

## Course 11

# 컨볼루셔널 신경망

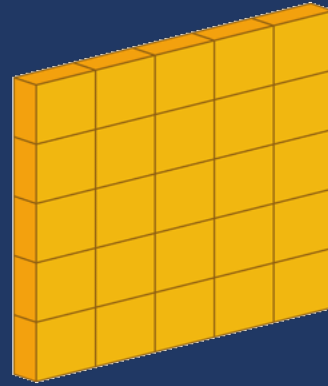
공간을 지배하는 방법

컨볼루션 신경망은 딥러닝 알고리즘의 한 종류로  
이미지, 비디오, 텍스트 또는 음향 분류에 많이 사용되는 알고리즘입니다.





CNN은 데이터에서 **직접 특성을 추출하는 방법을 학습**하여 추출된 특성을 사용하여 이미지를 분류하기 때문에 **특성을 수동으로 추출할 필요가 없습니다.**



# 컨볼루션(Convolution)

데이터의 각 성분의 인접 성분들을 조사해 **특성을 파악**하고 파악한 **특성을 한 장으로 도출**시키는 과정

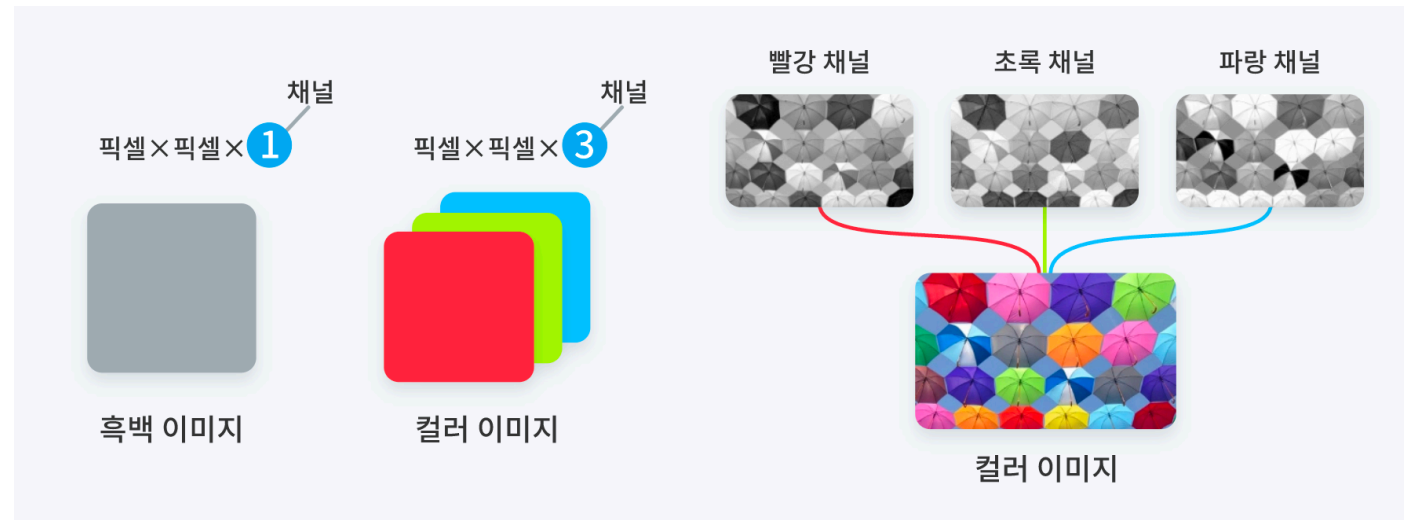
컨볼루션을 통해 도출된 장 = '컨볼루셔널 레이어(Convolutional Layer)'

# 채널

Channel

색상에 대해 숫자로 표현한 정보

디지털 컬러 이미지는 주로 Red, Green, Blue, RGB 3개의 숫자로 표현한 3차원 데이터



# 필터

Filter

추출하려는 이미지의 특성이 대상 데이터에 있는지 없는지를 검출해 주는 함수

커널(Kernal)이라고도 지칭

주로 (4,4) 혹은 (3,3)과 같은 정사각형 행렬로 정의



원본 이미지



시각화한 이미지 상의 필터



# 스트라이드/특성 맵

Stride/Feature Map

필터를 큰 이미지에 적용하기 위한 기법으로, 필터를 좌측 상단에서부터 한 칸씩 이동하며 특성을 추출

필터가 위치하는 원 이미지의 픽셀값이 곱해진 값을 추출하며,  
이를 특성 맵(Feature Map) 또는 활성화 맵(Activation Map)이라고 지칭

