CNN 기초

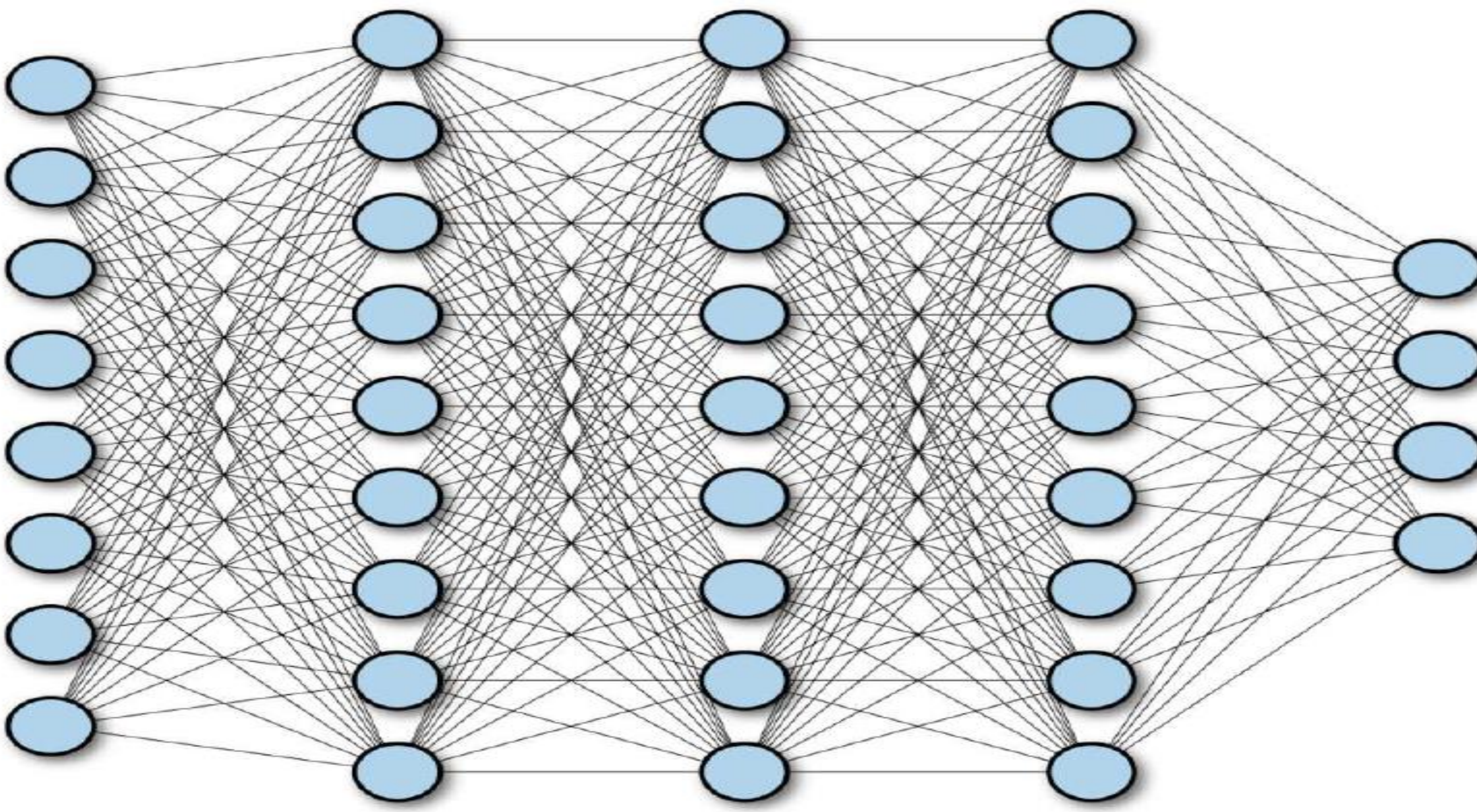
컴퓨터 비전를 중심으로

Dec 2020

합성곱 신경망 기초

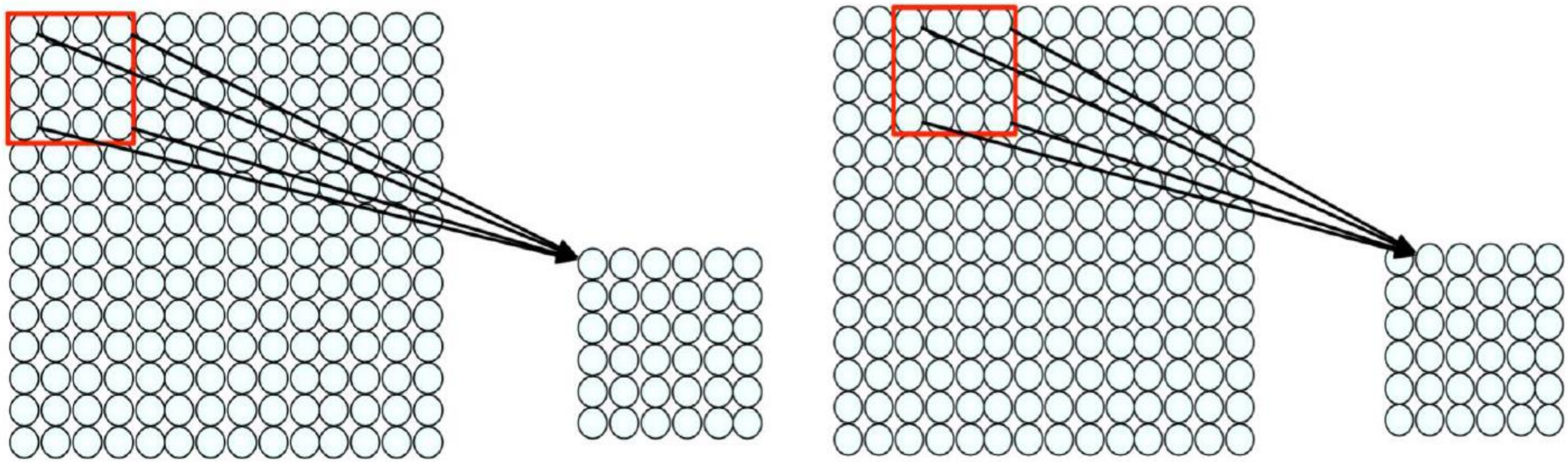
전결합 신경망

- 전결합 신경망 (Fully Connected: FC)



합성곱층의 이해

- 합성곱을 통한 공간적 특성추출



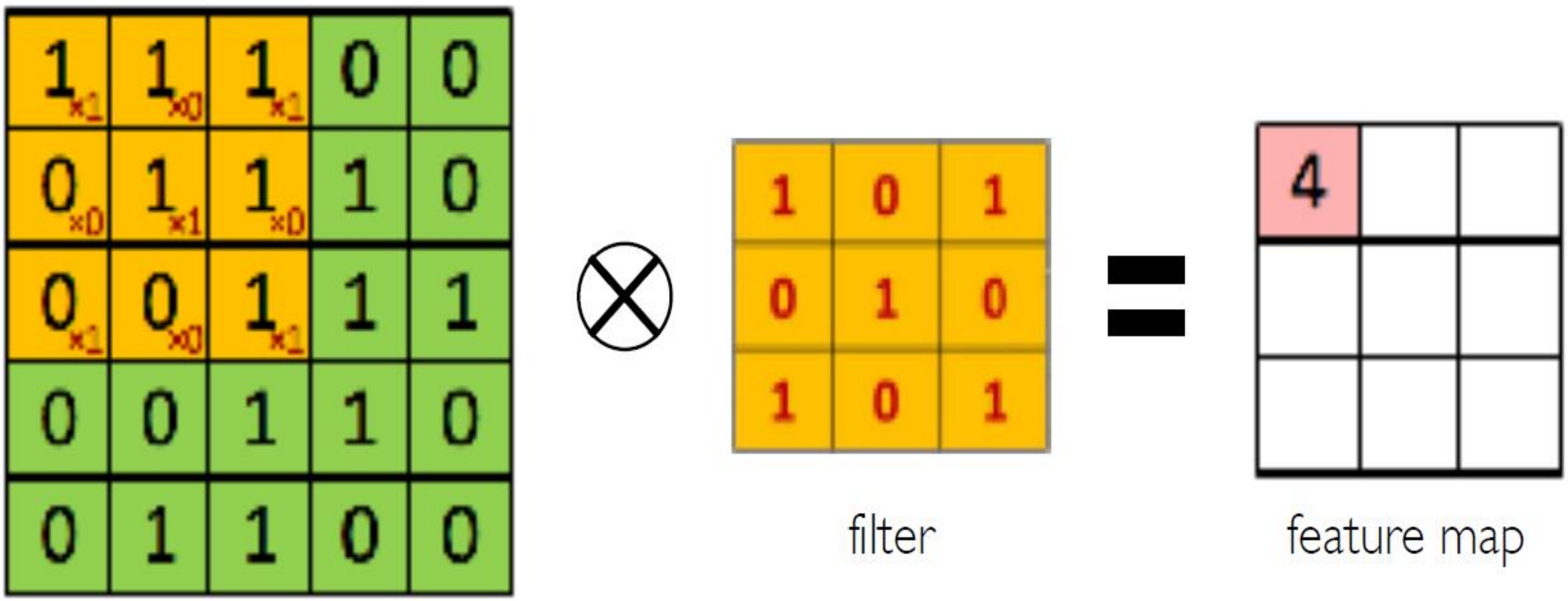
1) Apply a set of weights – a filter – to extract **local features**

2) Use **multiple filters** to extract different features

3) **Spatially share** parameters of each filter

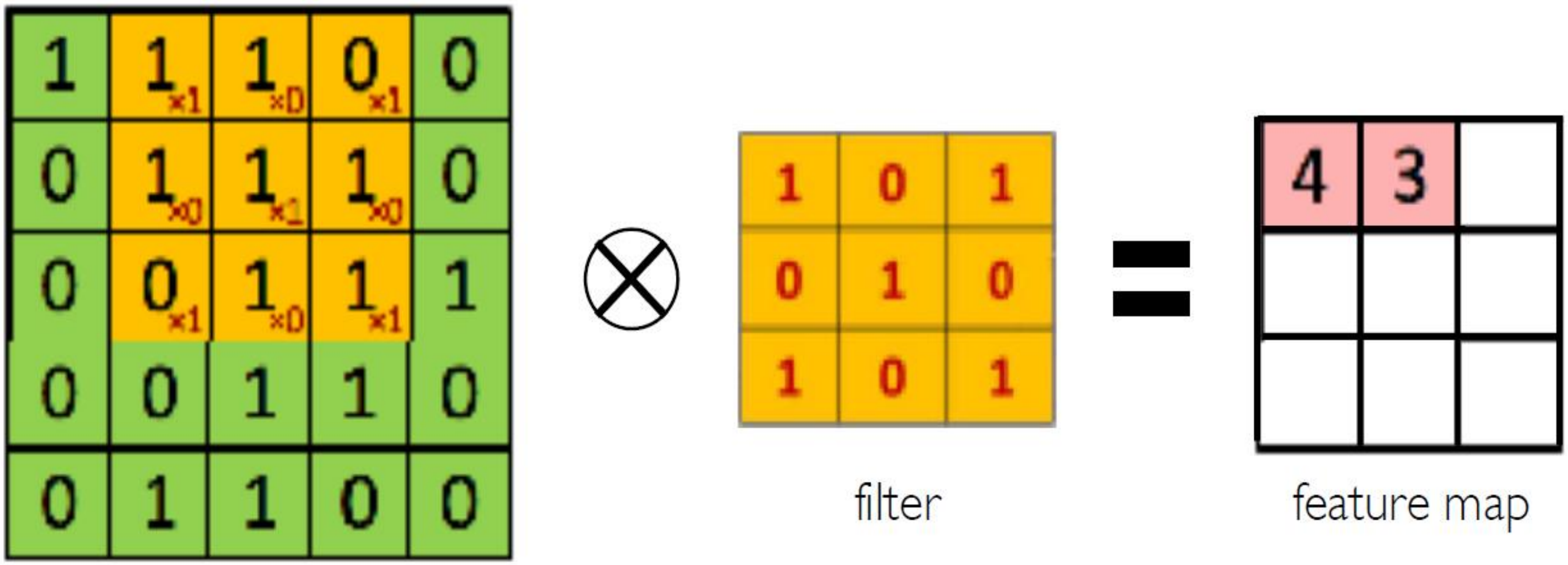
합성곱층의 이해

- 합성곱이란?



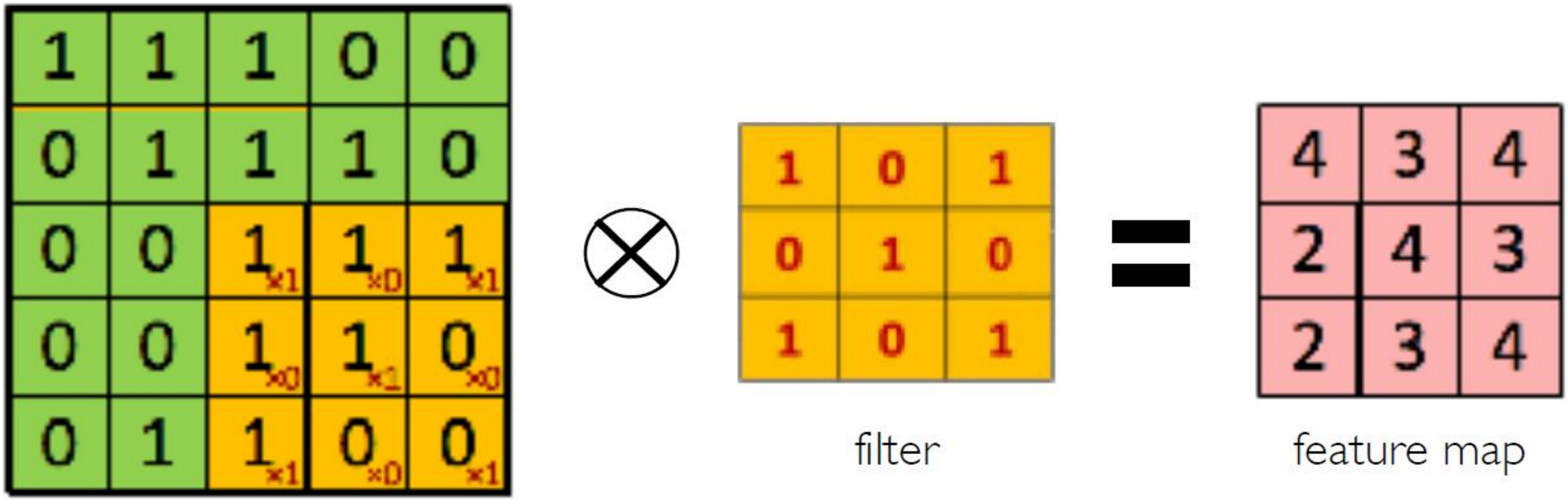
합성곱층의 이해

- 합성곱이란?



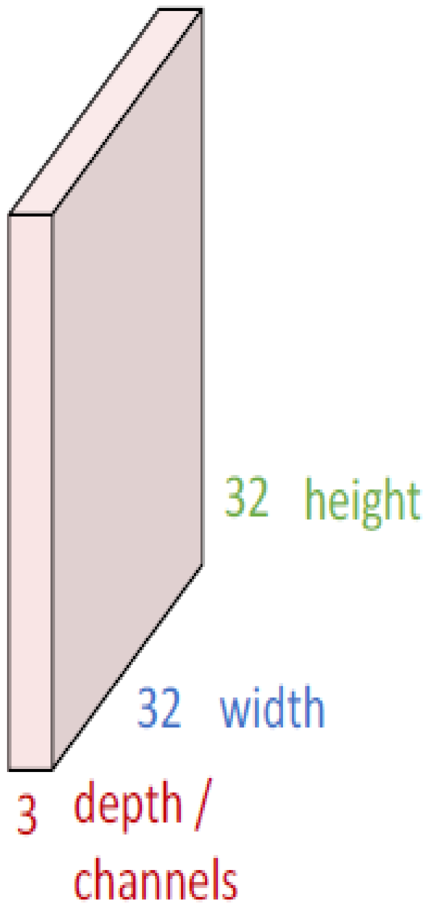
합성곱층의 이해

- 합성곱이란?



