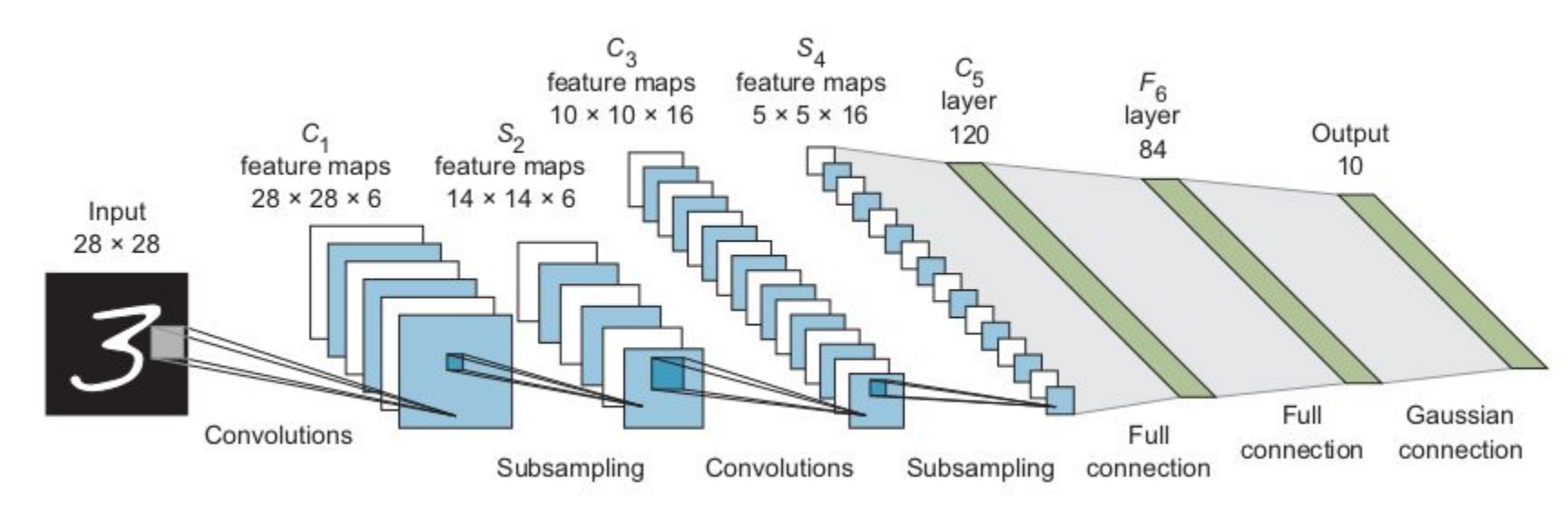
Network Architectures

LeNet architecture



Tanh as the activation function!

