DNN_05

Convolutional Neural Network-2

Shahrukh

Why Convolution?

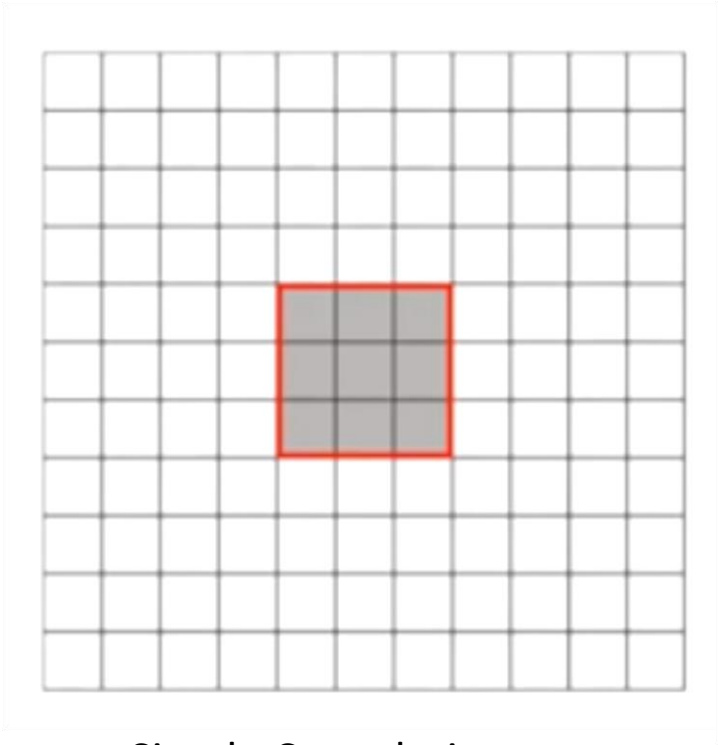
- Parameter sharing .
- Retain Spatial Information.
- Can be scaled for larger Input dimension.
- Can support multiple dimension of Input Image.

Dilated Convolution

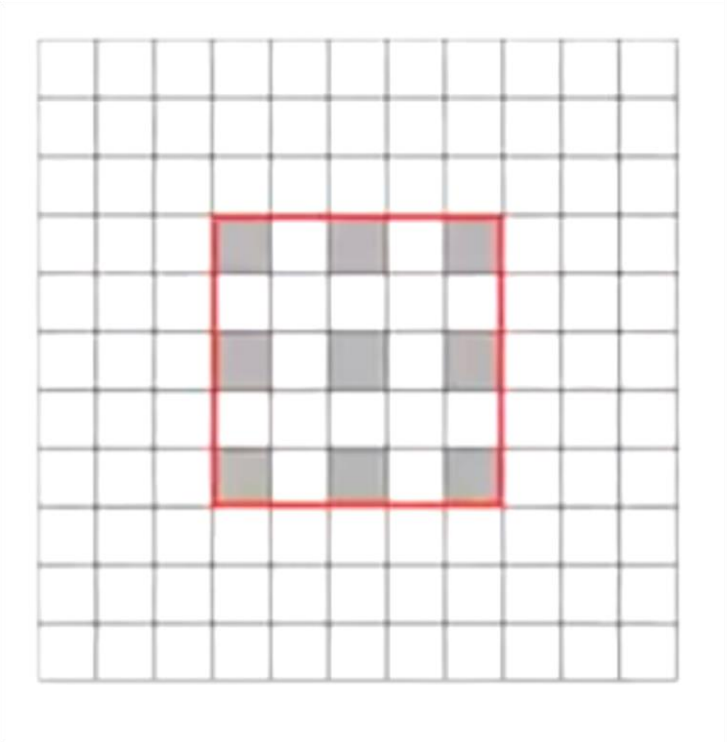
Dilated convolution is just a convolution applied to input with defined gaps. It is a way of increasing receptive field (global view) of a filter.

