
Convolutional Neural Network (CNN)

하 상 천

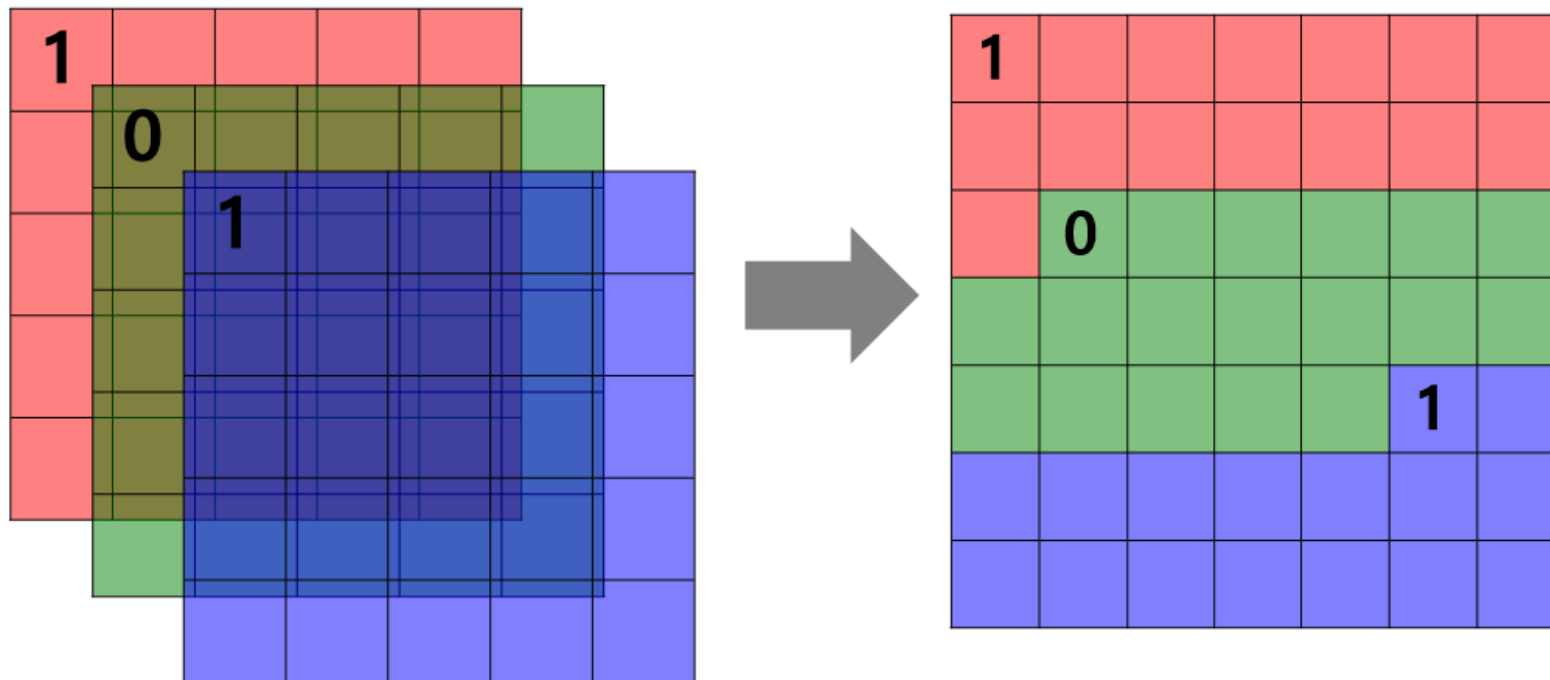
2020.08.17 (MON)

CNN의 등장 배경

▪ 전연결 신경망 (Fully Connected Neural Network, FCNN) 의 한계

Fully Connected Layer만으로 구성된 인공 신경망의 입력 데이터는 1차원 배열 형태로 한정된다
한 장의 컬러 사진은 RGB 값을 갖는 3차원 배열, 배치 모드에 사용되는 여러 장의 사진은 4차원 배열 형태

Image Data로 전연결 신경망을 학습시켜야 할 경우 3차원 데이터를 1차원으로 평면화 시키는 과정이 선행되어야 함
이 과정에서 공간 정보가 손실, 특징 추출과 학습이 비효율적



CNN의 등장배경

▪ 전연결 신경망 (Fully Connected Neural Network, FCNN) 의 한계

이미지의 공간 정보를 유지한 상태로 학습이 가능한 모델의 필요성이 대두되어
합성곱 신경망(Convolutional Neural Network)이 탄생

CNN은 기존의 FCNN에 대하여 다음과 같은 **차별점**을 가진다 :