IMAGENET

14,197,122 images, 21841 synsets indexed

Home Download Challenges About

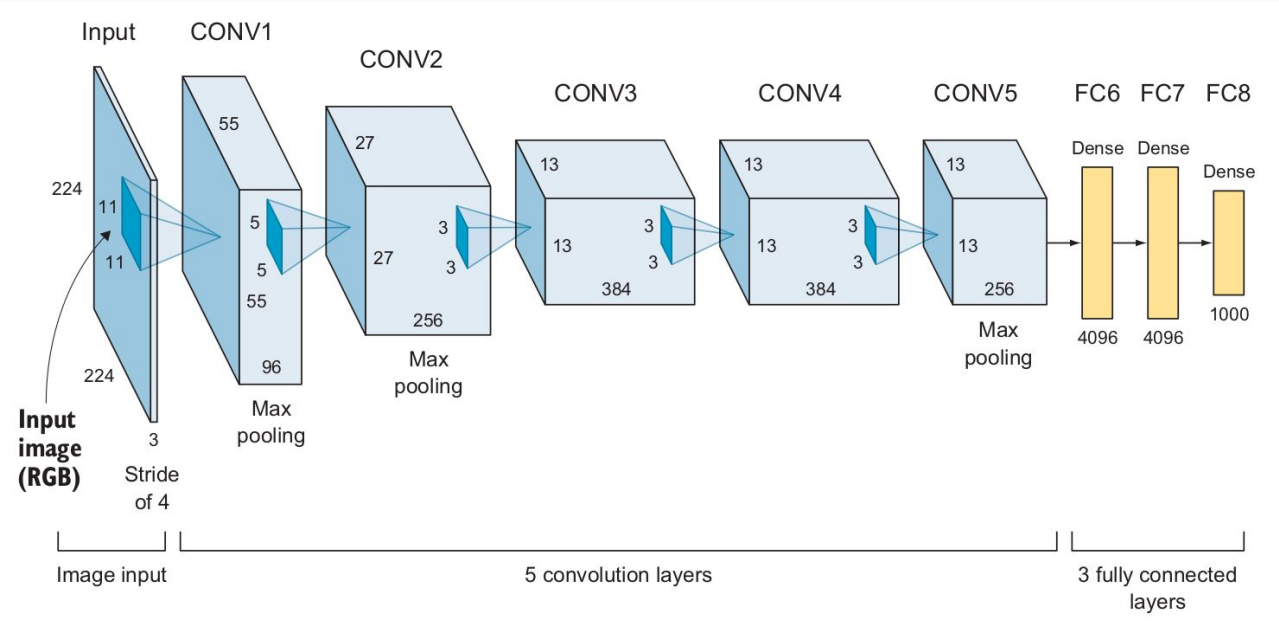
Not logged in. Login | Signup

ImageNet is an image database organized according to the WordNet hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. The project has been instrumental in advancing computer vision and deep learning research. The data is available for free to researchers for non-commercial use.

Mar 11 2021. ImageNet website update.

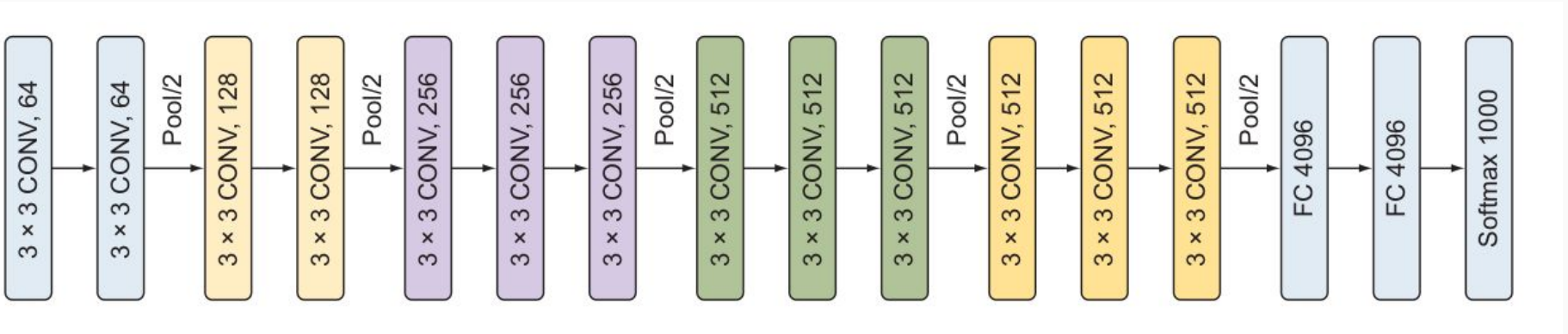
© 2020 Stanford Vision Lab, Stanford University, Princeton University imagenet.help.desk@gmail.com Copyright infringement

