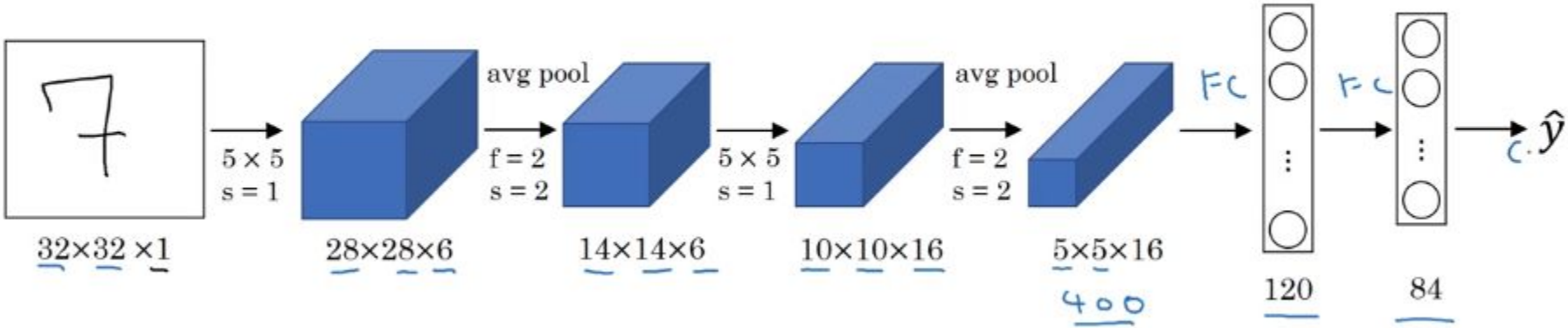
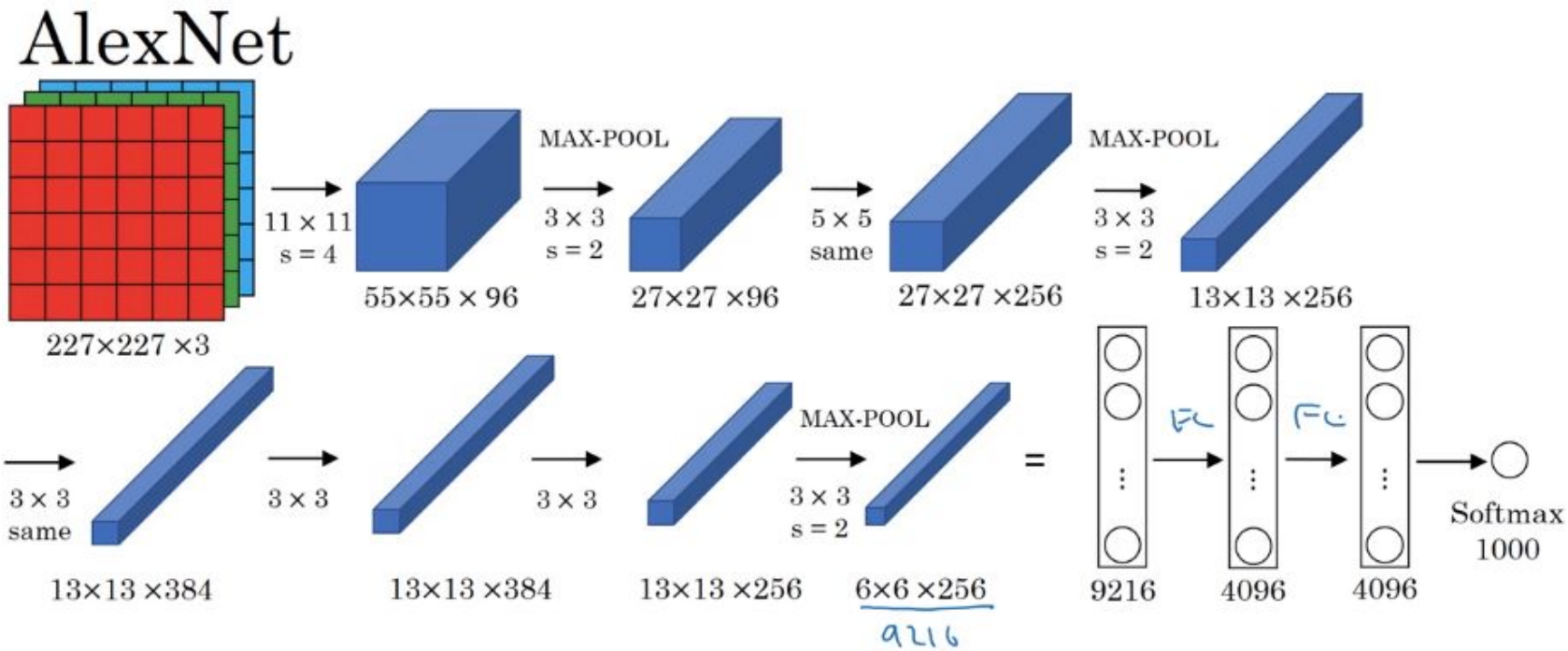
Deep CNN Architectures

