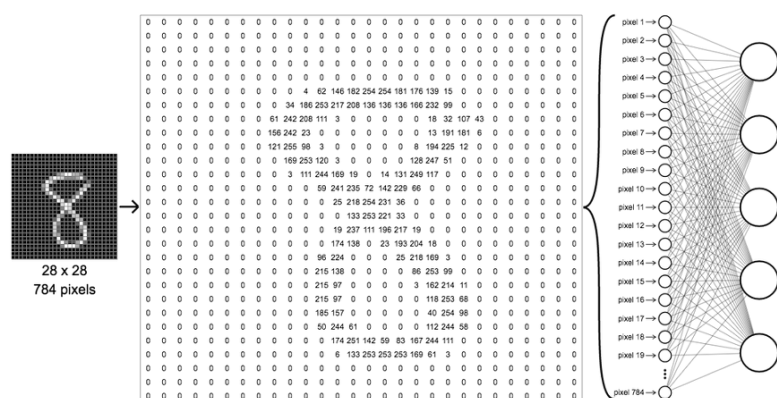


Vision Basic

1 이미지와 텐서의 관계

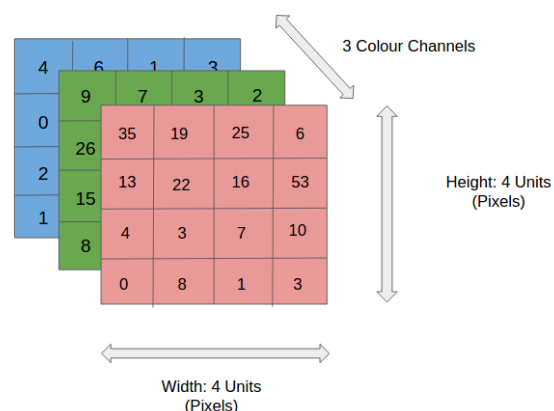
- 흑백 이미지

흑백 이미지는 2차원 tensor로 이루어져 있음



- 컬러 이미지

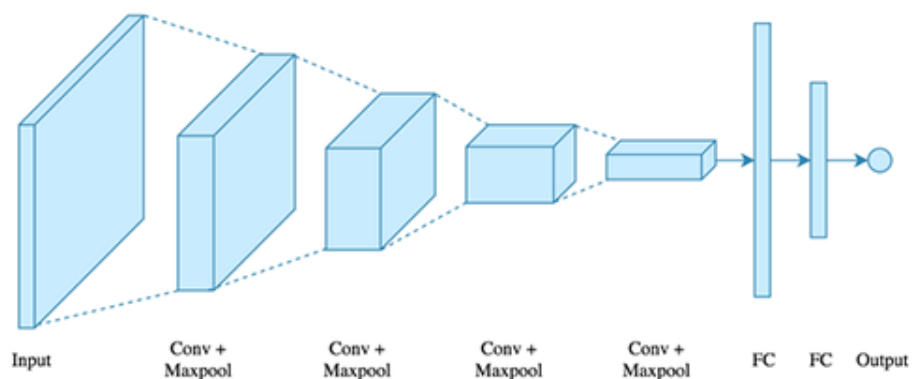
컬러 이미지는 R,G,B 3개의 행렬이 합쳐져 있는 형태로 3차원 tensor로 구성됨



CHW 또는 HWC 형식

2 Convolutional Neural Network(CNN)

- 순서 : (convolution layer → Pooling layer) * 반복 → FC layer



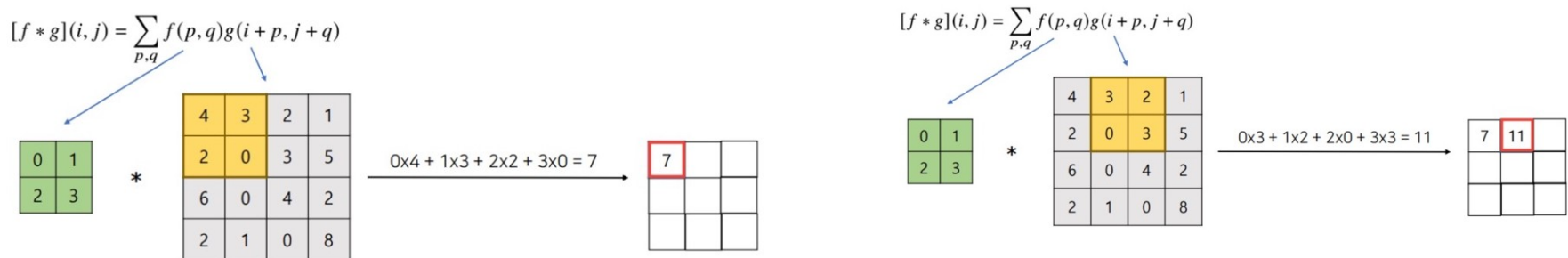
- 합성곱의 목적 : 사람이 보기에 동일한 이미지(살짝 변형되거나 찌그러진)에 대해 Fully connected layer로 이루어진 DNN으로 적합하면 기계는 동일한 이미지로 처리할 수 없다. → 이러한 공간적인 구조 정보의 유실을 막기 위해 합성곱 신경망을 사용한다.

i . Convolution layer

- 합성곱 연산을 통해 이미지의 특징을 추출한다.