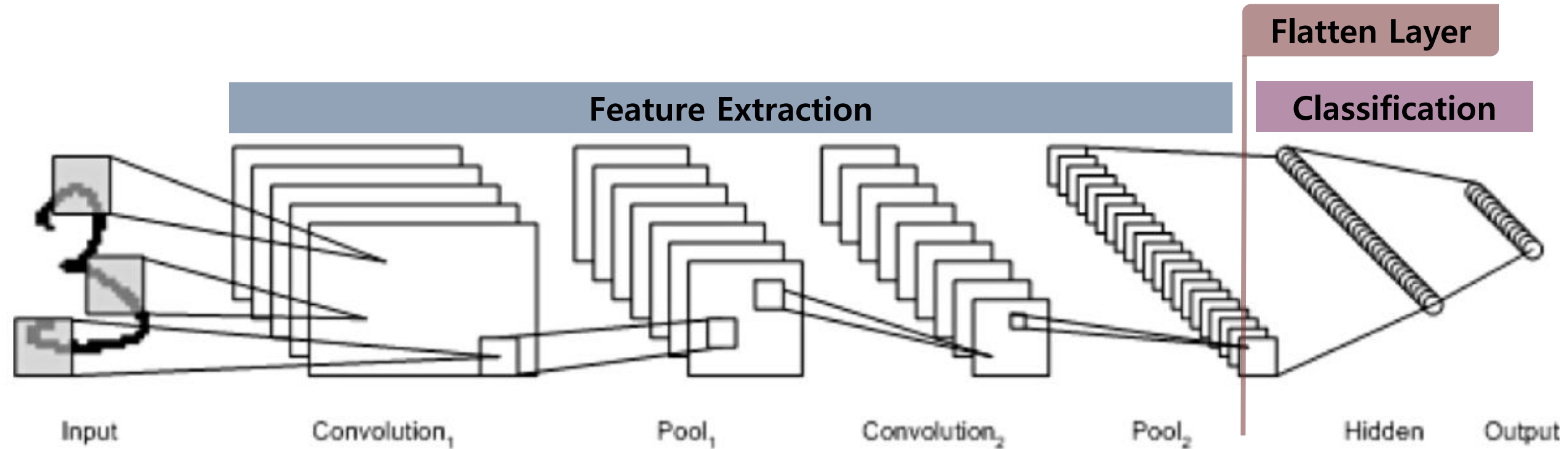
- 각 레이어의 입출력 데이터 형상 유지
- 이미지의 공간 정보를 유지하면서 인접 이미지와의 특징을 효과적으로 인식
- 복수의 필터로 이미지의 특징 추출 및 학습
- 추출한 이미지를 모으고 강화하는 Pooling Layer
- 필터를 공유 파라미터로 사용하므로, 일반 인공 신경망과 비교하여 학습 파라미터가 매우 적음

CNN의 기본 구조

특징 추출(Feature Extraction) 과 분류(Classification) 영역

특징 추출 영역 : Convolution Layer와 Pooling Layer를 겹쳐 쌓는 형태
분류 영역 : FC Layer로 구성

특징 추출 영역과 분류 영역 사이에 이미지 형태의 데이터를 배열 형태로 만드는 Flatten Layer 위치



CNN 용어 정리

- CNN에서 사용되는 주요 용어
 - 컨볼루션 (Convolution), 합성곱
 - 채널 (Channel)
 - 필터 (Filter)
 - 스트라이드 (Stride)
 - 패딩 (Padding)
 - 피쳐 맵 (Feature Map)
 - 풀링 (Pooling) 레이어

CNN 용어 정리

- CNN에서 사용되는 주요 용어
 - 컨볼루션 (Convolution), 합성곱
 - 채널 (Channel)
 - 필터 (Filter)
 - 스트라이드 (Stride)
 - 패딩 (Padding)
 - 피쳐 맵 (Feature Map)
 - 풀링 (Pooling) 레이어

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

Image

4		

Convolved
Feature

CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱

- 채널 (Channel)

- 필터 (Filter)

- 스트라이드 (Stride)

- 패딩 (Padding)

- 피쳐 맵 (Feature Map)

- 풀링 (Pooling) 레이어

- 컨볼루션 레이어의 뉴런은 입력 이미지의 모든 픽셀에 연결되는 것이 아니라 뉴런의 수용영역 (Receptive Field) 안에 있는 픽셀에만 연결

- 초반의 컨볼루션 레이어에서는 저수준 특성에 집중하고, 후반에는 고수준 특성으로 조합하도록 동작

CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱

- 채널 (Channel)

- 필터 (Filter)

- 스트라이드 (Stride)

- 패딩 (Padding)

- 피쳐 맵 (Feature Map)

- 풀링 (Pooling) 레이어

- 합성곱:

하나의 함수와 또 다른 함수를 반전 이동한 값을 곱한 다음,
구간에 대하여 적분하여 새로운 함수를 만드는 것

$$(f * g)(i, j) = \sum_{x=0}^{h-1} \sum_{y=0}^{w-1} f(x, y) g(i - x, j - y)$$

- 매우 유사한 연산으로 교차 상관(Cross-Correlation)이 있음

$$(f * g)(i, j) = \sum_{x=0}^{h-1} \sum_{y=0}^{w-1} f(x, y) g(i \oplus x, j \oplus y)$$