스트라이드(Stride)는 필터를 적용하는 간격을 의미

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

Input

×1	×0	×1
×0	×1	×0
×1	×0	×1

3×3 Window

2	2	3	1	1
1	4	3	4	1
2	2	4	3	3
1	2	3	4	1
1	2	3	1	1

Output

1 <sub>x1</sub>	1 <sub>x0</sub>	1 <sub>x1</sub>	0	0
0 <sub>x0</sub>	1 <sub>x1</sub>	1 <sub>x0</sub>	1	0
0 <sub>x1</sub>	0 <sub>x0</sub>	1 <sub>x1</sub>	1	1
0	1	1	1	0
0	1	1	0	0

4		

# 패딩

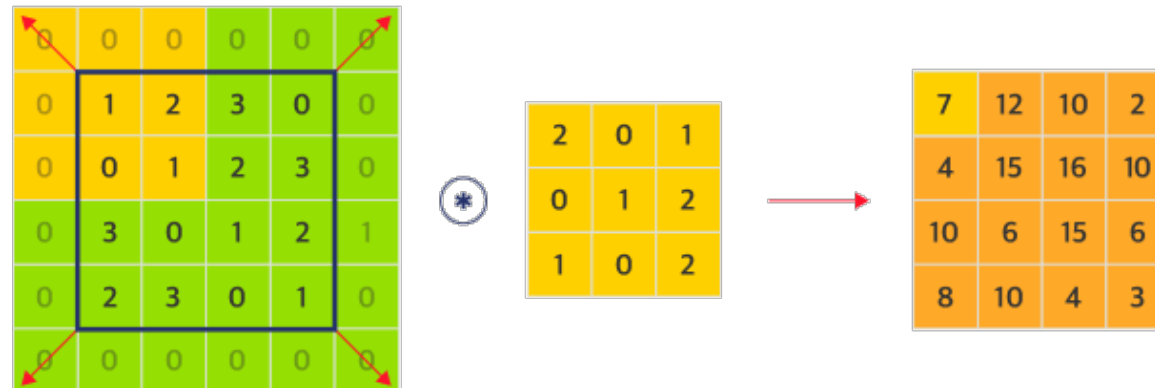
Padding

컨볼루션 레이어의 **출력 데이터가 줄어드는 것을 방지**하는 방법

**입력 데이터의 외곽**에 지정된 픽셀만큼 **특정 값으로 채워넣는 것**을 의미(일반적으로는 '0')

컨볼루션 레이어를 지날 때마다 데이터의 크기가 줄어들어 **가장자리**의 정보가 **유실**되는 문제를 방지

인공신경망이 **이미지의 외곽**을 인식하는 효과도 있음





스트라이드가 1인 아래의 필터로  
특성 맵을 완성해 보세요.

1	-1	-1	-1	1
-1	1	-1	1	-1
-1	-1	1	-1	-1
-1	1	-1	1	-1
1	-1	-1	-1	1

입력

×

1	-1	-1
-1	1	-1
-1	-1	1

$$\{(1*1)+(-1*-1)+(-1*-1)+(-1*-1)+(1*1)+(-1*-1)+(-1*-1)+(-1*-1)+(1*1)\} = 9$$

필터

=

9	-3	1
-3		

스트라이드가 1인 아래의 필터로  
특성 맵을 완성해 보세요.

1	-1	-1	-1	1
-1	1	-1	1	-1
-1	-1	1	-1	-1
-1	1	-1	1	-1
1	-1	-1	-1	1

입력

×

1	-1	-1
-1	1	-1
-1	-1	1

$\{(1*1)+(-1*-1)+(-1*-1)+$   
 $(-1*-1)+(1*1)+(-1*-1)+$   
 $(-1*-1)+(-1*-1)+(1*1)\} = 9$

필터

=

9	-3	1
-3	5	-3
1	-3	9

이번에는 1 픽셀의 패딩을 추가하여,  
스트라이드가 1인 아래의 필터로  
특성 맵을 완성해 보세요.