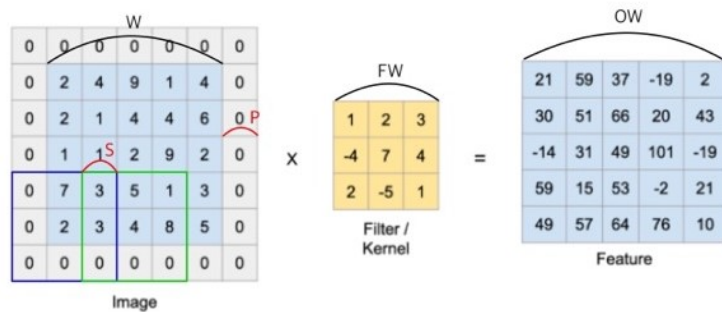
✓ 2D Convolution

$$[f * g](i, j) = \sum_{p, q} f(p, q)g(i + p, j + q)$$



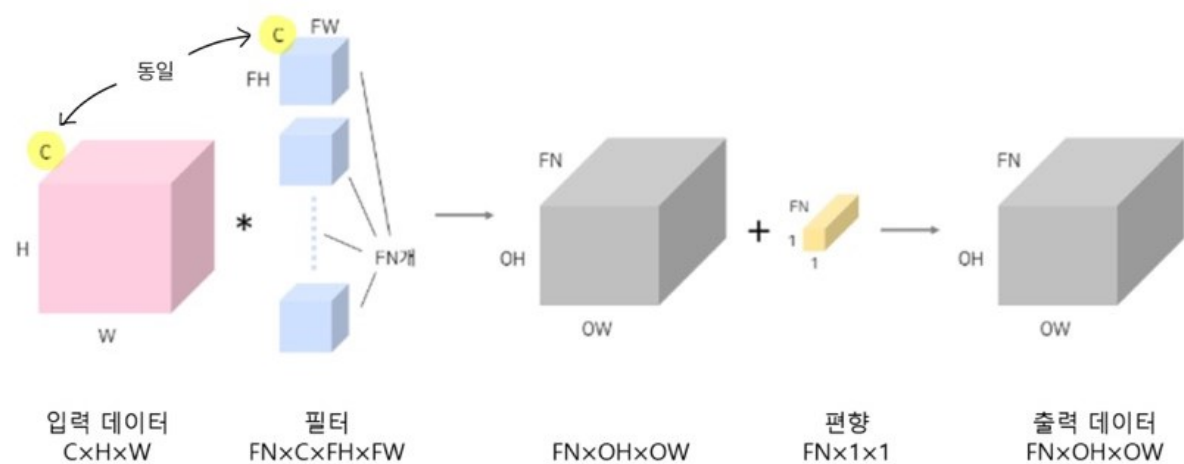
- 입력으로부터 커널을 사용해 합성곱 연산을 통해 나온 결과를 **특성 맵(feature map)**이라고 한다.

$$OW = \frac{W + 2P - FW}{S} + 1$$

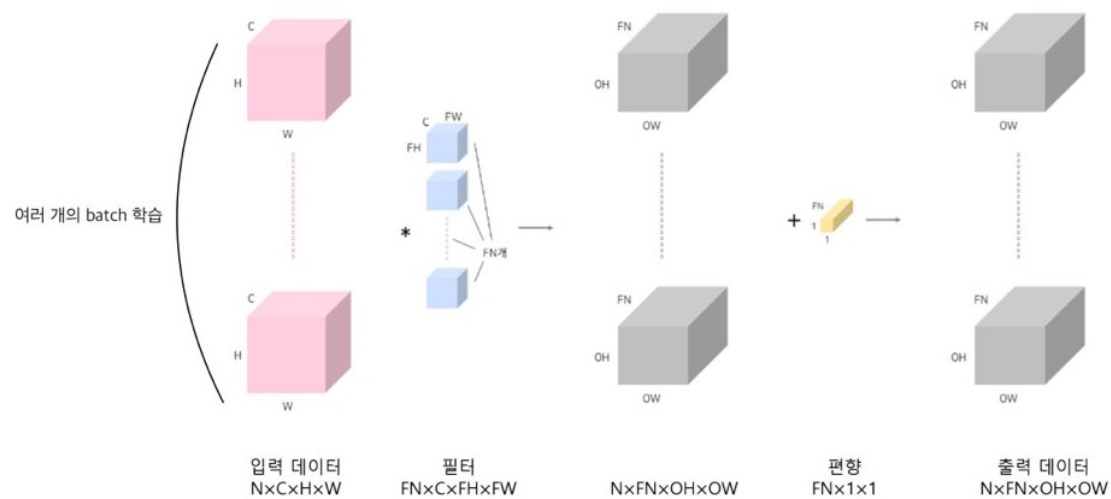


- OW(Output Width) : 출력 가로길이
- W(Width) : 입력 가로길이
- P(Padding) : 입력 주위에 0이 채워지는 줄의 개수
- FW(Filter Width) : 필터 가로길이
- S(Stride) : filter의 보폭

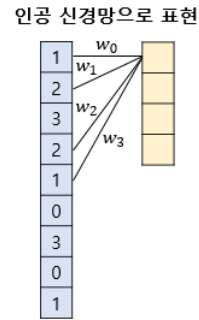
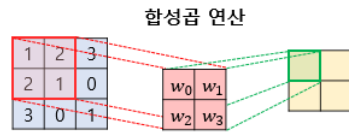
✓ 3D Convolution



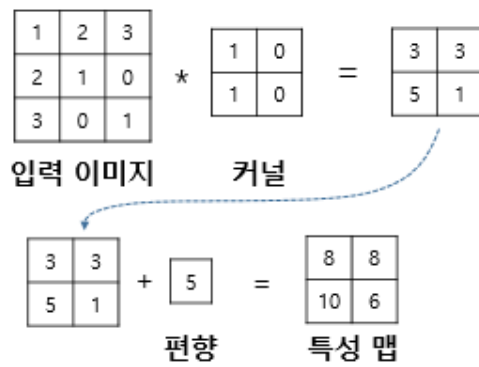
✓ 4D Convolution



✓ 가중치와 편향



합성곱 신경망에서 가중치는 **커널 행렬의 원소들**이다.



합성곱 신경망에 편향을 사용한다면 커널을 적용한 뒤에 더해진다.

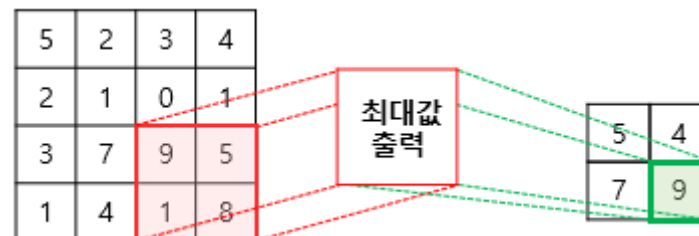
편향은 하나의 값에만 존재하며, 커널이 적용된 결과의 모든 원소에 더해진다.

가중치와 편향 계산이 끝난 convolution layer와 pooling layer 사이에 **activation layer(relu, sigmoid, ...)** 비선형 함수를 사용해 모델의 layer를 깊게 가져간다.

ii. Pooling layer

- 특성 맵을 **down sampling** 하여 특성 맵의 크기를 줄인다.

✓ maxpooling



stride(filter의 보폭)가 2일 때 각 필터 안에서의 최대값을 출력 원소로 갖는다. 2x2 크기 커널로 maxpooling 연산을 했을 때 특성맵이 절반의 크기로 다운샘플링 되었다.