LeNet - 5



Parameters: 60K

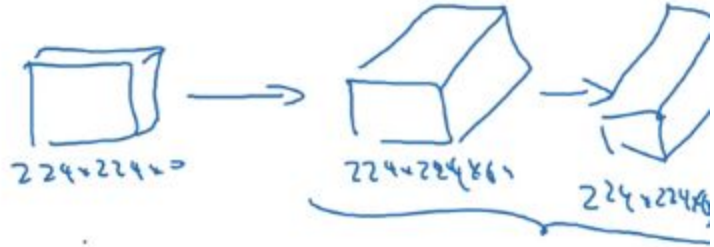


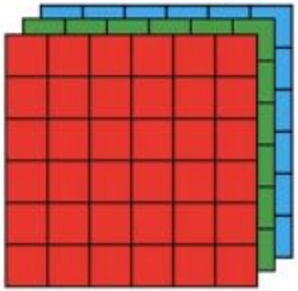
Parameters: 60,000K

VGG - 16

CONV = 3×3 filter, s = 1, same

MAX-POOL = 2×2 , s = 2





224×224 ×3

→ [CONV 64] ×2 → 224×224×64 → POOL → 112×112 ×64 → [CONV 128] ×2 → 112×112 ×128 → POOL → 56×56 ×128

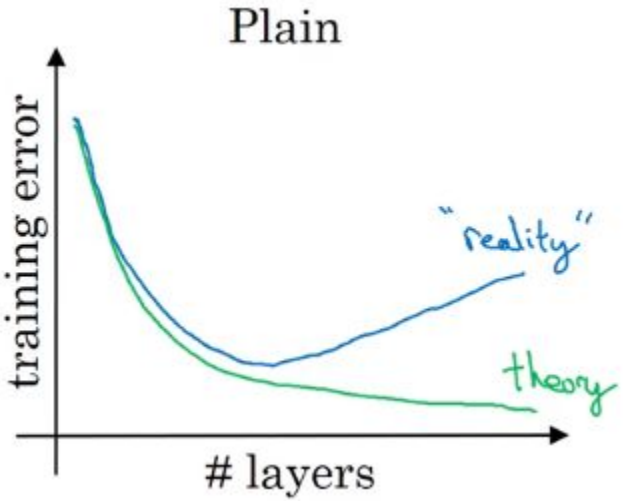
→ [CONV 256] ×3 → 56×56 ×256 → POOL → 28×28 ×256 → [CONV 512] × 3 → 28×28 ×512 → POOL → 14×14×512