AlexNet: The Tipping Point



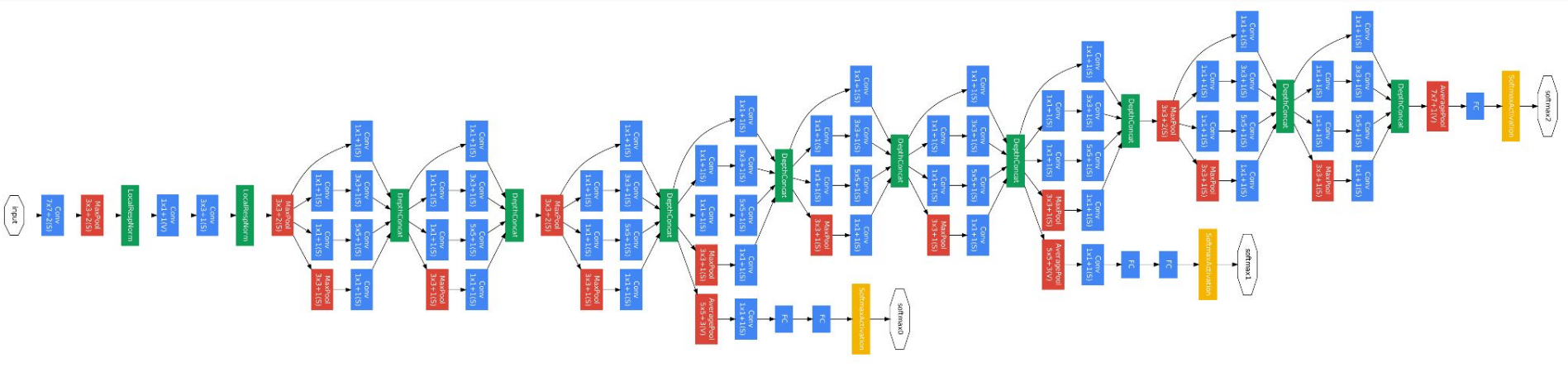
- Convolutional layers with the following kernel sizes: 11 × 11, 5 × 5, and 3 × 3
- Max pooling layers for images downsampling
- Dropout layers to avoid overfitting
- Unlike LeNet, ReLU activation functions in the hidden layers

