



## **NPTEL ONLINE CERTIFICATION COURSES**

**Course Name: Deep Learning**

**Faculty Name: Prof. P. K. Biswas**

**Department : E & ECE, IIT Kharagpur**

**Topic**

**Lecture 41: Popular CNN Models V**

## CONCEPTS COVERED

### Concepts Covered:

#### ☐ CNN

- ☐ AlexNet

- ☐ VGG Net

- ☐ Transfer Learning

- ☐ Challenges in Deep Learning

- ☐ GoogLeNet

- ☐ ResNet

- ☐ etc.

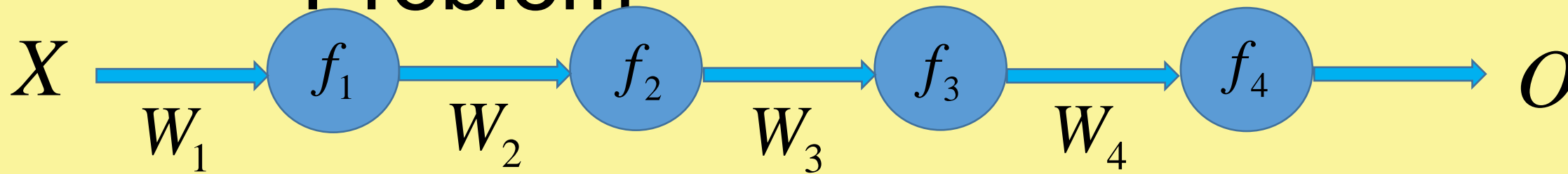


# Challenges

- ☐ Deep learning is data hungry.
- ☐ Overfitting or lack of generalization.
- ☐ Vanishing/Exploding Gradient Problem.
- ☐ Appropriate Learning Rate.
- ☐ Covariate Shift.
- ☐ Effective training.



# Vanishing Gradient Problem



$$\frac{\partial O}{\partial W_1} = X \cdot f_1' \cdot W_2 \cdot f_2' \cdot W_3 \cdot f_3' \cdot W_4 \cdot f_4'$$



# Vanishing Gradient Problem

- ❑ Choice of activation function: ReLU instead of Sigmoid.
- ❑ Appropriate initialization of weights.
- ❑ Intelligent Back Propagation Learning Algorithm.

