deeplearning.ai

Case Studies

Why look at case studies?

Outline

Classic networks:

- LeNet-5 ←
- AlexNet ←
- VGG ←

ResNet (152)

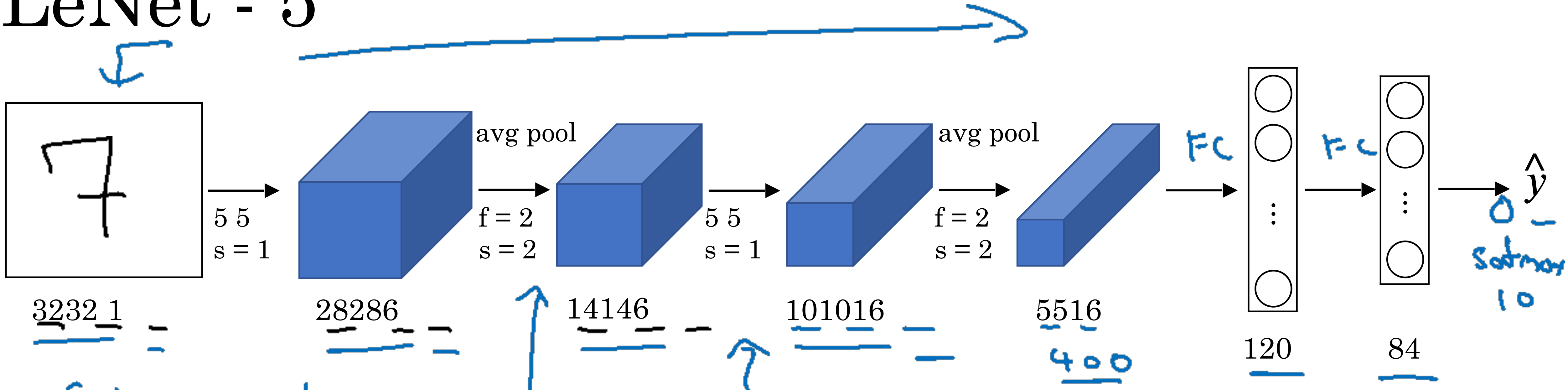
Inception

deeplearning.ai

Case Studies

Classic networks

LeNet - 5



60K parameters

n_H, n_W ↓ n_C ↑

nonlinearity after pooling

$n_H \times n_W \times n_C$

$f \times f \times n_C$

conv pool conv pool fc fc output

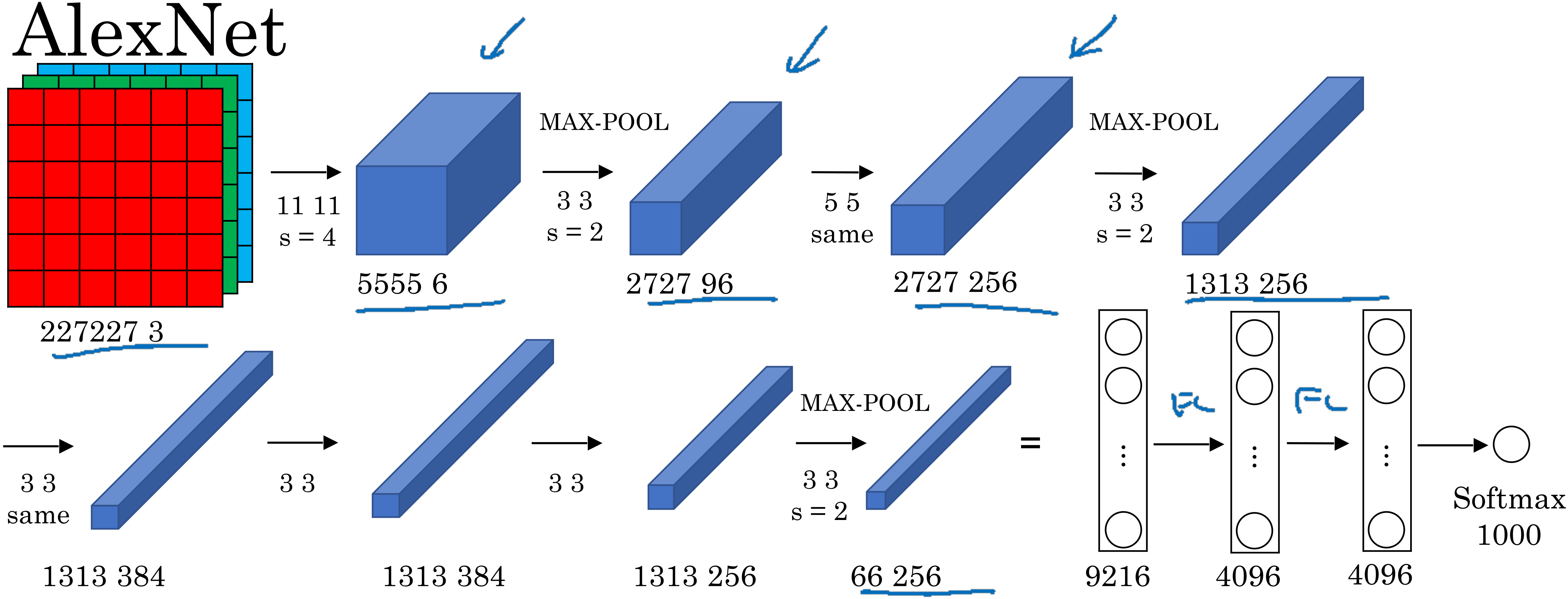
Advanced: sigmoid/tanh ReLU

II III.

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

Andrew Ng