- Increase Receptive Field of a Filter
- Effective when used in Semantic Segmentation
- Capturing fine Details by processing inputs in higher resolutions.
- Faster Runtime

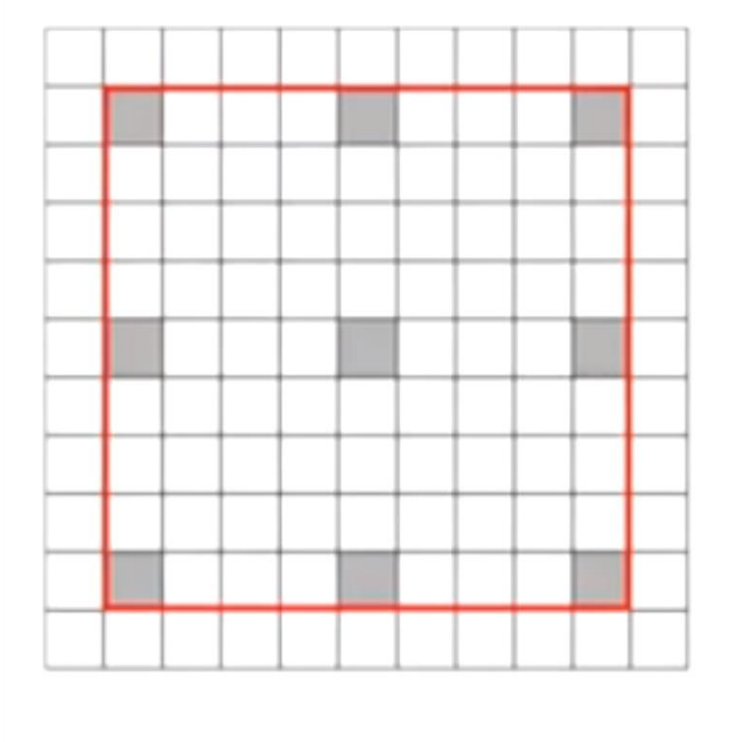
Dilated Convolution



Simple Convolution
Receptive Field : 3*3



Dilation rate l= 2
Receptive Field 5*5



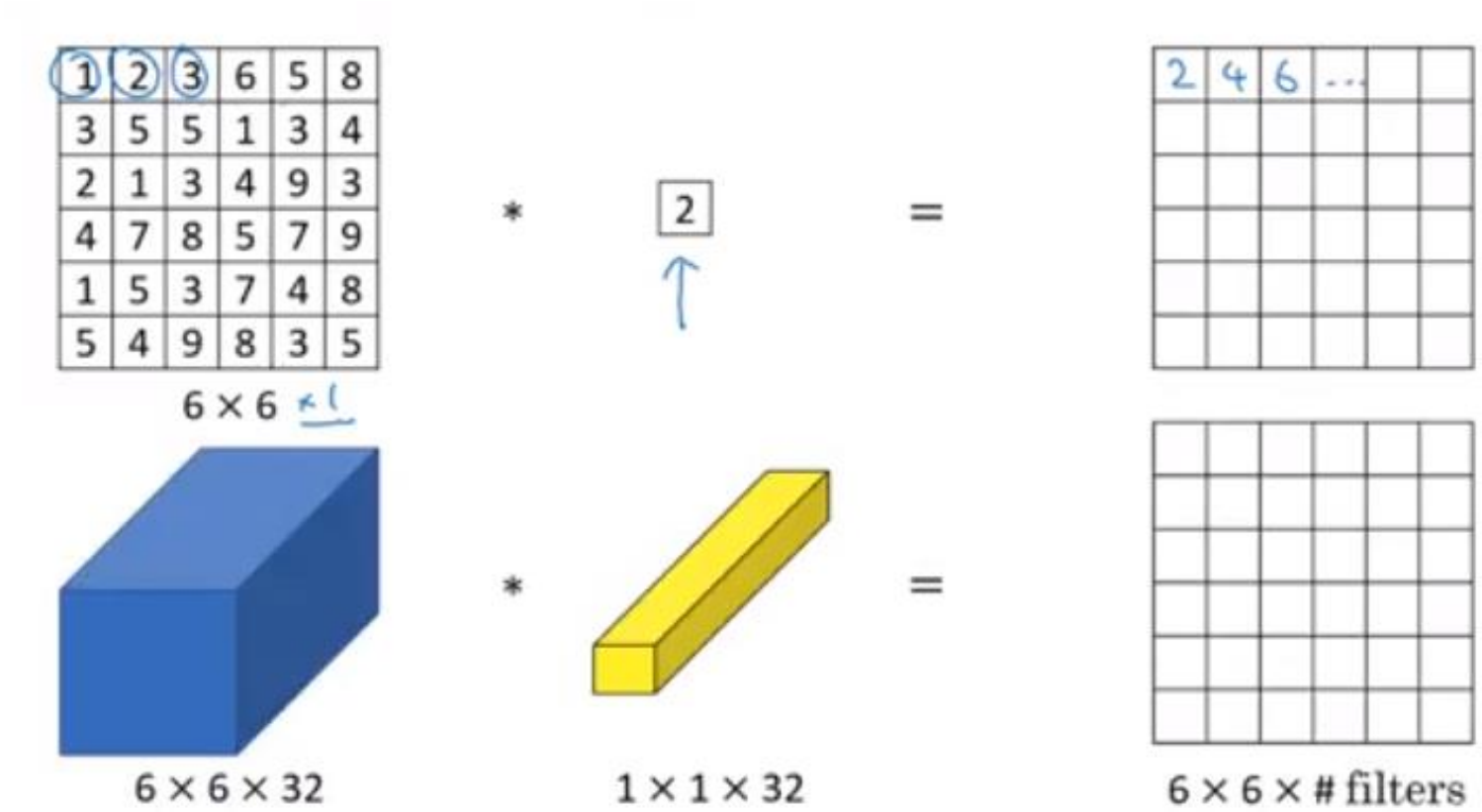