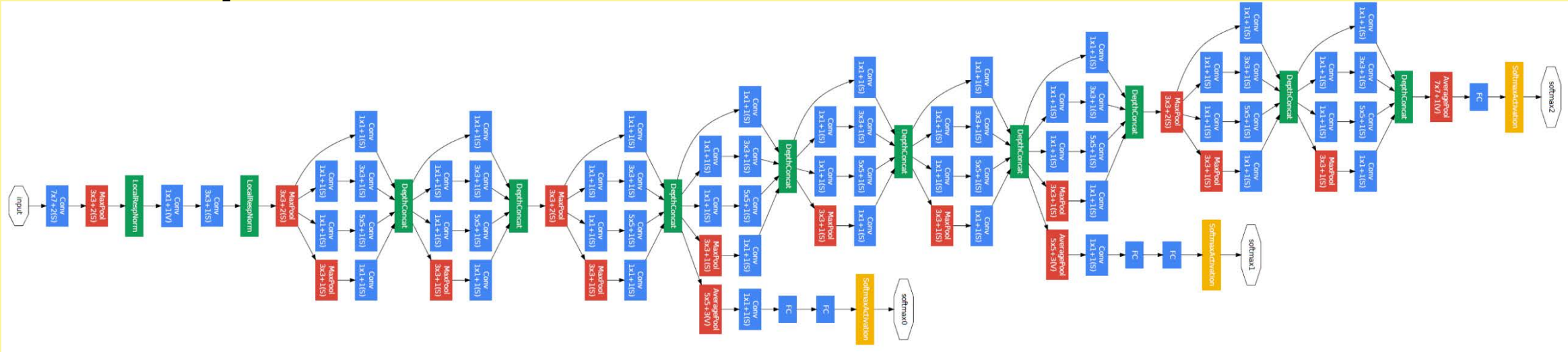


# GoogLeNet

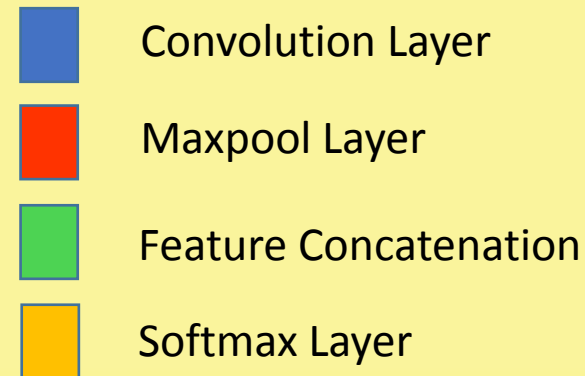
## ILSVRC 2014 Winner



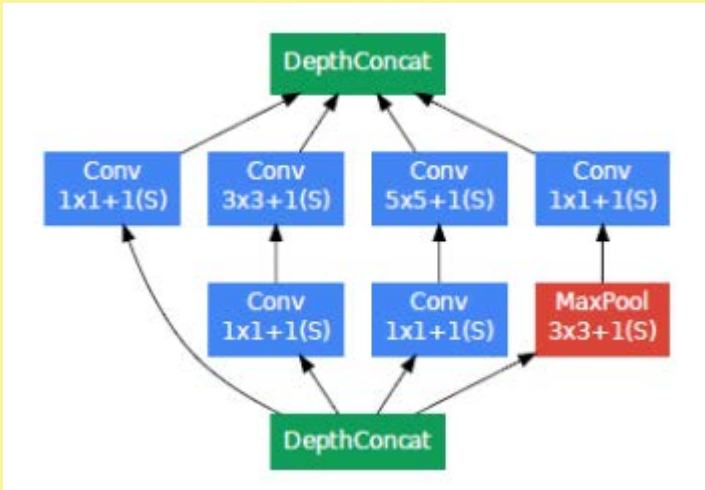
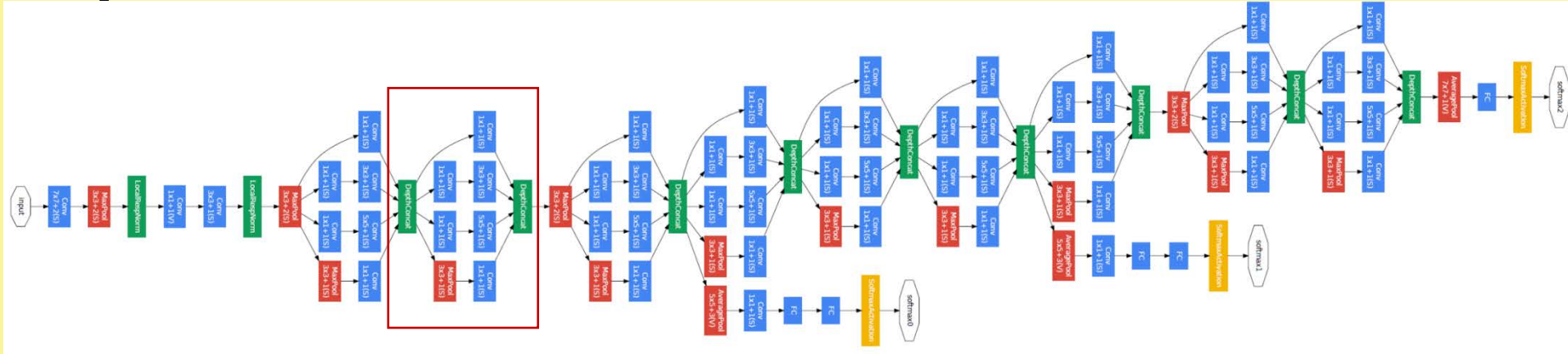