CNN 용어 정리

- CNN에서 사용되는 주요 용어

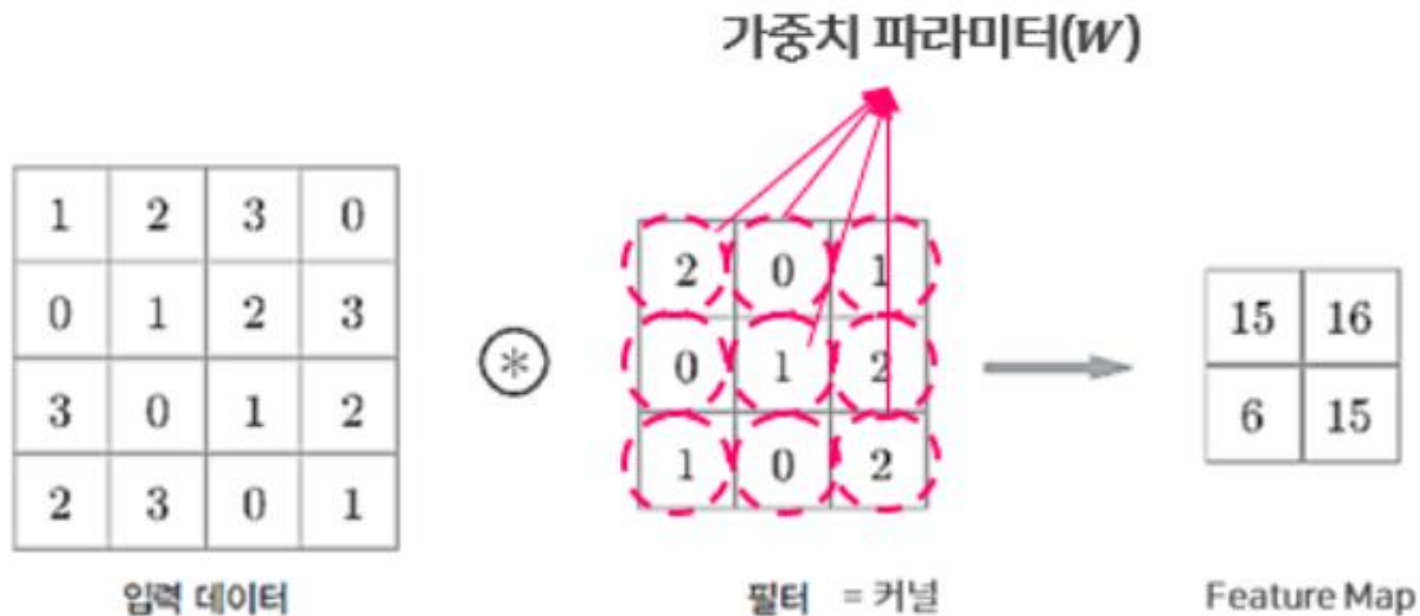
- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 합성곱과 교차 상관의 차이는 한 함수를 반전하는지 여부의 차이 뿐
- CNN에 합성곱을 적용하려면 필터의 값을 반전시켜 연산해야 함
- CNN의 목표는 필터의 값을 학습하는 것이므로 학습 단계와 추론 단계에서 일관된 필터 값을 사용한다면 반전 여부는 중요치 않음
- 때문에 필터를 반전시키는 데 걸리는 연산량과 시간 소요를 줄이기 위하여 Tensorflow를 비롯한 다른 딥러닝 프레임워크들은 합성곱 대신 교차 상관을 사용하여 컨볼루션 레이어를 구현

<https://tensorflow.blog/2017/12/21/convolution-vs-cross-correlation/>

CNN 용어 정리

- CNN에서 사용되는 주요 용어
 - 컨볼루션 (Convolution), 합성곱
 - 채널 (Channel)
 - 필터 (Filter)
 - 스트라이드 (Stride)
 - 패딩 (Padding)
 - 피쳐 맵 (Feature Map)
 - 풀링 (Pooling) 레이어



CNN 용어 정리

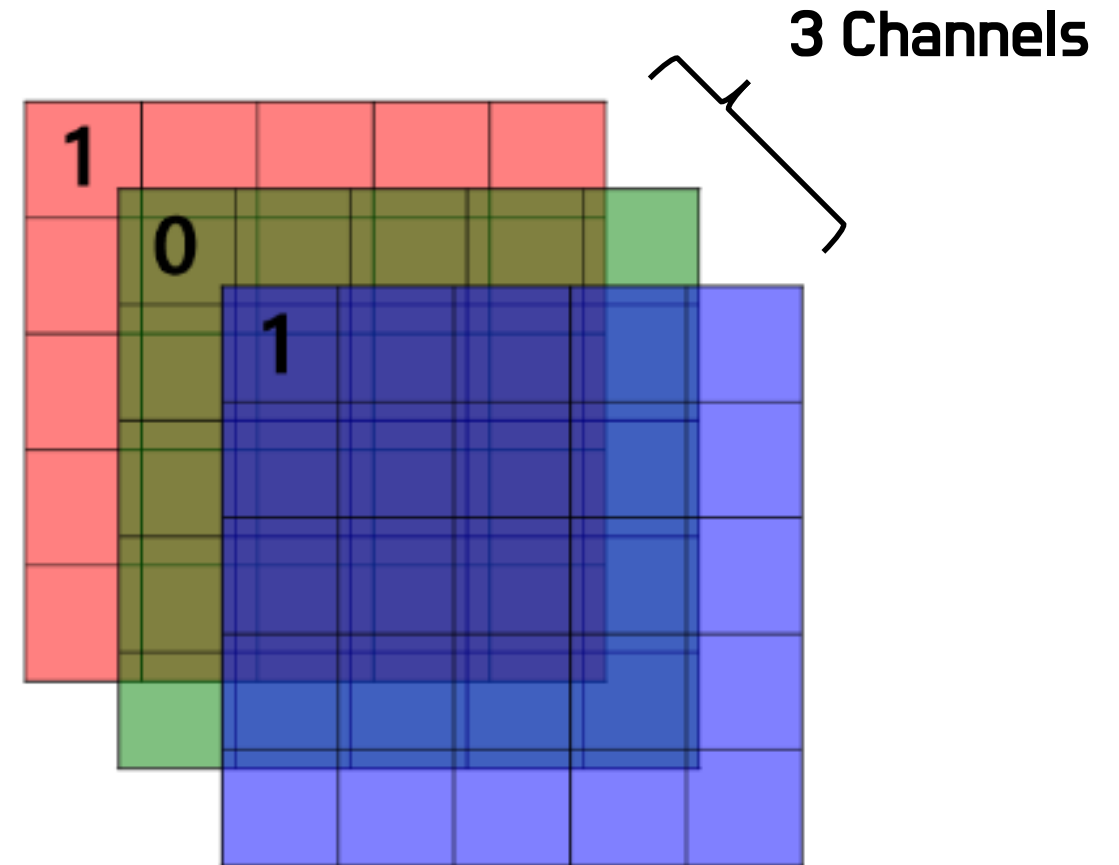
- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 앞서 언급했던 뉴런의 수용영역(Receptive Field)를 결정하는 것이 필터
- 커널(Kernel)이라고도 하며, 컨볼루션 레이어의 가중치 파라미터 W 에 해당
- 필터와 유사한 이미지의 영역을 강조하는 특성 맵 (Feature Map)을 출력하여 다음 레이어로 전달