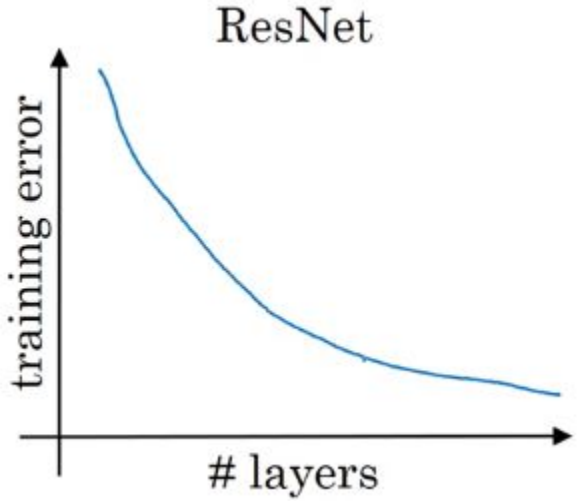
→ [CONV 512] ×3 → 14×14 ×512 → POOL → 7×7×512 → FC 4096 → FC 4096 → Softmax 1000

Parameters: 134,000K

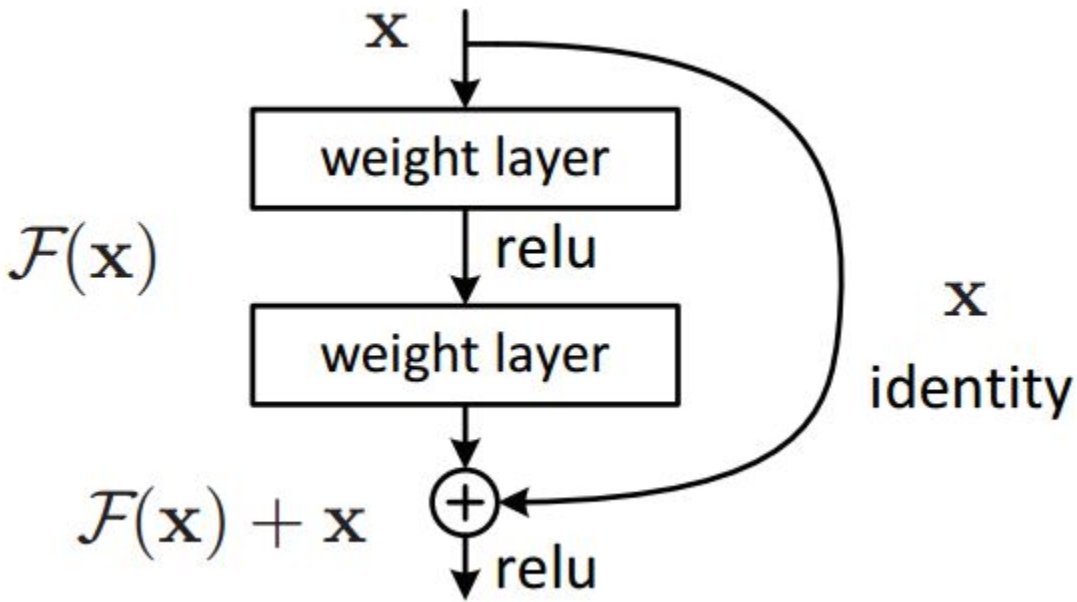
DICE ANALYTICS

Problem With Very Deep Networks

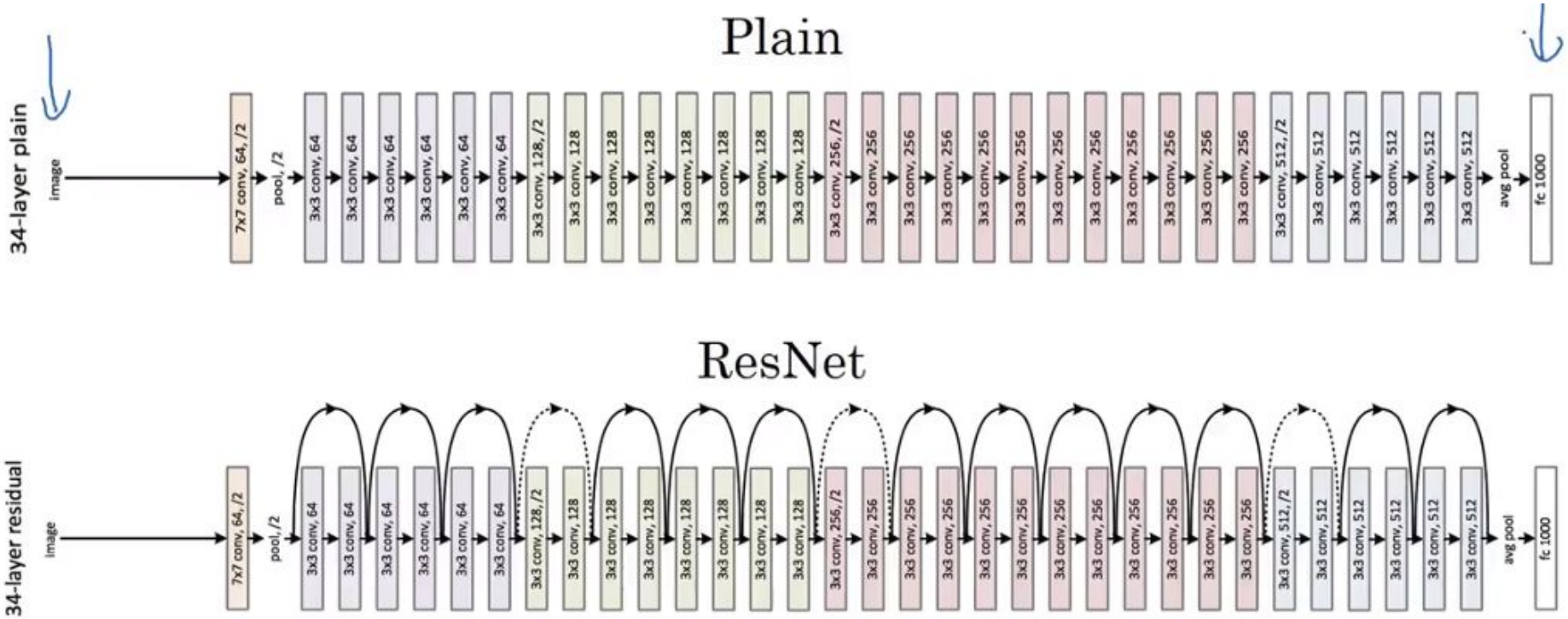




Residual Block