AlexNet



227227 3

11 11
s = 4

5555 6

MAX-POOL
3 3
s = 2

2727 96

5 5
same

2727 256

MAX-POOL
3 3
s = 2

1313 256

3 3
same

1313 384

3 3

1313 384

3 3

1313 256

MAX-POOL
3 3
s = 2

66 256

=

9216

FC

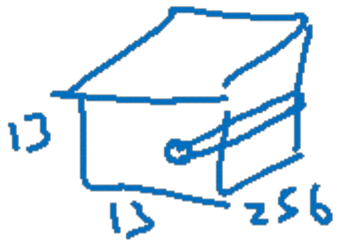
4096

FC

4096

Softmax
1000

9216



- Similar to LeNet, but much bigger.
- ReLU
- Multiple GPUs.
- Local Response Normalization (LRN)

~60M parameters

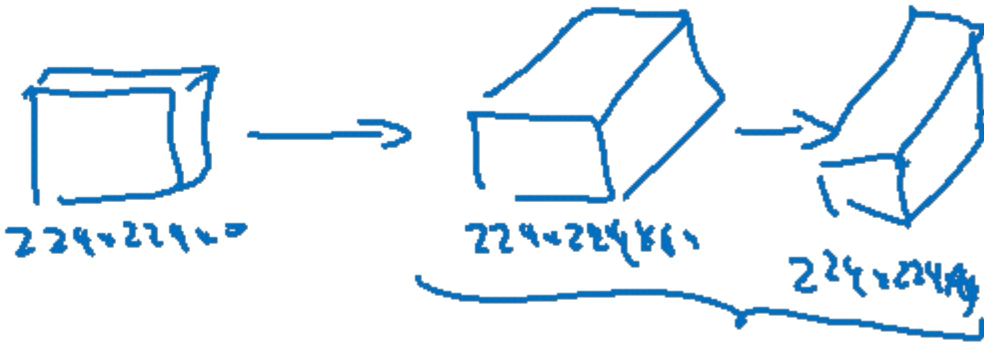
[Krizhevsky et al., 2012. ImageNet classification with deep convolutional neural networks]