CNN 용어 정리

- CNN에서 사용되는 주요 용어
 - 컨볼루션 (Convolution), 합성곱
 - 채널 (Channel)
 - 필터 (Filter)
 - 스트라이드 (Stride)
 - 패딩 (Padding)
 - 피쳐 맵 (Feature Map)
 - 풀링 (Pooling) 레이어



CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱

- 채널 (Channel)

- 필터 (Filter)

- 스트라이드 (Stride)

- 패딩 (Padding)

- 피쳐 맵 (Feature Map)

- 풀링 (Pooling) 레이어

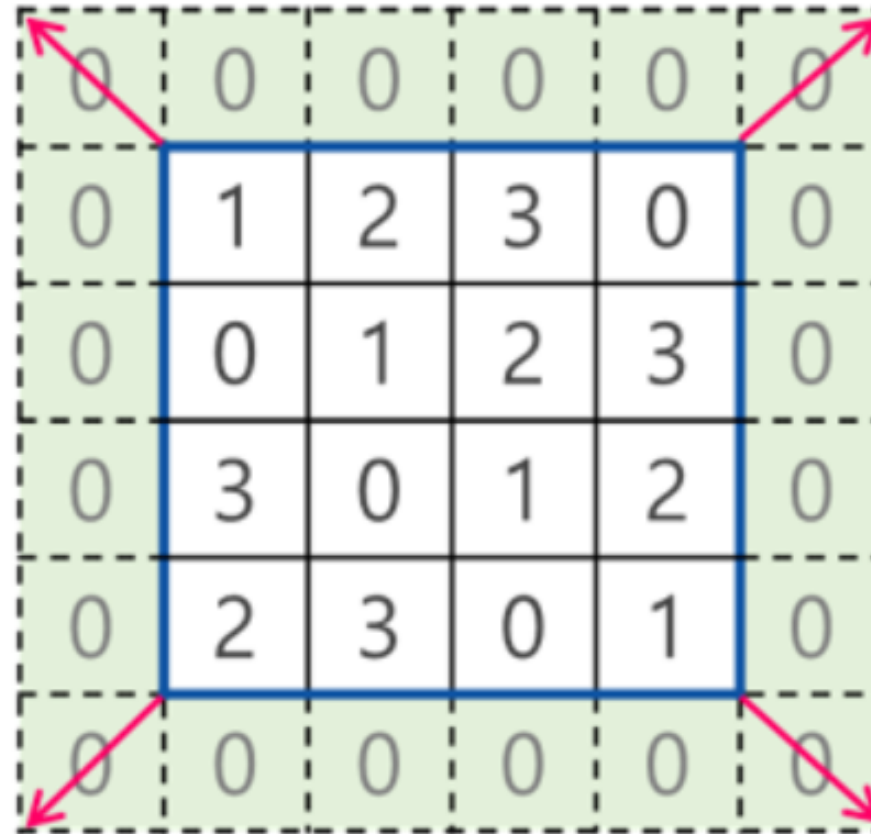
- 컬러 이미지는 3개 채널로 구성 (R,G,B)

- 흑백 이미지는 1개 채널

- 입력 데이터와 필터의 채널 수를 일치시켜야 함

CNN 용어 정리

- CNN에서 사용되는 주요 용어
 - 컨볼루션 (Convolution), 합성곱
 - 채널 (Channel)
 - 필터 (Filter)
 - 스트라이드 (Stride)
 - **패딩 (Padding)**
 - 피쳐 맵 (Feature Map)
 - 풀링 (Pooling) 레이어



1픽사리 zero- padding

CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 합성곱 연산 수행 전, 입력 데이터 주변을 특정 값으로 채워 늘리는 것
- 데이터의 공간적(Spatial) 크기는 컨볼루션 레이어를 지날 때 마다 작아지므로 가장자리의 정보가 사라지는 문제 발생
- 이 문제를 해결하기 위해 출력 데이터의 공간적 크기를 조절하는 것이 목적, 주로 입력 데이터의 크기와 동일하도록 조정
- 주로 0으로 채우는 Zero-Padding 사용

CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱

- 채널 (Channel)