Dilation rate l= 4
Receptive Field 9*9

L=1 -> Standard Convolution
L=2 -> Skipping one pixel per input
L=4 -> Skipping three pixels per input

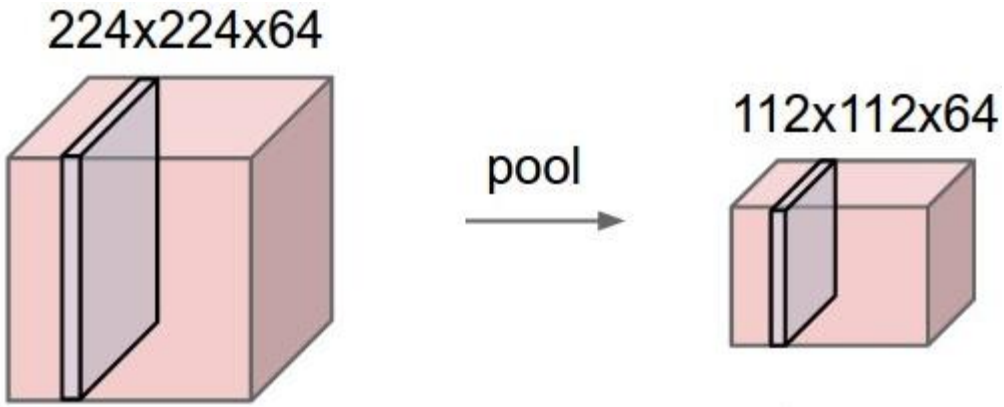
One by One [1 x 1] Convolution

- Also called as Network in Network
- Reduce Number of channel
- Can be used to replace Fully connected Layer
- **Reduce Computational Cost**

