

Week5_Convolutional Neural Network

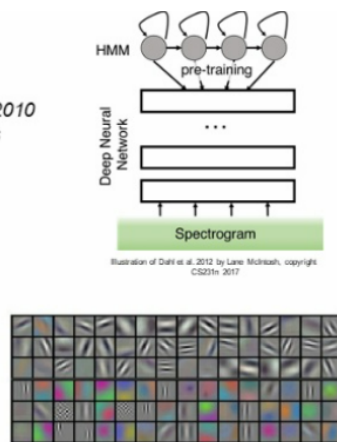
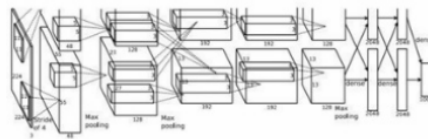
CNN-History

1. Mark 1 Perceptron
 - a. 단층 퍼셉트론 방식, W값을 조절해서 Train
2. Adaline/Madaline
3. Back-propagation 시작 (1986)
4. RBM-initialized & Fine tuning (2006) ⇒ weight 정밀 조정
5. AlexNet(2012)

First strong results

Acoustic Modeling using Deep Belief Networks
Abdel-rahman Mohamed, George Dahl, Geoffrey Hinton, 2010
Context-Dependent Pre-trained Deep Neural Networks for Large Vocabulary Speech Recognition
George Dahl, Dong Yu, Li Deng, Alex Acero, 2012

Imagenet classification with deep convolutional neural networks
Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton, 2012



Figures copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

CNN으로 ImageNet error rate를 급격히 줄이고, CNN이 발전하게 됨.

큰 영역의 특징에서 작은 영역의 특징 순으로 특징을 추출함

CNN 사용 영역

- Classification
- Retrieval

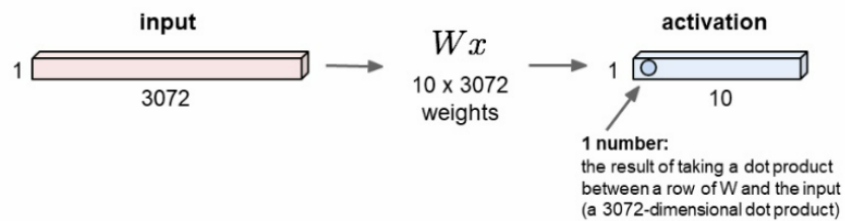
- Detection
- Segmentation
- Self-driving car(전용 GPU까지 탑재해서 더 빠른 속도로 구분)
- 등등...

Convolutional Neural Networks

-Fully Connected Layer

Fully Connected Layer

32x32x3 image -> stretch to 3072 x 1

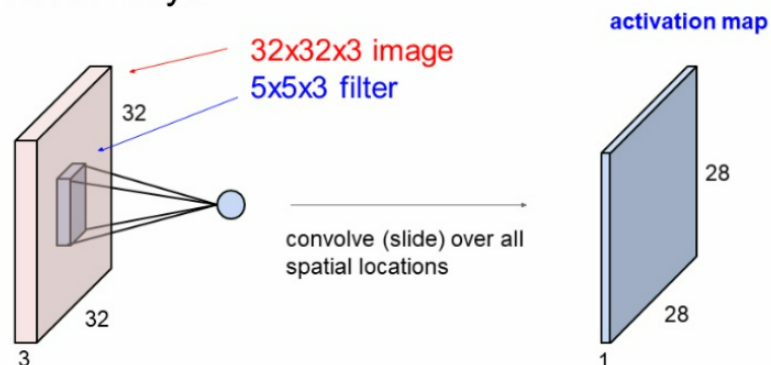


이전 FC에서는 input image를 stretch 해서 wx의 input 값으로 넣었음.