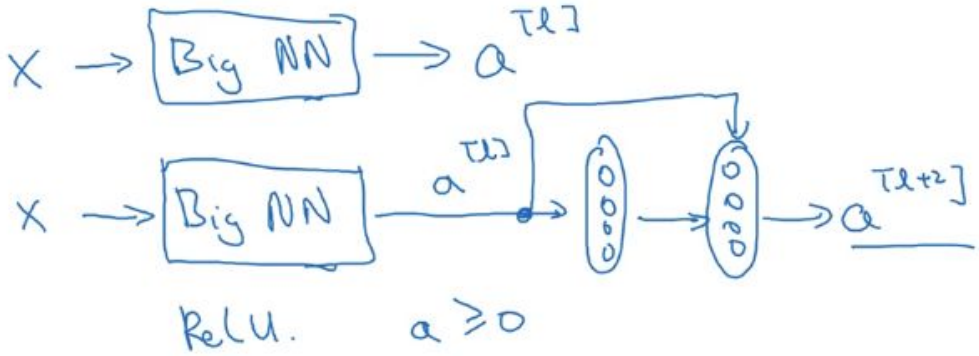


ResNet


Plain
34-layer plain
image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
avg pool
fc 1000
ResNet
34-layer residual
image
7x7 conv, 64, /2
pool, /2
3x3 conv, 64
3x3 conv, 128, /2
3x3 conv, 128
3x3 conv, 256, /2
3x3 conv, 256
3x3 conv, 512, /2
3x3 conv, 512
avg pool
fc 1000