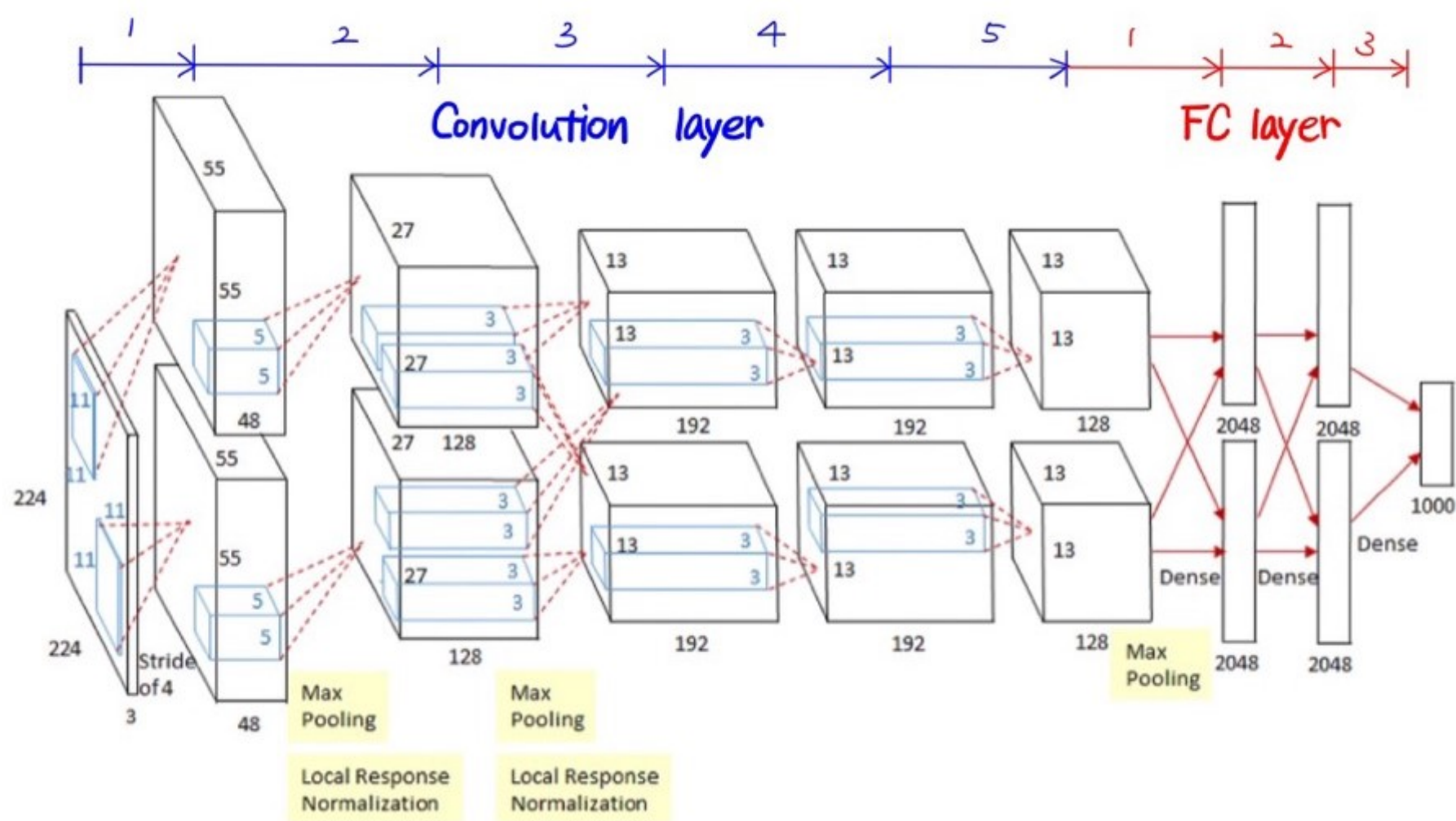
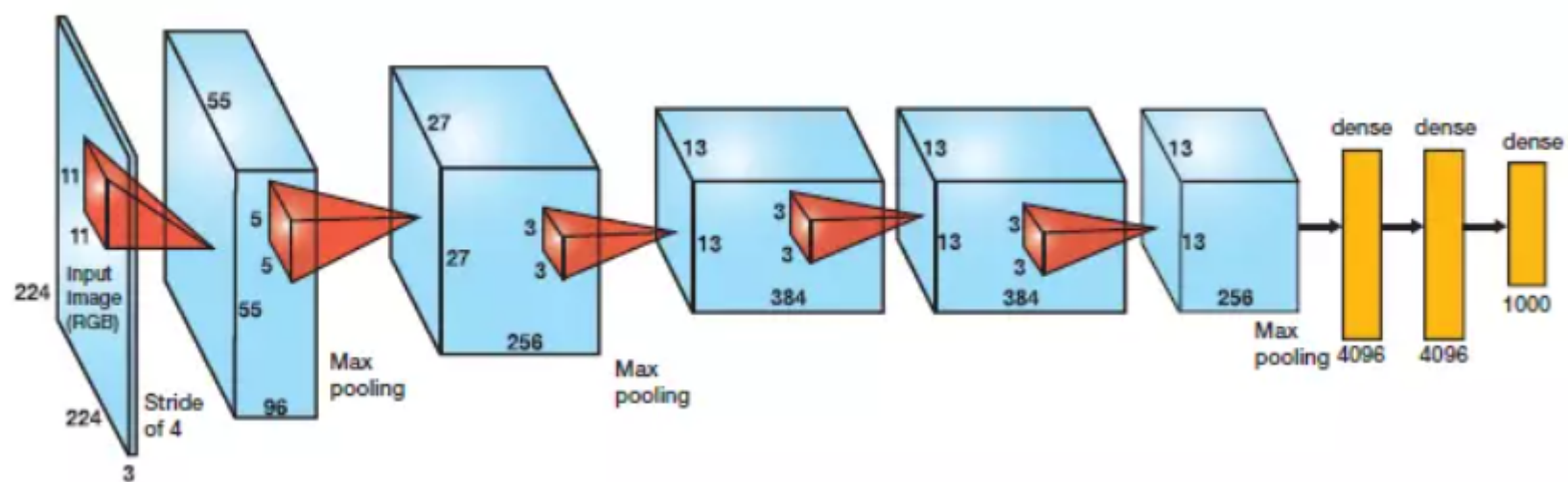
iii. FC layer

✓ Fully Connected layer(Dense layer)