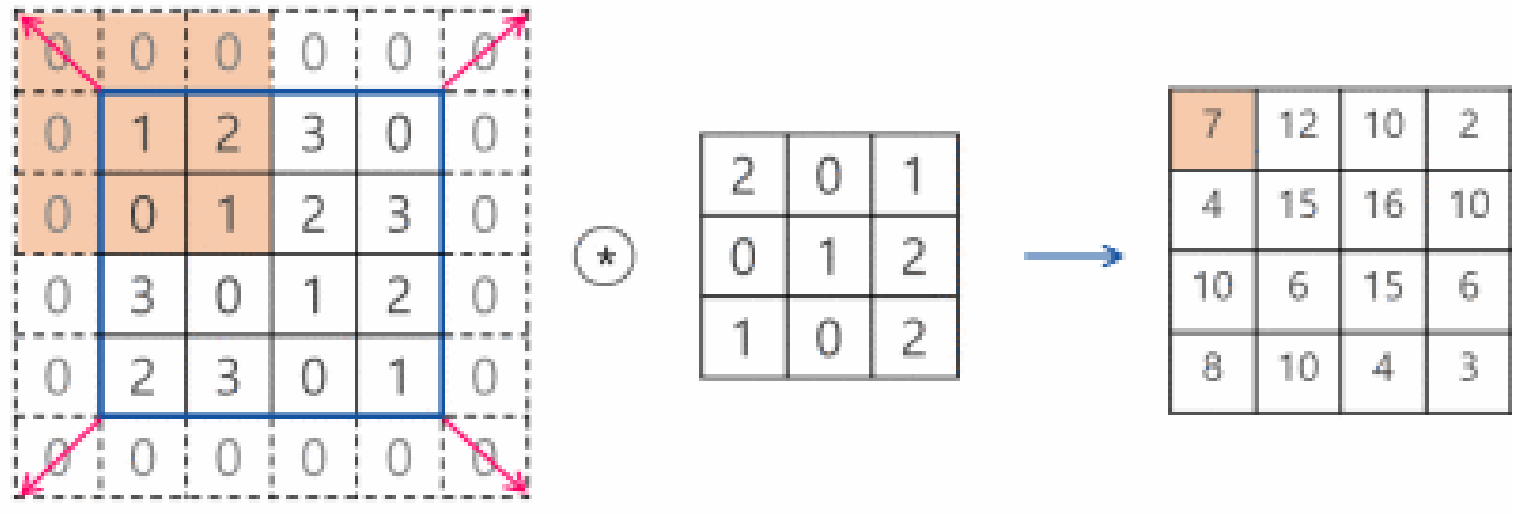
- 필터 (Filter)

- **스트라이드 (Stride)**

- 패딩 (Padding)

- 피쳐 맵 (Feature Map)

- 풀링 (Pooling) 레이어



CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 입력 데이터에 필터를 적용할 때 필터가 이동할 간격
- 출력 데이터의 크기 조절이 목적
- 보통 1과 같이 작은 값이 더 잘 작동

CNN 용어 정리

▪ CNN에서 사용되는 주요 용어

▪ 컨볼루션 (Convolution), 합성곱 *Max-pooling*

▪ 채널 (Channel)

▪ 필터 (Filter)

▪ 스트라이드 (Stride)

▪ 패딩 (Padding)

▪ 피쳐 맵 (Feature Map)

▪ 풀링 (Pooling) 레이어

1	2	1	0
0	1	2	3
3	0	1	2
2	4	0	1



2	

1	2	1	0
0	1	2	3
3	0	1	2
2	4	0	1



2	3
4	

Stride = 2

1	2	1	0
0	1	2	3
3	0	1	2
2	4	0	1



2	3

1	2	1	0
0	1	2	3
3	0	1	2
2	4	0	1



2	3
4	2

CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 컨볼루션 레이어의 출력을 입력으로 받아 출력 데이터의 크기를 줄이거나 특정 데이터를 강조하기 위해 사용
- Max Pooling, Average Pooling, Min Pooling 등의 방식이 있으나 이미지 처리에서는 Max Pooling을 주로 사용
- 보통 풀링의 window size와 stride는 같은 값으로 설정해서 모든 원소가 한번 씩만 연산에 참여하도록 한다.

CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱

- 채널 (Channel)

- 필터 (Filter)

- 스트라이드 (Stride)

- 패딩 (Padding)

- 피쳐 맵 (Feature Map)