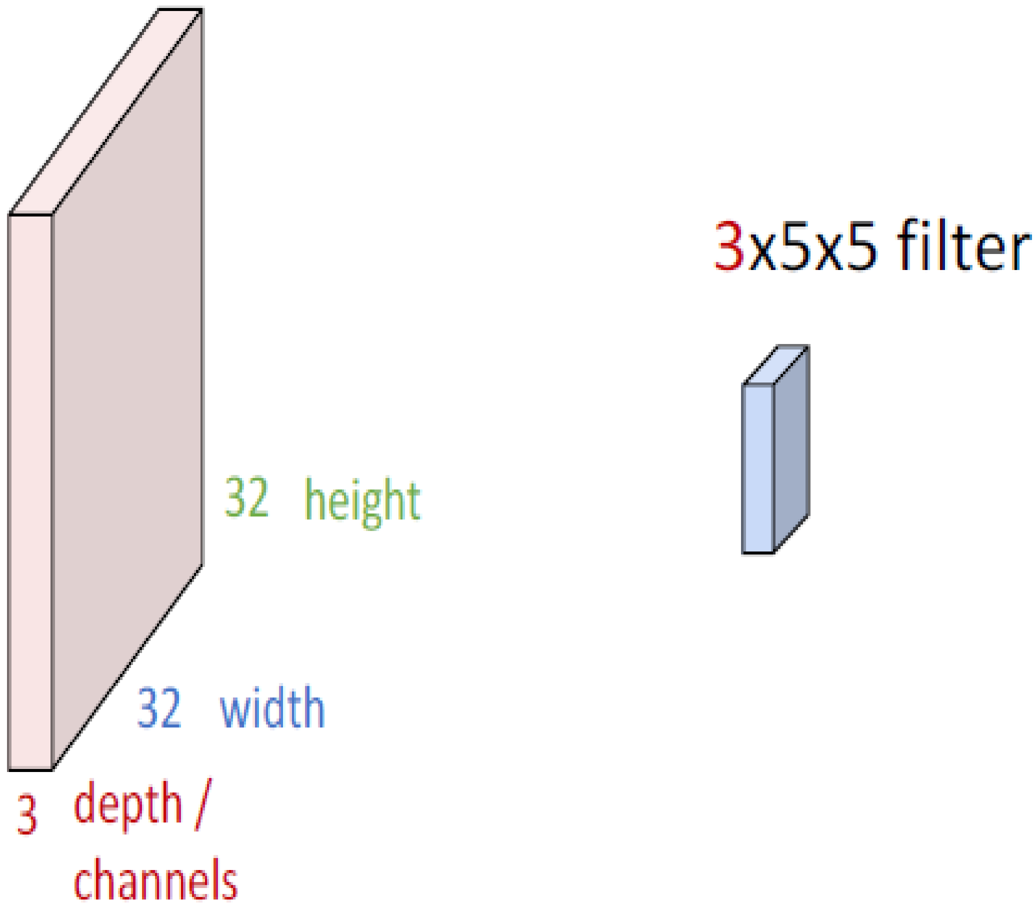
합성곱층의 이해

- 3x32x32 이미지



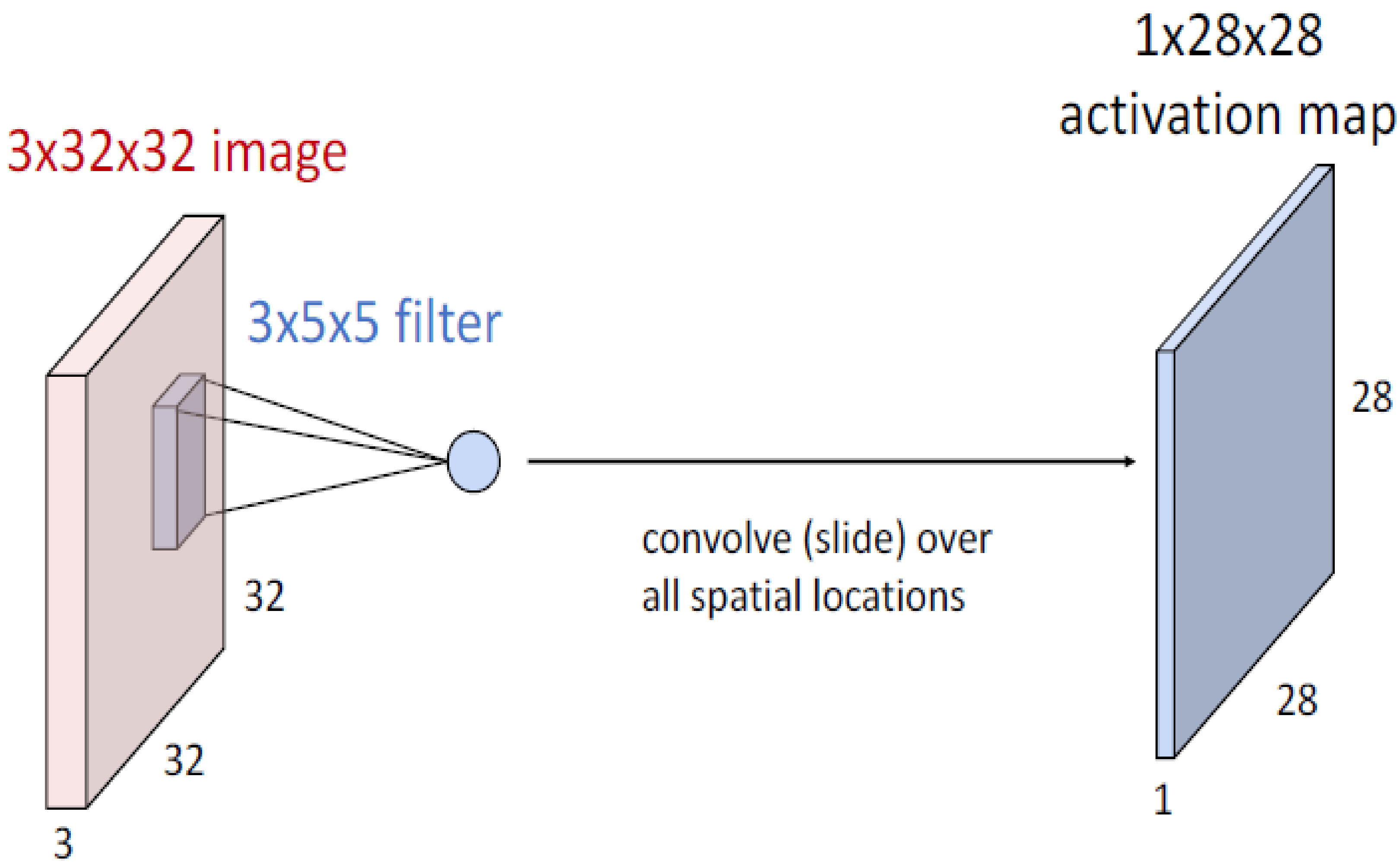
합성곱층의 이해

- 3x32x32 이미지



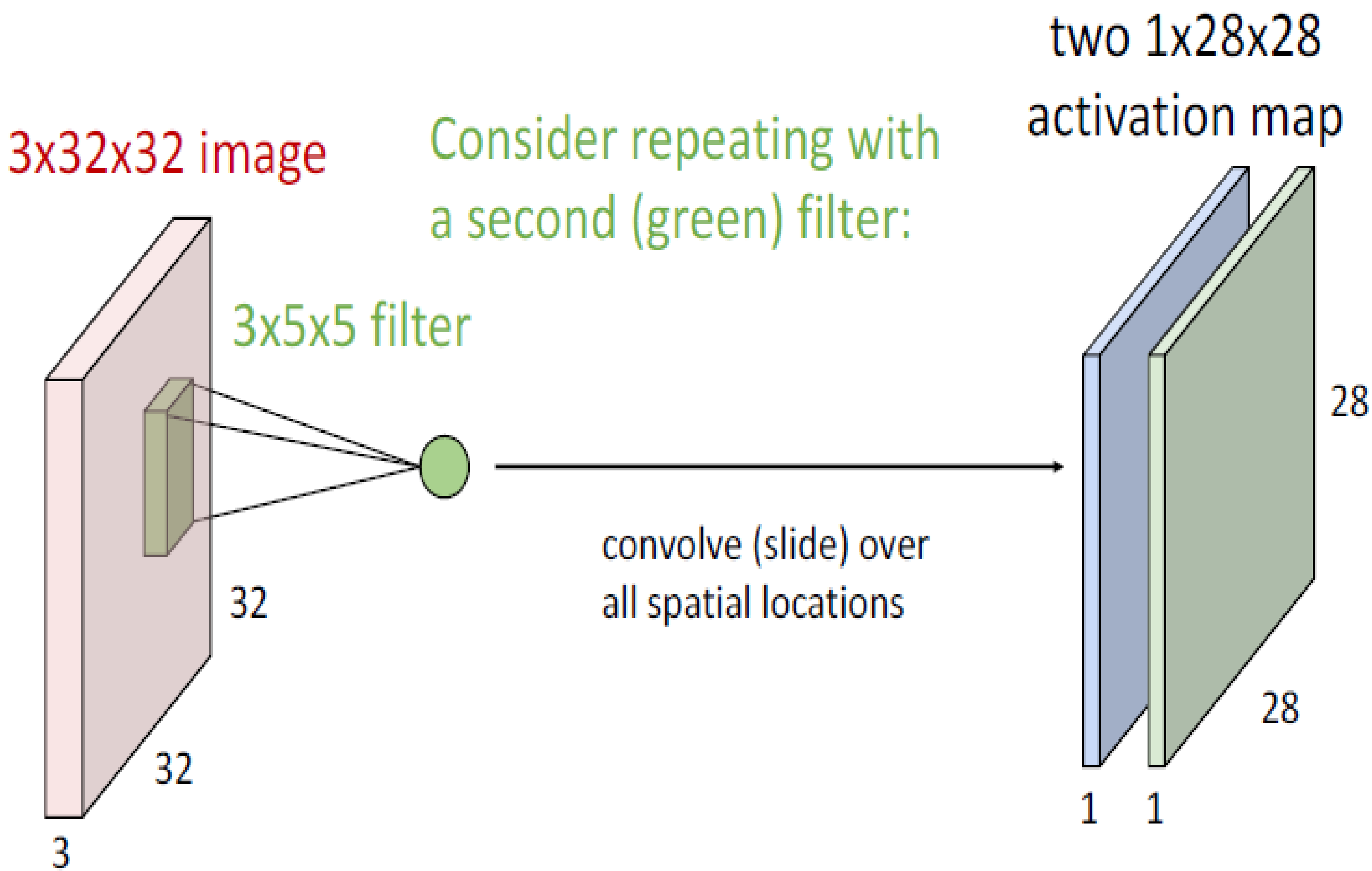
합성곱층의 이해

- 첫번째 활성 맵



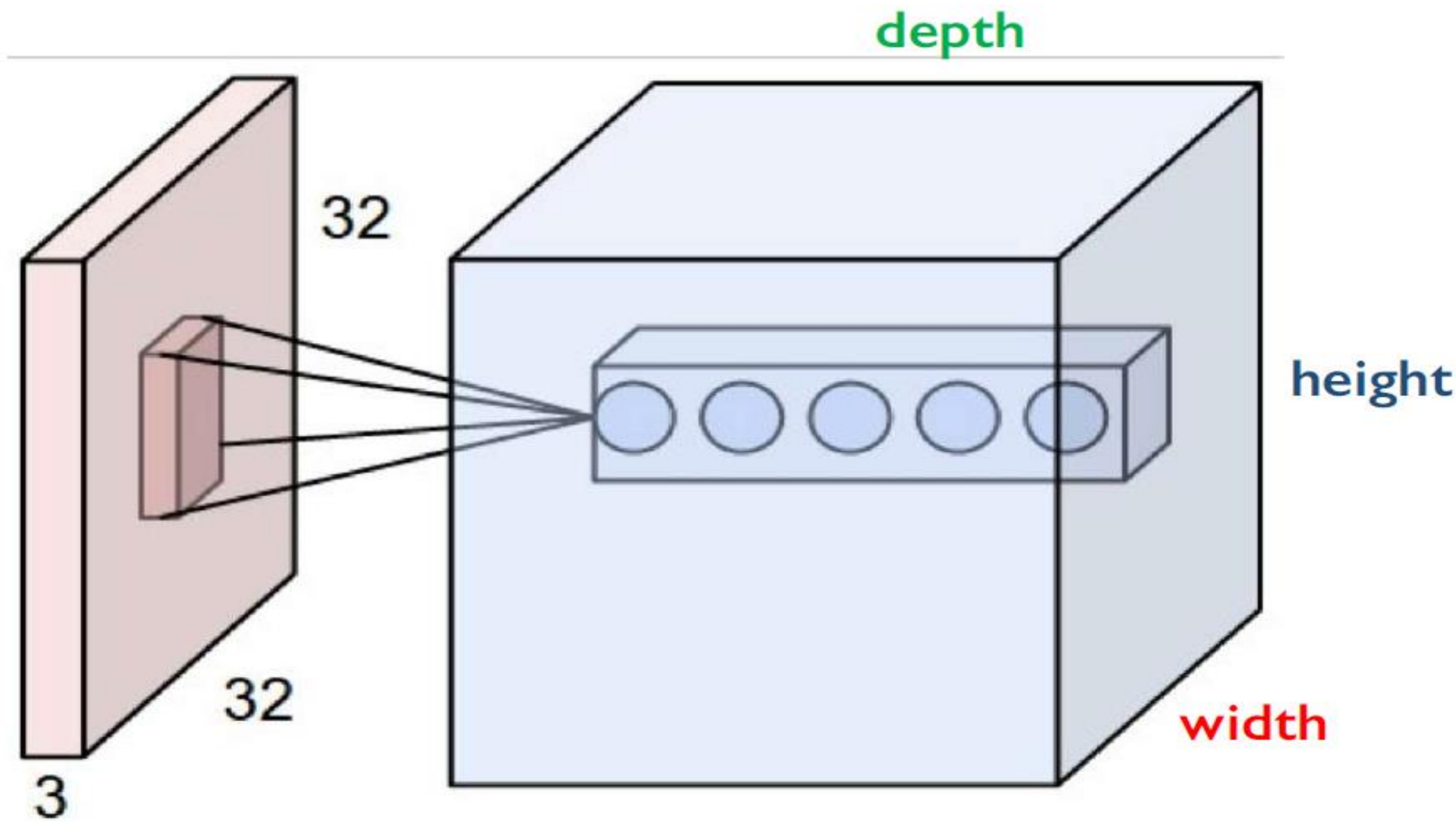
합성곱층의 이해

- 두번째 활성 맵



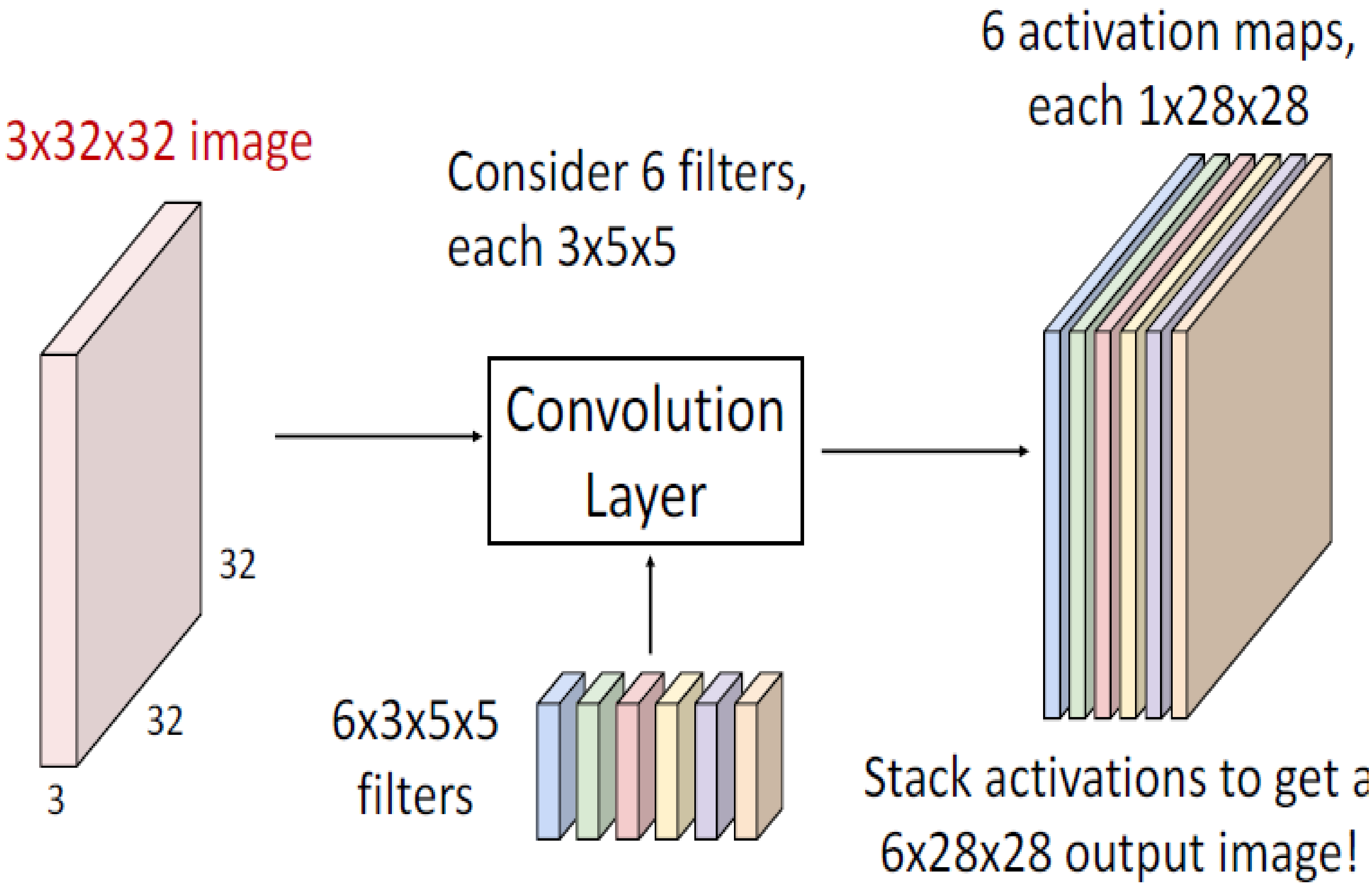
합성곱층의 이해

- 여기서 depth는 필터의 개수임.



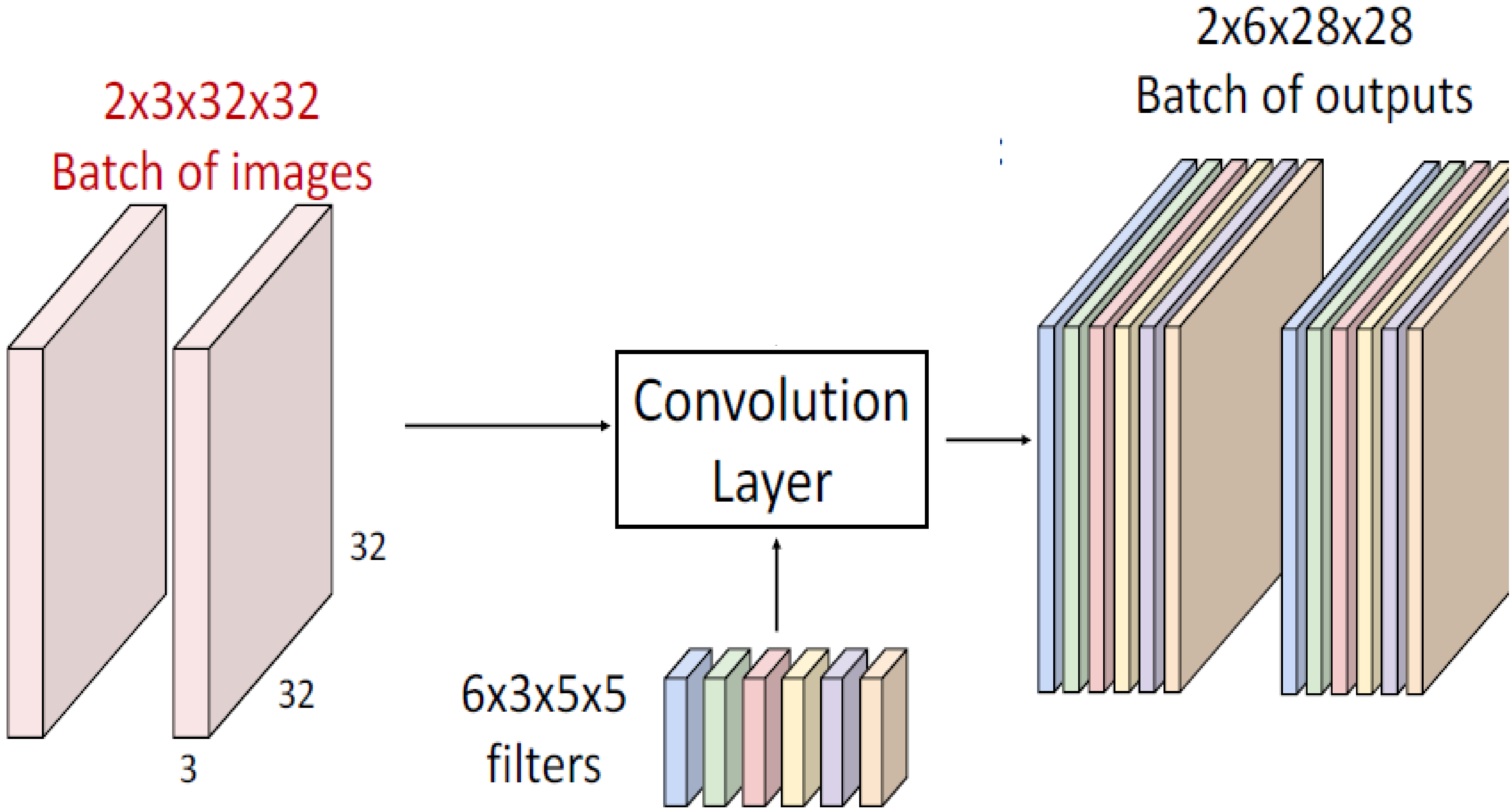
합성곱층의 이해

- 6개의 필터로 6개의 활성 맵을 만든다.



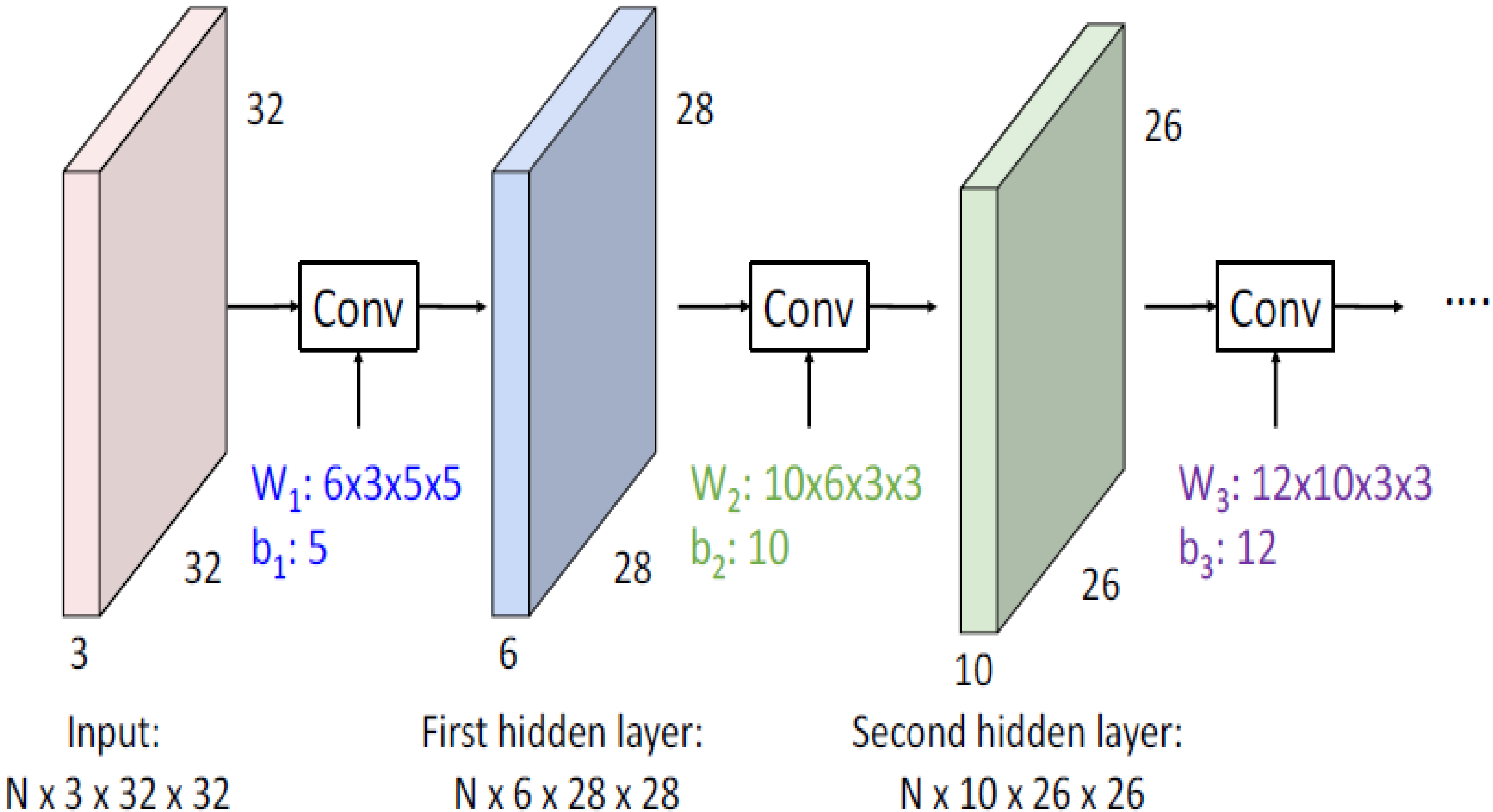
합성곱층의 이해

- 2개의 이미지에 대해 6개의 필터로 6개짜리 2세트의 활성맵을 만든다.



