Case Studies

Classic networks

deeplearning.ai

LeNet - 5

32×32×1 → 5×5 s=1 → 28×28×6 → avg pool f=2 s=2 → 14×14×6 → 5×5 s=1 → 10×10×16 → avg pool f=2 s=2 → 5×5×16 (400) → FC → 120 → FC → 84 → $\hat{y}$ softmax 10

Non-linearity after pooling
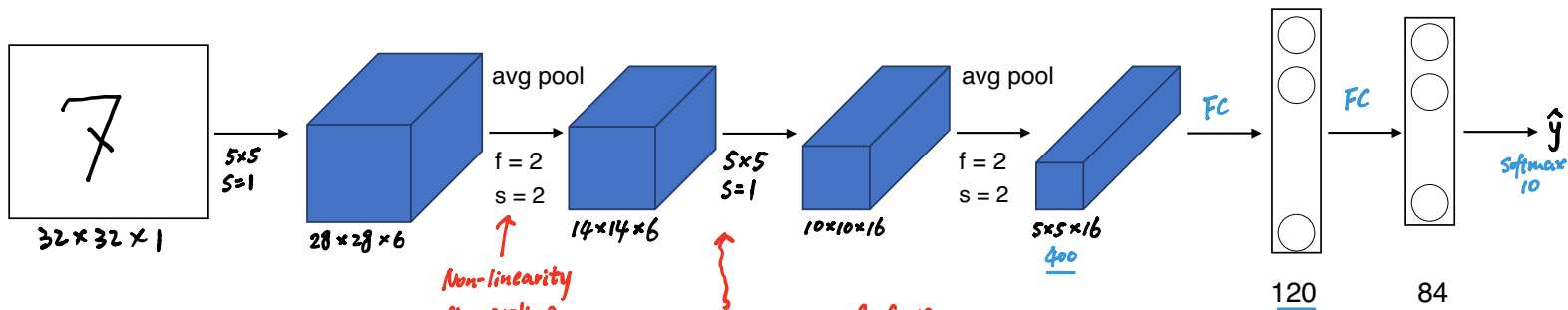
$n_H \times n_w \times \underline{n_c}$ $f \times f \times n_c$

60 k parameters

$n_H, n_w \downarrow \quad n_c \uparrow$

Conv pool Conv pool fc fc output.

Advanced: Sigmoid/tanh   ReLU

[LeCun et al., 1998. Gradient-based learning applied to document recognition]

Andrew Ng

AlexNet

227 × 227 × 3 → [11×11 S=4] → 55 × 55 × 96 → MAX-POOL [3×3 S=2] → 27 × 27 × 96 → [5×5 Same] → 27 × 27 × 96 → MAX-POOL [3×3 S=2] → 13 × 13 × 256