- 풀링 (Pooling) 레이어

1	2	0	7	1	0
0	9	2	3	2	3
3	0	1	2	1	2
2	4	0	1	0	1
6	0	1	2	1	2
2	4	0	1	8	1



CNN 용어 정리

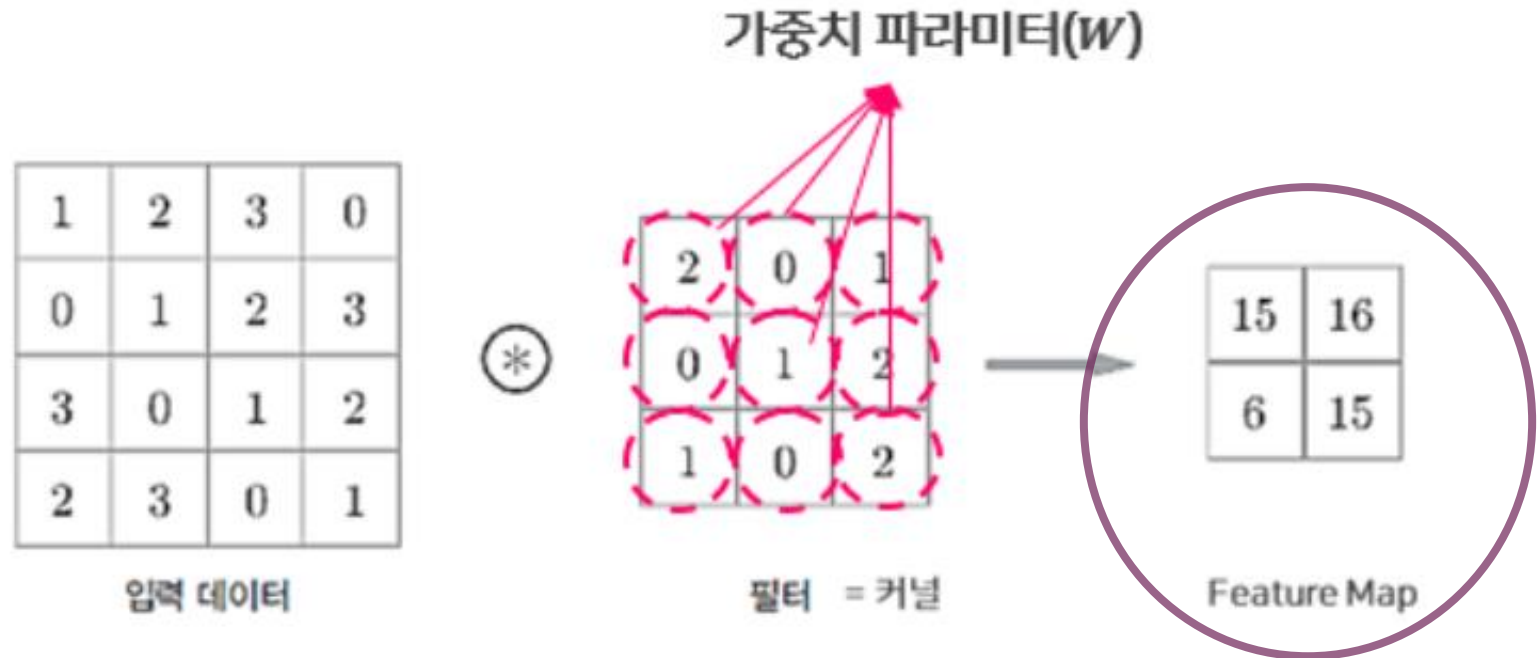
- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 풀링을 통해 계산한 이미지의 특징은 이미지 내의 위치에 대한 변화에 영향을 덜 받음
- ex) 눈(Eye)의 특징이 이미지 우상단과 좌하단에 있는 경우
- 이처럼 풀링을 통해 찾고자 하는 특징의 불변성 (Invariance)를 발견하여 공간적 변화를 극복할 수 있음
- 물론 처리되는 데이터의 양과 모델의 전체 매개변수의 수를 줄일 수 있는 기술적인 장점도 존재

CNN 용어 정리

- CNN에서 사용되는 주요 용어
 - 컨볼루션 (Convolution), 합성곱
 - 채널 (Channel)
 - 필터 (Filter)
 - 스트라이드 (Stride)
 - 패딩 (Padding)
 - 피쳐 맵 (Feature Map)
 - 풀링 (Pooling) 레이어



CNN 용어 정리

- CNN에서 사용되는 주요 용어

- 컨볼루션 (Convolution), 합성곱
- 채널 (Channel)
- 필터 (Filter)
- 스트라이드 (Stride)
- 패딩 (Padding)
- 피쳐 맵 (Feature Map)
- 풀링 (Pooling) 레이어

- 특성 맵(Feature Map)은 레이어의 출력 데이터
- 활성화 맵(Activation Map)이라고도 지칭
- 필터의 가중치 값을 기준으로 연산하여 필터와 유사한 이미지의 영역이 강조된 형태

CNN 레이어 출력 크기 계산

▪ 출력 크기 계산법

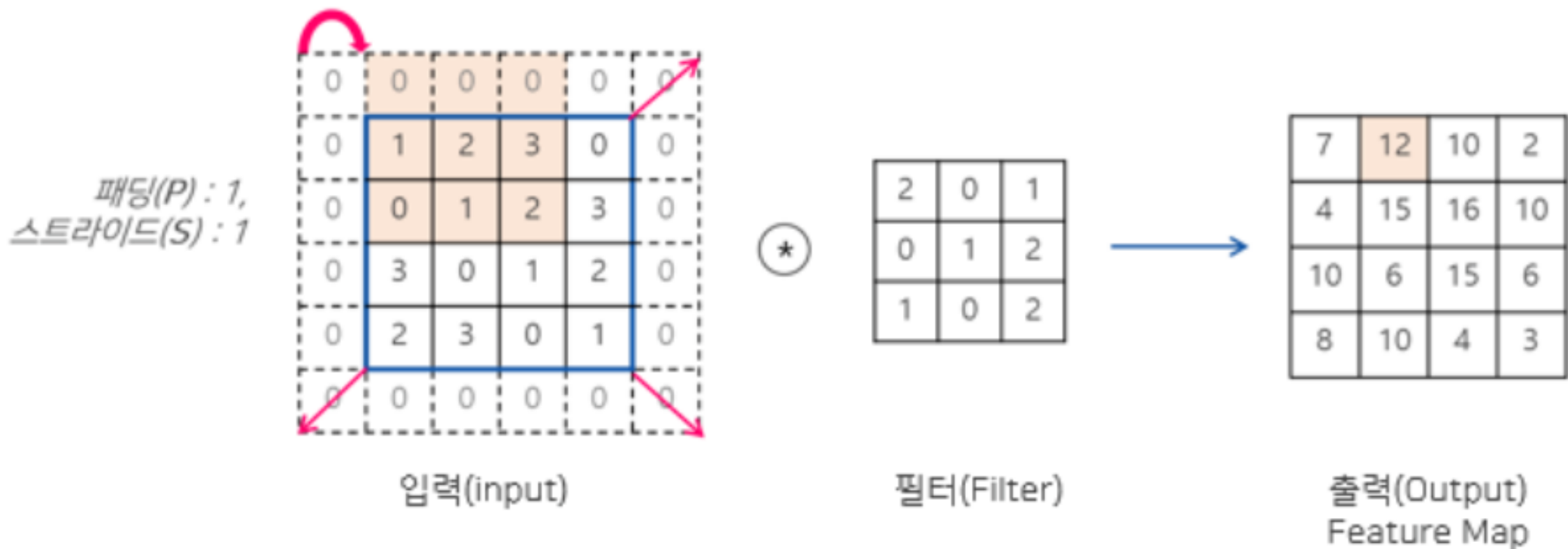
패딩과 스트라이드를 적용하고, 입력 데이터와 필터의 크기가 주어졌을 때 출력 데이터의 크기를 구하는 식