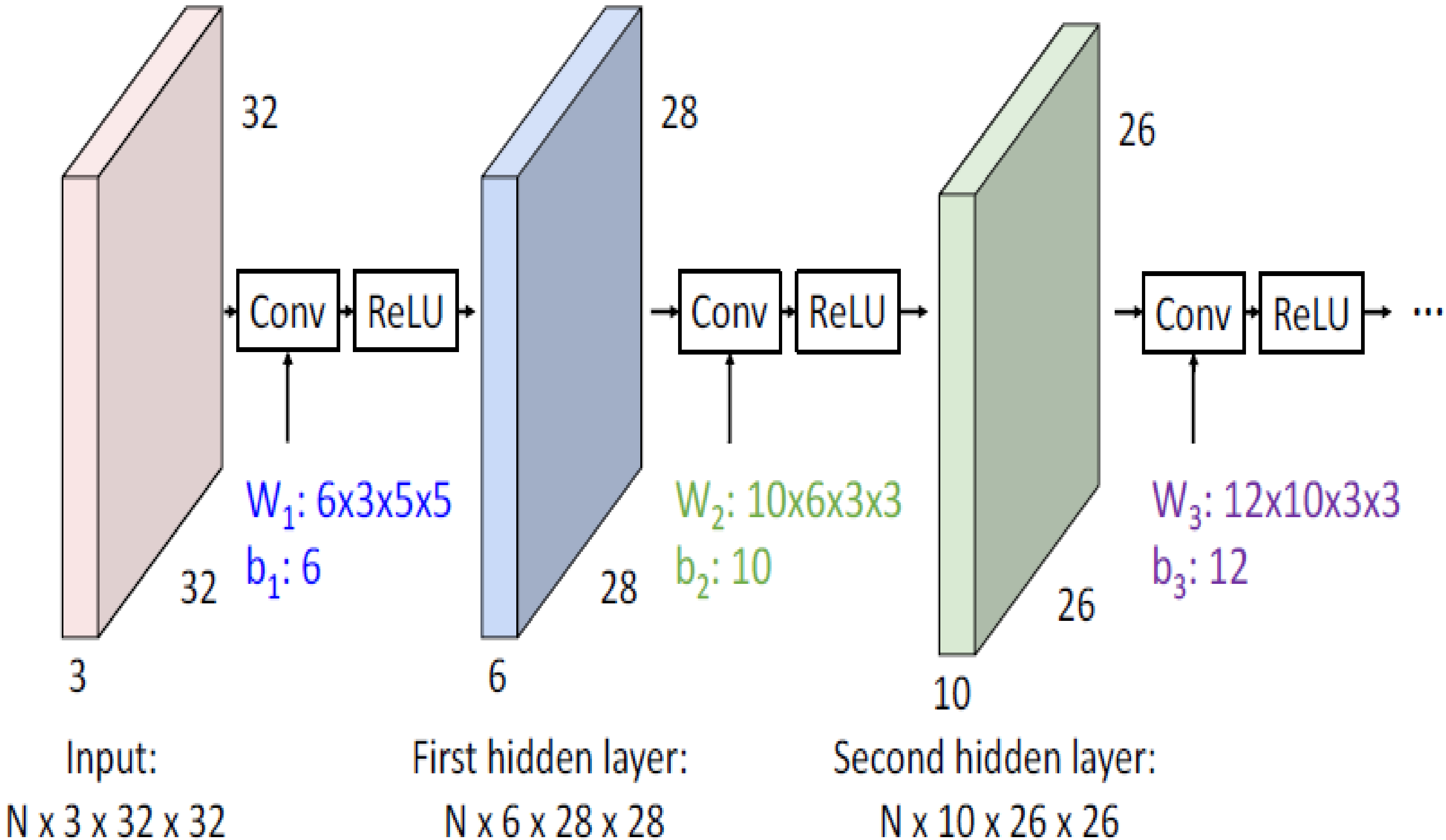
합성곱층의 이해

- 합성곱의 적층



합성곱층의 이해

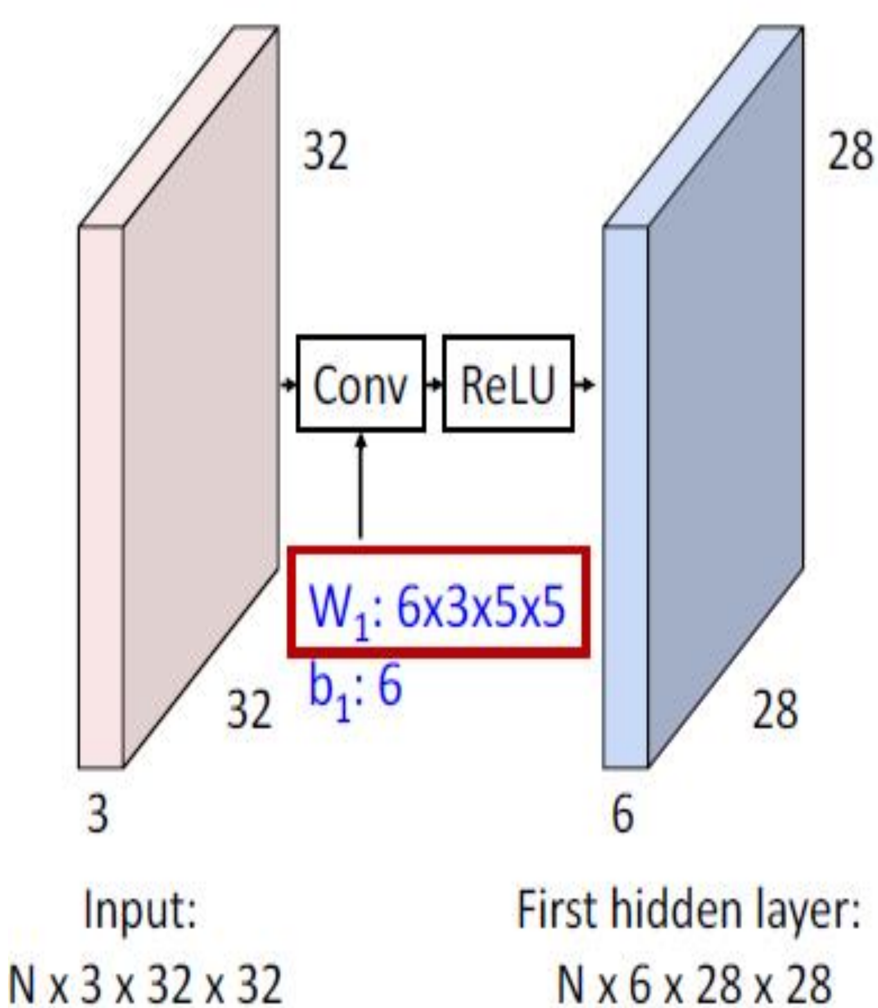
- RELU 층의 추가로 진정한 활성맵을 만든다.

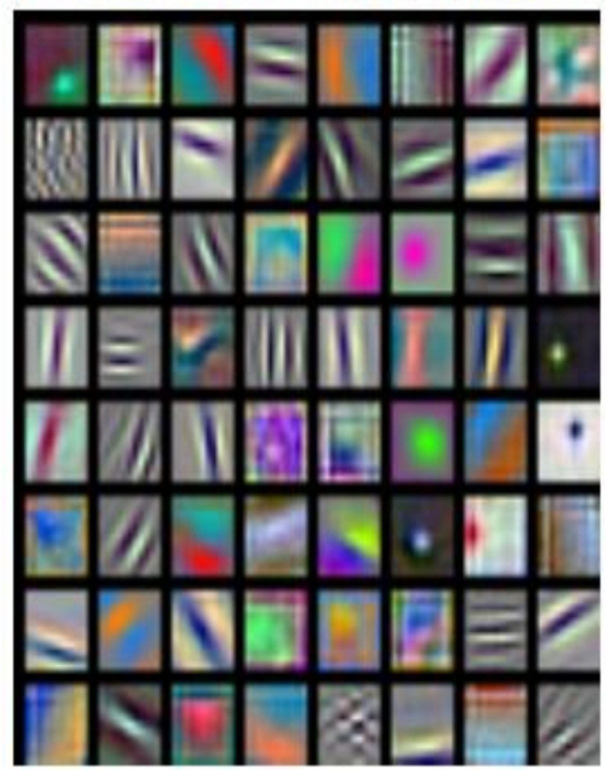


합성곱층의 이해

- 필터의 시각화 및 의미

첫번째 합성곱층은은 국지적 이미지 템플릿이다. (방향성 모서리, 색상 대조 등)

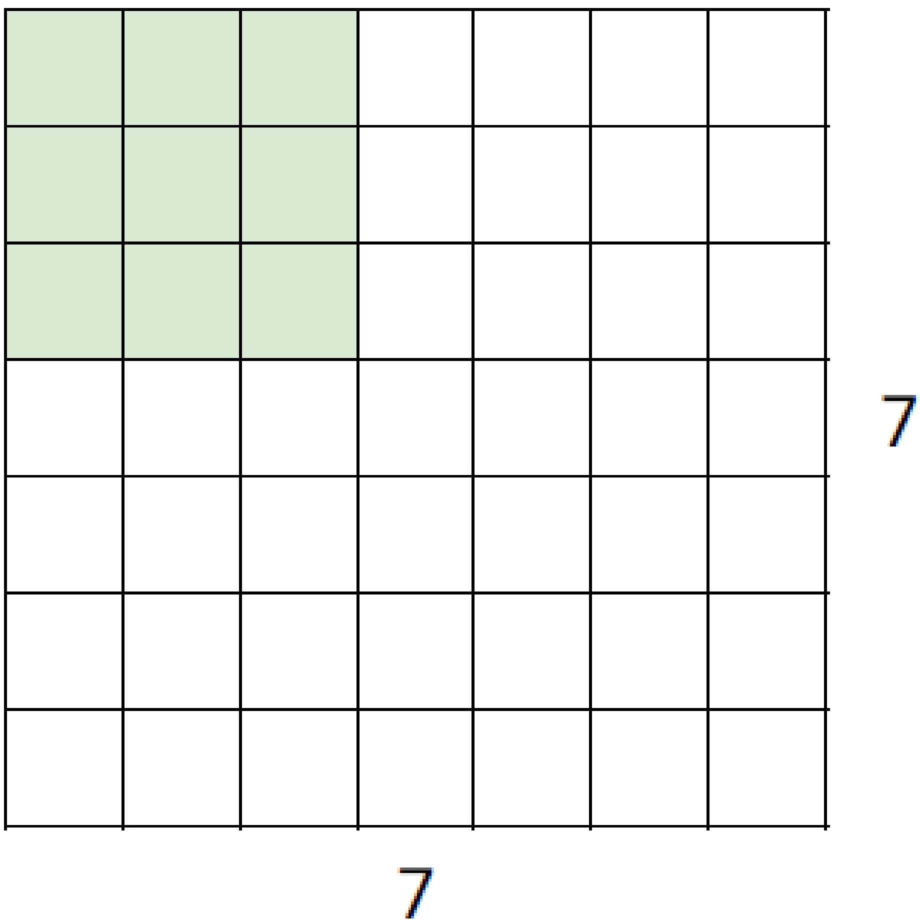




AlexNet: 64 filters, each 3x11x11

합성곱 신경망의 이해

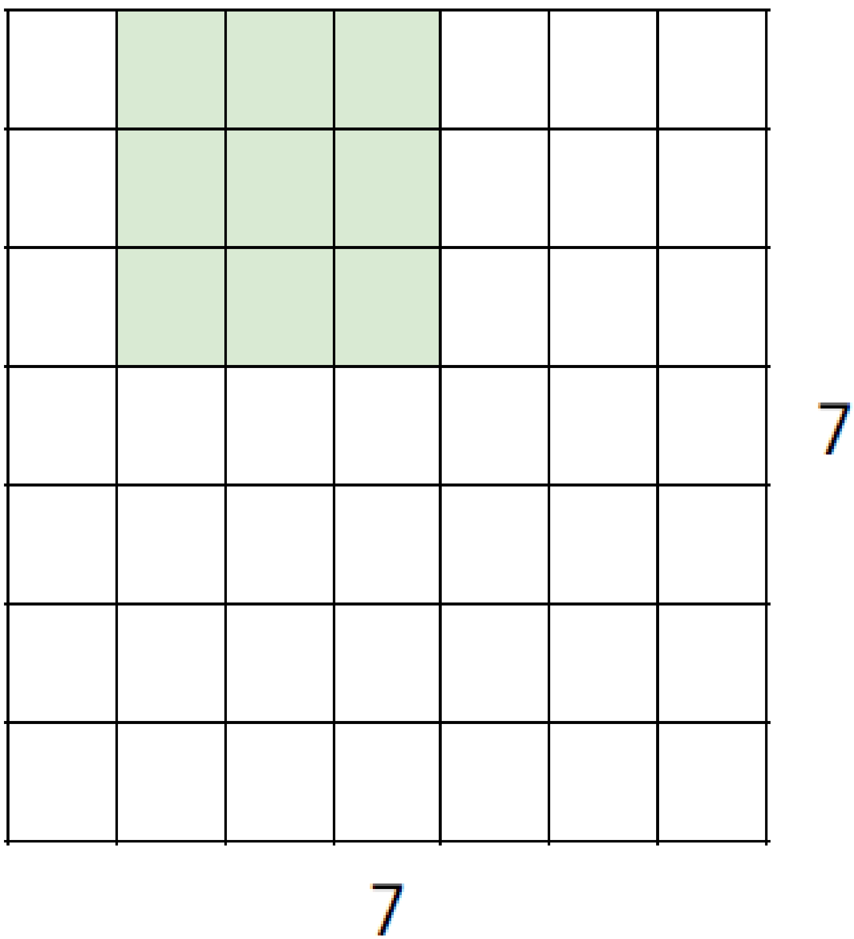
- 공간적 이해



Input: 7x7
Filter: 3x3

합성곱 신경망의 이해

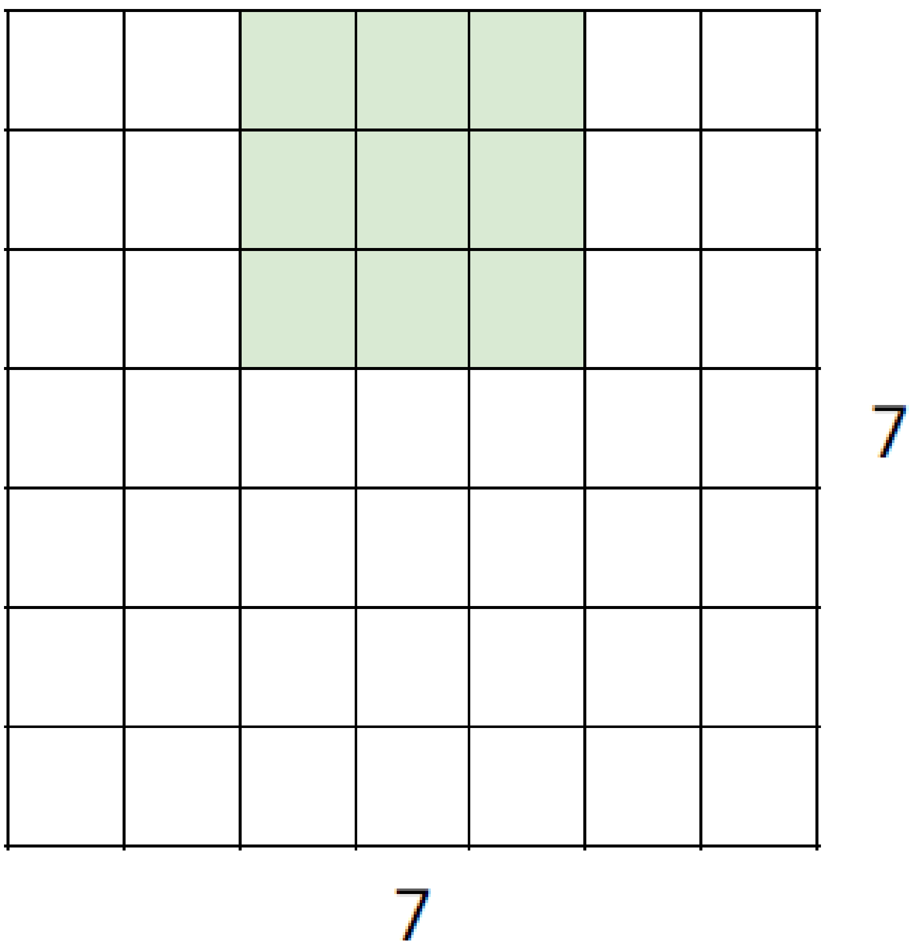
- 공간적 이해



Input: 7x7
Filter: 3x3

합성곱 신경망의 이해

- 공간적 이해

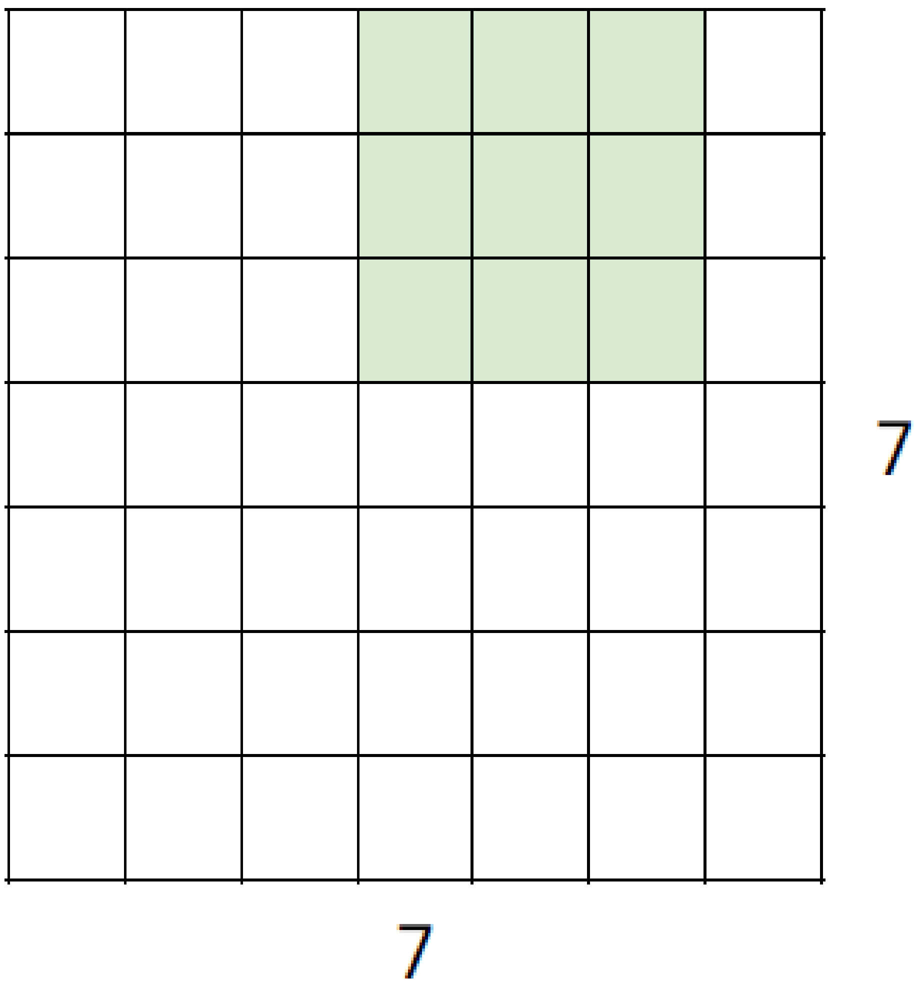


Input: 7x7

Filter: 3x3

합성곱 신경망의 이해

- 공간적 이해

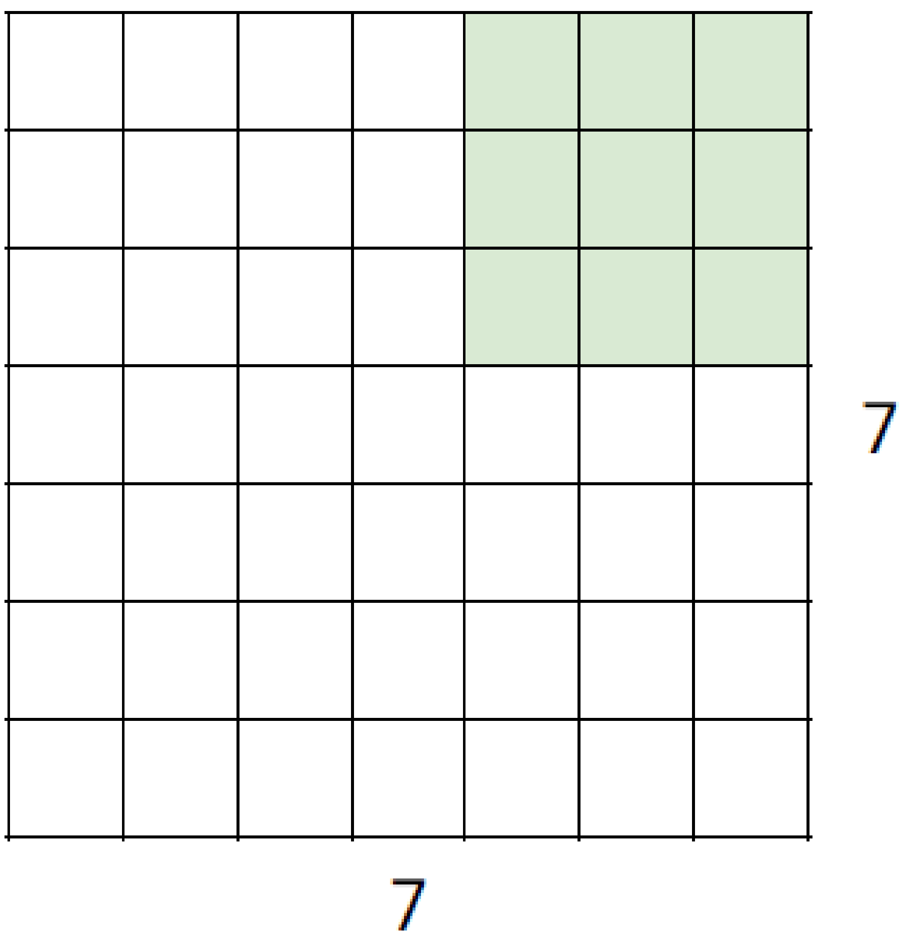


Input: 7x7

Filter: 3x3

합성곱 신경망의 이해

- 공간적 이해



Input: 7x7

Filter: 3x3

Output: 5x5

- 일반적으로 입력 크기가 W이고 필터크기가 K이면 출력은 W-k+1이다.

합성곱 신경망의 이해

- 패딩(padding)

0	0	0	0	0	0	0	0	0
0								0
0								0
0								0
0								0
0								0
0								0
0								0
0	0	0	0	0	0	0	0	0

Input: 7x7

Filter: 3x3

Output: 5x5

In general:

Input: W

Filter: K

Padding: P

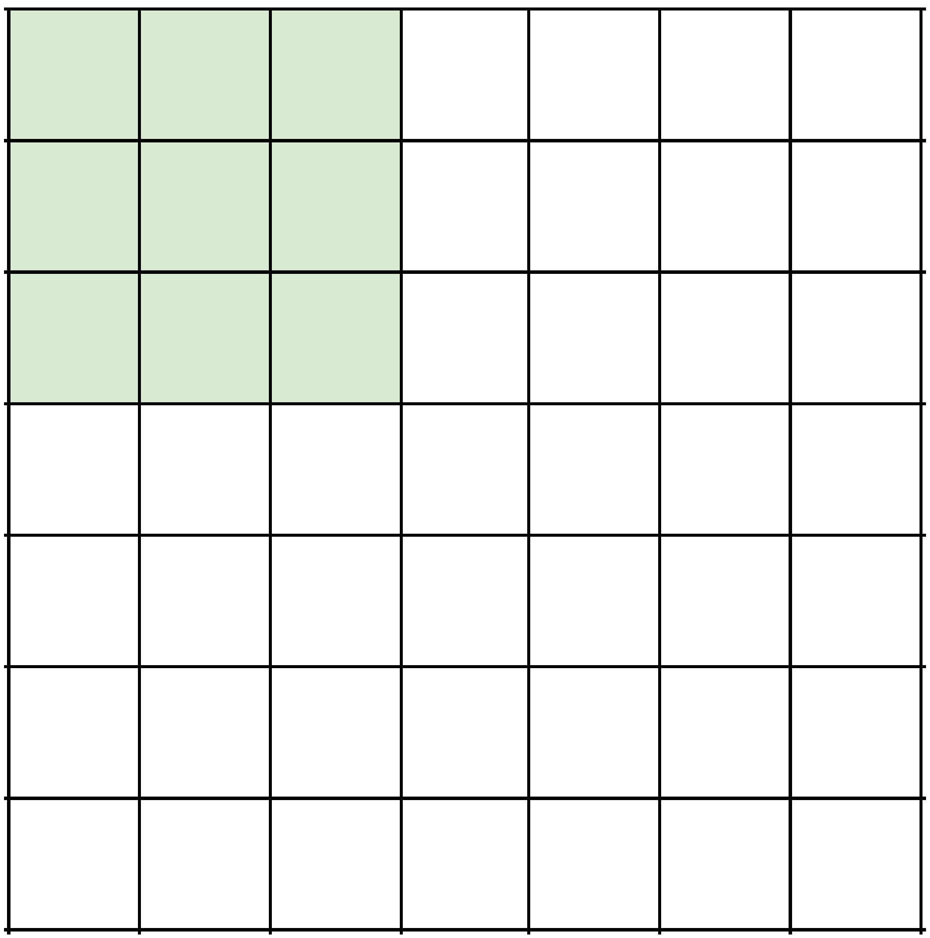
Output: W – K + 1 + 2P

Very common:

Set P = (K – 1) / 2 to make output have same size as input!