→ [3×3 Same] → 13 × 13 × 384 → [3×3] → 13 × 13 × 384 → [3×3] → 13 × 13 × 384 → MAX-POOL [3×3 S=2] → 6 × 6 × 256 (9216) = 9216 → FC → 4096 → FC → 4096 → Softmax 1000
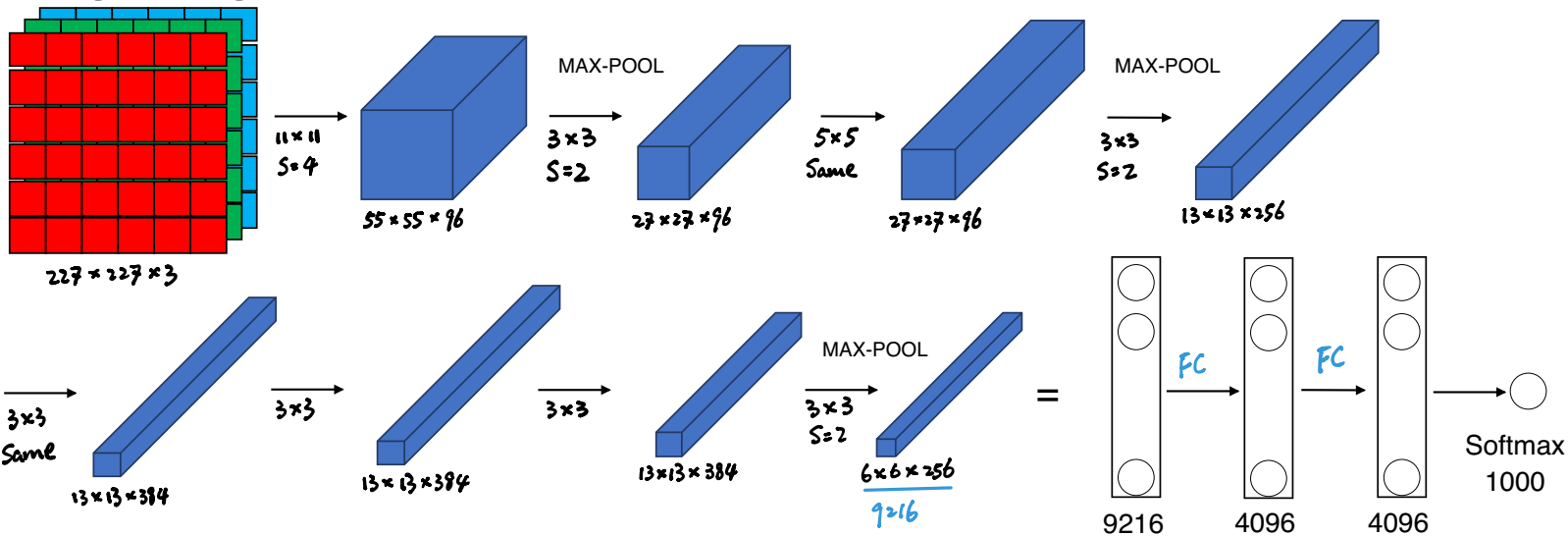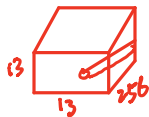
— Similarity to LeNet, but much bigger   ~ 60M parameters.

— ReLU

— Multiple GPUs

— Local Response Normalization. (LRN)   13 13 256
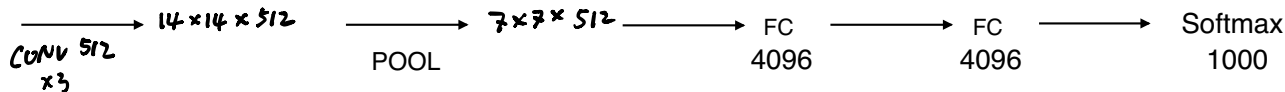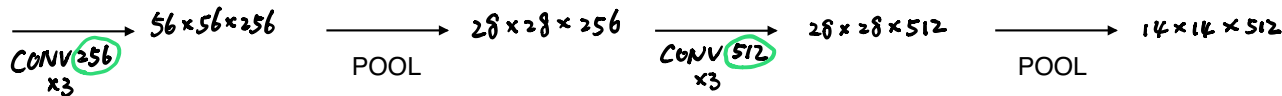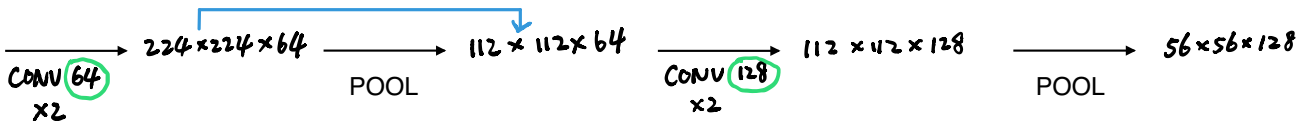
[Krizhevsky et al., 2012. ImageNet classification with deep convolutional neural networks]   Andrew Ng

VGG - 16 — 16 layers with weight.   VGG-19 Even bigger.

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

~138 M

[Simonyan & Zisserman 2015. Very deep convolutional networks for large-scale image recognition]

Andrew Ng