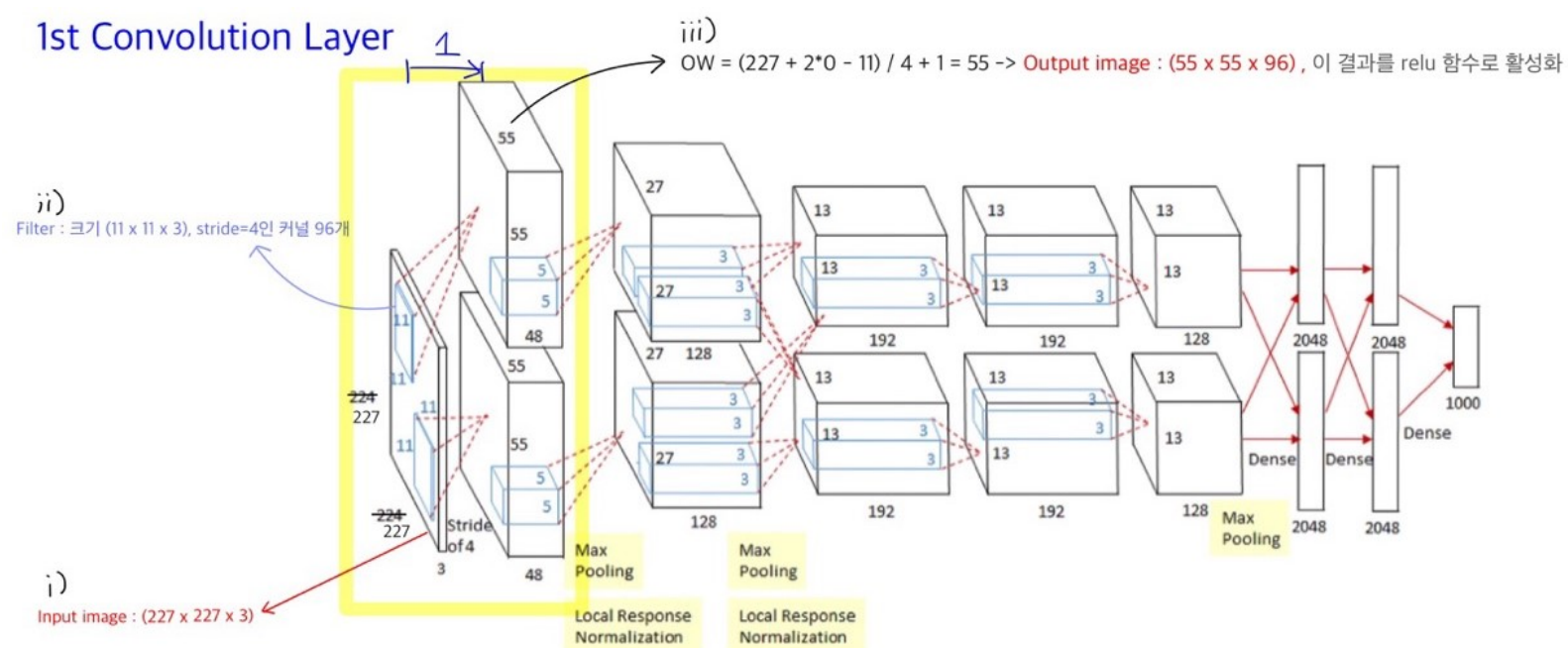
1차원 배열의 형태로 행렬을 평탄화한 다음, 이미지를 분류한다.

