모델 구조

Input   Conv1   Conv2   Conv3   Conv4   Conv5   Dense   Dense   Dense

227

227

55

55

3

96

11x11 Conv
Max Pooling

29

29

256

5x5 Conv
Max Pooling

13

13

384

3x3 Conv

13

13

384

3x3 Conv

13

13

256

3x3 Conv
Max Pooling

4096

4096

1000

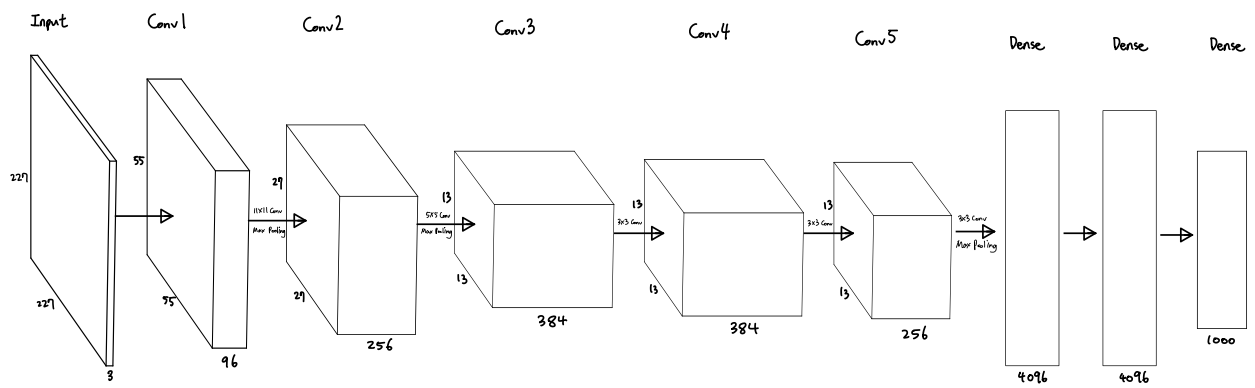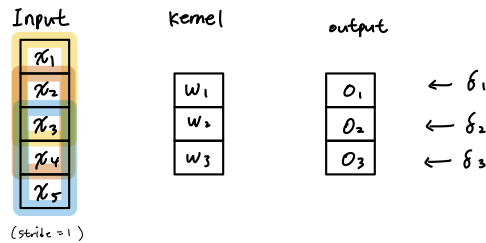## CNN 역전파

: 커널이 모든 입력데이터에 적용되므로 역전파 계산에도 Convolution 연산이 사용됨.

＊ 수식

$$\frac{\delta}{\delta x} = \frac{\delta}{\delta x} \int_{\mathbb{R}^d} f(y) g(x-y) \, dy$$

$$= \int_{\mathbb{R}^d} f(y) \frac{\delta g}{\delta x}(x-y) \, dy$$

$$= [f \cdot g'](x)$$

＊ 예시

Input	kernel	output
x_1	w_1	O_1 ← δ_1
x_2	w_2	O_2 ← δ_2
x_3	w_3	O_3 ← δ_3
x_4		
x_5		

(stride = 1)

$$\left.\begin{array}{l} O_1 = w_1 x_1 + w_2 x_2 + w_3 x_3 \\ O_2 = w_1 x_2 + w_2 x_3 + w_3 x_4 \\ O_3 = w_1 x_3 + w_2 x_4 + w_3 x_5 \end{array}\right\} \longrightarrow O_i = \sum_{j=1} w_j \cdot x_{i+j-1}$$

- 편미분값 : $\delta_1, \delta_2, \delta_3$
- Convolution의 최종 출력값 : L

$$\delta_1 = \frac{\partial L}{\partial O_1}, \quad \delta_2 = \frac{\partial L}{\partial O_2}, \quad \delta_3 = \frac{\partial L}{\partial O_3}$$

$$\frac{\partial L}{\partial w_1} = \delta_1 \frac{\partial O_1}{\partial w_1} + \delta_2 \frac{\partial O_1}{\partial w_2} + \delta_3 \frac{\partial O_1}{\partial w_3}$$

$$= \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3$$

$$\frac{\partial L}{\partial w_2} = \delta_1 x_2 + \delta_2 x_3 + \delta_3 x_4$$

$$\frac{\partial L}{\partial w_3} = \delta_1 x_3 + \delta_2 x_4 + \delta_3 x_5$$

$$\therefore \quad \frac{\partial L}{\partial w_i} = \sum_{j=1} \delta_j \, x_{i+j-1}$$