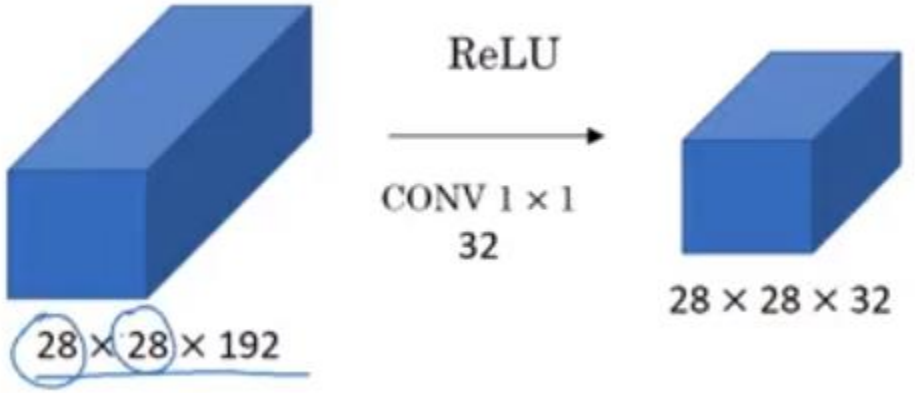
One by One [1 x 1] Convolution



One by One [1 x 1] Convolution



Pooling : Reduce the dimension (W, H)

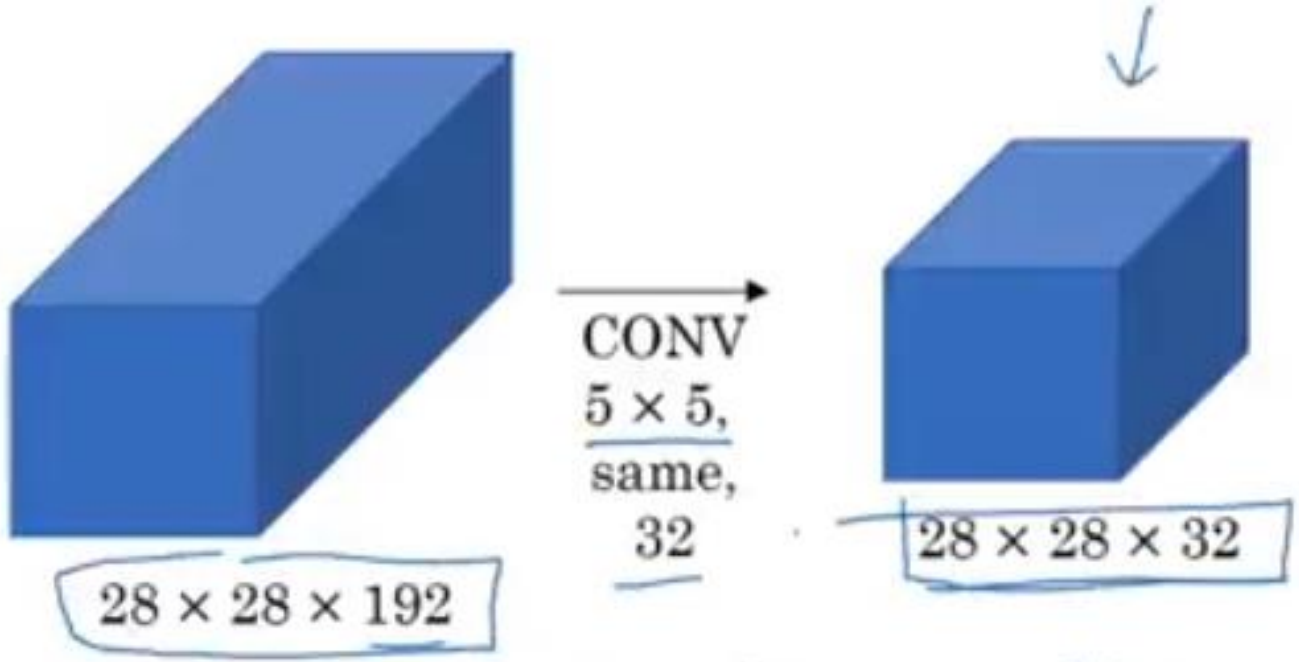


1 x 1 Convolution : Reduce number of channels

- Pooling layer Reduces the dimension (W and H)
- 1 x 1 Convolution reduces the number of channels