VGG

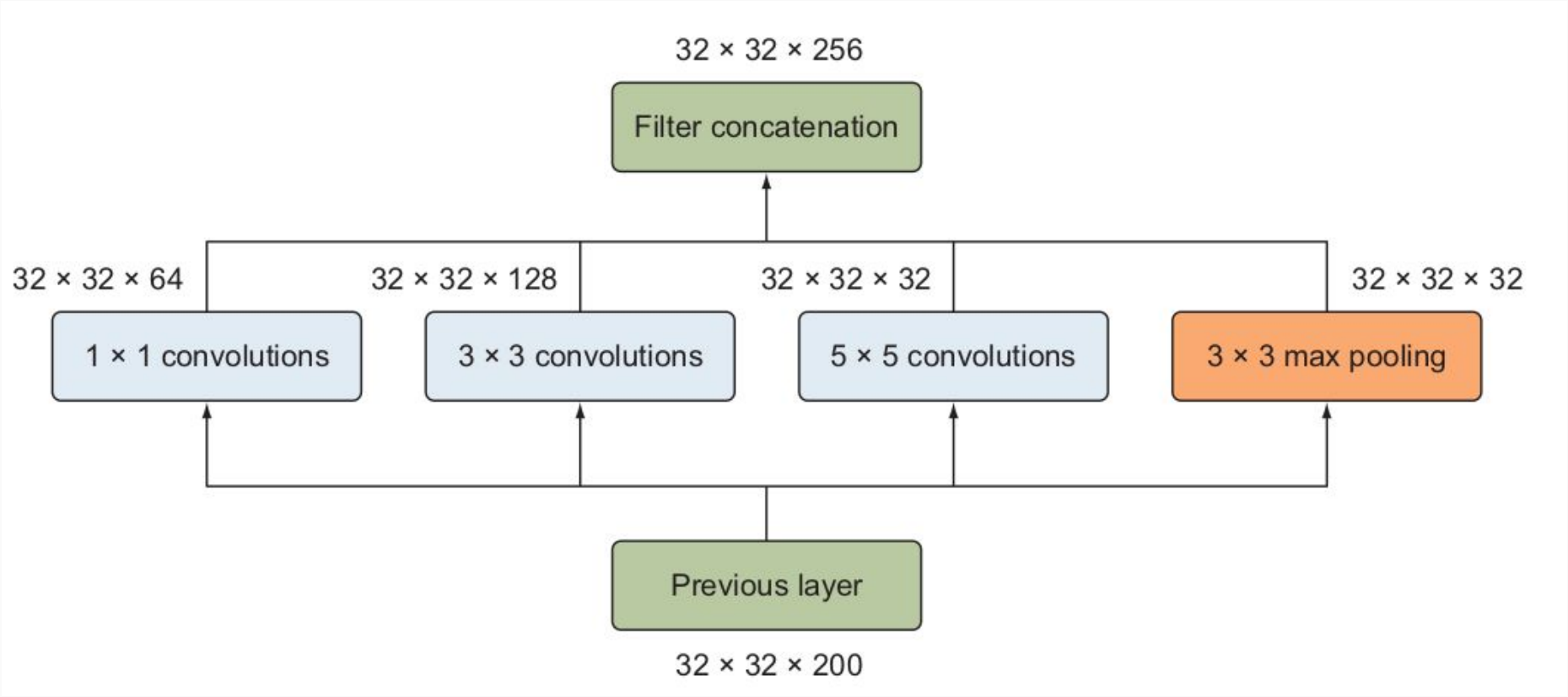

3 × 3 CONV, 64
3 × 3 CONV, 64
Pool/2
3 × 3 CONV, 128
3 × 3 CONV, 128
Pool/2
3 × 3 CONV, 256
3 × 3 CONV, 256
3 × 3 CONV, 256
Pool/2
3 × 3 CONV, 512
3 × 3 CONV, 512
3 × 3 CONV, 512
Pool/2
3 × 3 CONV, 512
3 × 3 CONV, 512
3 × 3 CONV, 512
Pool/2
FC 4096
FC 4096
Softmax 1000