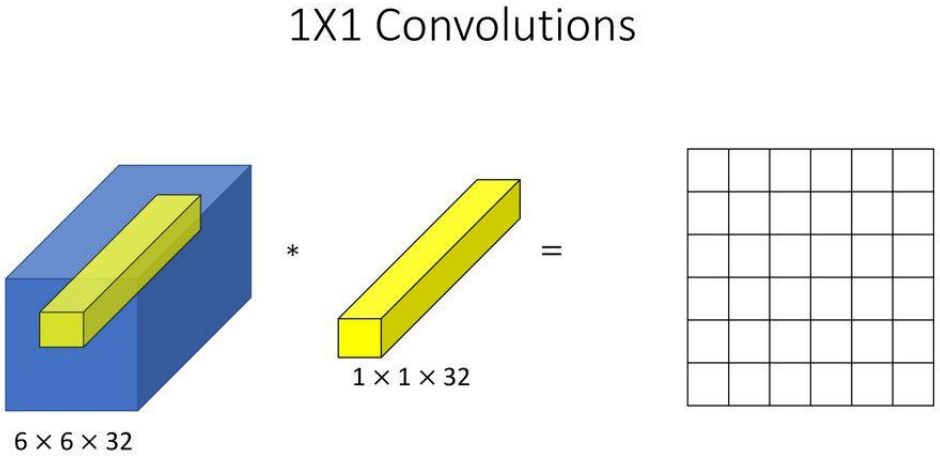
Why ResNets work?

It’s easy for a ResNet to learn identity function.