One by One [1 x 1] Convolution

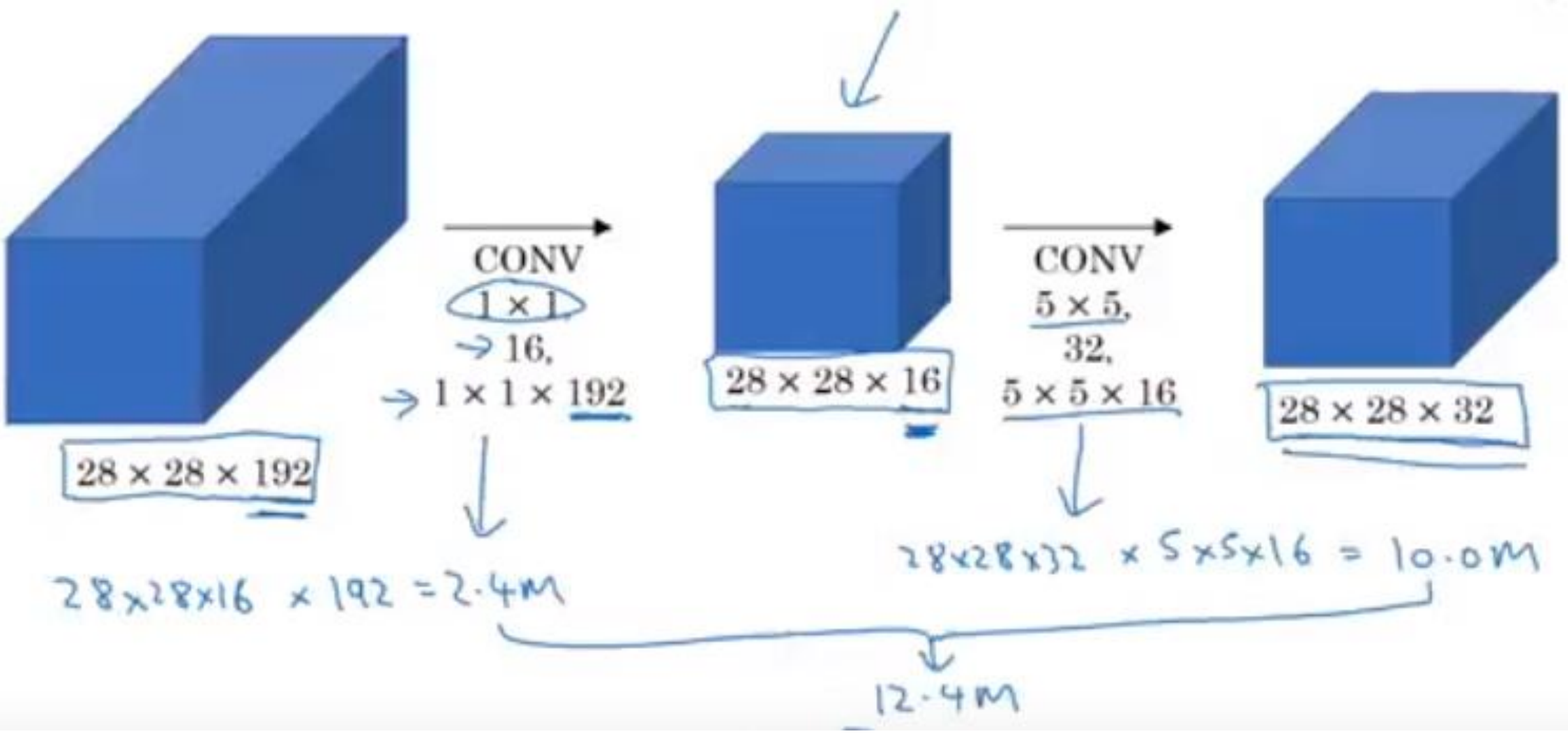
The problem of computational cost