WE NEED TO GO
DEEPER

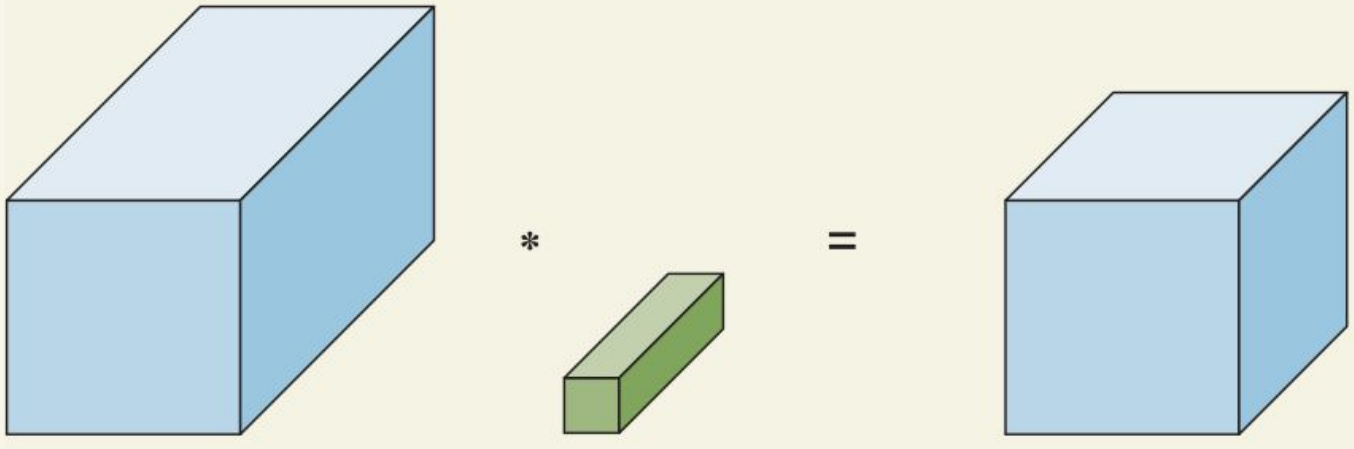
Inception and GoogLeNet



Inception Module: Naive Approach

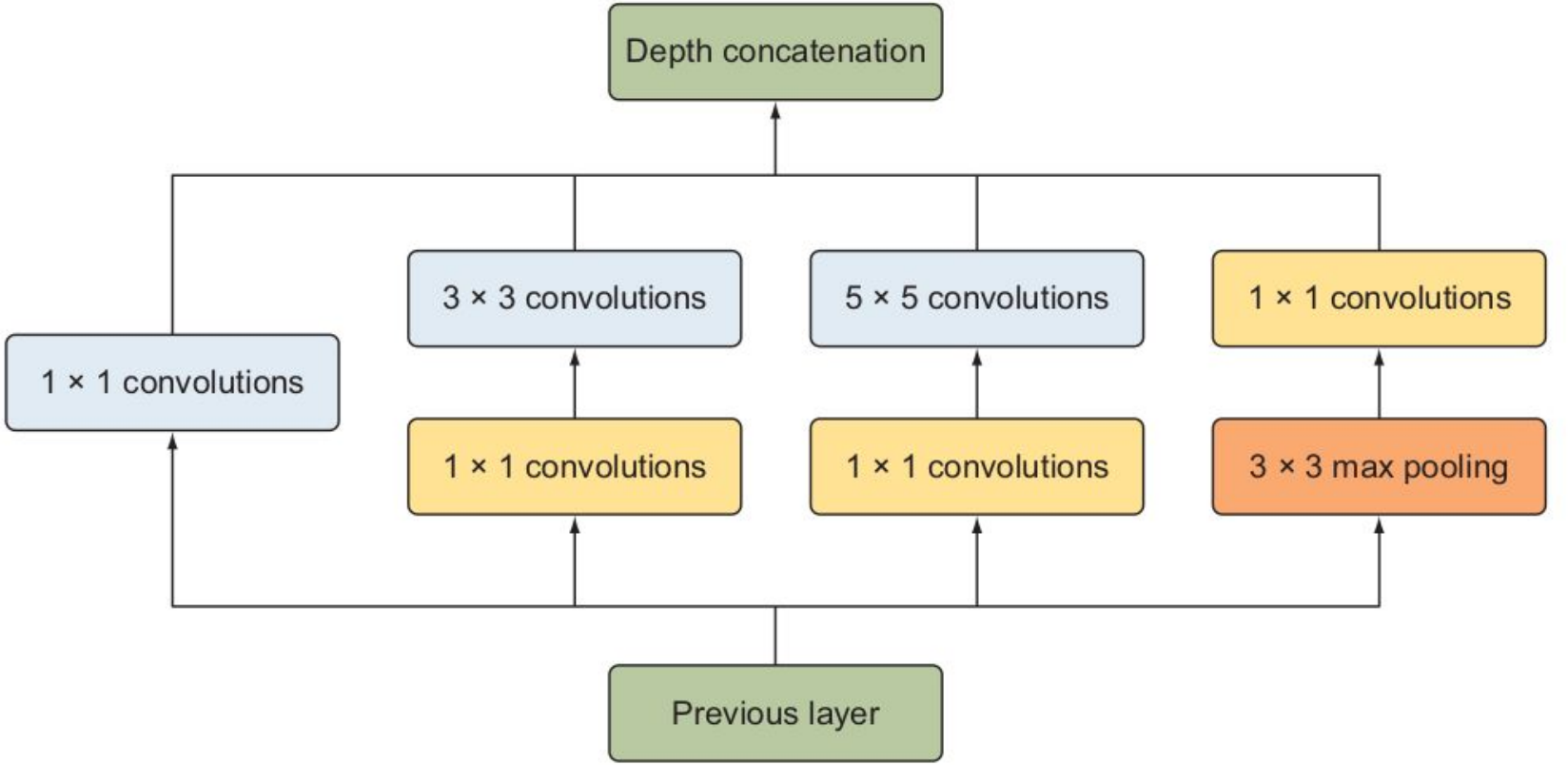


1 × 1 Convolutional (Bottleneck) Layers

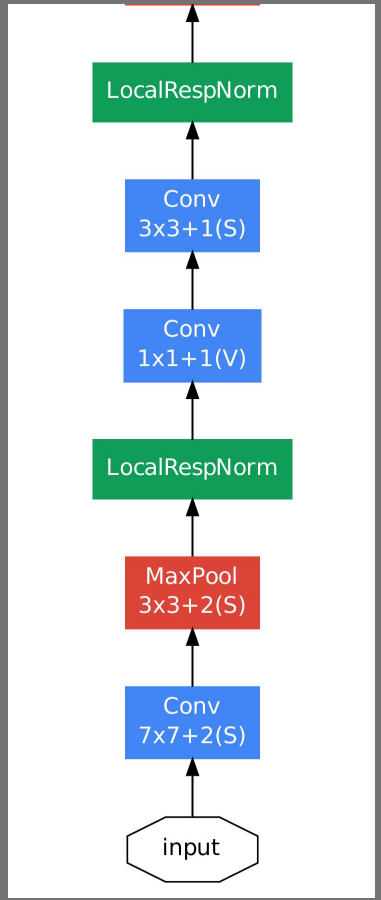


1 × 1 conv layers preserve the spatial dimensions but change the depth.