0	0	0	0	0	0	0
0	1	-1	-1	-1	1	0
0	-1	1	-1	1	-1	0
0	-1	-1	1	-1	-1	0
0	-1	1	-1	1	-1	0
0	1	-1	-1	-1	1	0
0	0	0	0	0	0	0

입력

×

1	-1	-1
-1	1	-1
-1	-1	1

필터

=

4	-2	2	-2	2
-2	9	-3	1	-2
2	-3	5		

$\{(0*1)+(0*-1)+(0*-1)+$   
 $(0*-1)+(1*1)+(-1*-1)+$   
 $(0*-1)+(-1*-1)+(1*1)\} = 4$

이번에는 1 픽셀의 패딩을 추가하여,  
스트라이드가 1인 아래의 필터로  
특성 맵을 완성해 보세요.

0	0	0	0	0	0	0
0	1	-1	-1	-1	1	0
0	-1	1	-1	1	-1	0
0	-1	-1	1	-1	-1	0
0	-1	1	-1	1	-1	0
0	1	-1	-1	-1	1	0
0	0	0	0	0	0	0

입력

×

1	-1	-1
-1	1	-1
-1	-1	1

필터

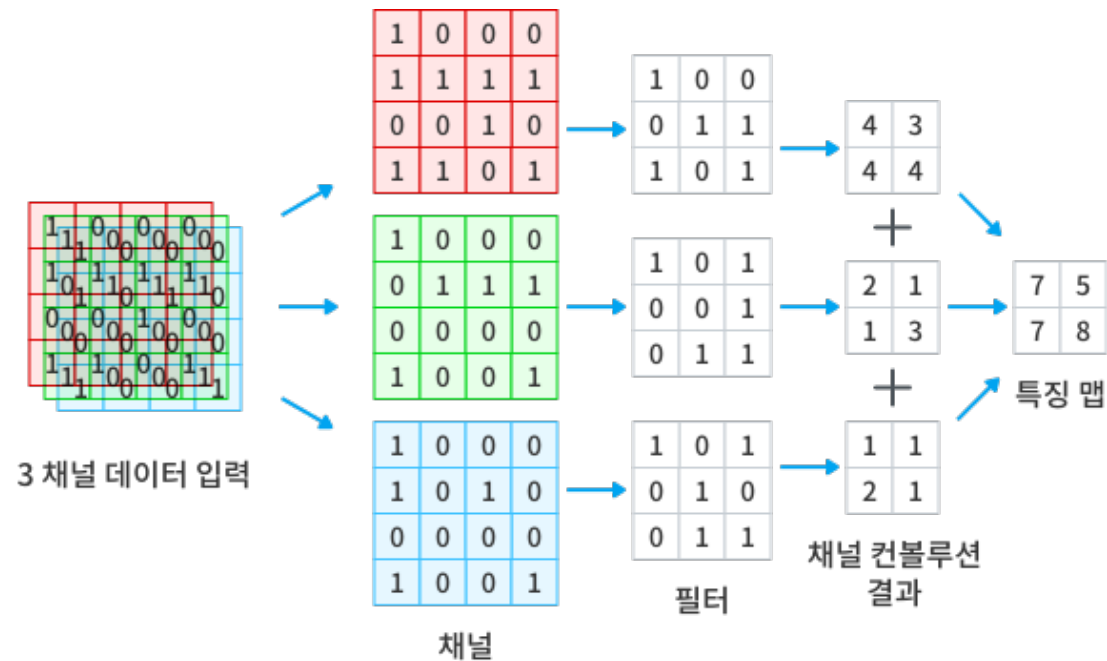
=

4	-2	2	-2	2
-2	9	-3	1	-2
2	-3	5	-3	2
-2	1	-3	9	-2
2	-2	2	-2	4

출력

$\{(0*1)+(0*-1)+(0*-1)+(0*-1)+(1*1)+(-1*-1)+(0*-1)+(-1*-1)+(1*1)\} = 4$

입력 데이터가 **여러 채널을 가질 경우** 필터는 **각 채널을 순회**하며  
 컨볼루션을 계산한 후, **채널별 특성 맵**을 만듭니다.  
 그리고 각 **채널의 특성 맵을 합산**하여 **최종 특성 맵으로 반환**합니다.