참고사이트 <https://wikidocs.net/62306>

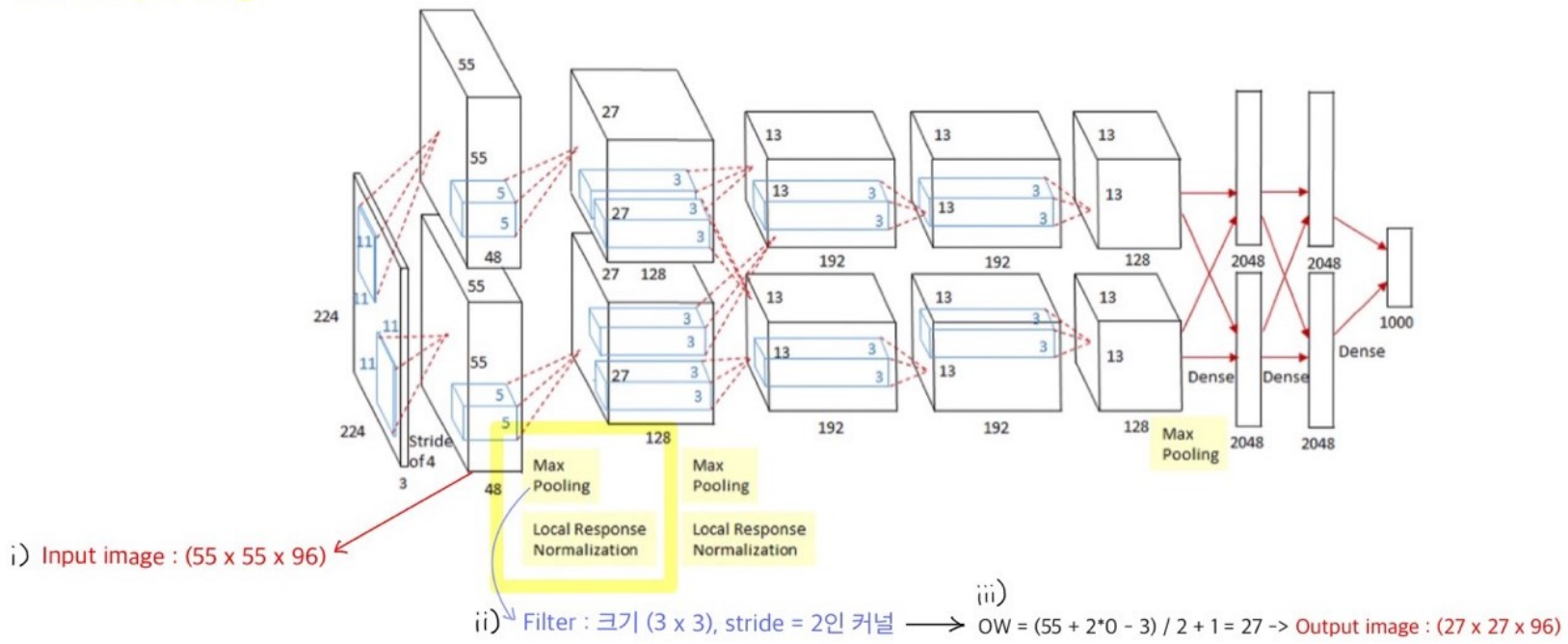
3 AlexNet model



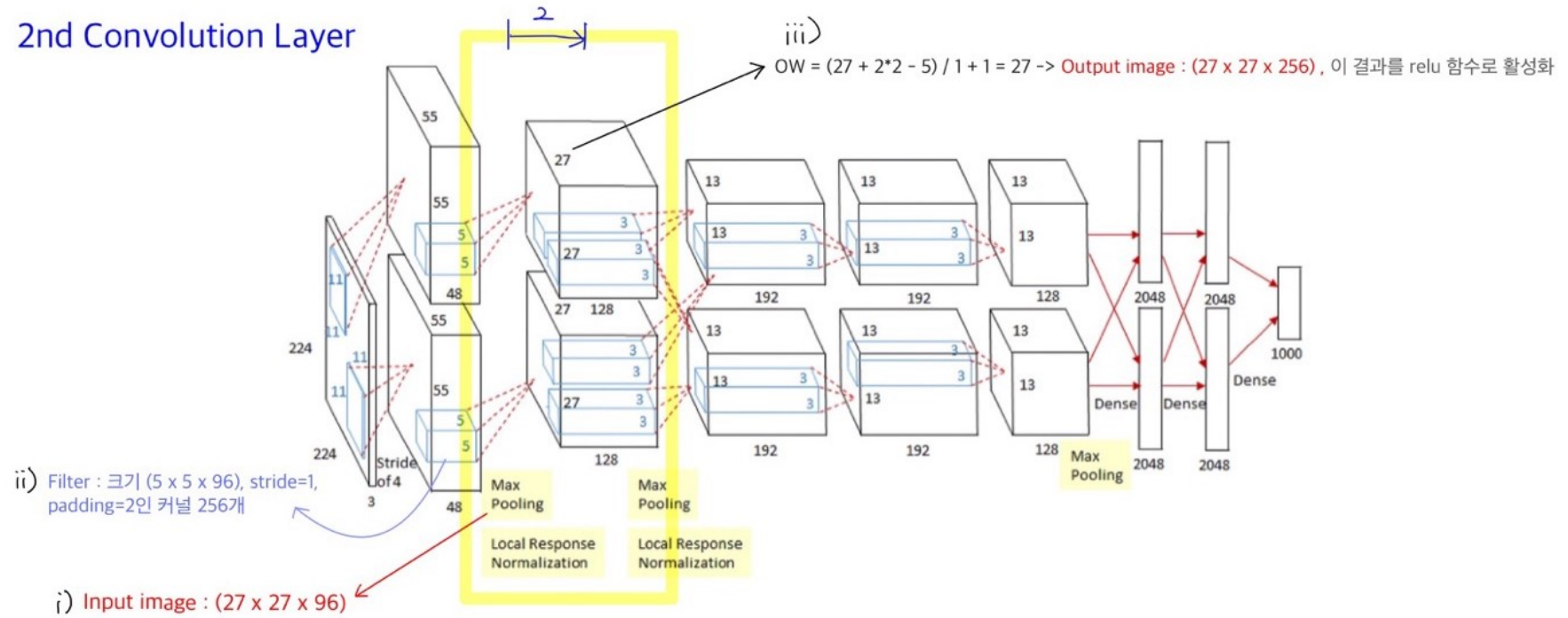
1st Convolution Layer