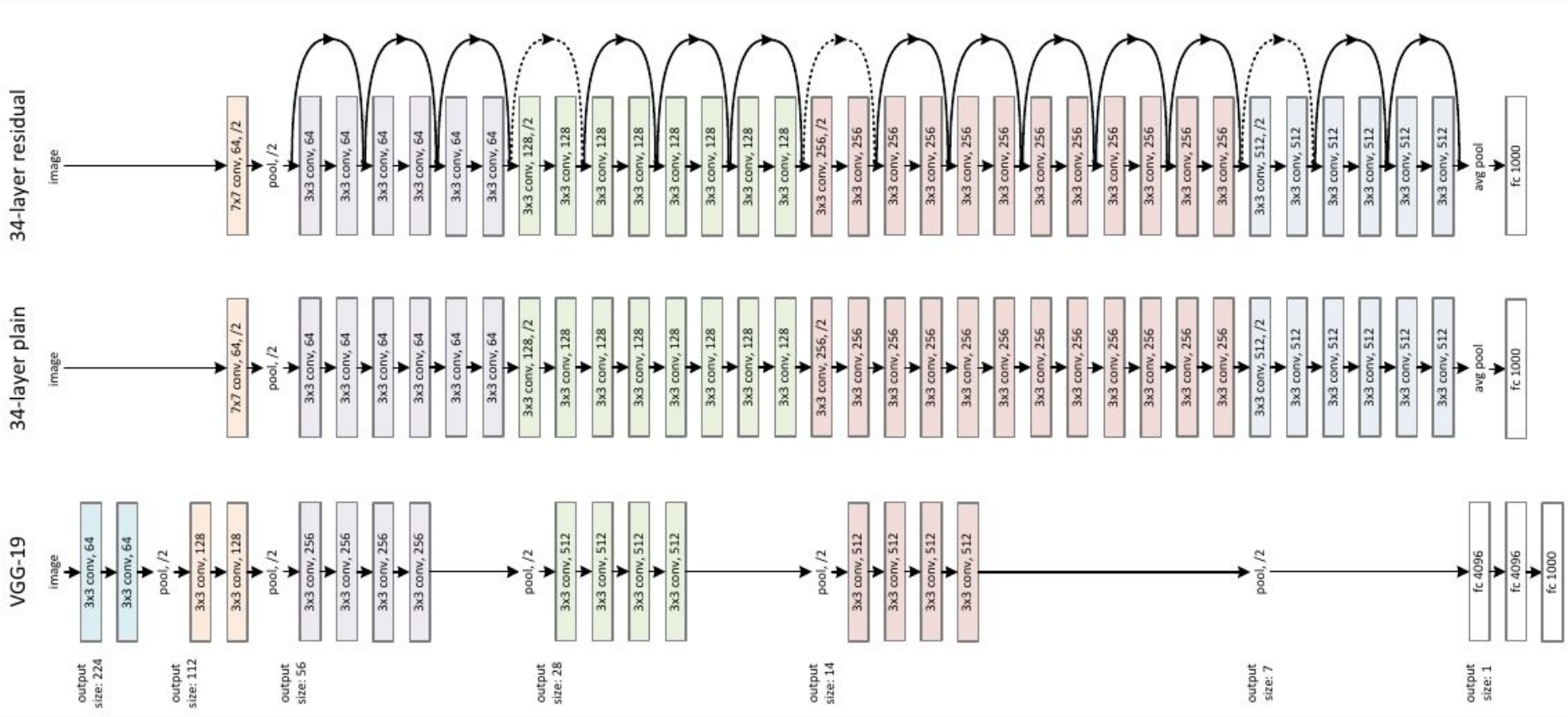
Inception Module



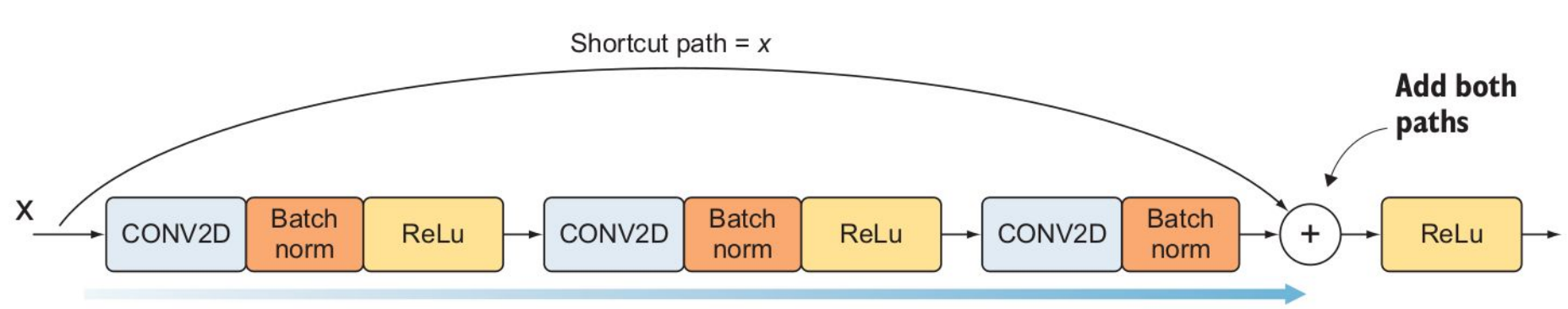
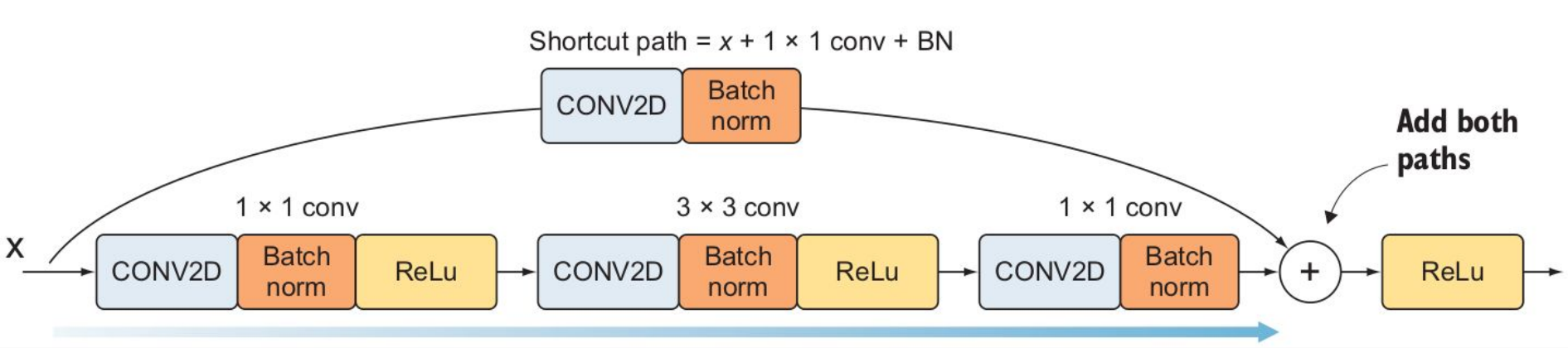
Inception and GoogLeNet



ResNet


VGG-19
34-layer plain
34-layer residual
image
3x3 conv, 64
pool, /2
output size: 224
output size: 112
output size: 56
output size: 28
output size: 14
output size: 7
output size: 1
fc 4096
fc 1000
7x7 conv, 64, /2
3x3 conv, 128
3x3 conv, 256
3x3 conv, 512
avg pool

Residual Block: Same input and output



Residual Block