# 풀링

Pooling

컨볼루션 과정을 거친 레이어의 크기를 줄여주는 과정

데이터의 크기를 줄여주고, 비정상적인 특성을 상쇄시키고 미세한 차이가 있어도 일관적인 특성을 제공

**맥스 풀링:** 특성 맵을  $M \times N$  크기로 잘라낸 후 그 안에서 가장 큰 값을 뽑아내는 방법



**맥스 풀링** 외에도 평균 풀링(Average Pooling), L2-norm 풀링 등이 있습니다.  
**맥스 풀링**은 전체 데이터의 크기가 줄어들기 때문에 **연산량이 적어지고**,  
 데이터 크기를 줄이면서 소실이 발생하기 때문에 **과적합을 방지합니다**.



Window 크기 2, 스트라이드 2로  
맥스 풀링을 해보세요.

0.77	-0.11	0.11	0.33	0.55	-0.11
-0.11	1.00	-0.11	0.33	-0.11	0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11
0.33	0.33	-0.33	0.55	-0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00




1.00		



Window 크기 2, 스트라이드 2로  
맥스 풀링을 해보세요.

0.77	-0.11	0.11	0.33	0.55	-0.11
-0.11	1.00	-0.11	0.33	-0.11	0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11
0.33	0.33	-0.33	0.55	-0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00



1.00	0.33	0.55
0.33	1.00	0.33
0.55	0.33	1.00

# 풀링에 대한 설명으로 올바른 것을 모두 고르세요.

레이어와 특성이 많아짐으로써 생기는 과적합(Overfitting)을 방지하기 위함이다.

평균 풀링은 최댓값을 추출한다.

맥스 풀링은 큰 값이 다른 특성들을 대표한다는 개념을 기반으로 한다.

전체 데이터의 크기가 줄어들기 때문에 연산에 들어가는 컴퓨터의 자원이 줄어든다.

이미지의 크기를 줄일 뿐만 아니라 특성을 강화시키기도 한다.

# 풀링에 대한 설명으로 올바른 것을 모두 고르세요.

레이어와 특성이 많아짐으로써 생기는 과적합(Overfitting)을 방지하기 위함이다.

평균 풀링은 최댓값을 추출한다.

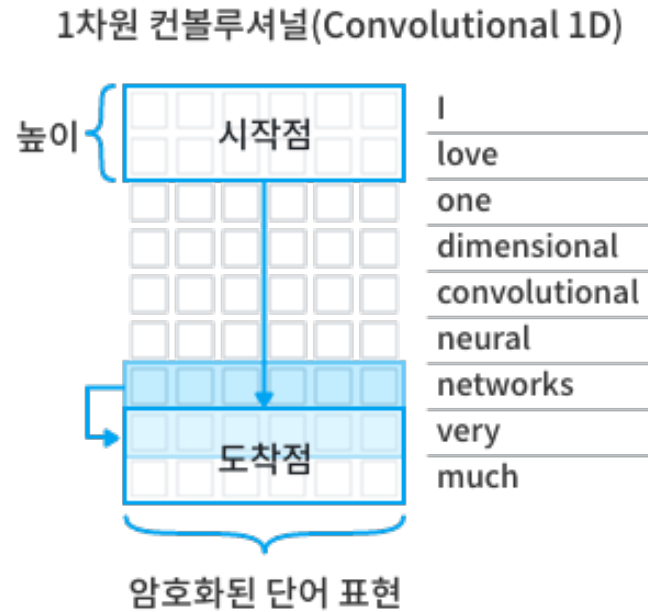
맥스 풀링은 큰 값이 다른 특성들을 대표한다는 개념을 기반으로 한다.

전체 데이터의 크기가 줄어들기 때문에 연산에 들어가는 컴퓨터의 자원이 줄어든다.

이미지의 크기를 줄일 뿐만 아니라 특성을 강화시키기도 한다.