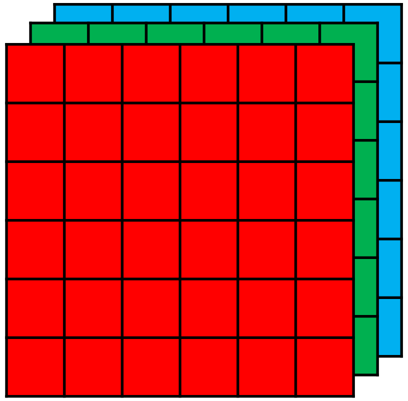
VGG - 16

VGG-19

CONV = 33 filter, s = 1, same

MAX-POOL = 22 , s = 2





224224 3

→ [CONV 64] 2 → 22422464 → POOL → 112112 64 → [CONV 128] 2 → 112112 128 → POOL → 5656 128

→ [CONV 256] 3 → 5656 256 → POOL → 2828 256 → [CONV 512] 3 → 2828 512 → POOL → 1414512

→ [CONV 512] 3 → 1414 512 → POOL → 77512 → FC 4096 → FC 4096 → Softmax 1000

$n_H, n_W \downarrow$ $n_C \uparrow$ ~138M

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

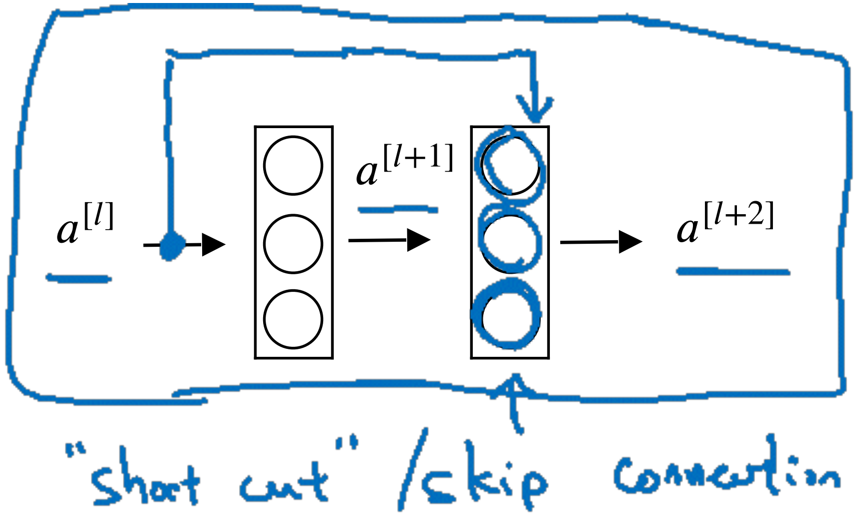
Andrew Ng

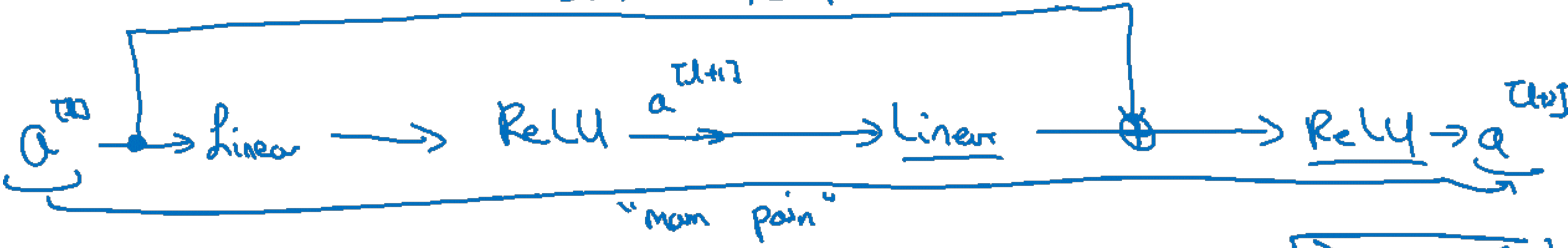
deeplearning.ai

Case Studies

Residual Networks (ResNets)