- ❖ 22 Layers with parameters
- ❖ 27 Layer including Maxpool layers



†



## Inception Module

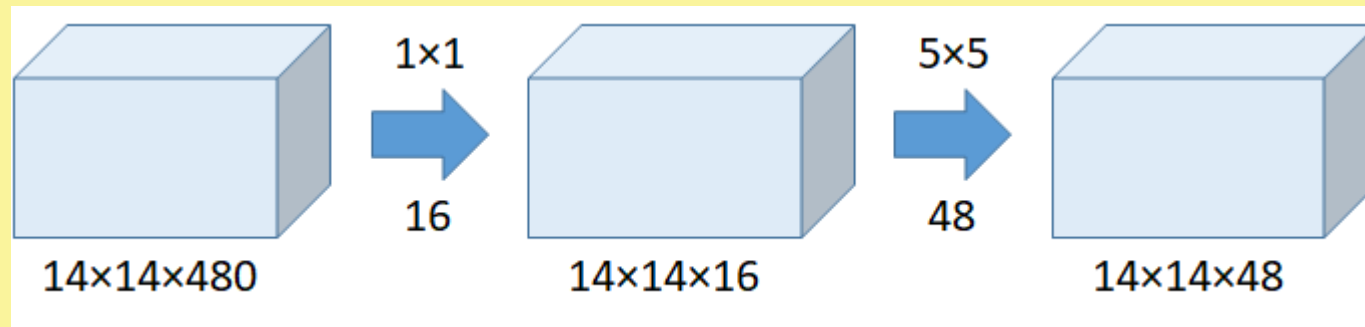


# Inception Module

- ❑ Computing  $1\times 1$ ,  $3\times 3$ , and  $5\times 5$  convolutions within the same module of the network.
- ❑ Covers a bigger area, at the same time preserves fine resolution for small information on the images.
- ❑ Use different convolution kernels of different sizes in parallel from the most accurate detailing ( $1\times 1$ ) to a bigger one ( $5\times 5$ ).
- ❑  $1\times 1$  convolution also reduces computation.



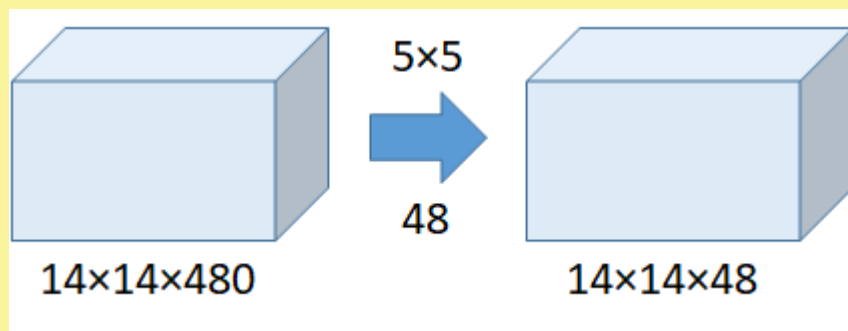
# Inception Module



Number of operations for  $1 \times 1 = (14 \times 14 \times 16) \times (1 \times 1 \times 480) = 1.5\text{M}$

Number of operations for  $5 \times 5 = (14 \times 14 \times 48) \times (5 \times 5 \times 16) = 3.8\text{M}$

Total number of operations =  $1.5\text{M} + 3.8\text{M} = 5.3\text{M}$



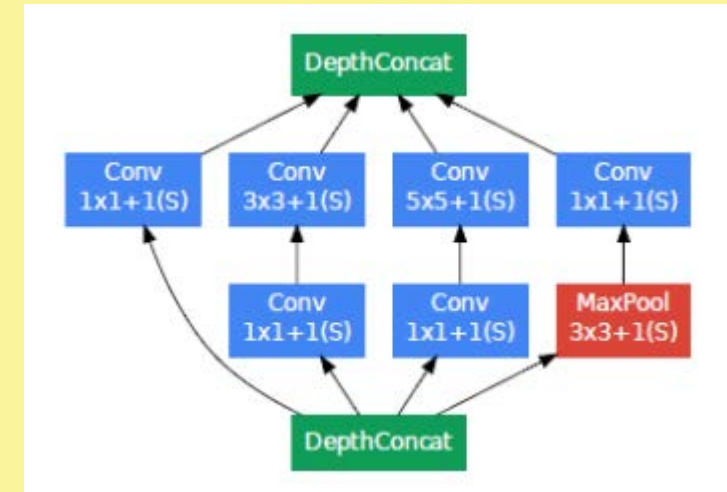
Number of operations =  $(14 \times 14 \times 48) \times (5 \times 5 \times 480) = 112.9\text{M}$