합성곱 신경망의 이해

- 스트라이드(Stride)



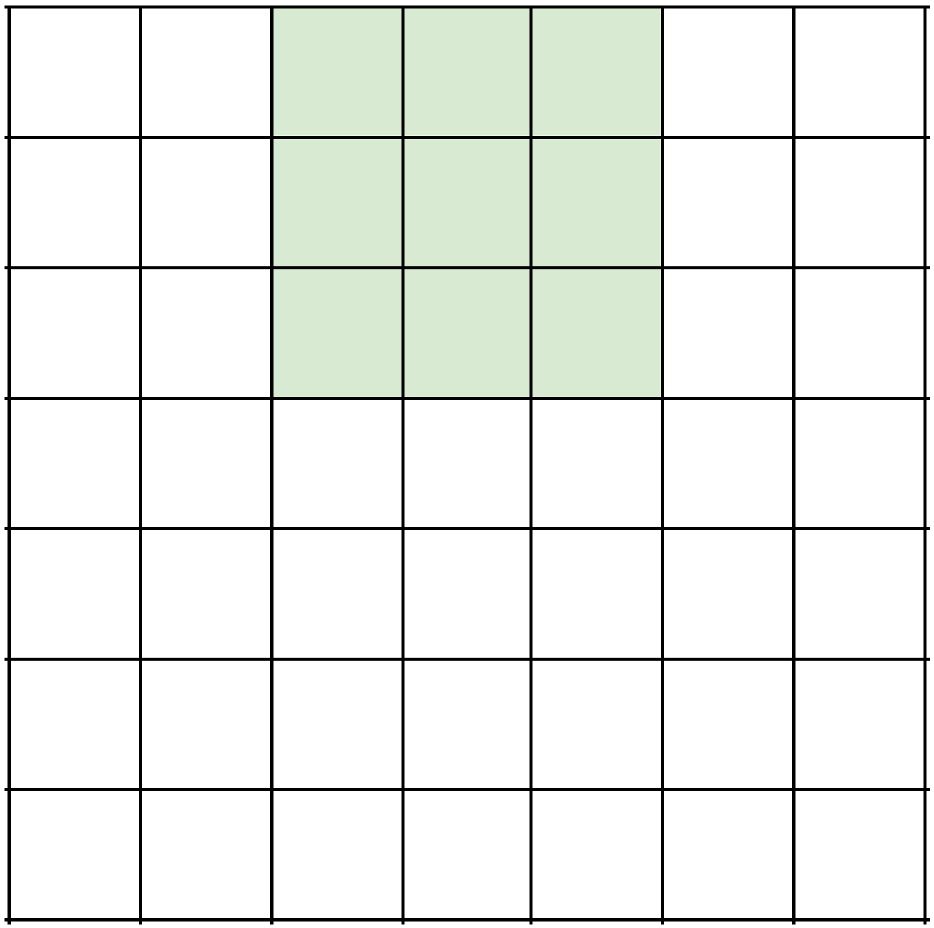
Input: 7x7

Filter: 3x3

Stride: 2

합성곱 신경망의 이해

- 스트라이드(Stride)



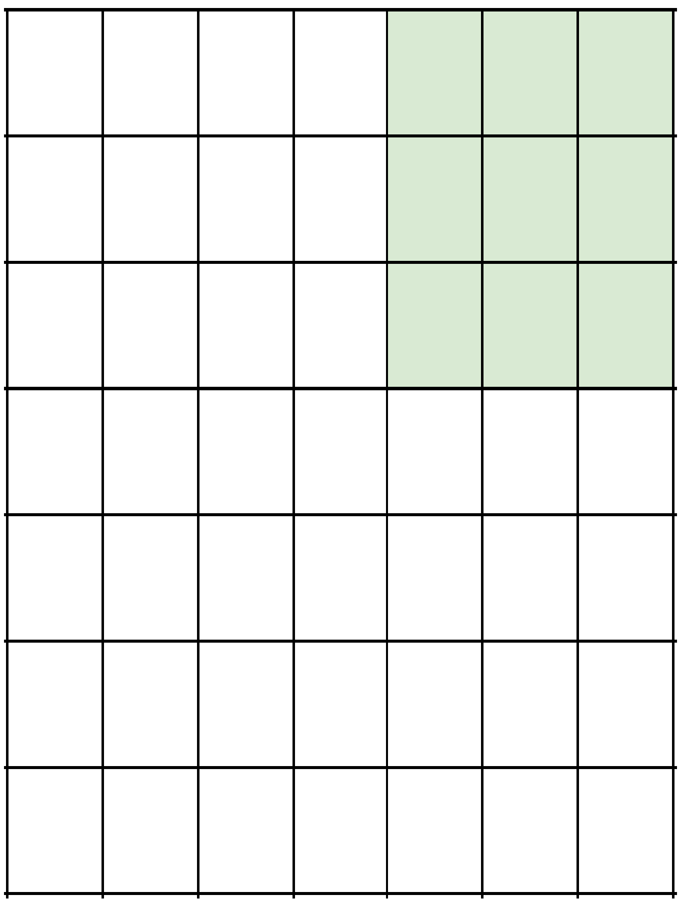
Input: 7x7

Filter: 3x3

Stride: 2

합성곱 신경망의 이해

- 스트라이드(Stride)



Input: 7x7

Filter: 3x3

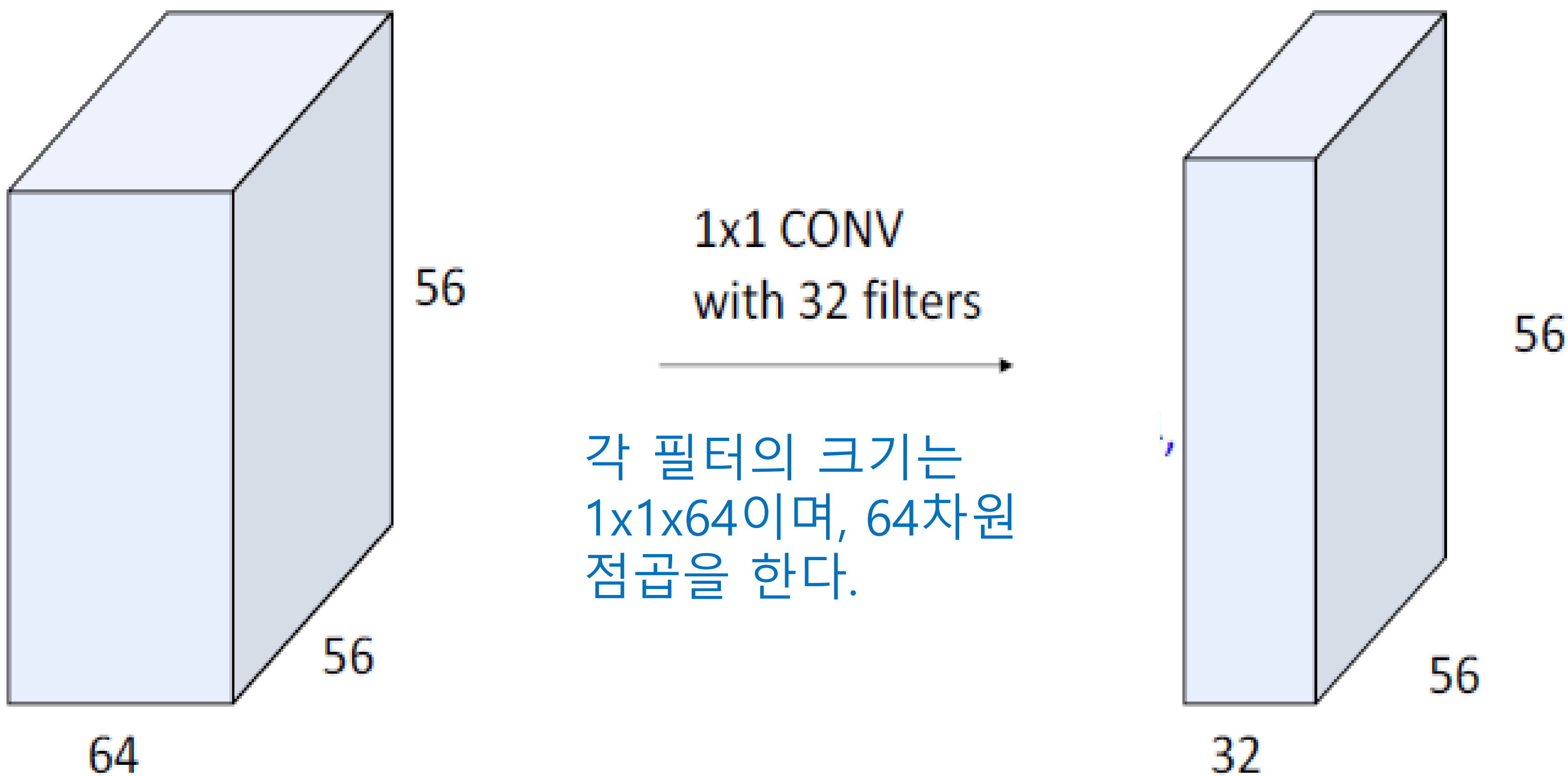
Stride: 2

Output: 3x3

- 일반적으로 입력 크기가 W이고 필터크기가 K이면 패딩이 P 그리고 Stride가 S이면 출력은 (W-K+2P)/S+1이다.

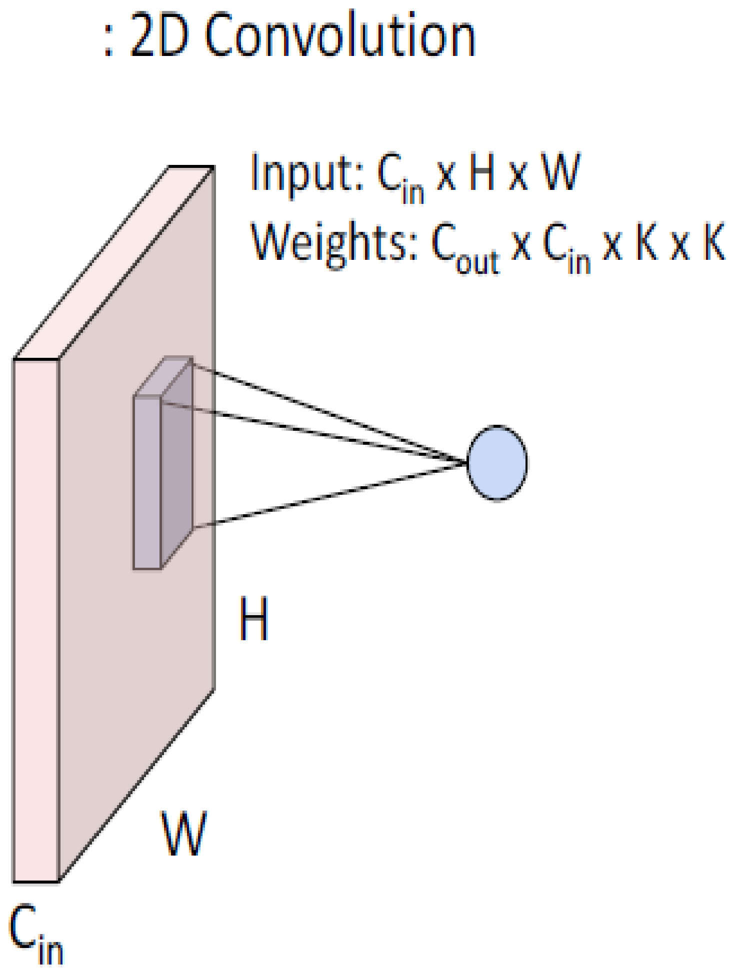
합성곱 신경망의 이해

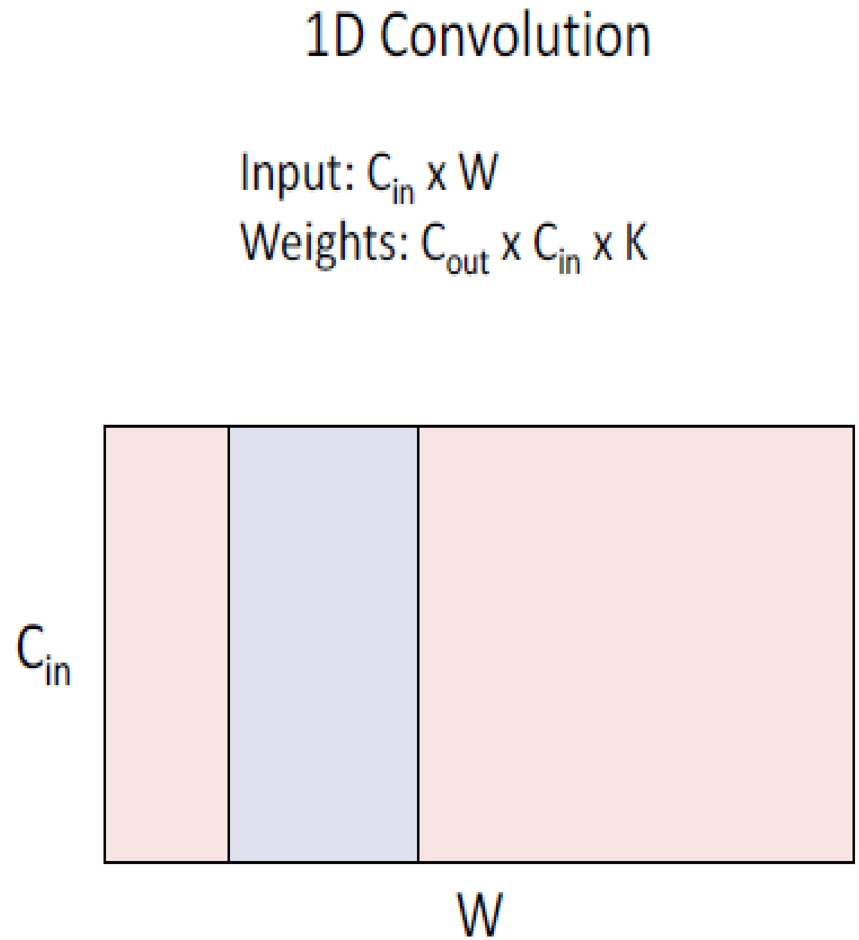
- 1x1 합성곱



합성곱 신경망의 이해

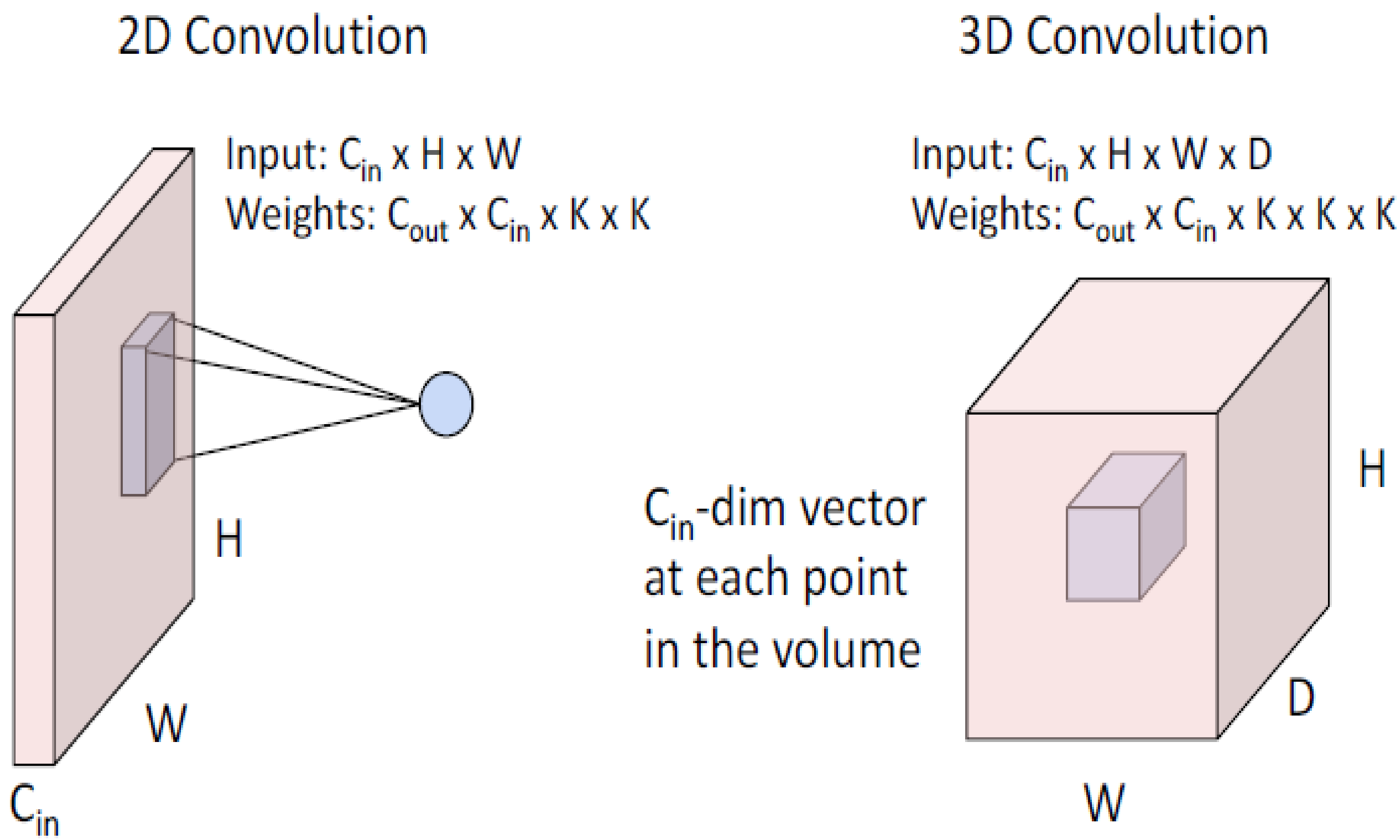
- 2D 대 1D 합성곱





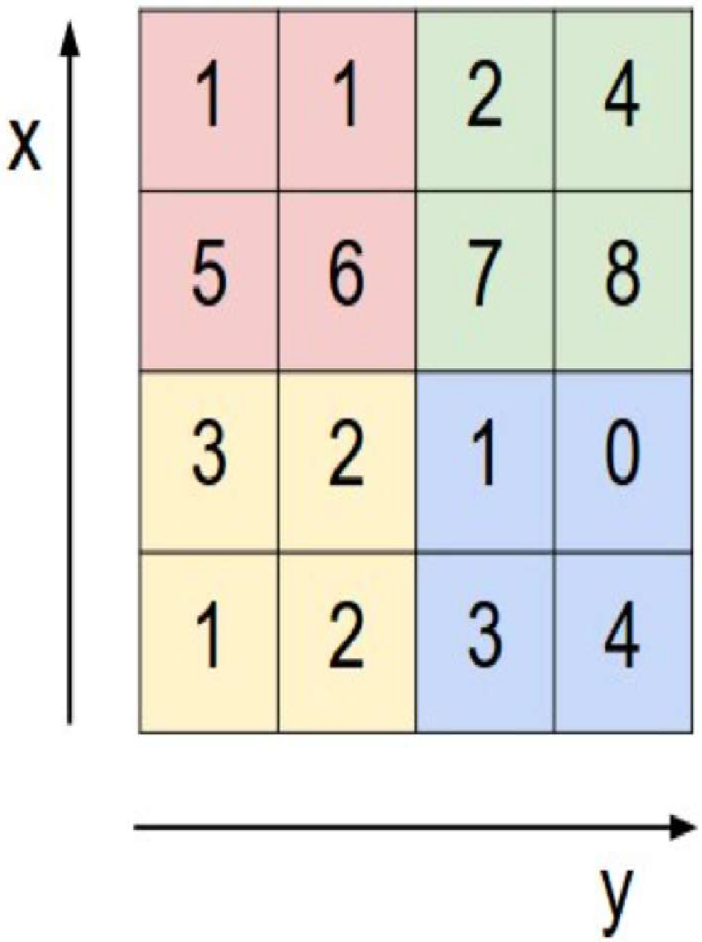
합성곱 신경망의 이해

- 2D 대 3D 합성곱



풀링층의 이해

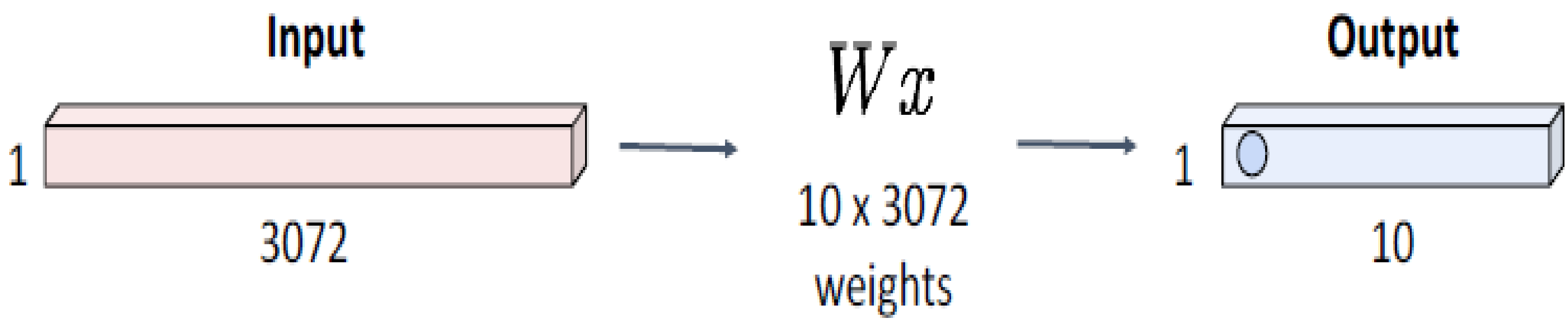
- 풀링: 다운샘플링



1. **차원의 축소**
2. **작은 변화에 대한 불변성**
3. **학습할 파라미터 없다.**

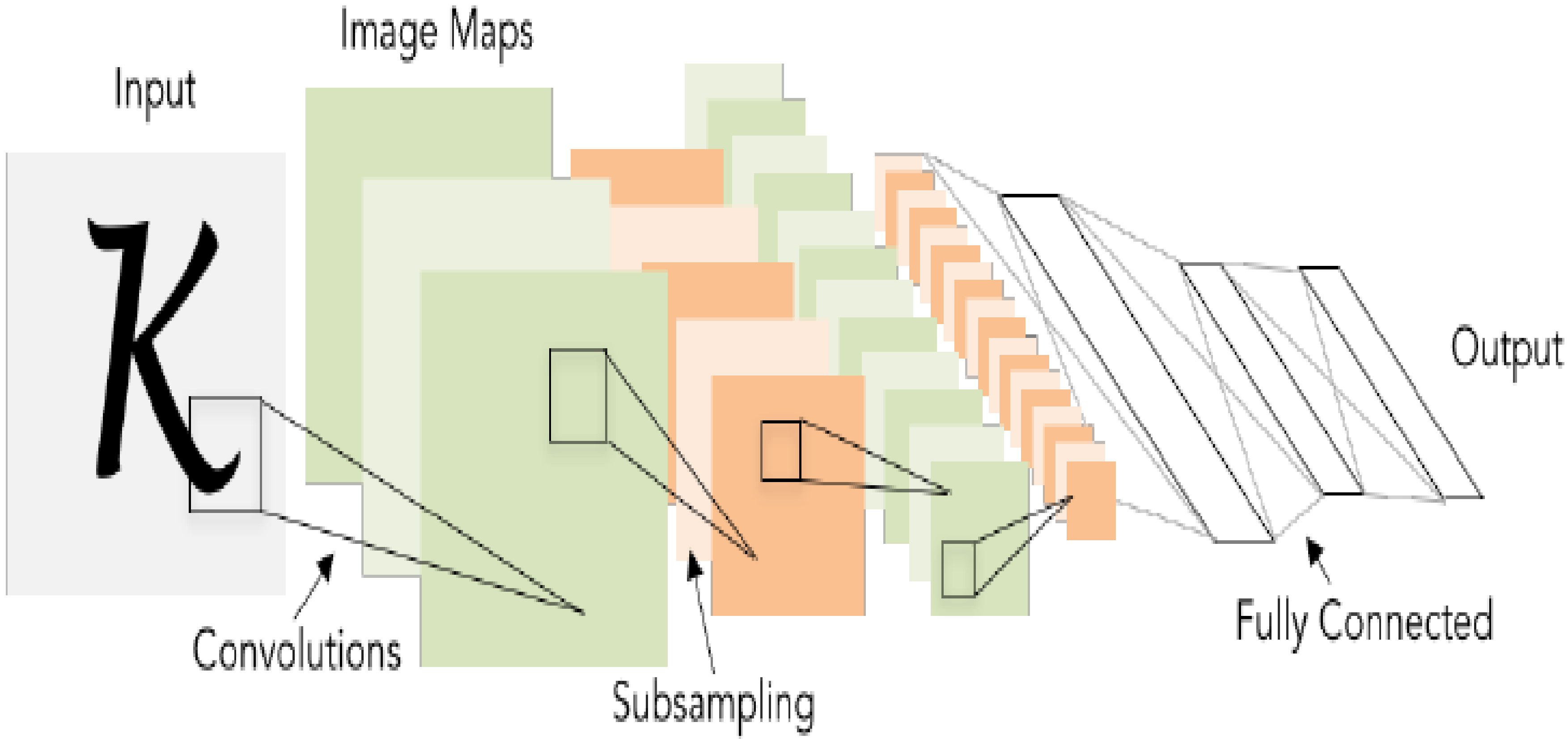
완전결합층 (Fully Connected Layer: FC)