$$(OH, OW) = \left(\frac{H + 2P - FH}{S} + 1, \frac{W + 2P - FW}{S} + 1 \right)$$

- (H, W) : 입력 데이터의 Height, Width Size
- (FH, FW) : 필터의 Height, Width Size
- S : 스트라이드 (Stride)
- P : 패딩 (Padding)
- (OH, OW) : 출력 데이터의 Height, Width Size

CNN 레이어 출력 크기 계산

- 출력 크기 계산법



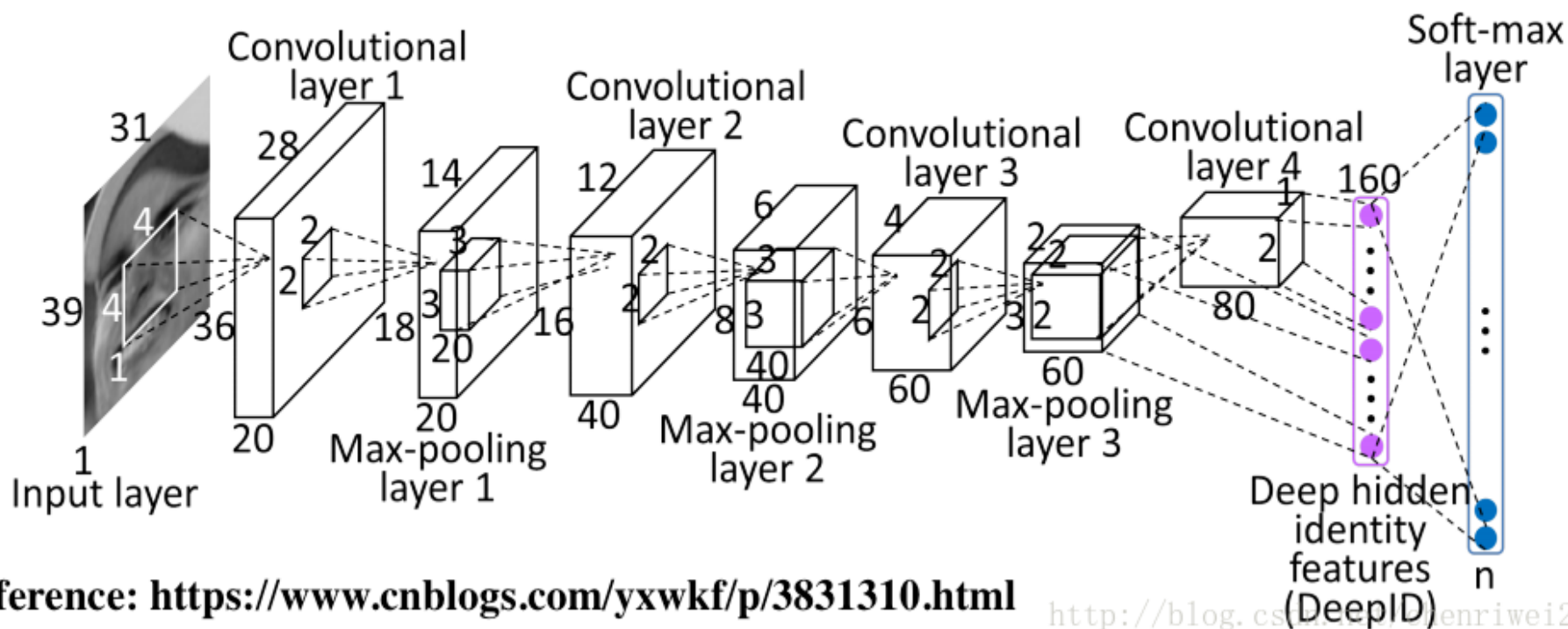
$$(OH, OW) = \left(\frac{4 + 2 \times 1 - 3}{1} + 1, \frac{4 + 2 \times 1 - 3}{1} + 1 \right) = (4, 4)$$