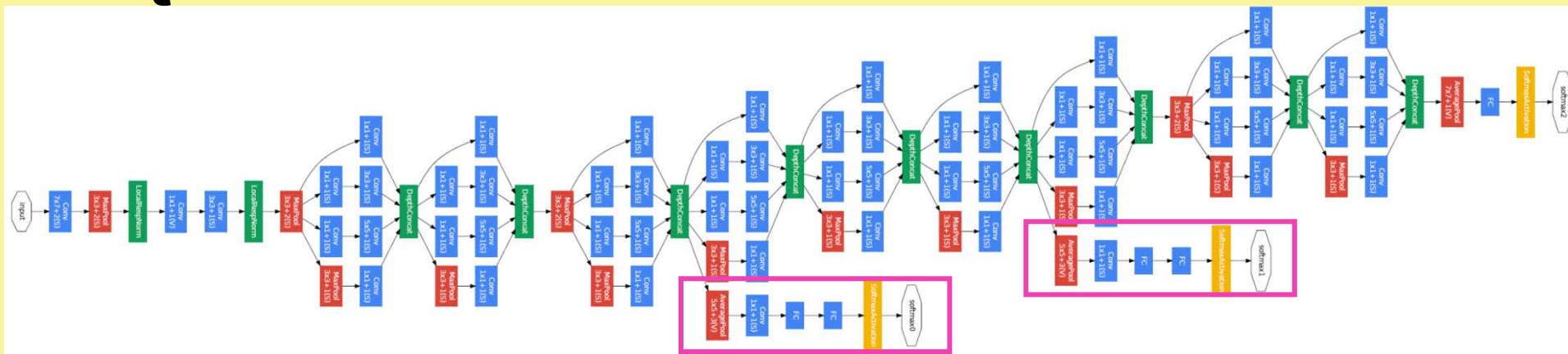
<https://medium.com/coinmonks/paper-review-of-googlenet-inception-v1-winner-of-ilsvlc-2014-image-classification-c2b3565a64e7>

# Inception Module

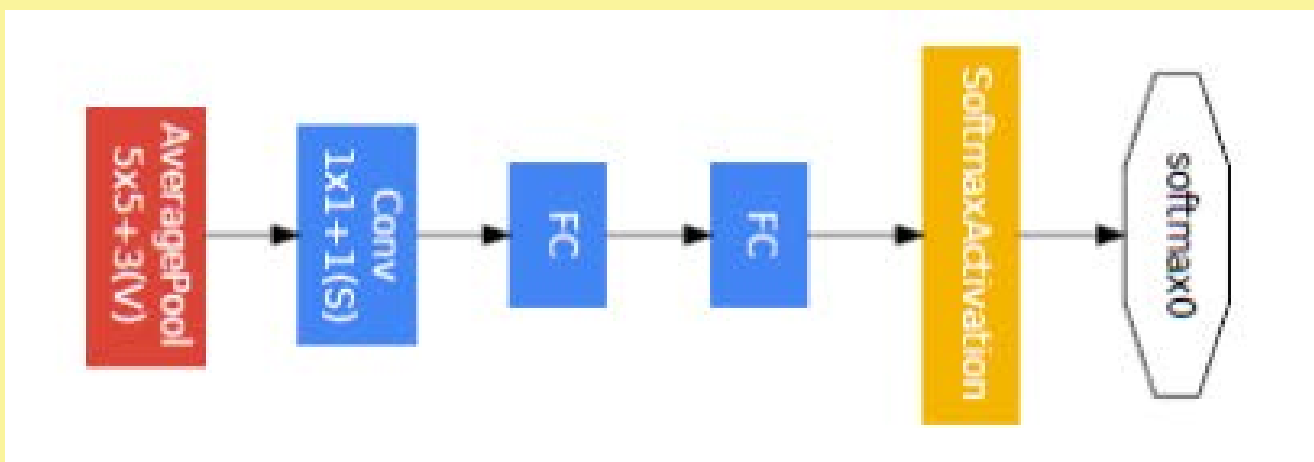
- ❑ Outputs of these filters are then stacked along the channel dimension.
- ❑ Multi-level feature extractor.
- ❑ There are 9 such inception modules.
- ❑ Top-5 error rate of less than 7 %.



# GoogLeNet



Auxiliary Classifier



# Auxiliary Classifier

- ❑ Due to large depth of the network, ability to propagate gradient back through all the layers was a concern.
- ❑ Auxiliary Classifiers are smaller CNNs put on top of middle Inception modules.
- ❑ Addition of auxiliary classifiers in the middle exploits the discriminative power of the features produced by the layers in the middle.



# Auxiliary Classifier

- ❑ During training, loss of Auxiliary classifiers are added to the total loss of the network.
- ❑ Losses from Auxiliary classifiers were weighted by 0.3.
- ❑ Auxiliary classifiers are discarded at Inference time.







## **NPTEL ONLINE CERTIFICATION COURSES**

*Thank  
you*