- Fatten -> FC



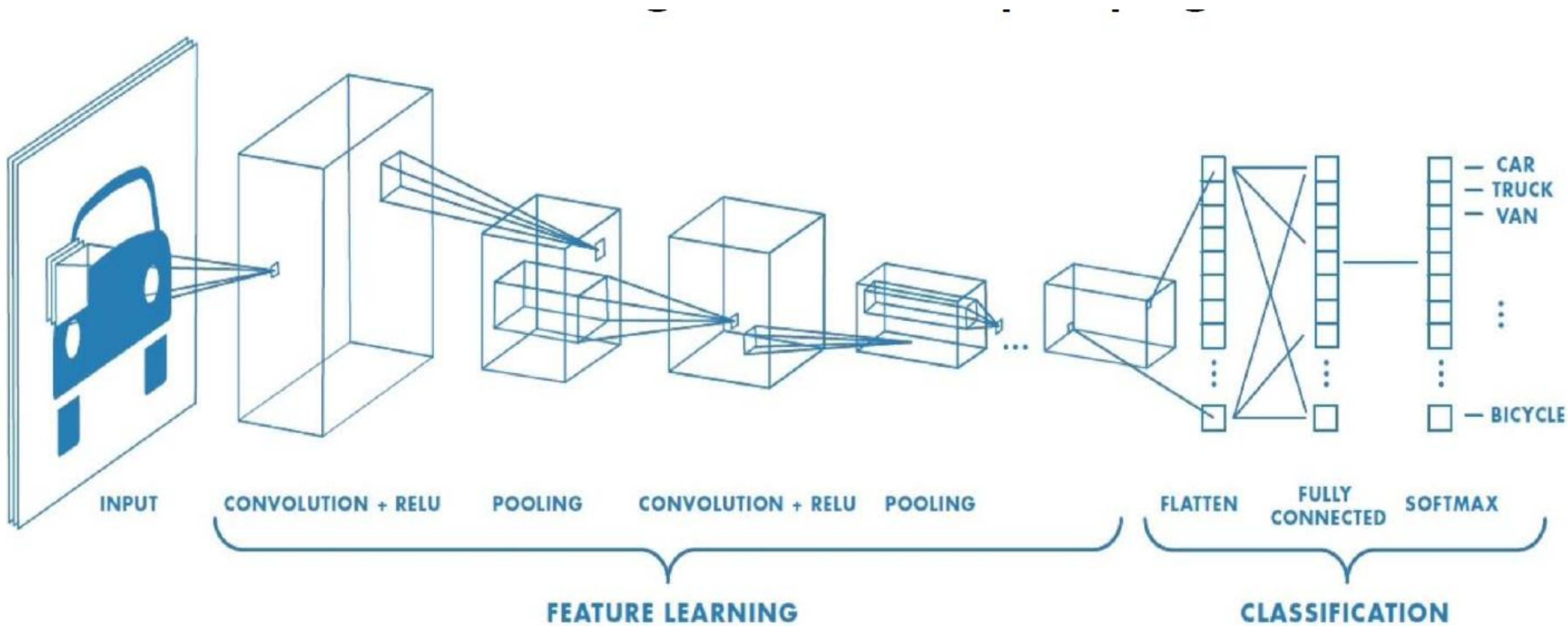
합성곱 신경망의 이해

- 고전적 구조: [Conv, ReLU, Pool]xN, flatten, [FC, ReLU]XN, FC



합성곱 신경망의 이해

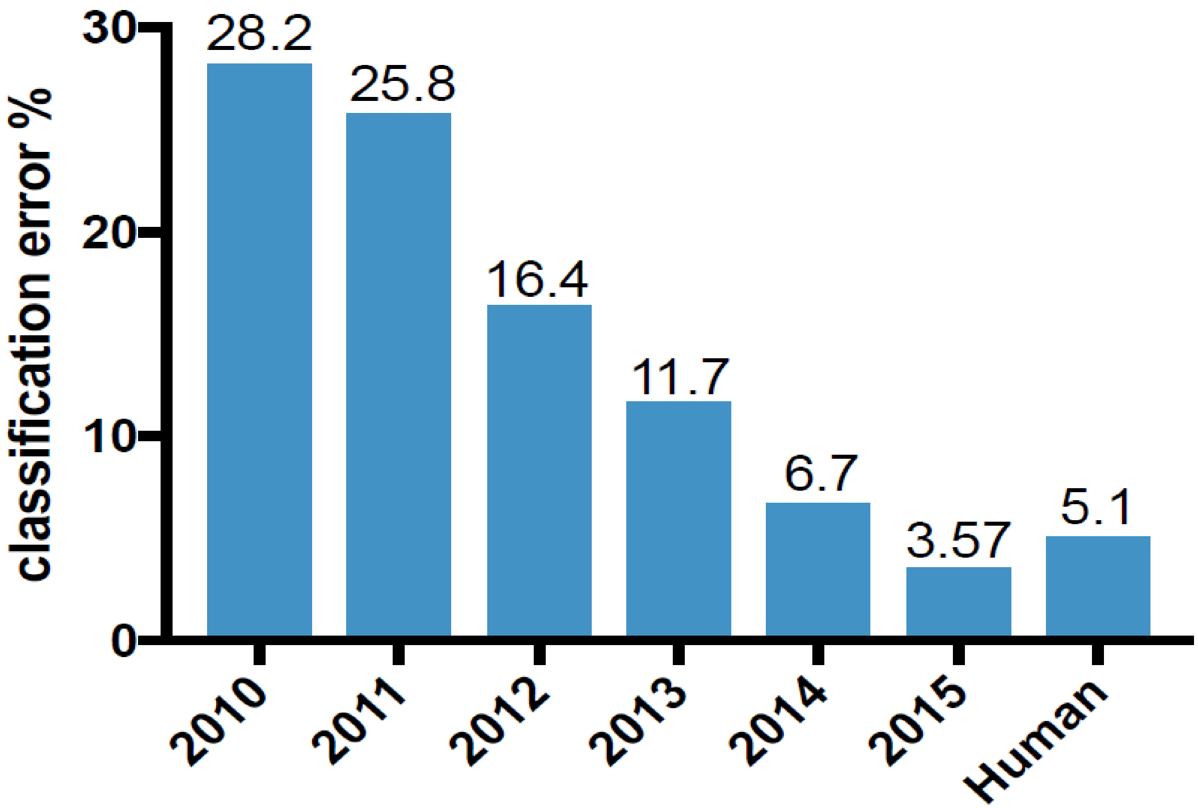
- Softmax와 Cross-Entropy 비용함수를 통한 분류



$$\mathrm{softmax}(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} \qquad J(\boldsymbol{\theta}) = \sum_i y^{(i)} \log\left(\hat{\boldsymbol{y}}^{(i)}\right)$$

합성곱 신경망의 발전

- ImageNet 대회



2012: AlexNet. First CNN to win.
- 8 layers, 61 million parameters

2013: ZFNet
- 8 layers, more filters

2014: VGG
- 19 layers

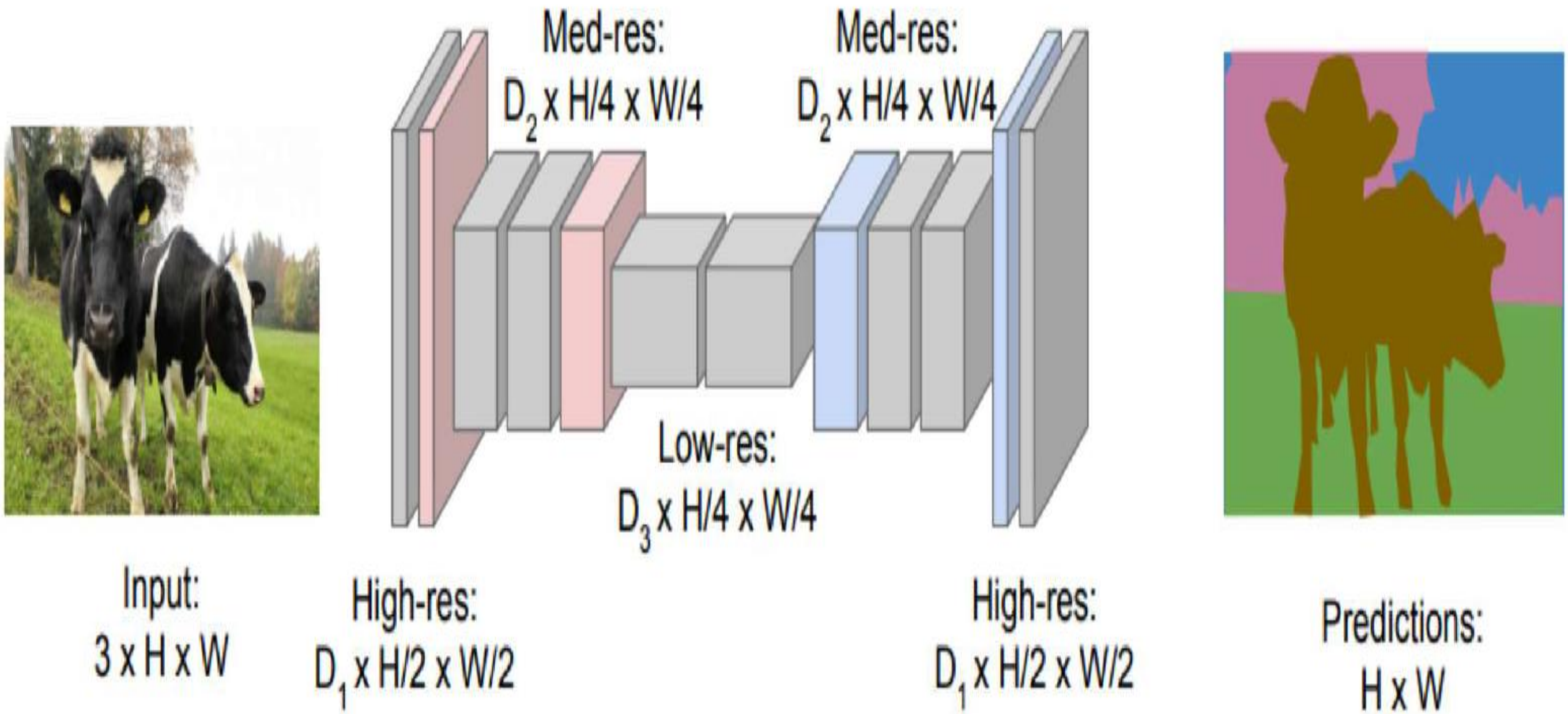
2014: GoogLeNet
- "Inception" modules
- 22 layers, 5million parameters

2015: ResNet
- 152 layers

업샘플링 (전치 합성곱의 이해)

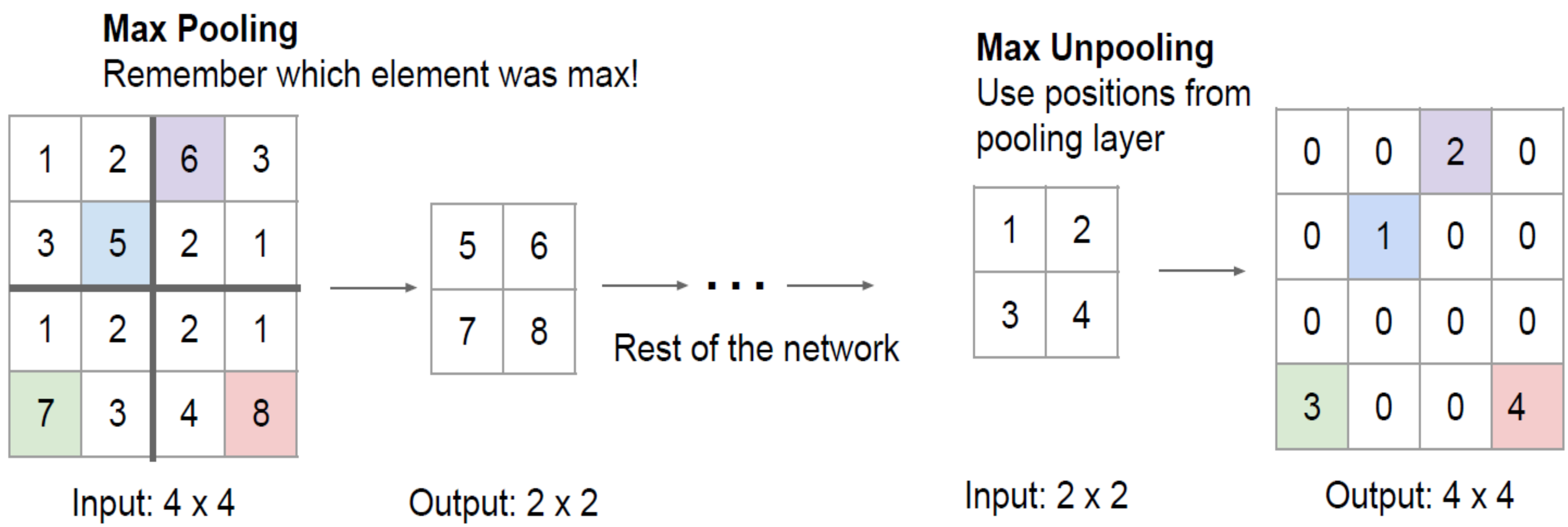
업샘플링

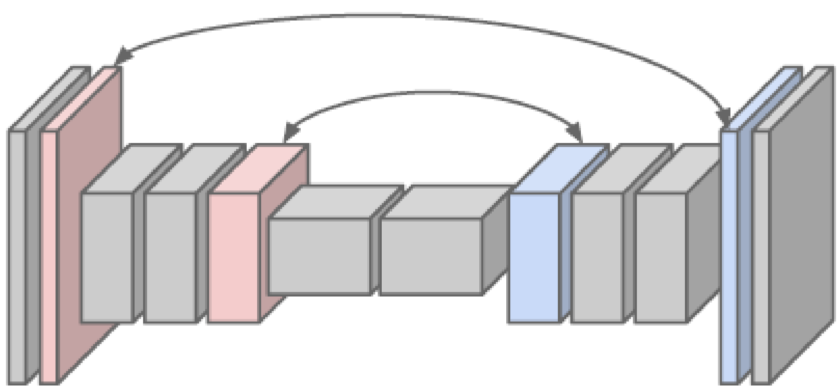
- ImangeNet 대회: 이미지 분할



업샘플링

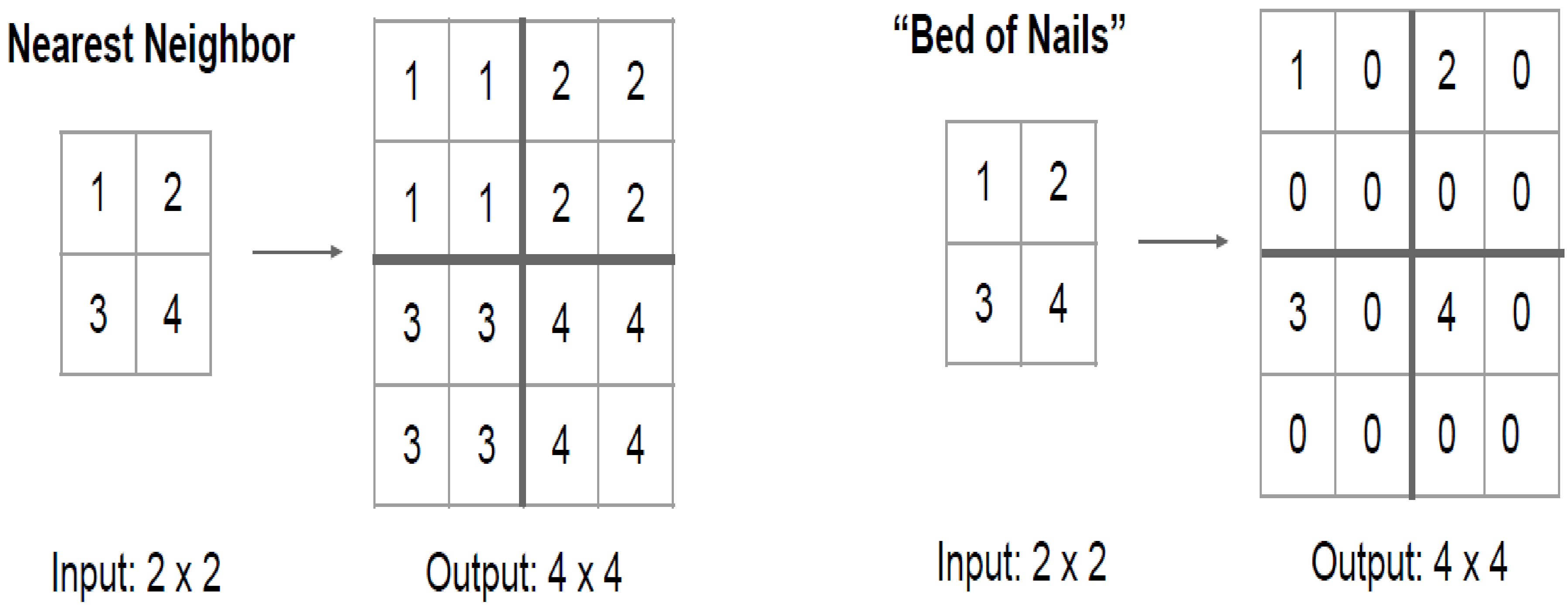
- Max Unpooling: 다운샘플링과 업샘플링이 대칭이 되도록 Up&Un샘플링





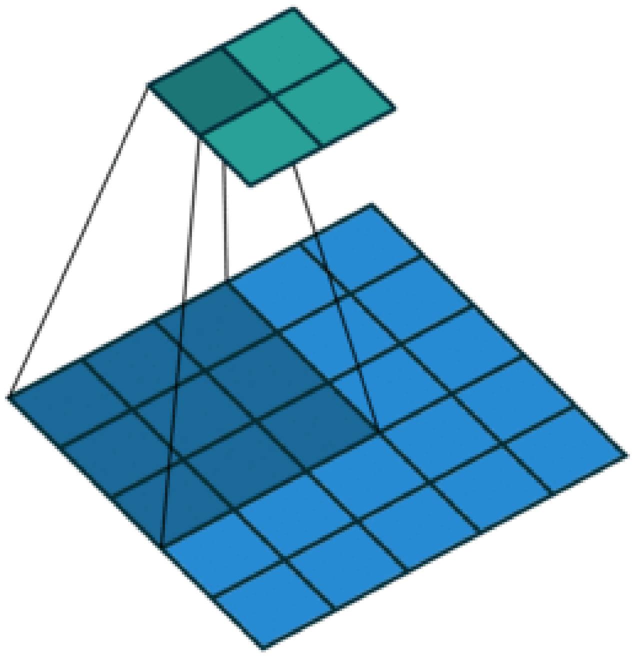
업샘플링

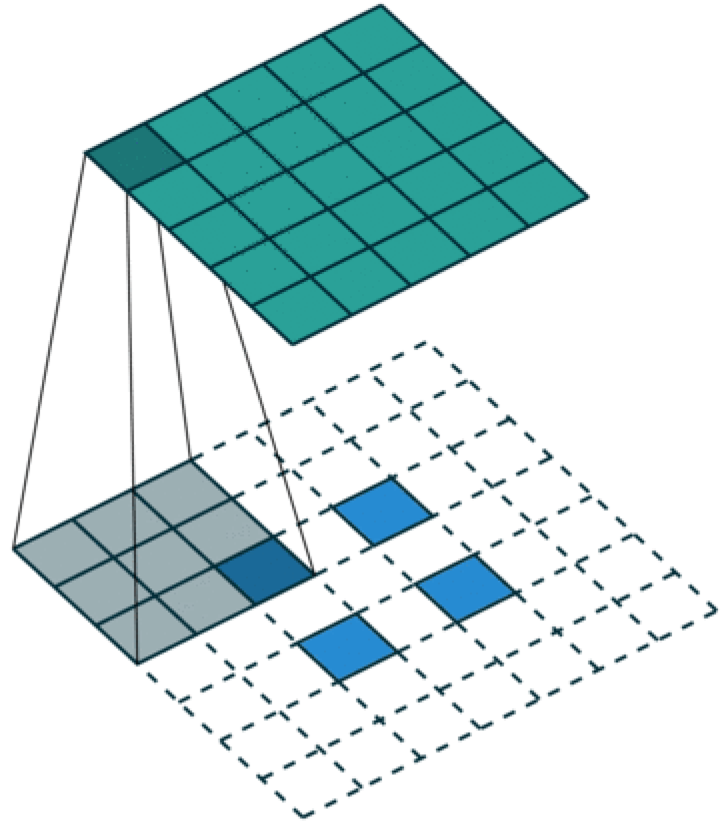
- 언풀링 (풀링을 푸는 것)