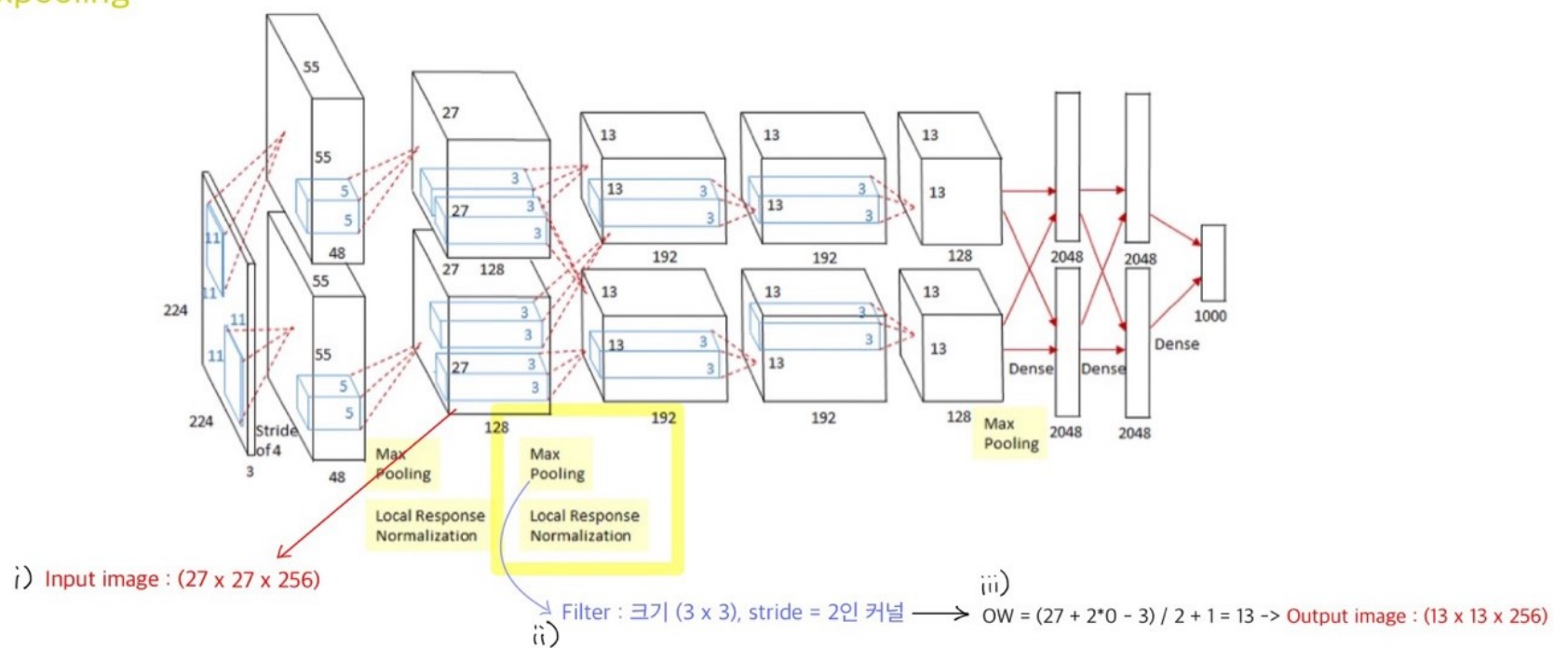
1st Maxpooling



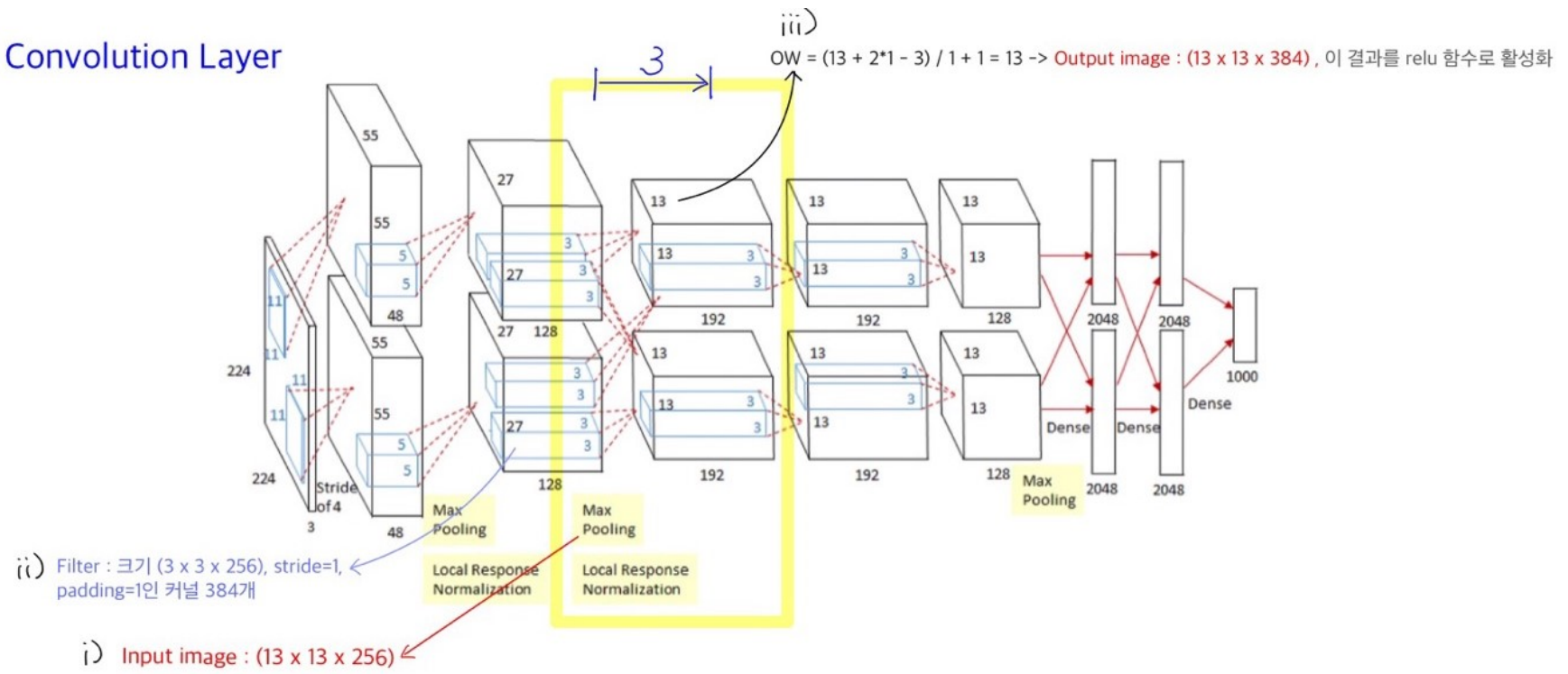
2nd Convolution Layer



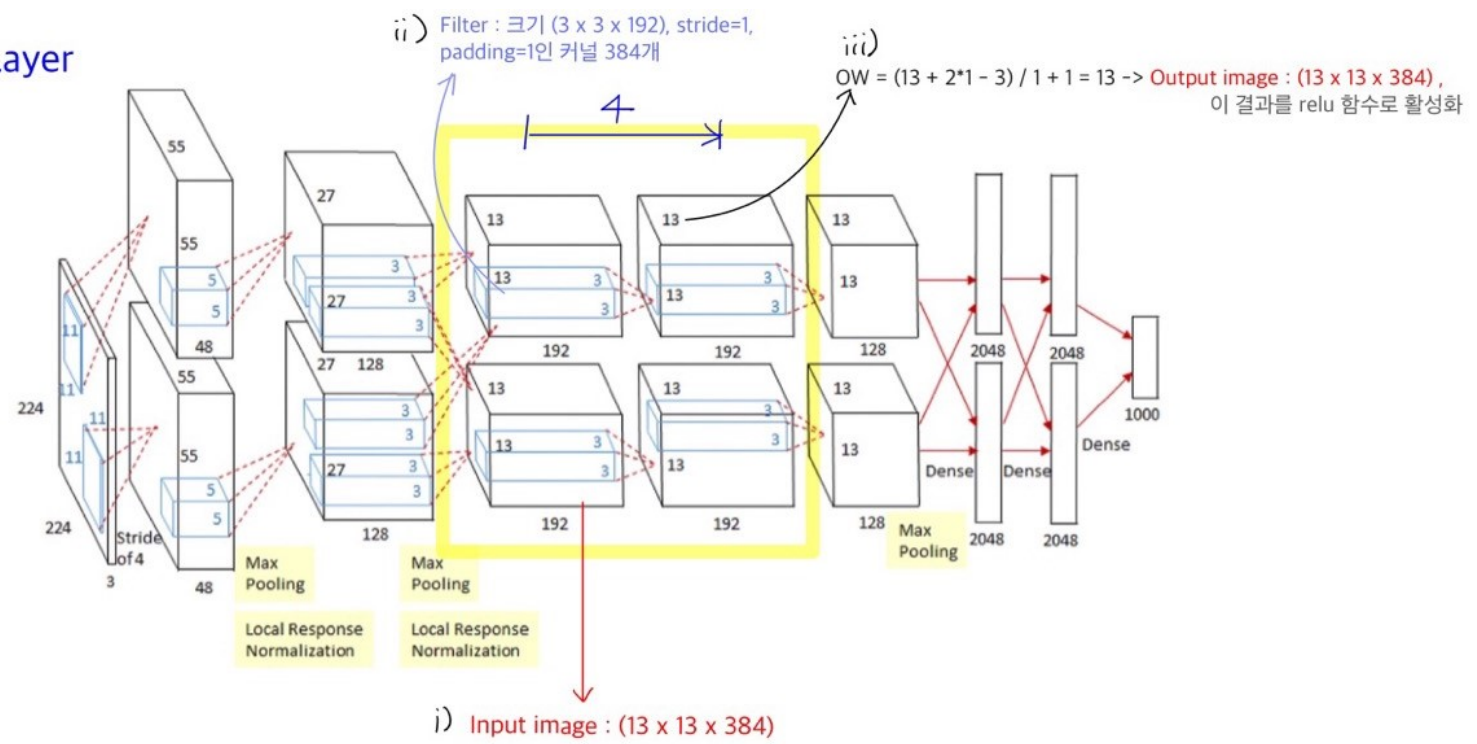