전체 input 값에 W전체 연산

-Convolution Layer

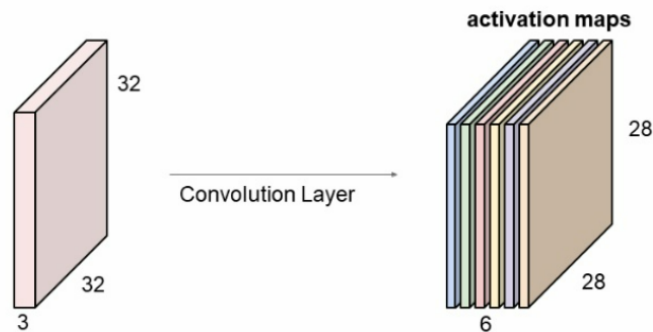
Convolution Layer



32x32x3(RGB)의 input이 있으면 filter의 크기(5x5x3)에 해당하는 지역에서 하나의 값을 뽑아낸다.

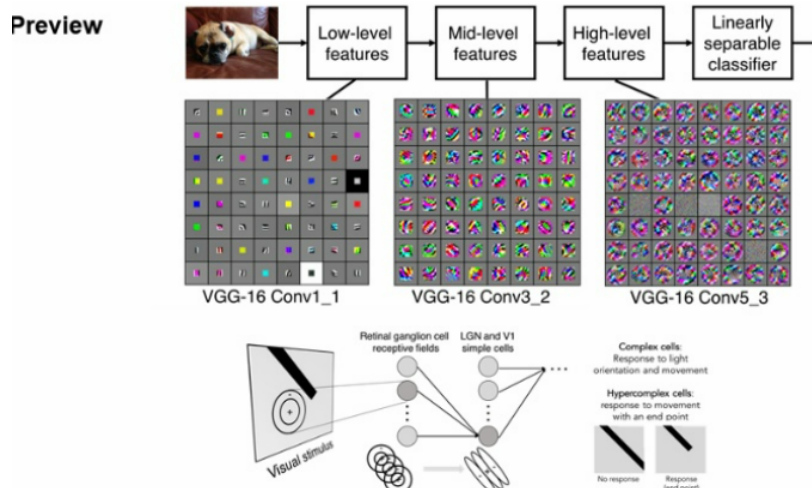
32x32x3 input에 5x5x3 convolution layer를 연산하면 28x28x1의 하나의 activation map이 나오게 된다.

For example, if we had 6 5x5 filters, we'll get 6 separate activation maps:



We stack these up to get a "new image" of size 28x28x6!

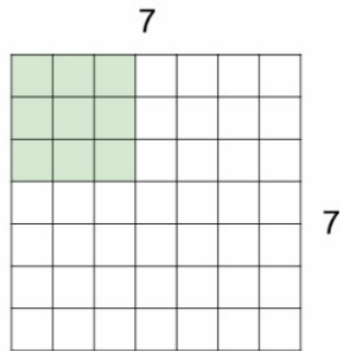
여러 filter를 사용해서 여러개의 activation map 만들기 ⇒ depth값에 변화를 줄 수 있음



시각화를 하면 하위계층으로 갈수록 더 다양한 영역에서 feature를 뽑아내는 것을 알 수 있다.

-Stride

A closer look at spatial dimensions:



7x7 input (spatially) assume 3x3 filter applied **with stride 3?**

doesn't fit!
cannot apply 3x3 filter on 7x7 input with stride 3.

activation map을 만들기 위해 conv filter를 움직일 때 stride값을 기준으로 움직인다. 위 예시의 32x32x3 → 28x28x1 은 stride=1값을 기준으로 움직였을 때.

이때 그림처럼 stride값을 잘못 설정하면 전체 input값에 fit하지 않는다.

$$\text{OUTPUT SIZE} = (N - F) / \text{stride} + 1$$

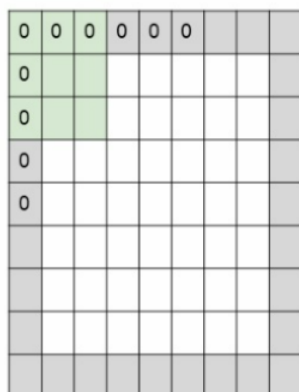
(N : input size, F : filter size)

위 식이 정수로 떨어지지 않으면 stride 값이 잘못된 것.