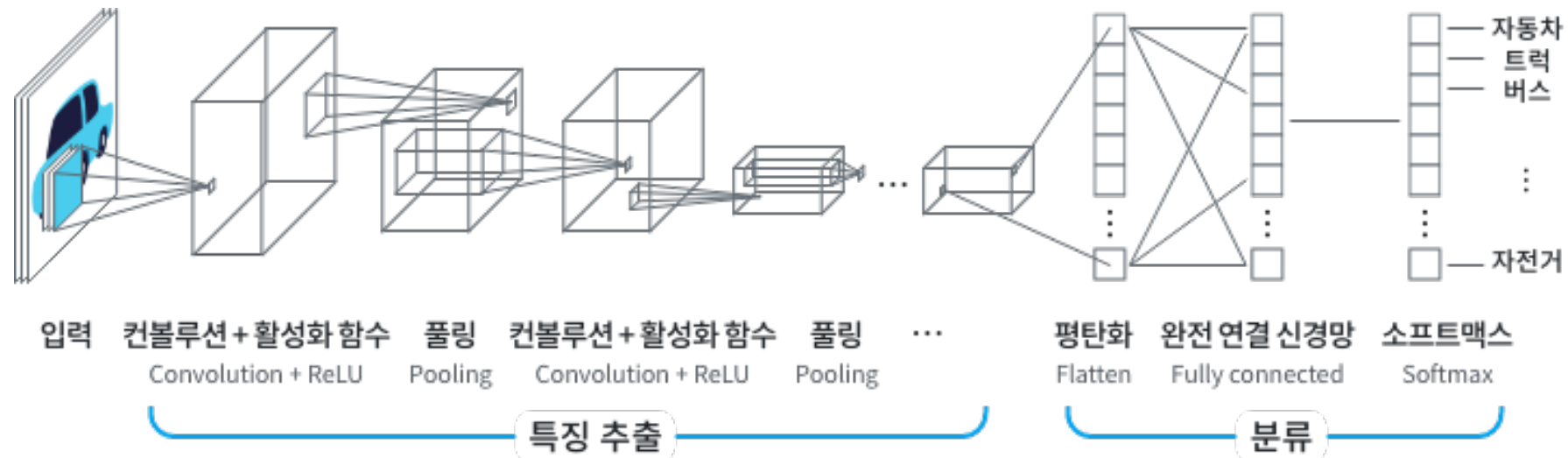


평균 풀링은 평균값을 추출합니다.



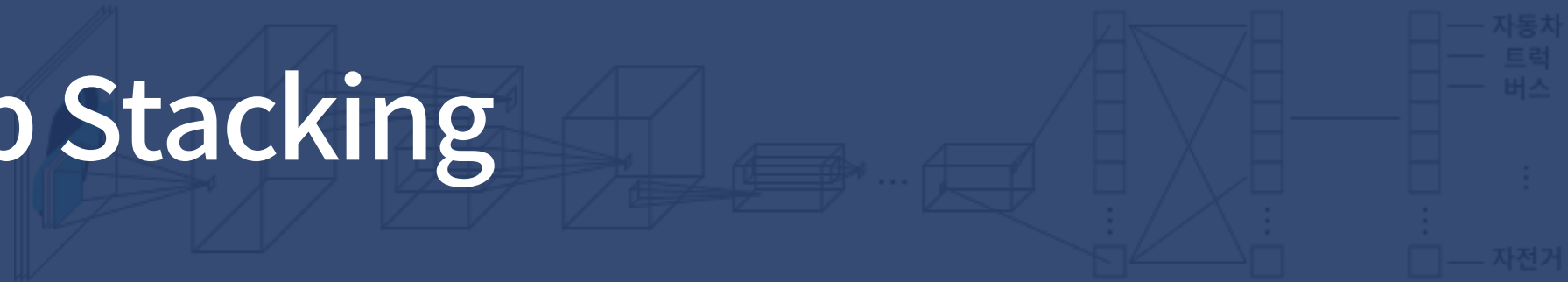
텍스트는 1차원 컨볼루셔널 레이어가 사용됩니다.  
필터의 크기를 변경하여 한 번에 몇 개의 단어를 볼 것인지 지정하고  
스트라이드를 조정하면 문장의 특성을 추출할 수 있습니다.