2nd Maxpooling



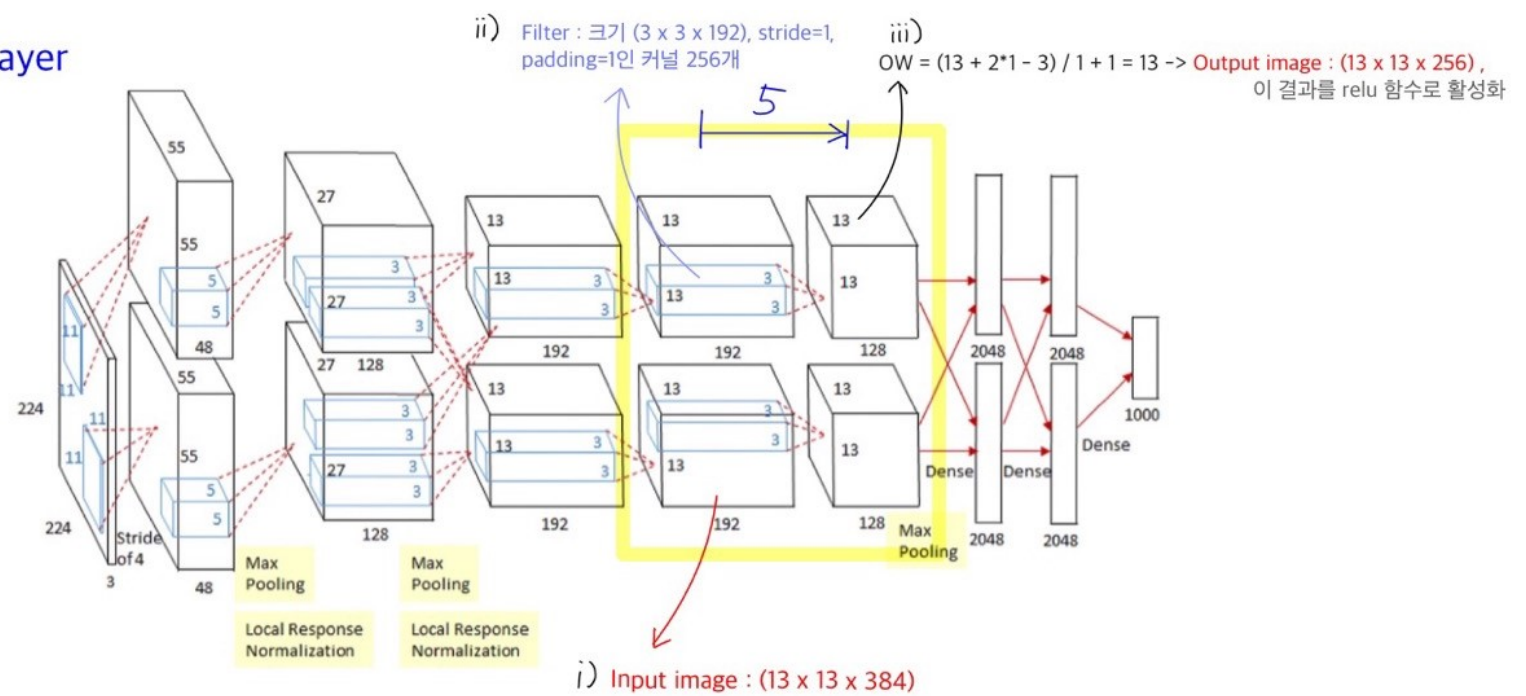
3rd Convolution Layer



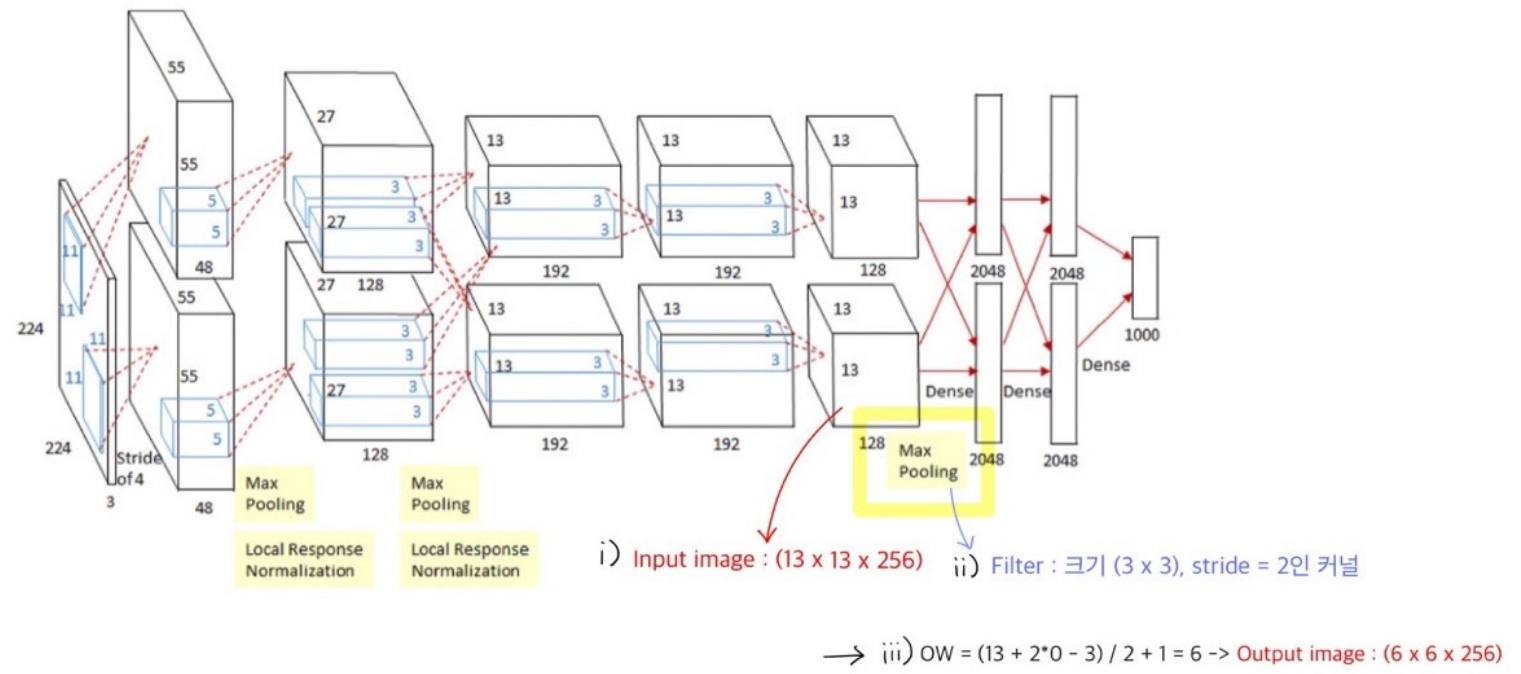
4th Convolution Layer