업샘플링: 전치 합성곱

- 원합성곱과 전치합성곱

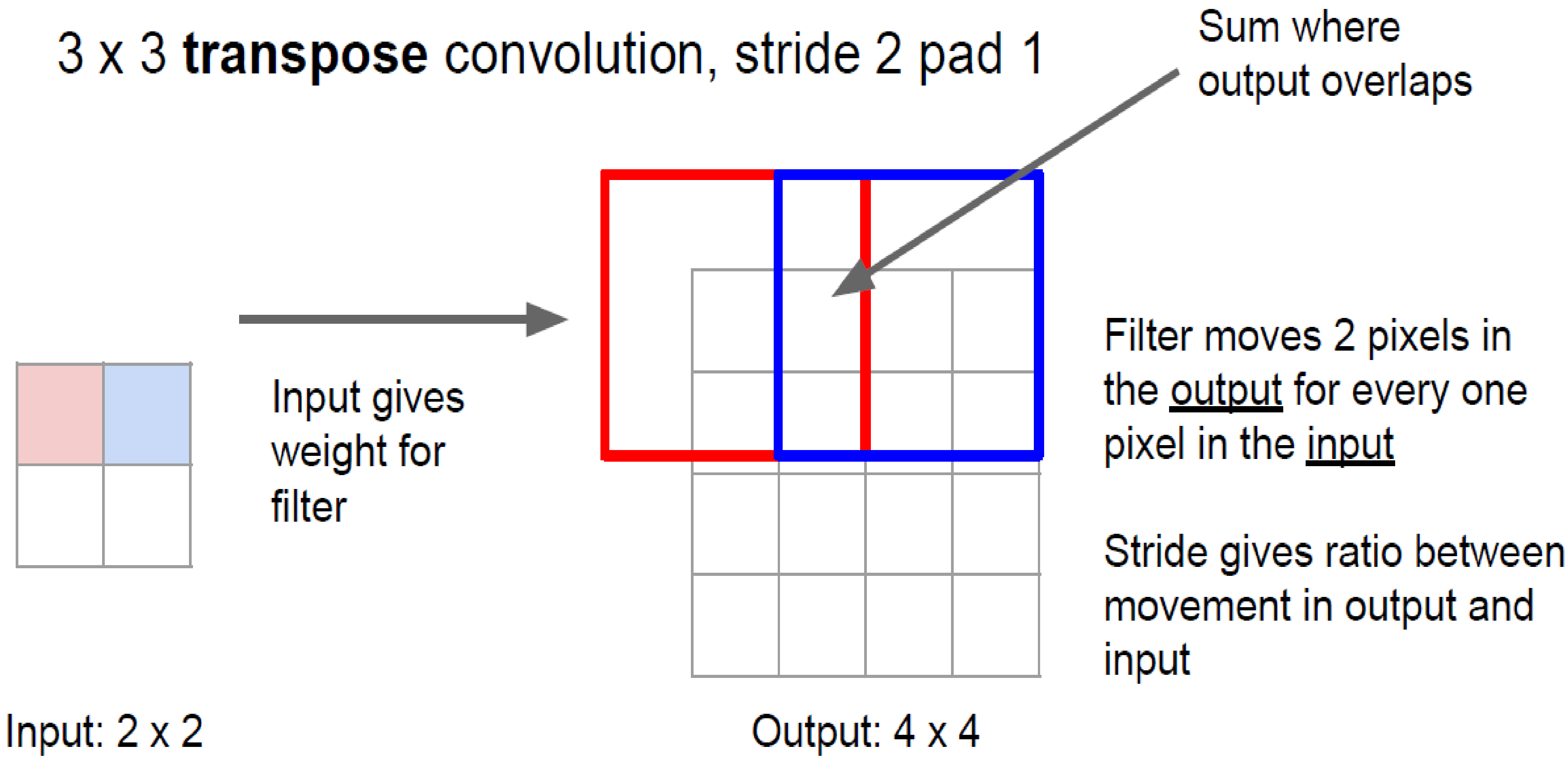




- 원합성곱(3x3 필터의 0 패딩, 스트라이드 2 합성곱
- 2x2의 출력에서 5x5의 입력으로 되돌리는 전치합성곱

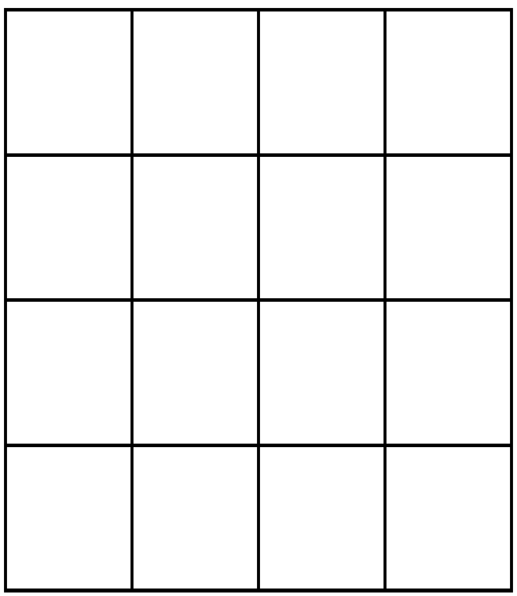
업샘플링

- 전치 업샘플링의 여러이름: transposed convolution, Deconvolution, Upconvolution, Fractionally strided convolution, Backward strided convolution

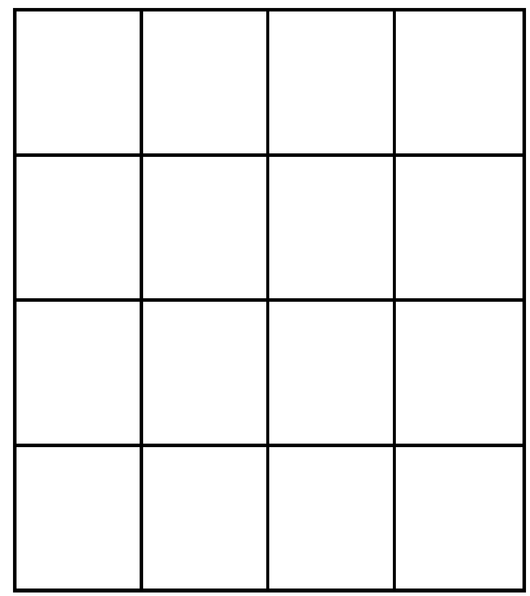


업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 1, 패드 1



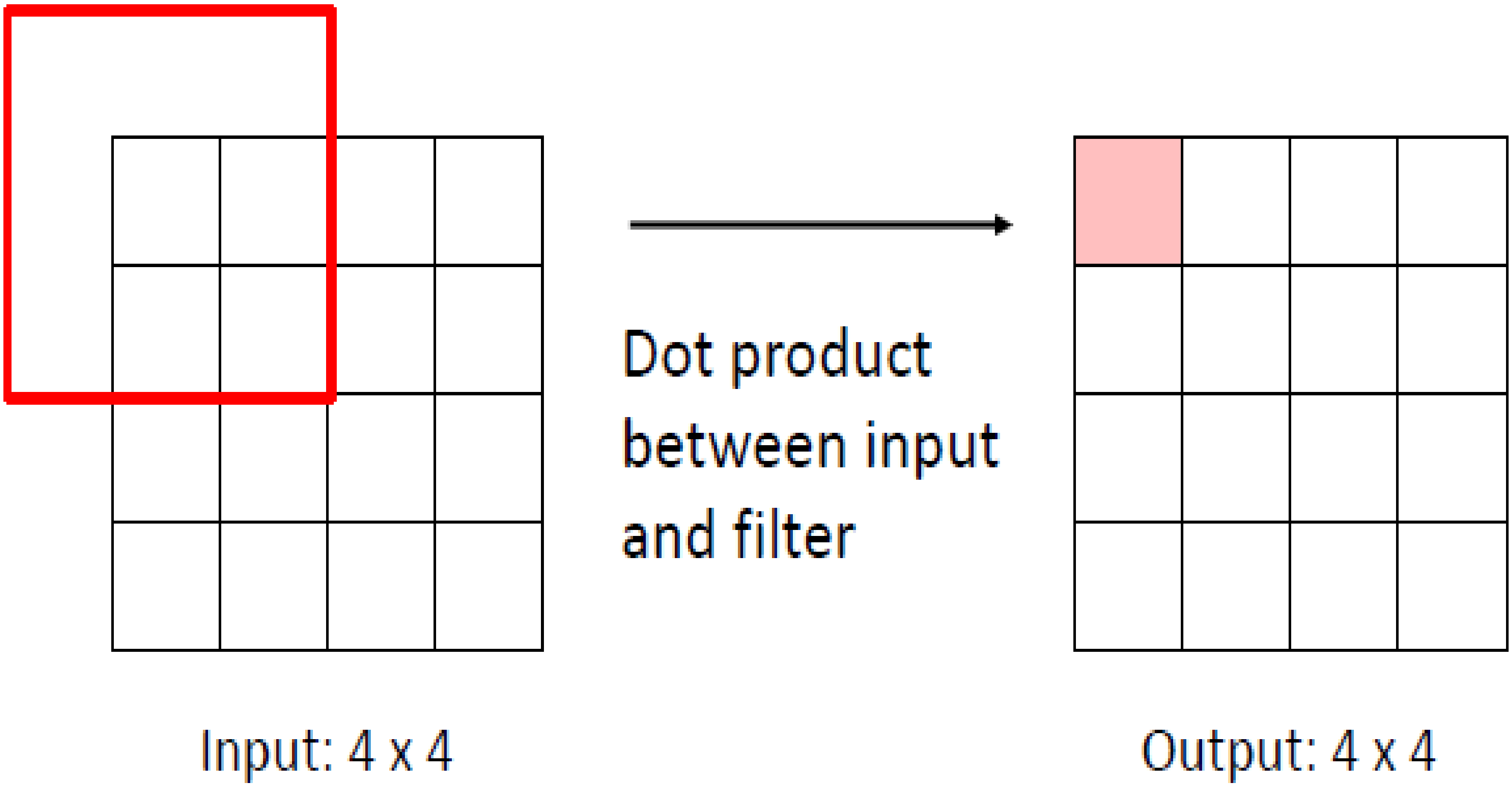
Input: 4 x 4



Output: 4 x 4

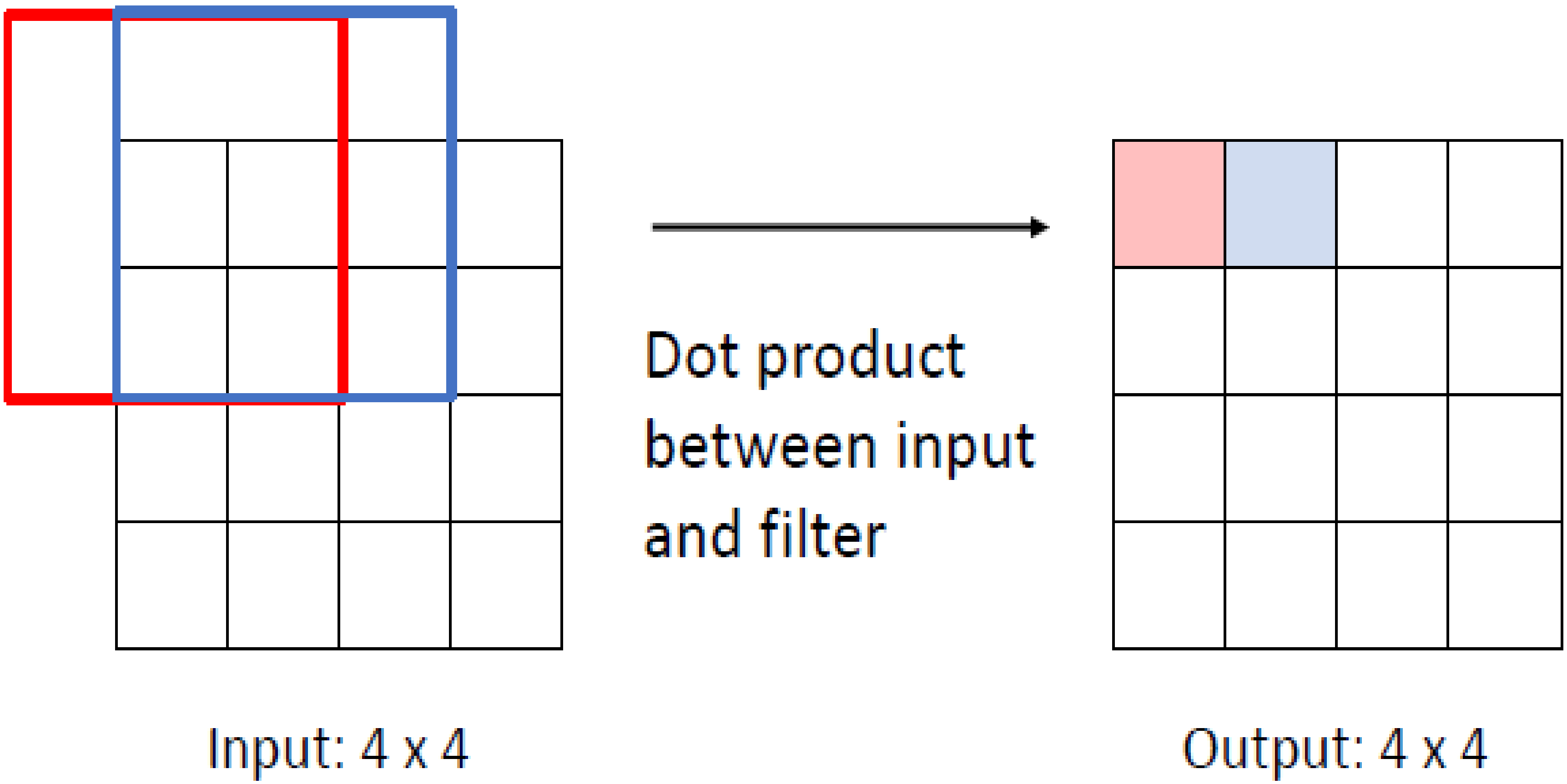
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 1, 패드 1



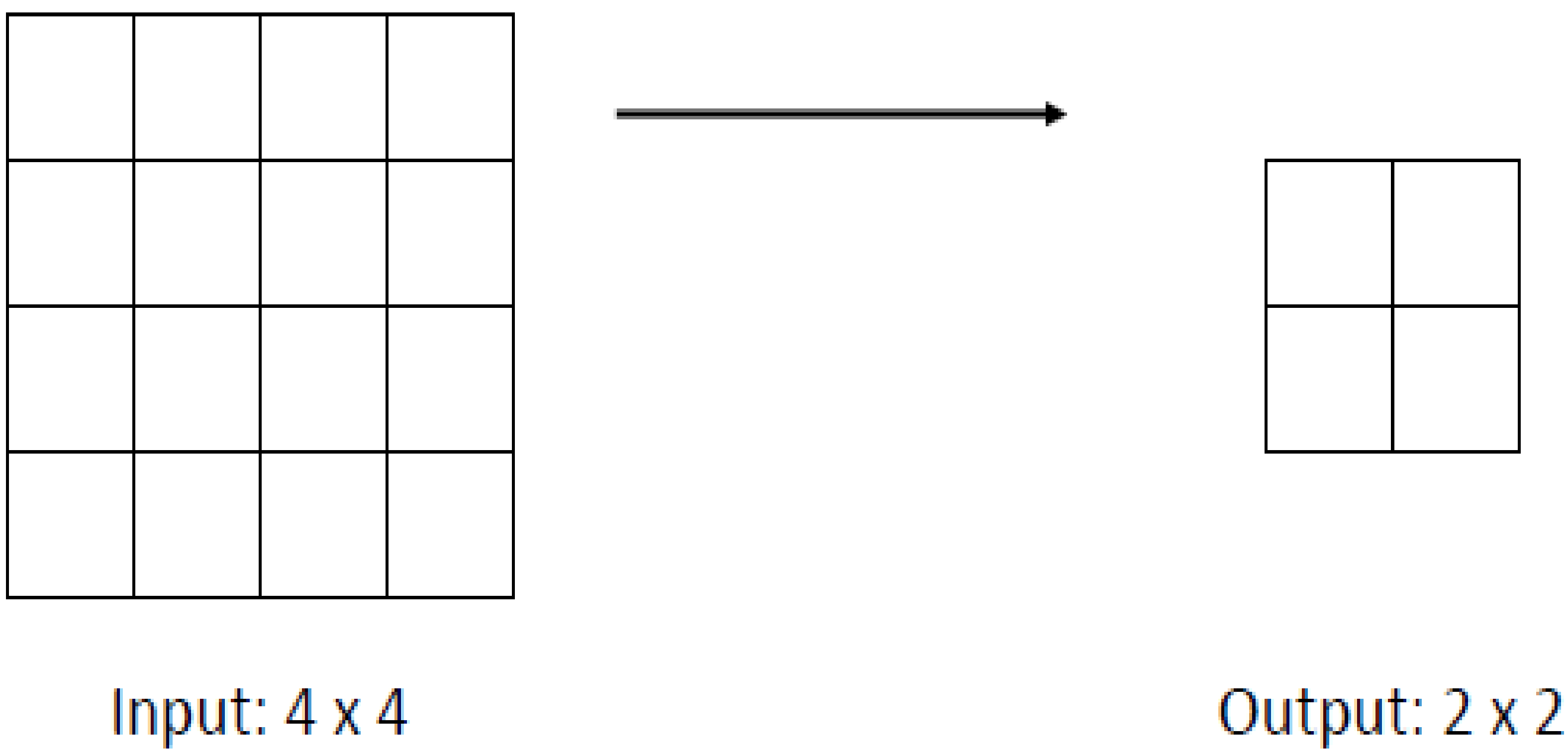
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 1, 패드 1



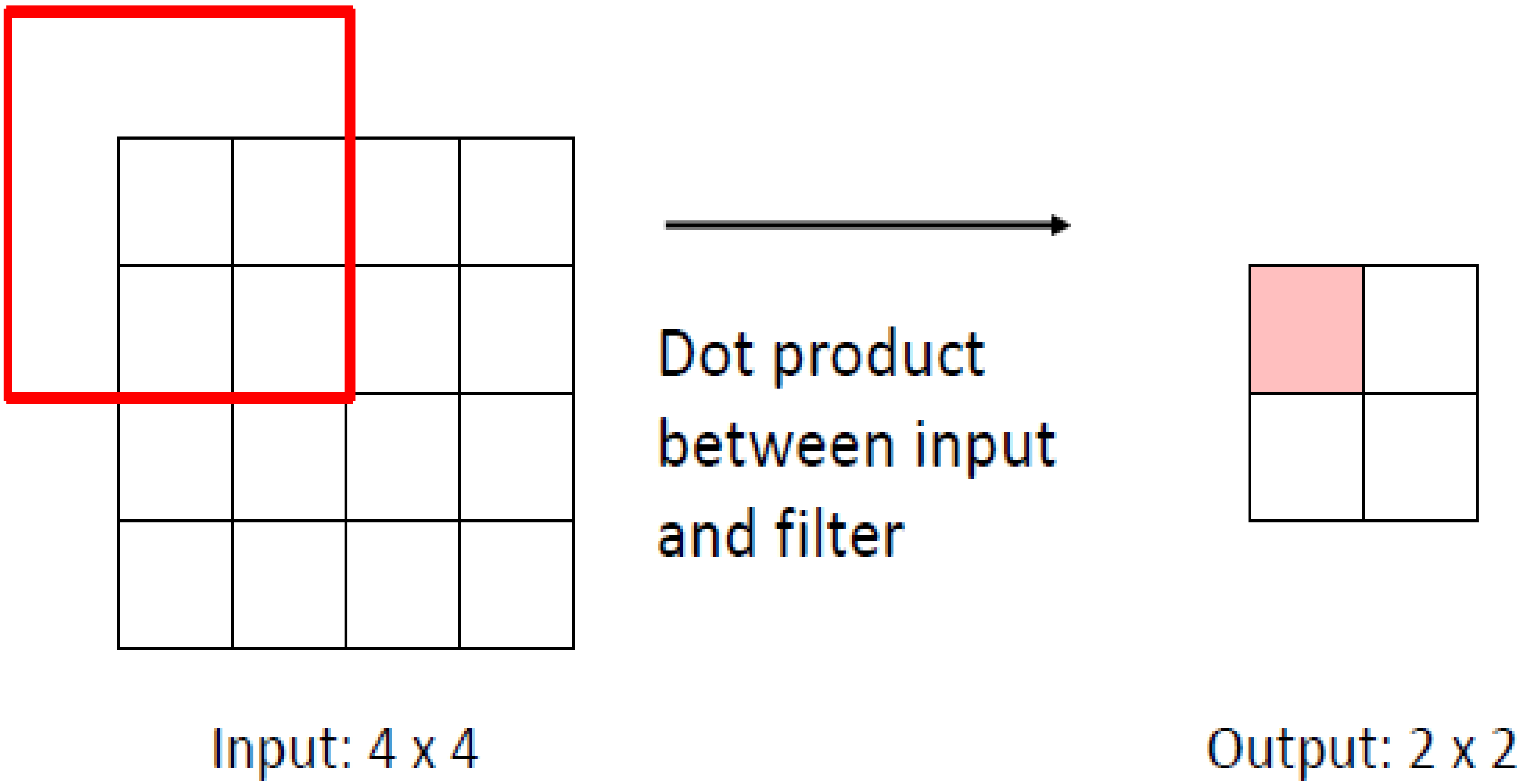
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2, 패드 1



