Alexnet CNN
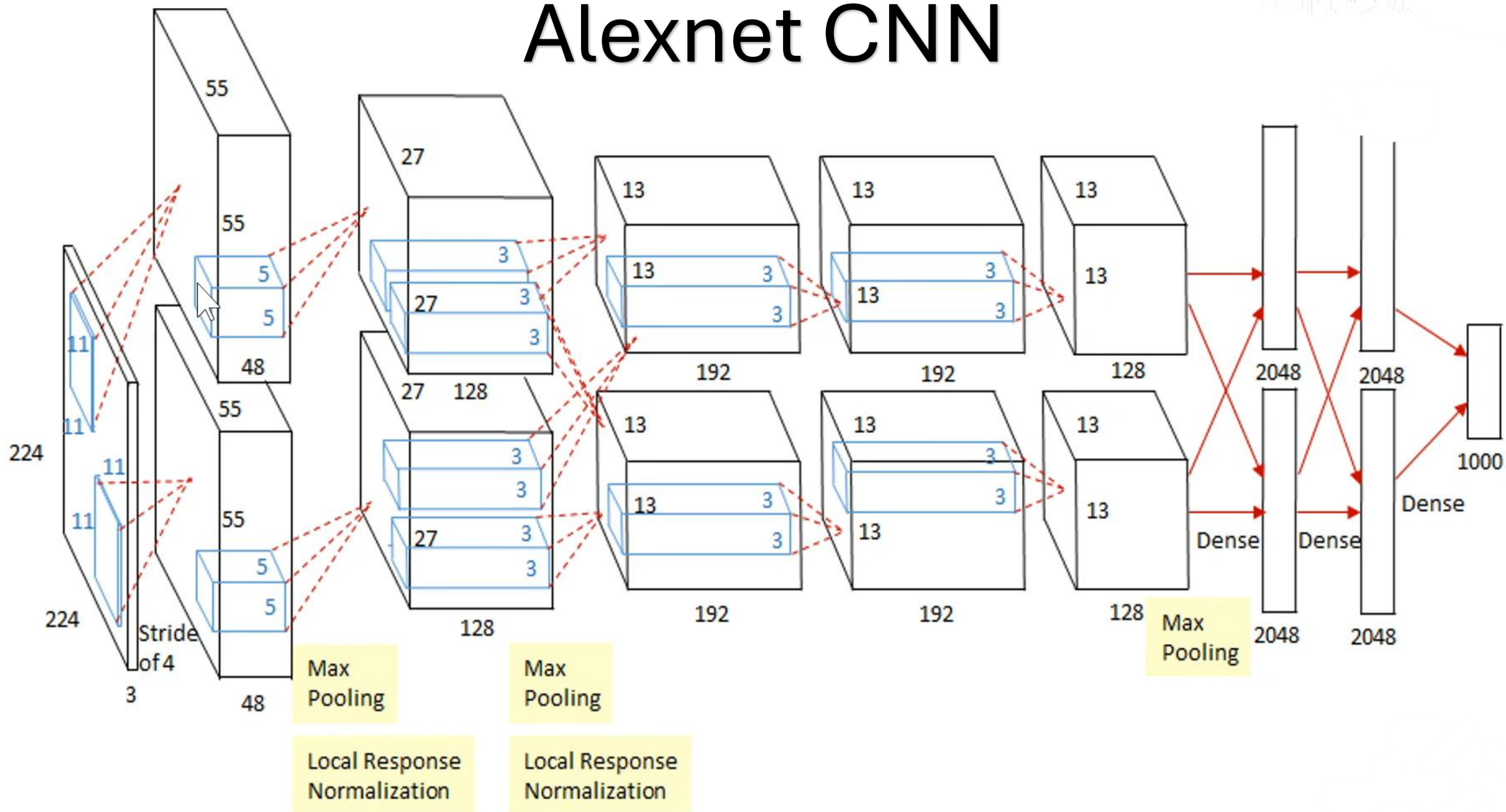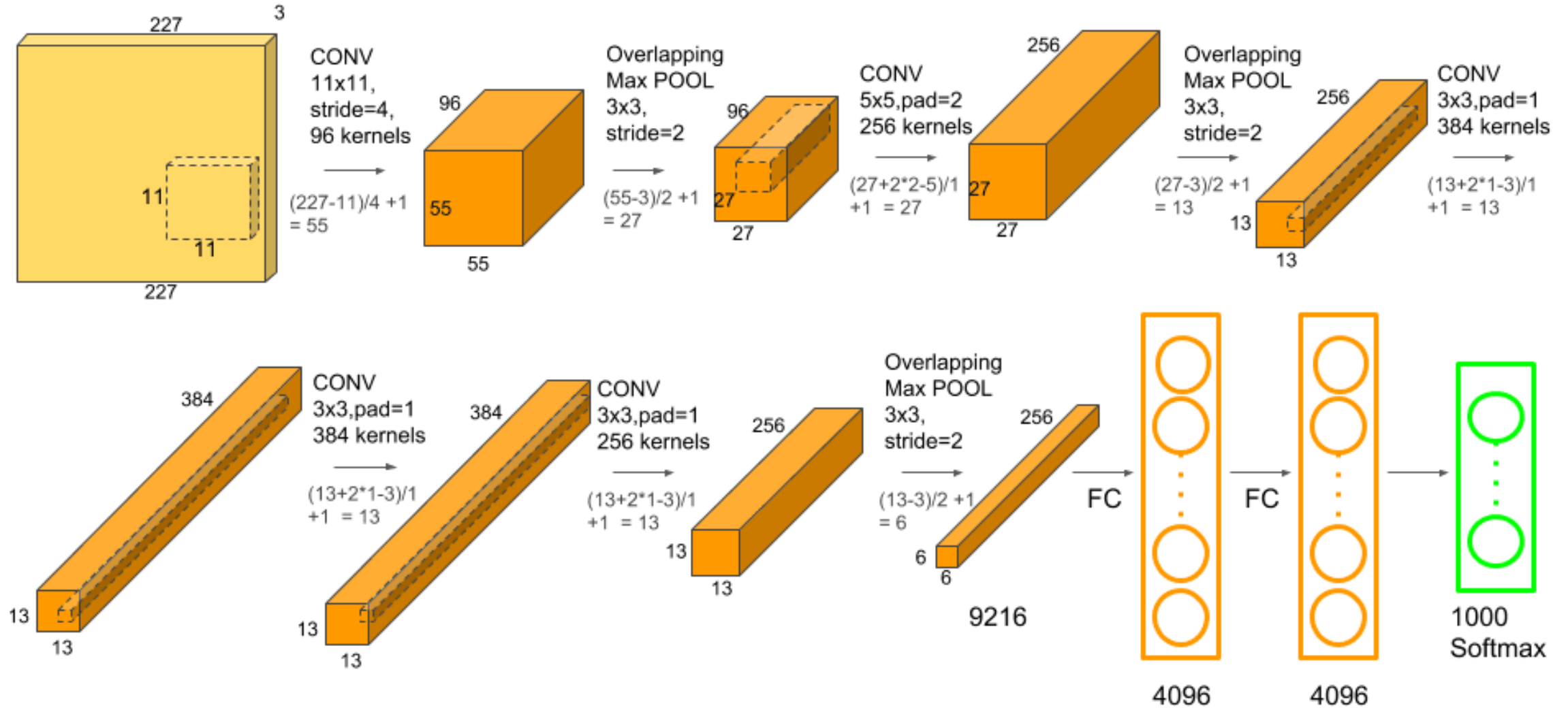
Paper Outcomes