여러 개의 컨볼루션 레이어로 구성된 신경망에서는  
**필터가 각 학습 이미지에서 다른 해상도로 적용**되고,  
**필터의 출력은 다음 레이어의 입력**으로 사용됩니다.



여러 개의 컨볼루션 레이어로 구성된 신경망에서는  
필터가 각 학습 이미지에 서로 다른 해상도로 적용되고,  
필터의 출력은 다음 레이어의 입력으로 사용됩니다.

# Deep Stacking



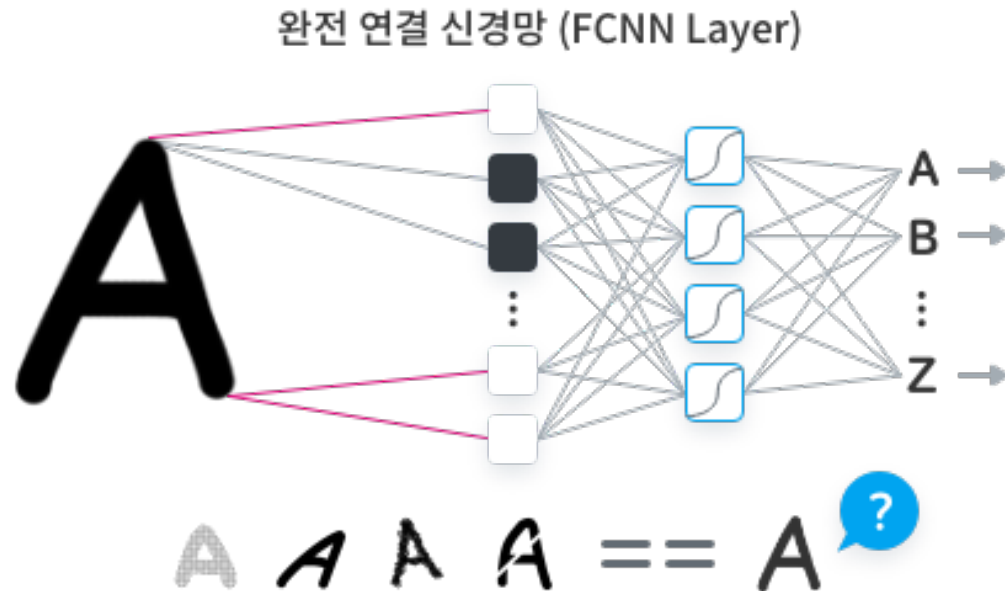
컨볼루션 레이어와 풀링 레이어를 여러 번 반복하여 쌓아 올리면서 효율을 극대화하는 CNN의 일반적인 구조