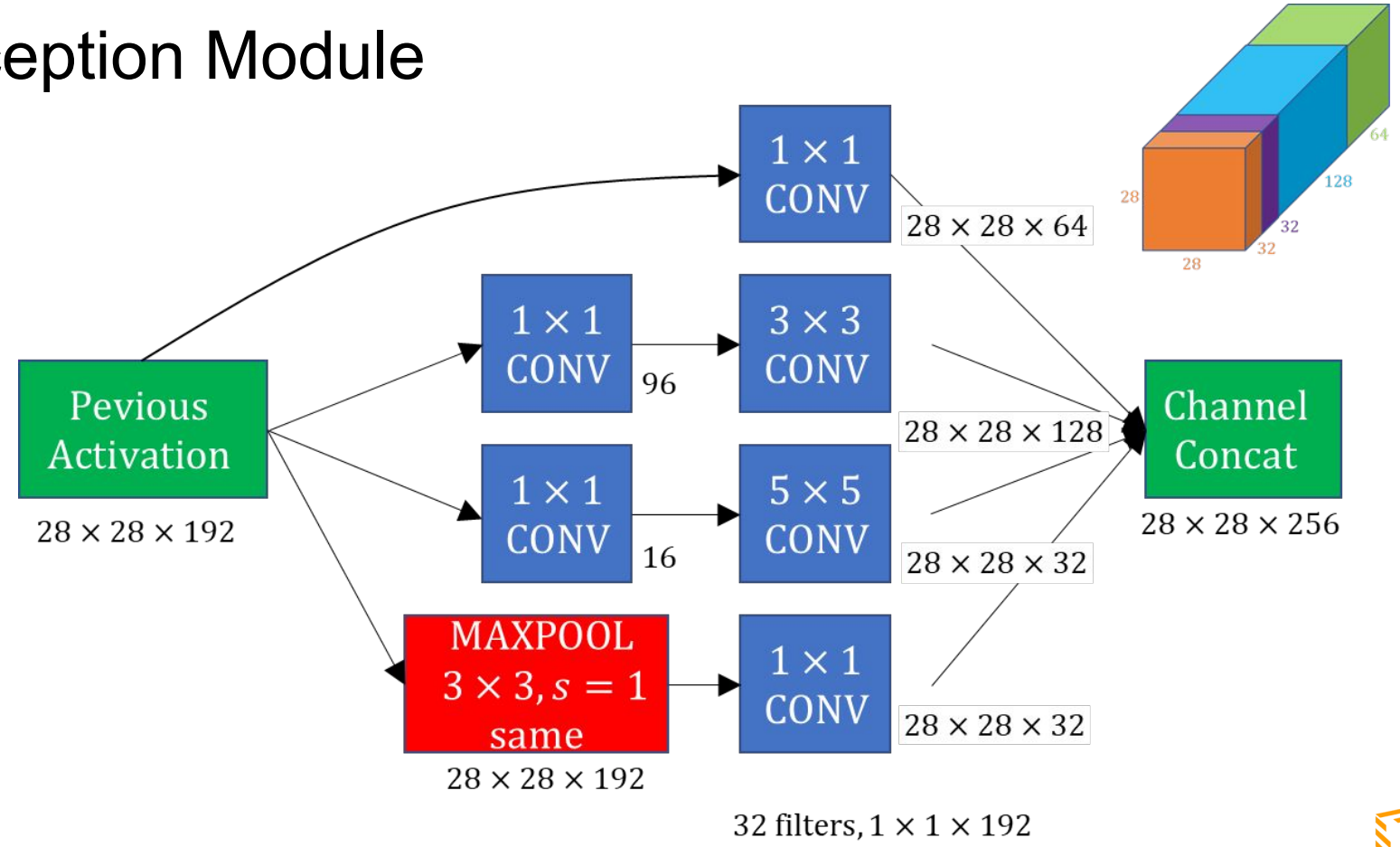


1x1 Convolution

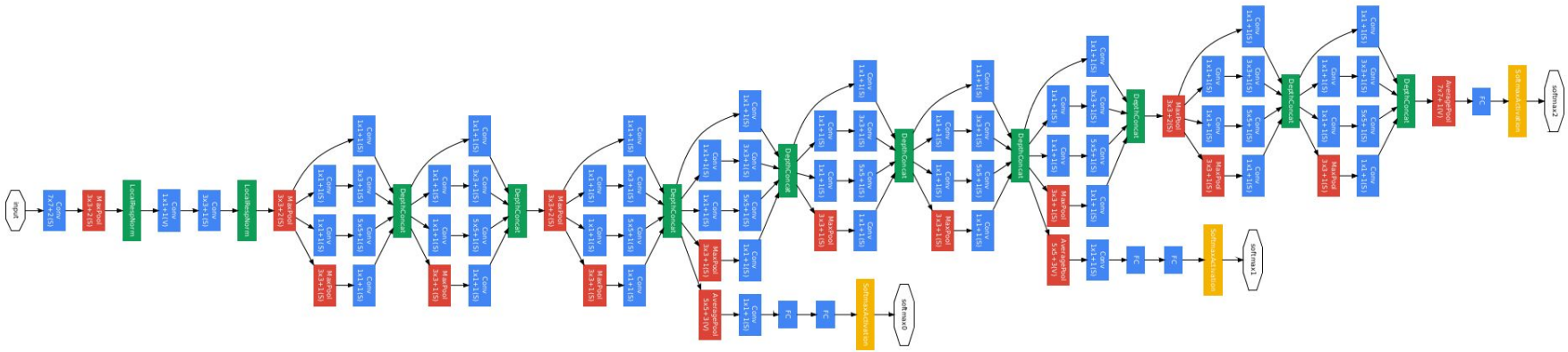
- Essentially a dense neural network layer
- Acts as a bottleneck layer to reduce depth



Inception Module



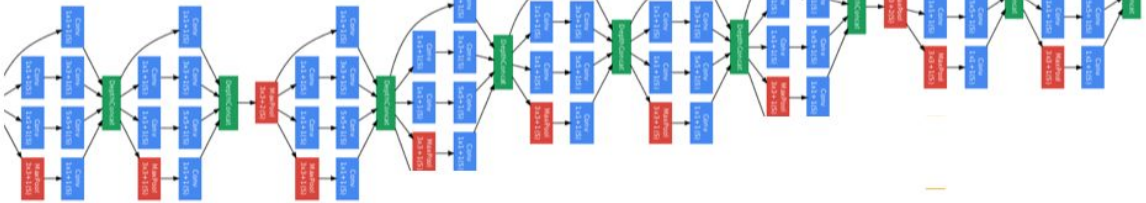
GoogLeNet / Inception Network



DICE ANALYTICS

inception modules

stem



output classifier

auxiliary classifiers