1. The ReLU non-linearity is applied to the output of every convolutional and fully-connected layer.
2. We applied this normalization after applying the ReLU nonlinearity in certain layers.
3. If we set s < z, we obtain overlapping pooling. This is what we use throughout our network, with s = 2 and z = 3.
4. The first convolutional layer filters the 224×224×3 input image with 96 kernels of size 11×11×3 with a stride of 4 pixels
5. The second convolutional layer takes as input the (response-normalized and pooled) output of the first convolutional layer and filters it with 256 kernels of size $5 \times 5 \times 48$.
6. The third, fourth, and fifth convolutional layers are connected to one another without any intervening pooling or normalization layers. The third convolutional layer has 384 kernels of size $3 \times 3 \times 256$ connected to the (normalized, pooled) outputs of the second convolutional layer.
7. The fourth convolutional layer has 384 kernels of size $3 \times 3 \times 192$, and the fifth convolutional layer has 256 kernels of size $3 \times 3 \times 192$. The fully-connected layers have 4096 neurons each.
8. We use dropout in the first two fully-connected layers [...]

Important Formula $= \frac{(n+2p-f)}{s}+1,$

n = size of image, p = padding, f = filter, s = stride

CONV 11x11, stride=4, 96 kernels

$(227-11)/4 +1 = 55$

Overlapping Max POOL 3x3, stride=2

$(55-3)/2 +1 = 27$

CONV 5x5, pad=2 256 kernels

$(27+2*2-5)/1 +1 = 27$

Overlapping Max POOL 3x3, stride=2

$(27-3)/2 +1 = 13$

CONV 3x3, pad=1 384 kernels

$(13+2*1-3)/1 +1 = 13$

CONV 3x3, pad=1 384 kernels

$(13+2*1-3)/1 +1 = 13$

CONV 3x3, pad=1 256 kernels

$(13+2*1-3)/1 +1 = 13$

Overlapping Max POOL 3x3, stride=2

$(13-3)/2 +1 = 6$

9216

FC

4096

FC

4096

1000 Softmax

Network architecture

- The network consists of 5 *Convolutional* layers and 3 *Fully Connected* Layers
- *Max Pooling* is applied Between the layers:
  - 1conv-2conv
  - 2conv-3conv
  - 5conv-1fc
- Before Max Pooling a normalization technique is applied.