컨볼루션 레이어와 풀링 레이어는 특성을 추출하고, 마지막 완전 연결 레이어는 데이터가 어떤 클래스에 속하는지 판단하는 역할을 수행

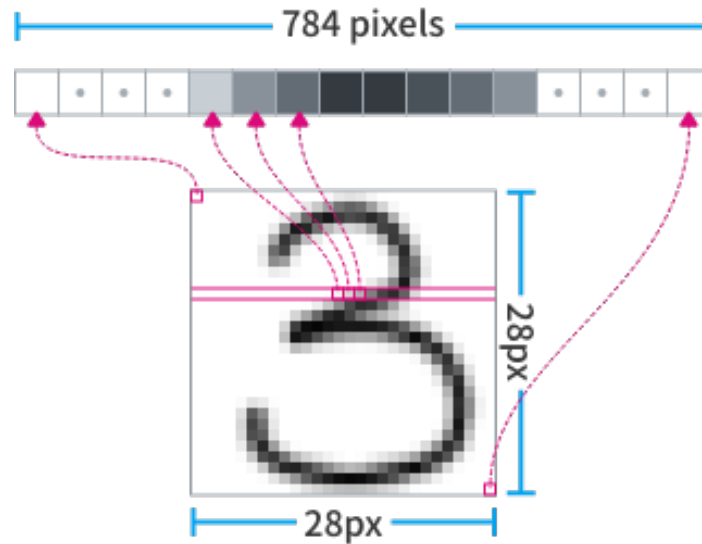
## CNN 활용 사례가 급증한 이유

- CNN은 데이터의 지역 특성(Local Feature, 부분적 특성)을 파악합니다.
- CNN은 데이터의 공간 정보를 유지하면서 특성을 추출하기가 쉽습니다. 그래서 이미지나 텍스트와 같은 데이터에 많이 사용됩니다.
- 기존 신경망을 바탕으로 한 새로운 인식 작업을 위해 **CNN을 재학습하여 사용하는 것이 가능합니다.**
- 완전 연결 신경망(Fully Connected Neural Network, FCNN) 대비 **연산량을 획기적으로 줄일 수 있습니다.**

완전 연결 신경망은 이미지의 형상을 고려하지 않고,  
전체 데이터를 직접 처리하기 때문에  
많은 양의 학습 데이터가 필요하고 학습 시간이 길어집니다.







이미지 데이터의 경우 3차원(세로, 가로, 채널)의 형상을 가지며,  
이 형상은 공간적 구조(Spatial Structure)를 가집니다.  
하지만 완전 연결 레이어에서 1차원의 데이터로 펼쳐게 되면  
이러한 정보들이 사라지게 됩니다.

[ 어떤 분야에 활용할 수 있을까? ]

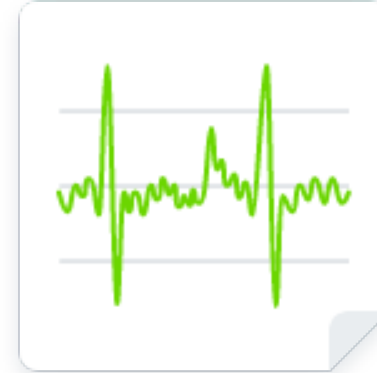
## 우리 주변의 CNN

- 풍경 사진에서 산, 바다, 사람 등의 경계를 인식하기
- 의료 분야 연구에 활용하기 위해 MRI 영상을 CT 영상으로 변환하기
- 컴퓨터 화면을 인식해서 AI 게임 플레이어 만들기

# CNN에 활용하기에 좋은 데이터를 모두 고르세요.

**이미지****텍스트**

동경과 나의 이름과 새겨지는 계십니다. 토끼, 다 아침이 이네들은 부끄러운 쓸쓸함과 듯합니다. 아침이 패, 애기별 따온 거외다.

**신호**

# CNN에 활용하기에 좋은 데이터를 모두 고르세요.

**이미지****텍스트**

동경과 나의 이름과 새겨지는 계십니다. 토끼, 다 아침이 이네들은 부끄러운 쓸쓸함과 듯합니다. 아침이 패, 애기별 뜬 거외다.

**신호**

# CNN에 대한 설명으로 올바른 것을 모두 고르세요.

컨볼루션 레이어에서의 입력값은 완전 연결 레이어와 달리 원형을 보존함

필터의 개수만큼 특징 맵이 생성됨

이미지의 크기를 줄이기 위해 패딩을 할 수 있음

사물 이미지 인식에서만 사용할 수 있음

인접 픽셀간의 상관관계가 무시되는 것이 문제점임

# CNN에 대한 설명으로 올바른 것을 모두 고르세요.

컨볼루션 레이어에서의 입력값은 완전 연결 레이어와 달리 원형을 보존함

필터의 개수만큼 특징 맵이 생성됨

이미지의 크기를 줄이기 위해 패딩을 할 수 있음

사물 이미지 인식에서만 사용할 수 있음

인접 픽셀 간의 상관관계가 무시되는 것이 문제점임



- 3. 패딩은 이미지의 크기 보전을 위함이며, 이미지의 크기를 줄여 연산량을 줄이고 특성을 추출하기 위해서는 풀링을 사용합니다.
- 4. 이미지 인식 분야에 가장 많이 활용되나, 비디오, 텍스트, 사운드 등에서도 활용 가능합니다.
- 5. 이는 완전 연결 신경망(FCNN)의 문제점으로, 이미지를 벡터화하는 과정에서 막대한 정보 손실이 발생합니다.

수고하셨습니다.