-Paddle

Stride값이 커질수록 activation map의 사이즈가 작아진다. ⇒ 이미지 사이즈가 작아짐 이를 방지하기 위해 zero paddle을 사용한다.

In practice: Common to zero pad the border



e.g. input 7x7
3x3 filter, applied with **stride 1**
pad with 1 pixel border => what is the output?

7x7 output!
in general, common to see CONV layers with stride 1, filters of size FxF, and zero-padding with (F-1)/2. (will preserve size spatially)
e.g. F = 3 => zero pad with 1
F = 5 => zero pad with 2
F = 7 => zero pad with 3

Filter의 크기에 따른 output size가 줄어들지 않는 paddle size가 있는데

$$F = 3 \Rightarrow 1$$

$$F = 5 \Rightarrow 2$$

$$F = 7 \Rightarrow 3$$

다음과 같다.

-관련 공식들

Summary. To summarize, the Conv Layer:

- Accepts a volume of size $W_1 \times H_1 \times D_1$
- Requires four hyperparameters:
 - Number of filters K ,
 - their spatial extent F ,
 - the stride S ,
 - the amount of zero padding P .
- Produces a volume of size $W_2 \times H_2 \times D_2$ where:
 - $W_2 = (W_1 - F + 2P)/S + 1$
 - $H_2 = (H_1 - F + 2P)/S + 1$ (i.e. width and height are computed equally by symmetry)
 - $D_2 = K$
- With parameter sharing, it introduces $F \cdot F \cdot D_1$ weights per filter, for a total of $(F \cdot F \cdot D_1) \cdot K$ weights and K biases.
- In the output volume, the d -th depth slice (of size $W_2 \times H_2$) is the result of performing a valid convolution of the d -th filter over the input volume with a stride of S , and then offset by d -th bias.

Input volumn : 32x32x3

10 5x5 filters withh stride = 1, pad = 2

Q1) Output volumn size? = $(32 + 4 - 5)/ 1 + 1 = 32$

$\Rightarrow 32 \times 32 \times 10$

Q2) # of parameters in this layers?

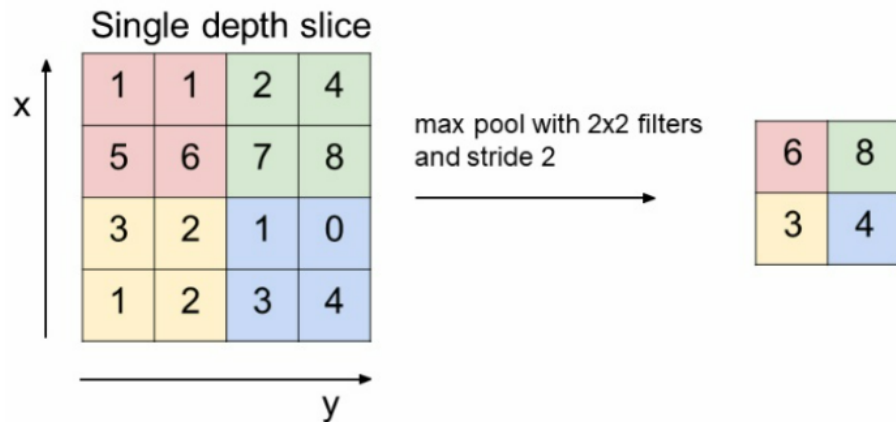
각 필터당 $5 * 5 * 3 + 1 = 76$ 개의 파라미터가 필요하다.(+1은 bias 값)

$\Rightarrow 76 * 10 = 760$ 개

-Pooling Layer

: 특정 특징값(주요 특징값)을 유지하면서 이미지 사이즈를 줄여준다.

1) MAX POOLING



Maxpooling은 pooling size에 해당하는 값 중 가장 큰 값만을 유지시킨다.

-공식

Pooling size 를 P라고 하면 output size는

$$\text{OUTPUTsize} = (I - F) / \text{stride} + 1$$