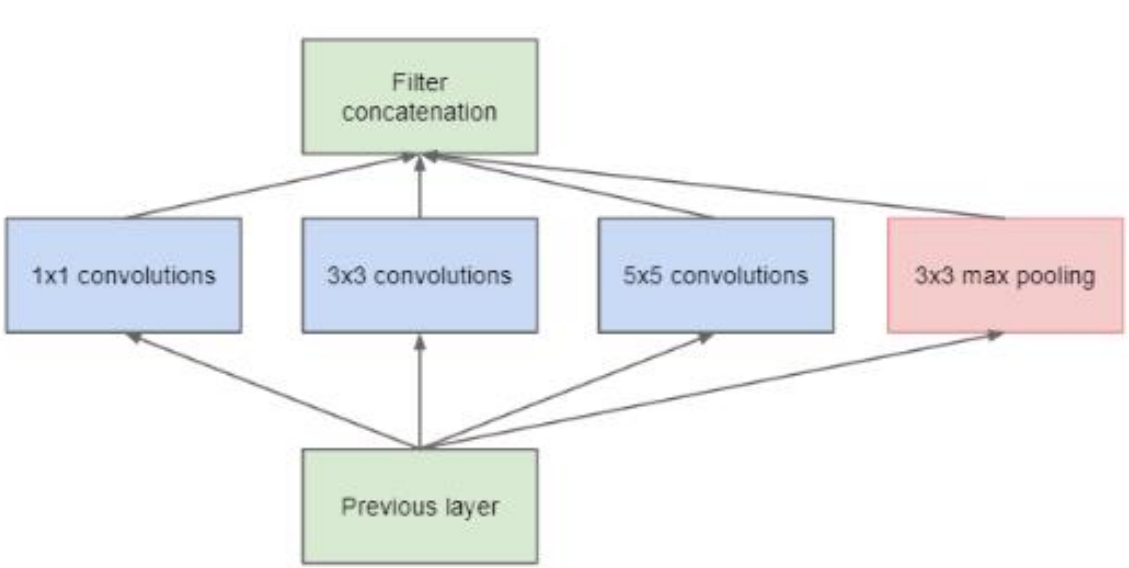
32 filters. filter one 5×5×192.

28×28×32 × 5×5×192 = 120M.

