- ▶ 일반적으로 3 x 3의 필터는 zero pad 1
- ▶ 일반적으로 5 x 5의 필터는 zero pad 2
- ▶ 일반적으로 7 x 7의 필터는 zero pad 3
- ▶ Padding 지정
  - (1) Valid Padding : padding을 하지 않음.
  - (2) Same Padding : output image가 input image와 크기가 동일

# 04 합성곱 계층(Convolutional Layer) – 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

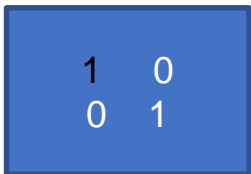


FCL

파라미터

## 04 합성곱 계층(Convolutional Layer) – 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 풀링(Pooling) 계층 - 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.

## 05 풀링(Pooling) 계층

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

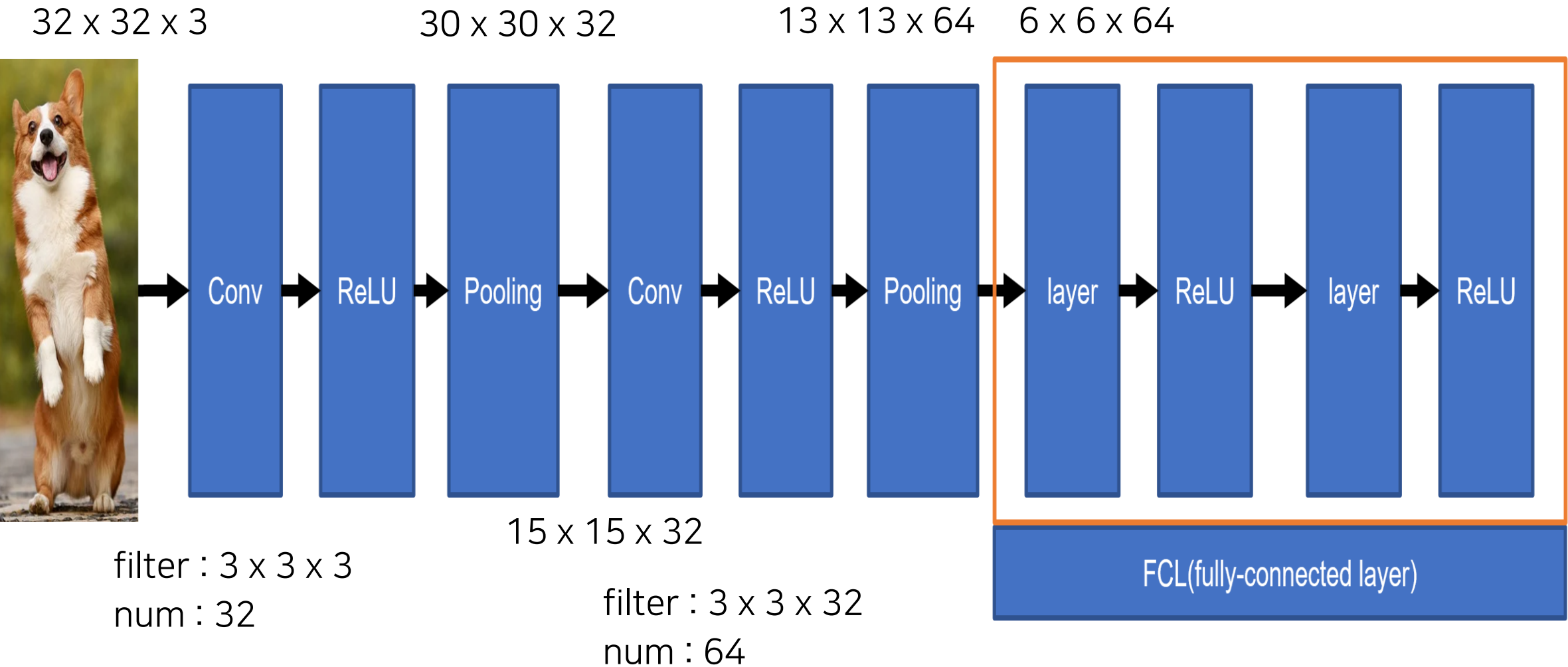
## 06 합성곱 연산층과 완전연결층(FCL)의 차이

▶ 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(Convolutional Neural Network, CNN) 요약



Pooling filter 2x2 이용

## 08 사전 훈련된 네트워크 - 알아보기

- (1) 작은 이미지 데이터 셋에서 딥러닝을 적용하는 일반적이고 매우 효과적인 방법.
- (2) 사전 훈련된 네트워크(pretrained network)는 일반적으로 대규모 이미지 분류 문제를 위해 대량의 데이터셋에서 미리 훈련되어 저장된 네트워크
- (3) 1400만개의 레이블된 이미지와 1000개의 클래스로 이루어진 데이터셋(ImageNet)

## 08 사전 훈련된 네트워크 - 종류

(1) VGG

(2) ResNet

(3) Inception-ResNet

(4) Xception

## 08 사전 훈련된 네트워크 - 사용하는 두가지 방법

### (1) 특성 추출(feature extraction)

사전에 학습된 네트워크의 표현을 사용하여 새로운 샘플에서 흥미로운 특성을 뽑아낸다.  
즉, 합성곱 기반층은 사전 훈련 네트워크 가중치를 이용하고 분류기 부분을 훈련 시킨다.

### (2) 미세 조정(fine tuning)

미세 조정은 특성 추출에 사용했던 동결 모델의 상위 층 몇 개를 동결에서 해제하고 모델에 새로 추가한 층과 함께 훈련하는 것.

## 09 과대 적합을 줄이는 방법 - 데이터 증식

- (1) 컴퓨터 비전에서 과대 적합을 줄이기 위한 강력한 방법.
- (2) 데이터 증식은 기존 훈련 샘플로부터 더 많은 훈련 데이터를 생성하는 방법입니다.
- (3) 그럴듯한 이미지를 생성하도록 여러가지 랜덤 변환(각도, 좌우변환 등)을 적용하여 샘플을 늘린다.
- (4) ImageDataGenerator 등의 클래스를 활용.