Residual block





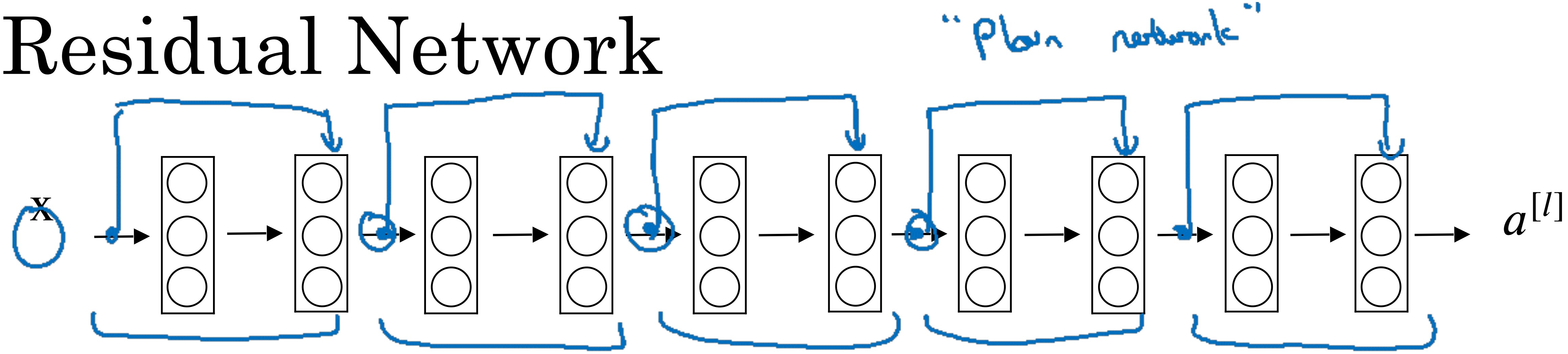
$$z^{[l+1]} = W^{[l+1]} a^{[l]} + b^{[l+1]} \quad a^{[l+1]} = g(z^{[l+1]}) \quad \underline{z^{[l+2]} = W^{[l+2]} a^{[l+1]} + b^{[l+2]}} \quad \boxed{\cancel{a^{[l+2]} = g(z^{[l+2]})}}$$