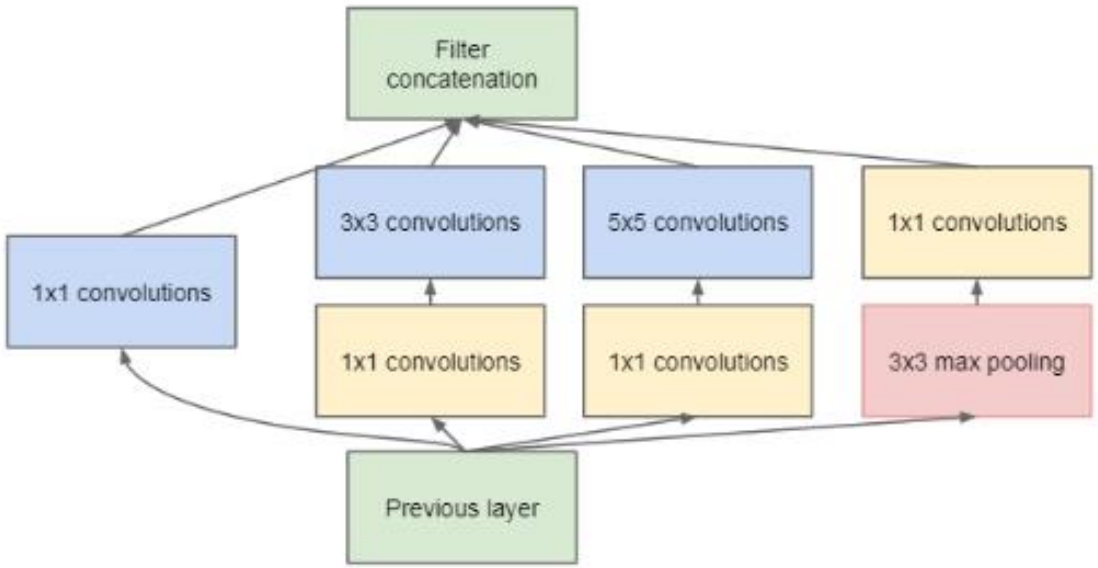
One by One [1 x 1] Convolution



One by One [1 x 1] Convolution in Inception



(a) Inception module, naïve version



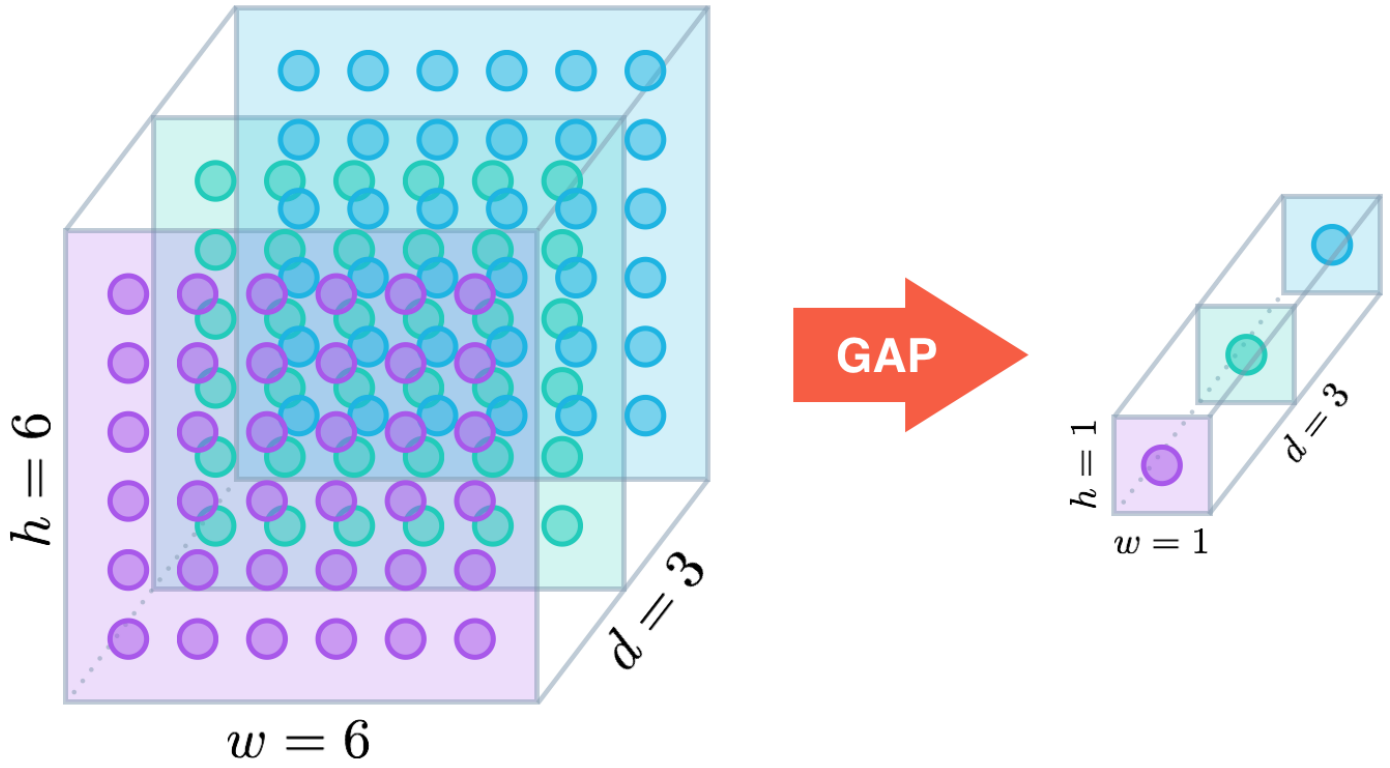
(b) Inception module with dimension reductions