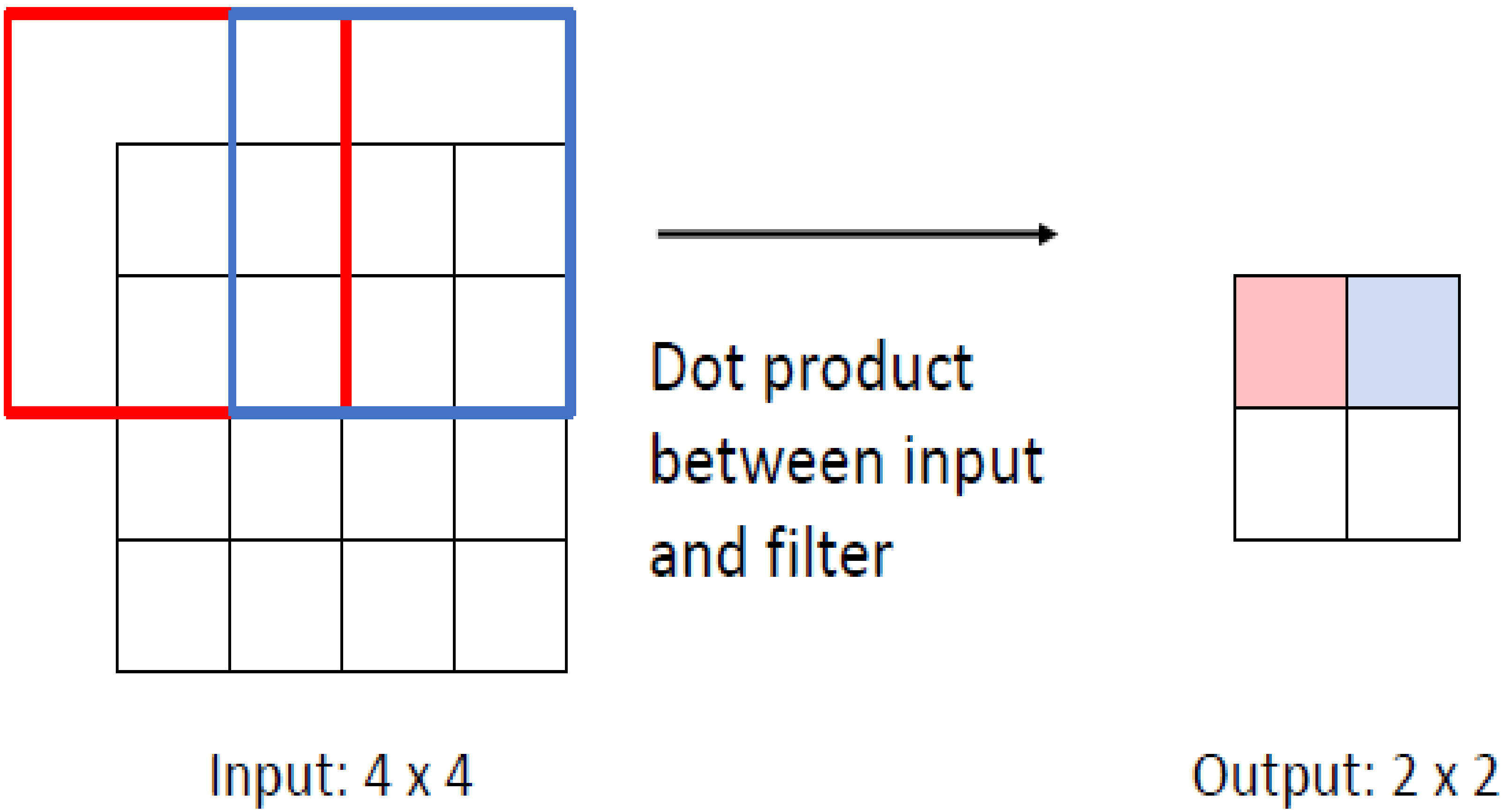
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2, 패드 1



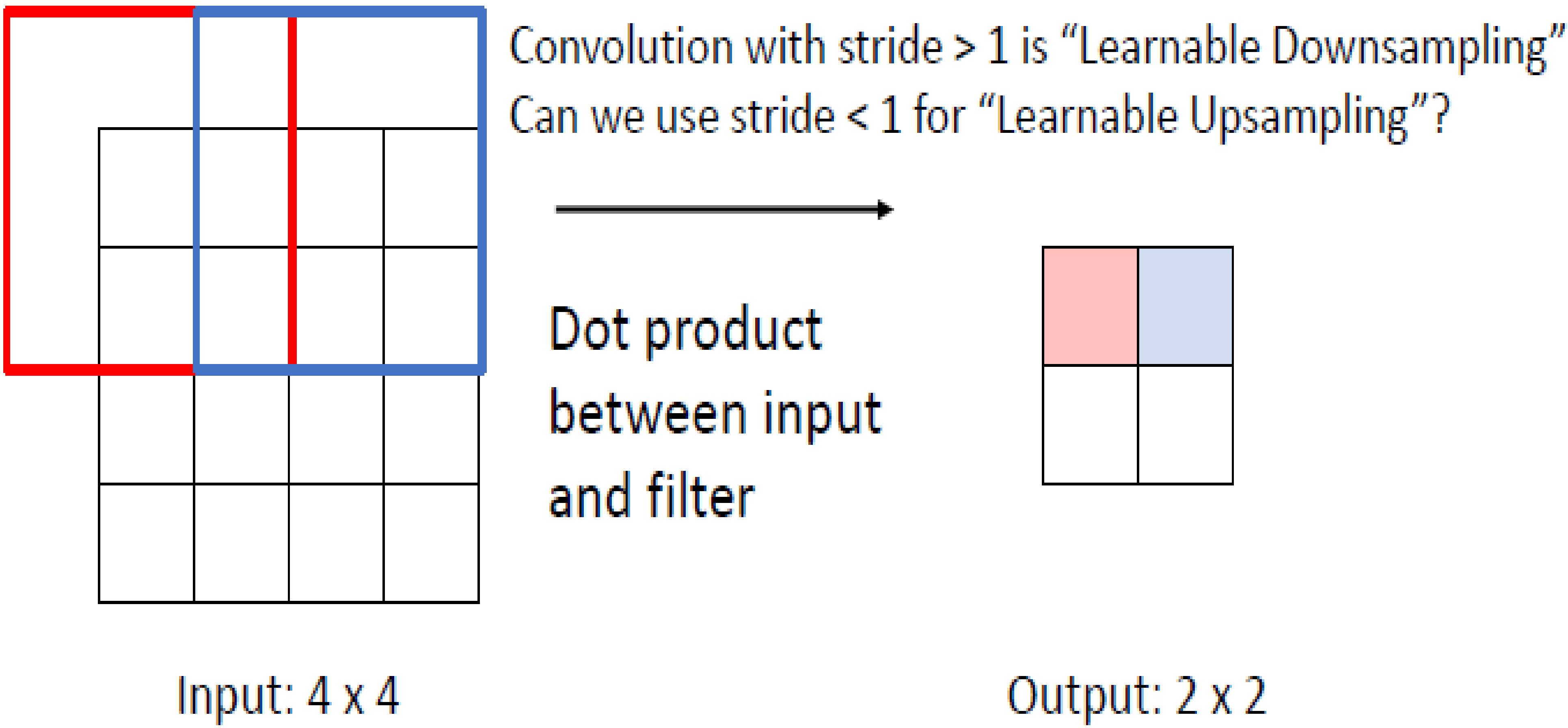
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2, 패드 1



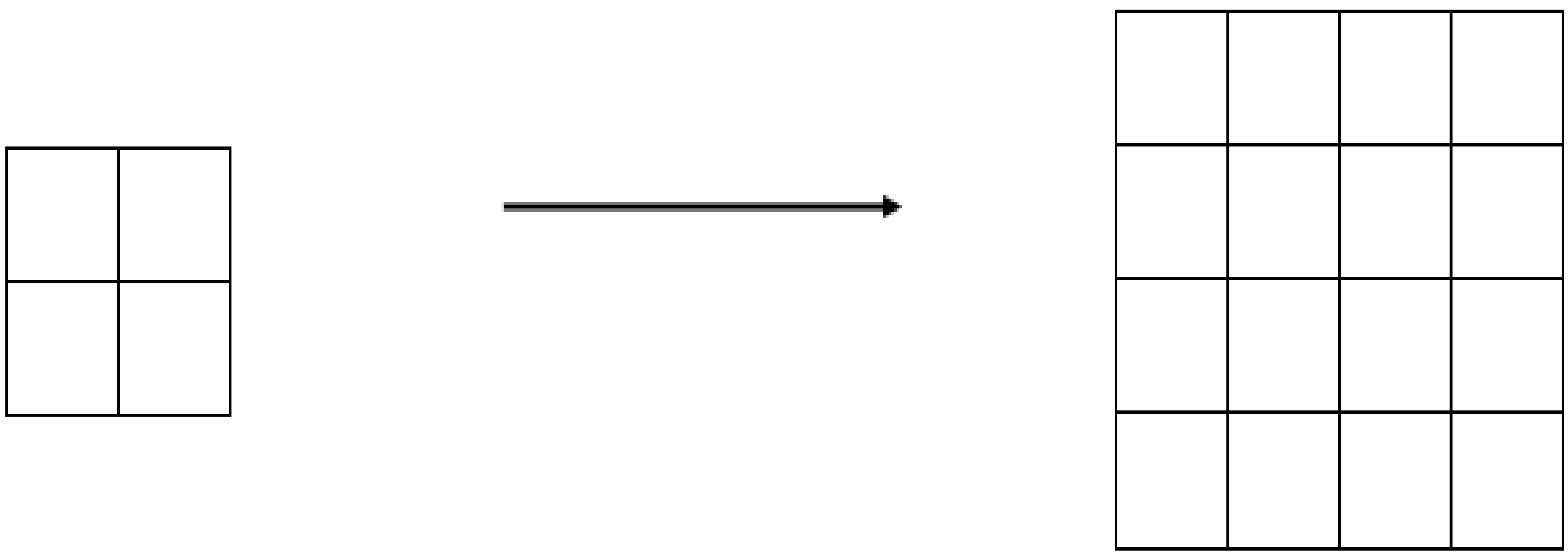
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2, 패드 1



업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2

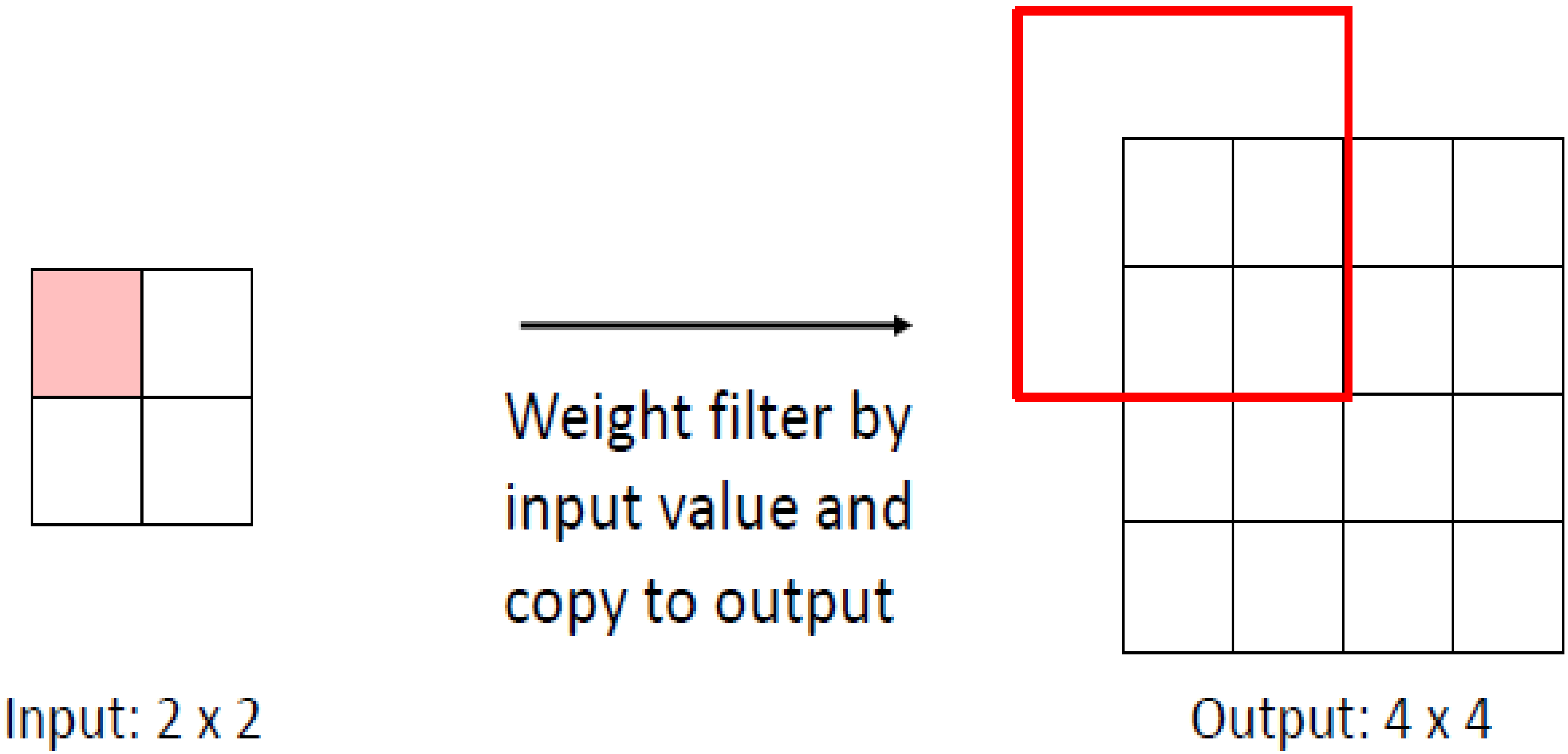


Input: 2 x 2

Output: 4 x 4

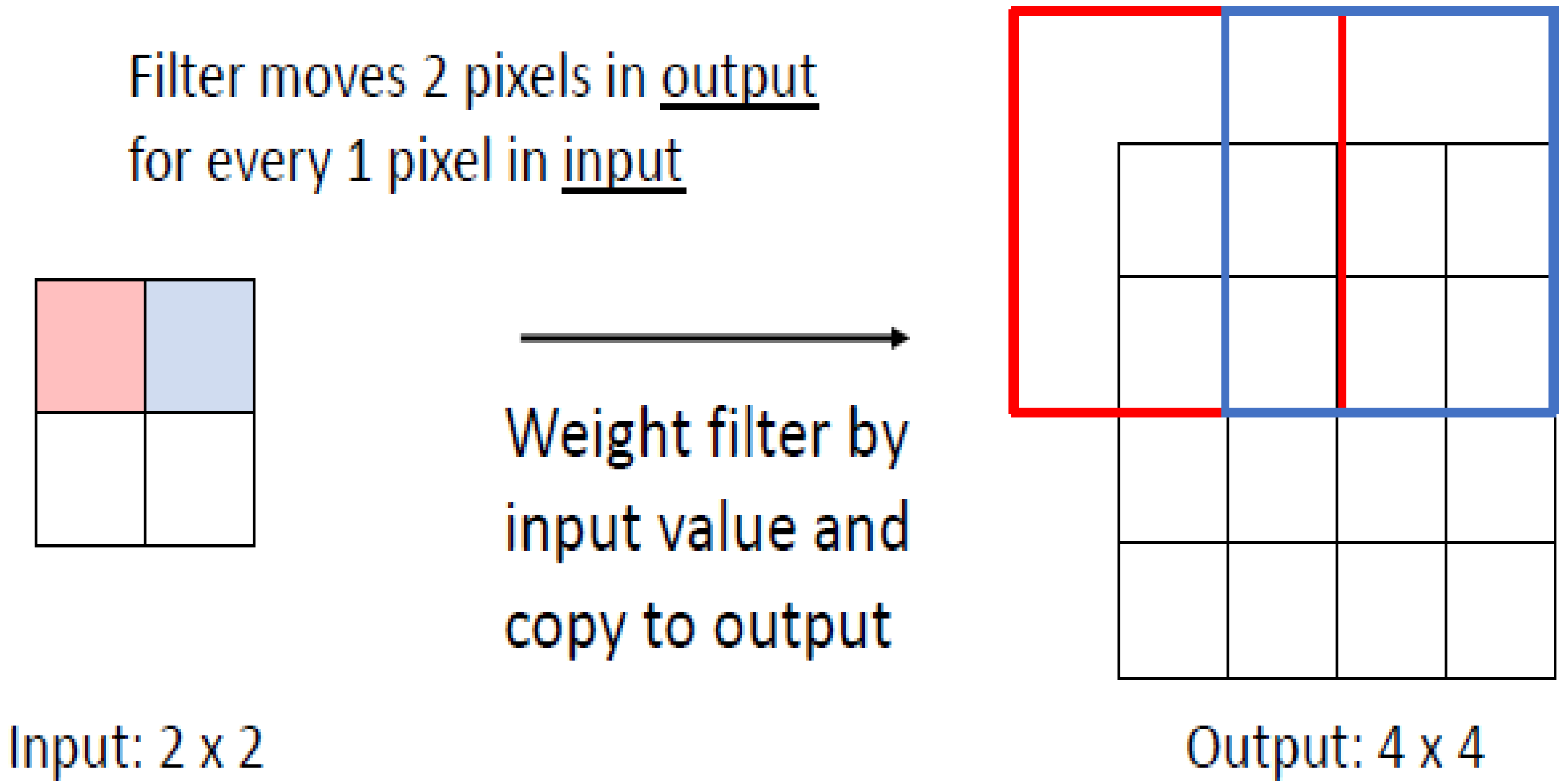
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2



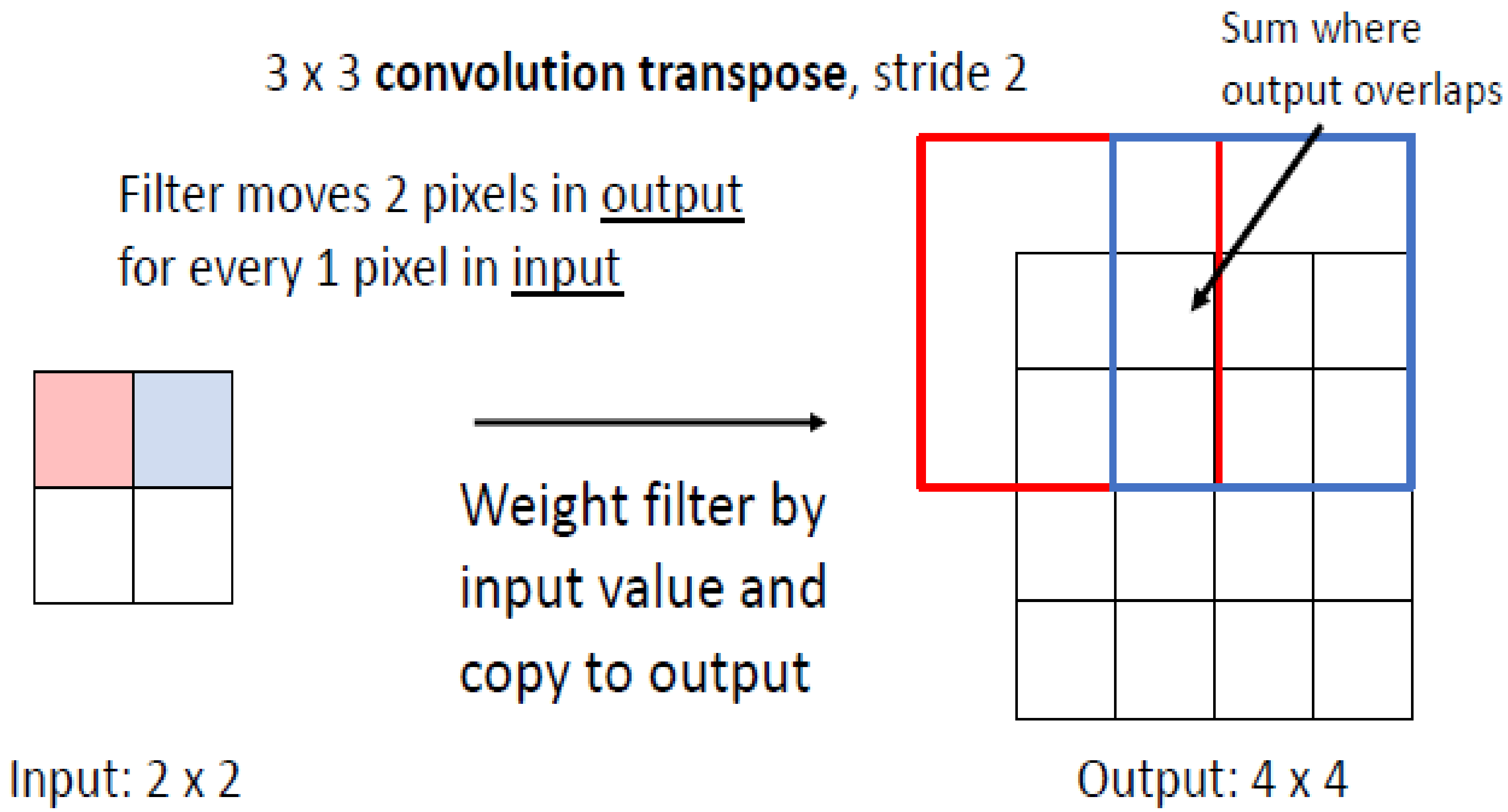
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2



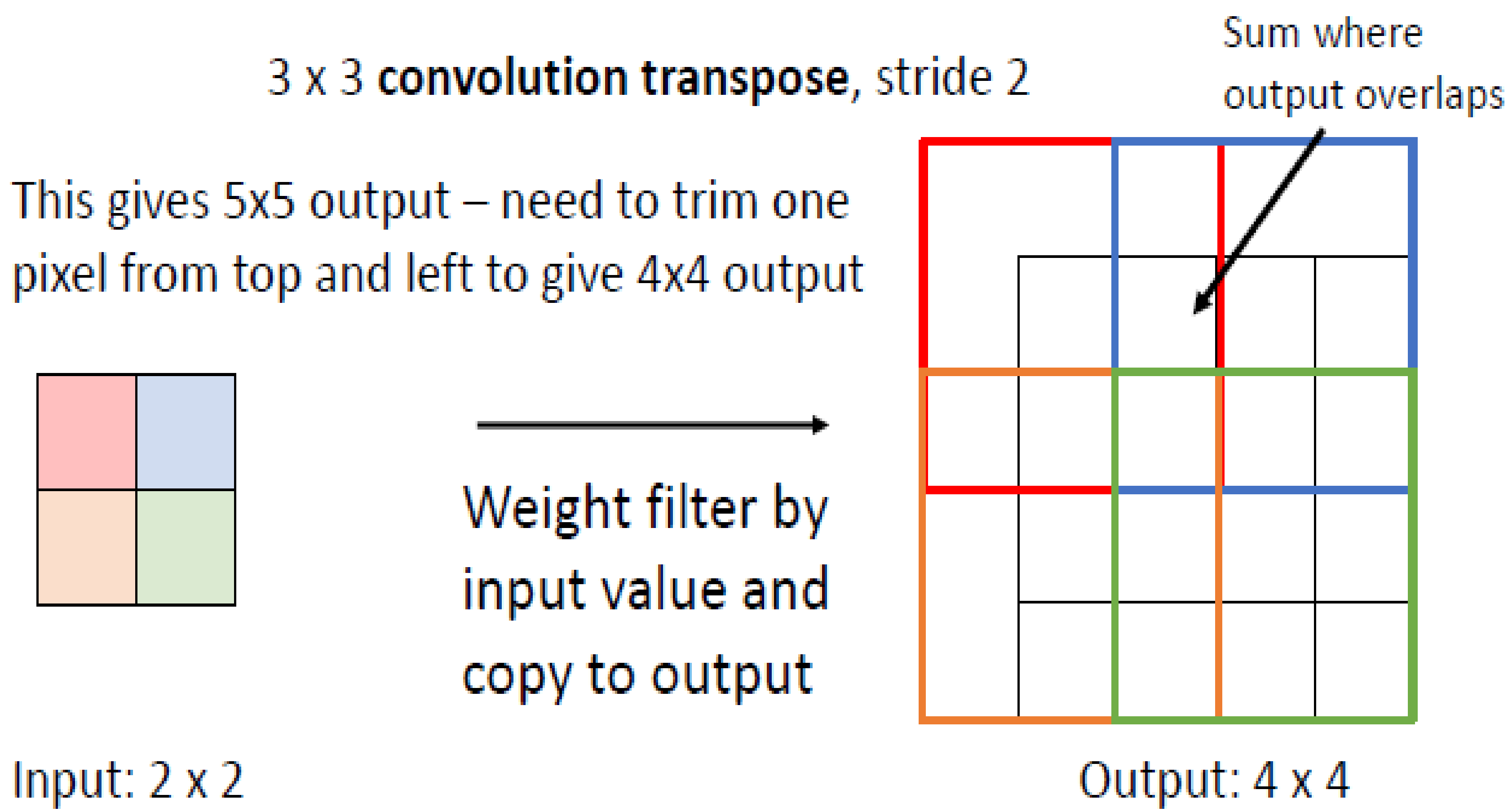
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2