CNN과 FCNN의 비교

▪ 학습 파라미터 수의 관점에서 비교

1. CNN

4개의 컨볼루션 레이어, 39x31x1 크기의 입력 데이터, 100개의 클래스로 분류



Reference: <https://www.cnblogs.com/yxwxf/p/3831310.html>

<http://blog.csdn.net/enriwei2>

CNN과 FCNN의 비교

▪ 학습 파라미터 수의 관점에서 비교

1. CNN

4개의 컨볼루션 레이어, 39x31x1 크기의 입력 데이터, 100개의 클래스로 분류

layer	input channel	Filter	output channel	Stride	Pooling	활성함 수	Input Shape	Output Shape	파라미터 수
Convolution Layer 1	1	(4, 4)	20	1	X	relu	(39, 31, 1)	(36, 28, 20)	320
Max Pooling Layer 1	20	X	20	2	(2, 2)	X	(36, 28, 20)	(18, 14, 20)	0
Convolution Layer 2	20	(3, 3)	40	1	X	relu	(18, 14, 20)	(16, 12, 40)	7,200
Max Pooling Layer 2	40	X	40	2	(2,2)	X	(16, 12, 40)	(8, 6, 40)	0
Convolution Layer 3	40	(2, 2)	60	1	1	relu	(8, 6, 40)	(6, 4, 60)	21,600

Max Pooling Layer 3	60	X	60	(2, 2)	60	X	(6, 4, 60)	(3, 2, 60)	0
Convolution Layer 4	60	(2, 2)	80	1	1	relu	(3, 2, 60)	(2, 1, 80)	19,200
Flatten	X	X	X	X	X	X	(2, 1, 80)	(160, 1)	0
fully connected Layer	X	X	X	X	X	softmax	(160, 1)	(100, 1)	160,000
합계	X	X	X	X	X	softmax	(160, 1)	(100, 1)	208,320

컨볼루션 레이어의 학습 파라미터 수 : $C \times FH \times FW \times FN$

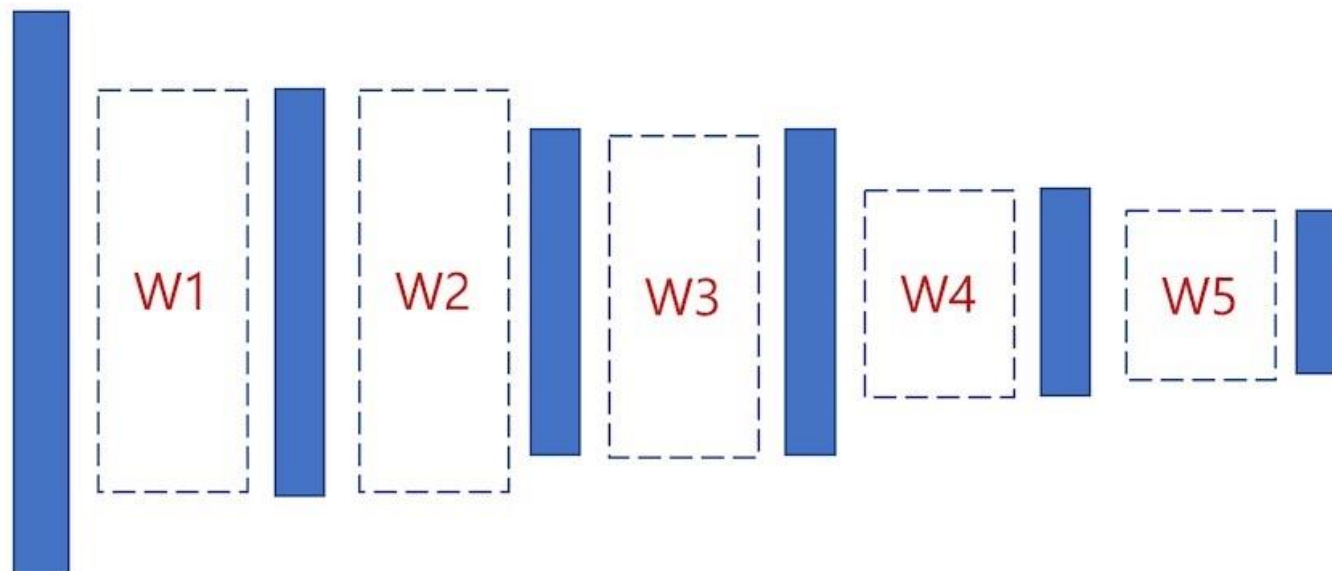
CNN과 FCNN의 비교

- 학습 파라미터 수의 관점에서 비교

2. FCNN

4개의 은닉층(Hidden Layer), 1209x1(39x31x1) 크기의 입력 데이터, 100개의 클래스로 분류

Input layer	Layer 1	Layer 2	Layer 3	Layer 4	Output Layer
(1209, 1)	(600, 1)	(300, 1)	(300, 1)	(150, 1)	(100, 1)



<http://taewan.kim>

CNN과 FCNN의 비교

- 학습 파라미터 수의 관점에서 비교

2. FCNN

4개의 은닉층(Hidden Layer), 1209x1(39x31x1) 크기의 입력 데이터, 100개의 클래스로 분류

레이어	입력 노드	출력 노드	Weight Shape	파라미터 수
Layer 1	1209	600	(1209,600)	725,400
Layer 2	600	300	(600,300)	180,000
Layer 3	300	300	(300,300)	90,000
Layer 4	300	150	(300,150)	45,000
Output	150	100	(150,100)	15,000
합계				1,055,400

CNN과 FCNN의 비교

- 학습 파라미터 수의 관점에서 비교
 - FCNN의 총 파라미터는 100만 개가 넘어감
 - CNN의 경우 20만개로 5배 이상의 차이
 - 은닉층이 깊어질 경우 이 차이는 더 급격히 늘어남

때문에 CNN은 FCNN과 비교하여 학습이 쉽고, 네트워크 처리 속도가 빠름