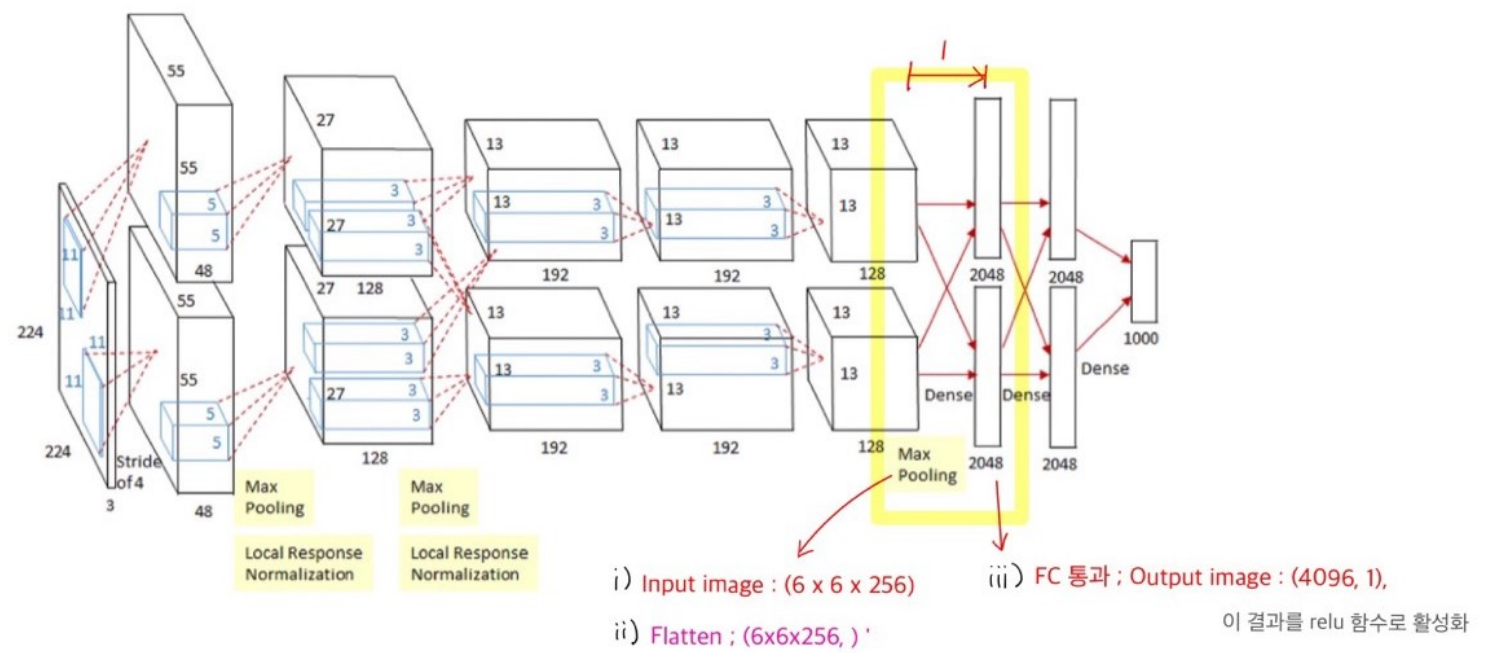
5th Convolution Layer



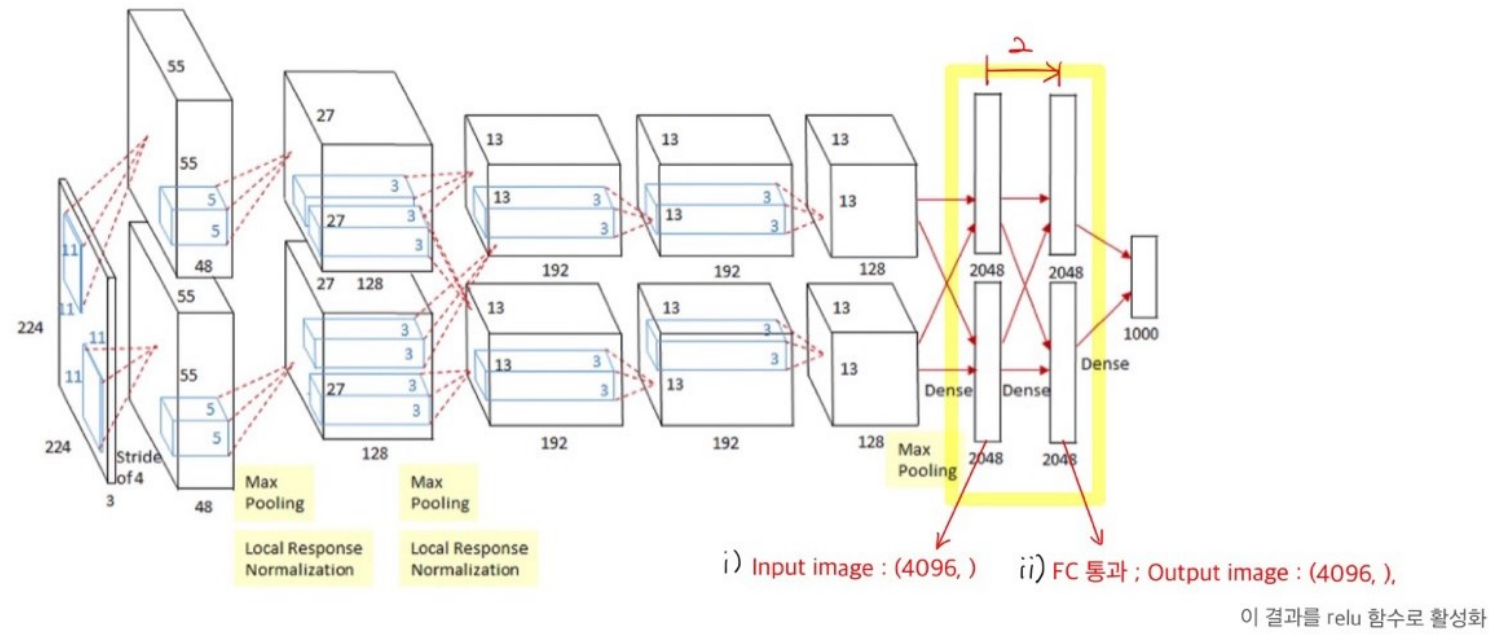
3rd Maxpooling



1st FC layer



2nd FC layer



3rd FC layer