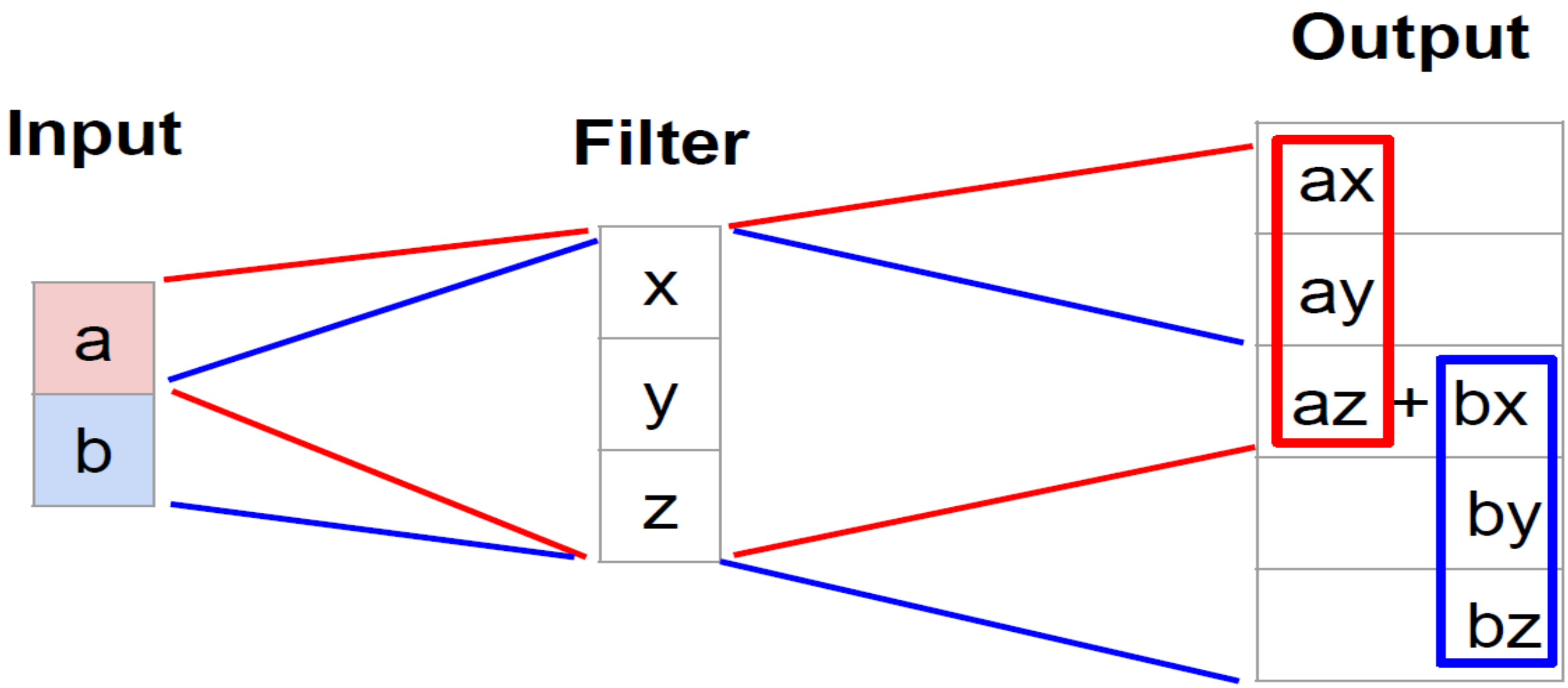
업샘플링: 전치 합성곱

- 3x3 합성곱, 스트라이드 2



업샘플링

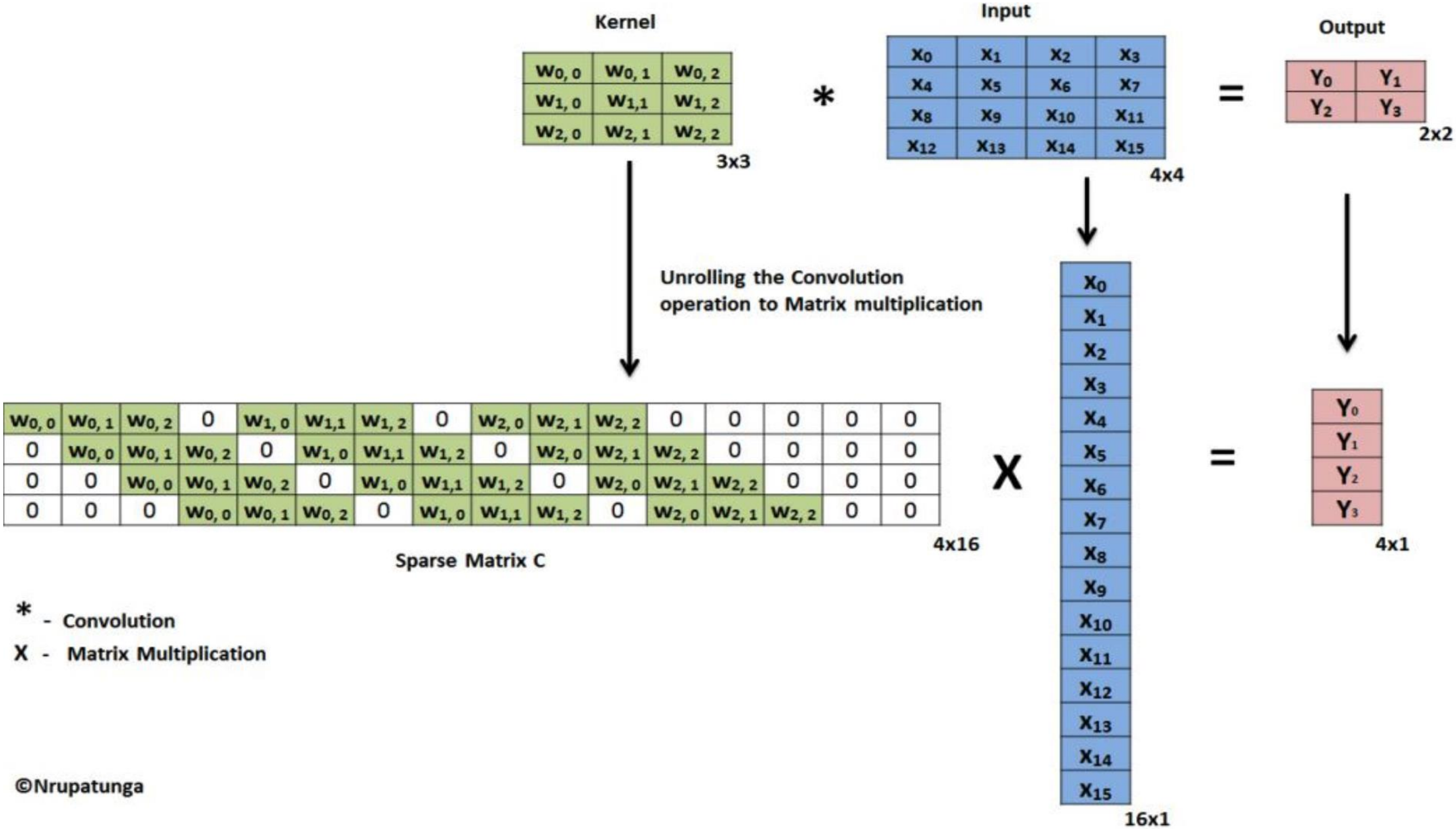
- Transpose 업샘플링: 1d 예시



- 위의 경우, 3X1 필터로 stride=2인 경우로 정확하게 2배 (2x2=4)로 만들기 위해서는 output의 한 픽셀을 제거할 필요가 있음.

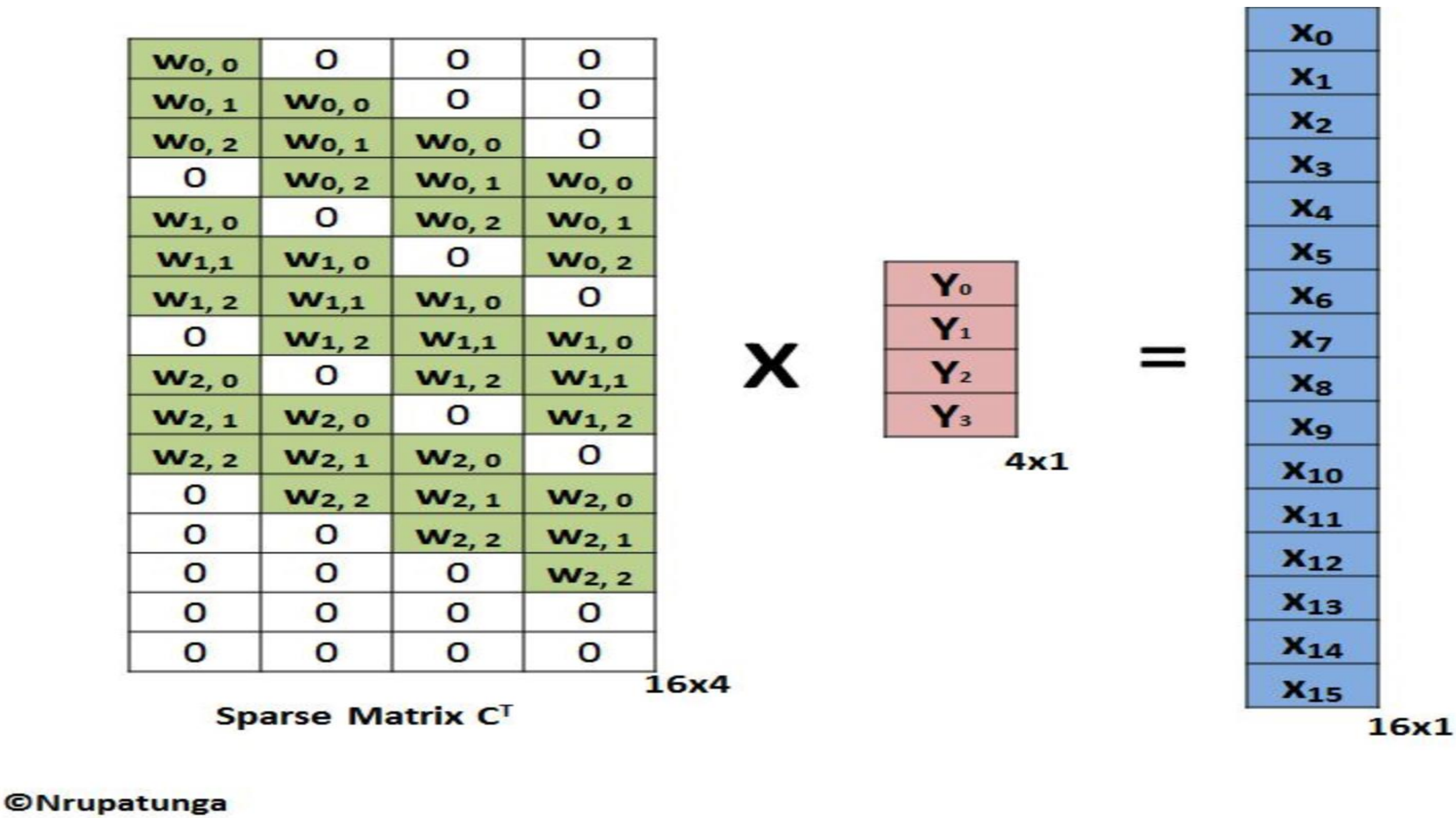
전치 업샘플링

- 왜 전치인가?



전치 업샘플링

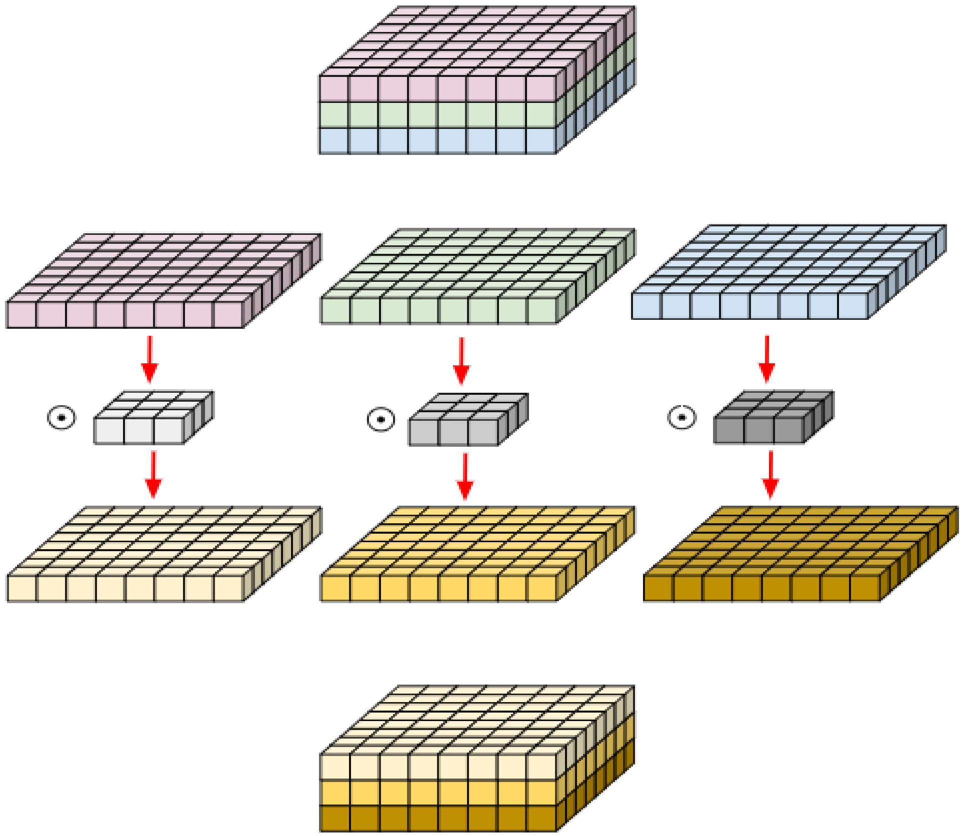
- 왜 전치인가?



채널별 개별 포인트 합성곱

채널별 합성곱

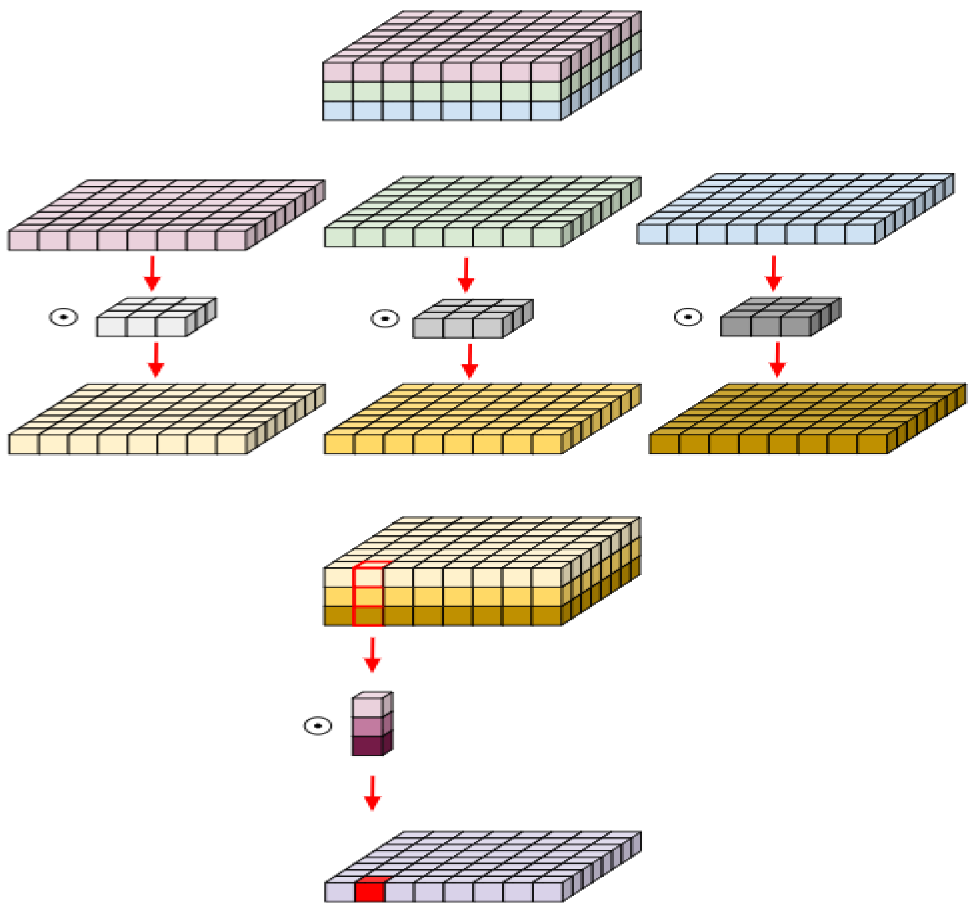
- 채널별 합성곱 (Depthwise Convolution)



- 채널별로 합성곱 실행 즉 채널방향이 아니라 공간 방향으로만 합성곱실행

채널별 합성곱 + 채널방향 합성곱

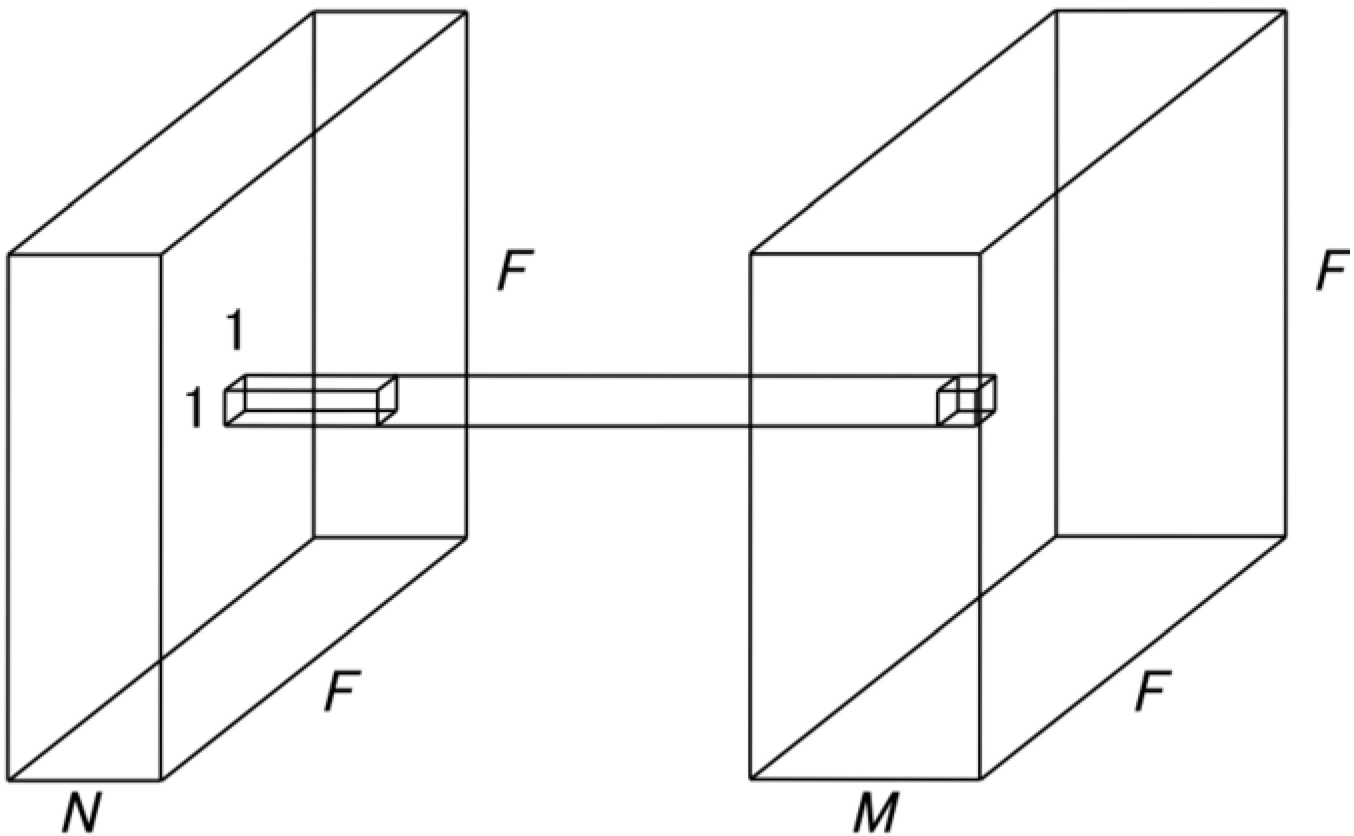
- 채널별 합성곱 이후 채널방향 개별 포인트 합성곱 (Depthwise Separable Convolution)



- 1단계에서는 공간방향으로 2단계에서는 채널방향으로 합성곱 실행한다.

개별 포인트 합성곱

- 개별 포인트 합성곱 (Pointwise Convolution)



- 채널 방향으로 압축을 해 채널 수를 줄인다.