1x1 convolutions in GoogLeNet

Global Average Pooling

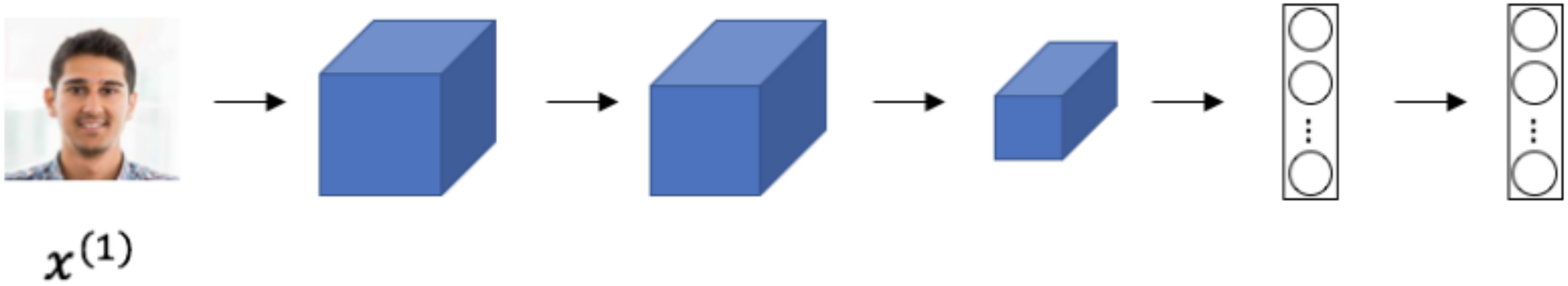
- Traditional we use fully connected layers in CNN as the final layer and apply softmax over it.
- The fully connected layers are prone to overfitting and heavily depend on dropout regularization
- global average pooling is itself a structural regularizer, which natively prevents overfitting for the overall structure.

Global Average Pooling

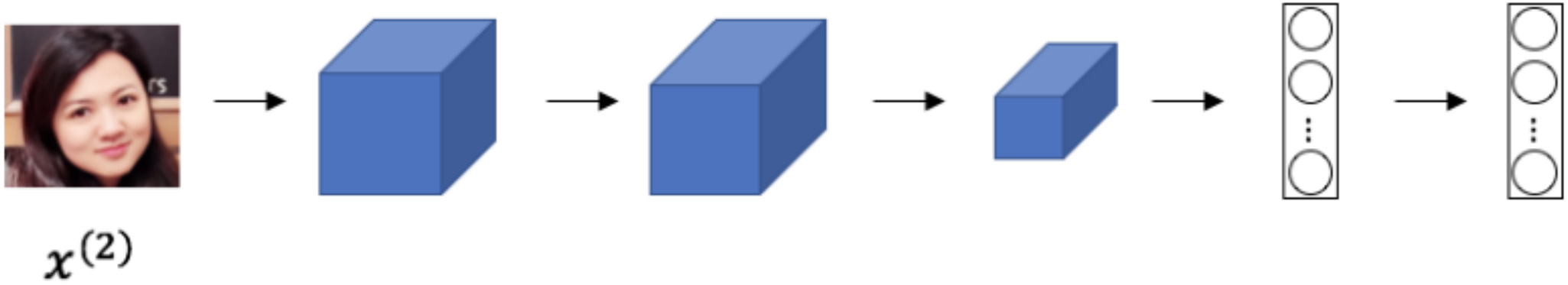



```
[convolutional]
batch_normalize=1
filters=1024
size=3
stride=1
pad=1
activation=leaky

[shortcut]
from=-3
activation=linear

[avgpool]

[convolutional]
filters=1000
size=1
stride=1
pad=1
activation=linear

[softmax]
groups=1
```

Siamese Network(One-Shot Learning)



$x^{(1)}$



$x^{(2)}$

Siamese Network

Triplet Loss

$$\mathcal{L} = max(d(a, p) - d(a, n) + margin, 0)$$

Triplet Loss Funciton

Reference

- DeepLearning.ai