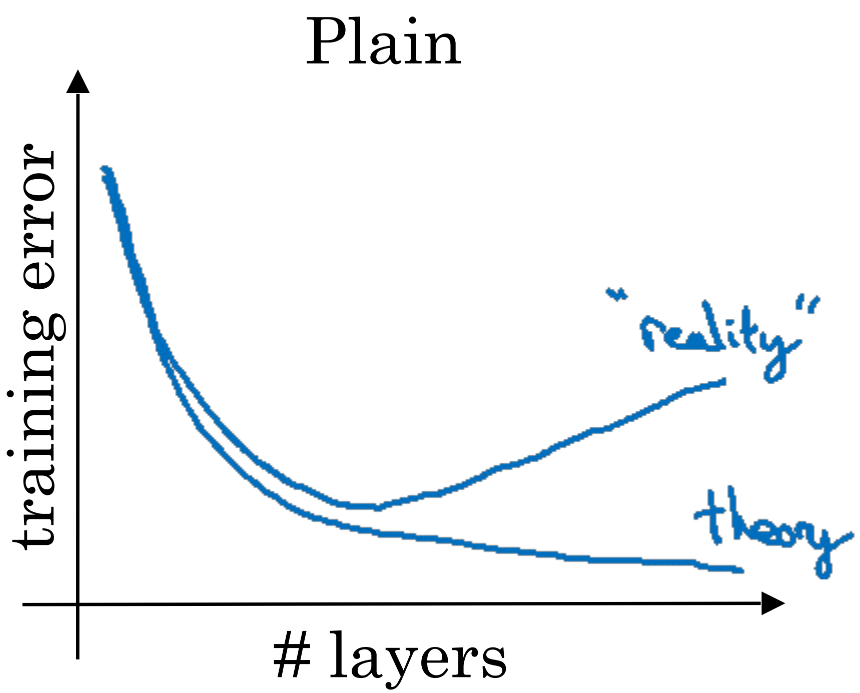
$$a^{[l+2]} = g(z^{[l+2]} + a^{[l]})$$

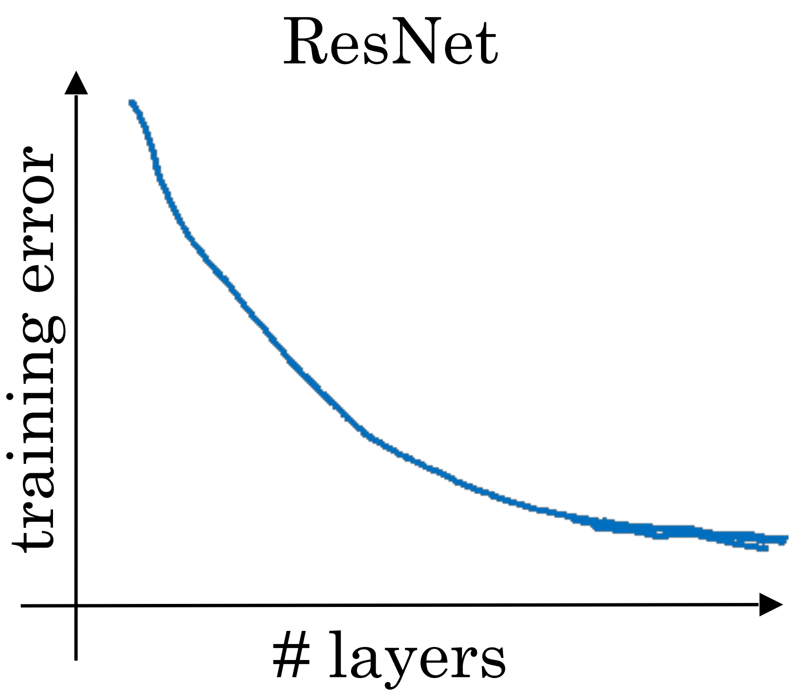
[He et al., 2015. Deep residual networks for image recognition]

Andrew Ng

Residual Network







[He et al., 2015. Deep residual networks for image recognition]

Andrew Ng

deeplearning.ai

Case Studies

Network in Network and 11 convolutions

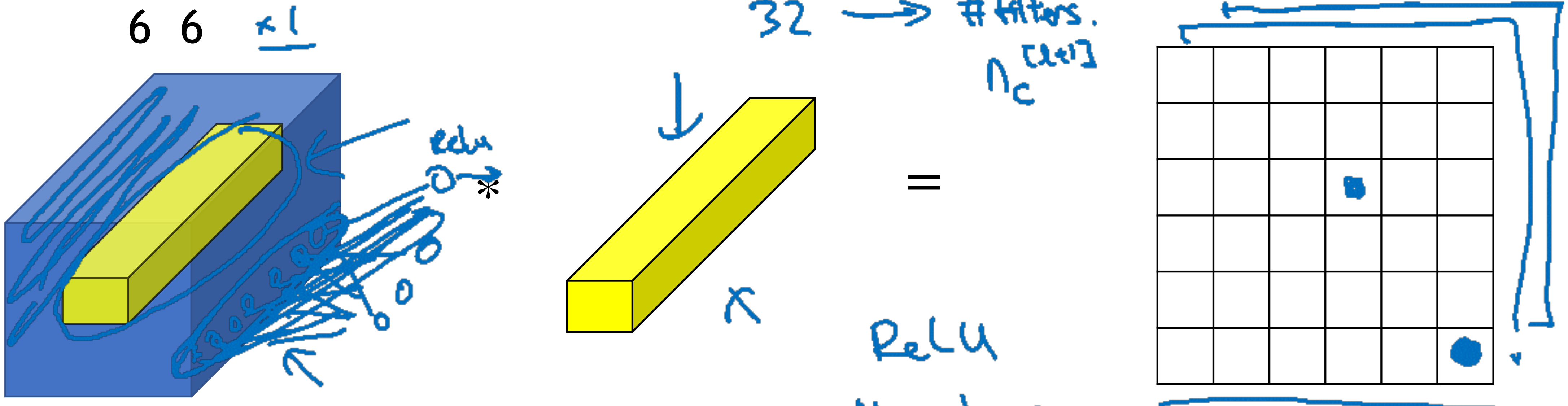
Why does a 1 1 convolution do?

$$\begin{array}{|c|c|c|c|c|c|}\hline 1&2&3&6&5&8\\\hline 3&5&5&1&3&4\\\hline 2&1&3&4&9&3\\\hline 4&7&8&5&7&9\\\hline 1&5&3&7&4&8\\\hline 5&4&9&8&3&5\\\hline\end{array} \;*\; \boxed{2} \;=\; \begin{array}{|c|c|c|c|c|c|}\hline 2&4&6&\cdots&&\\\hline &&&&&\\\hline &&&&&\\\hline &&&&&\\\hline &&&&&\\\hline &&&&&\\\hline\end{array}$$

6 6 $\underline{\times 1}$

32 → # filters.

$n_c^{[l+1]}$



ReLU

Network in Network

$\underline{6\ \ 6}\ \underline{32}$

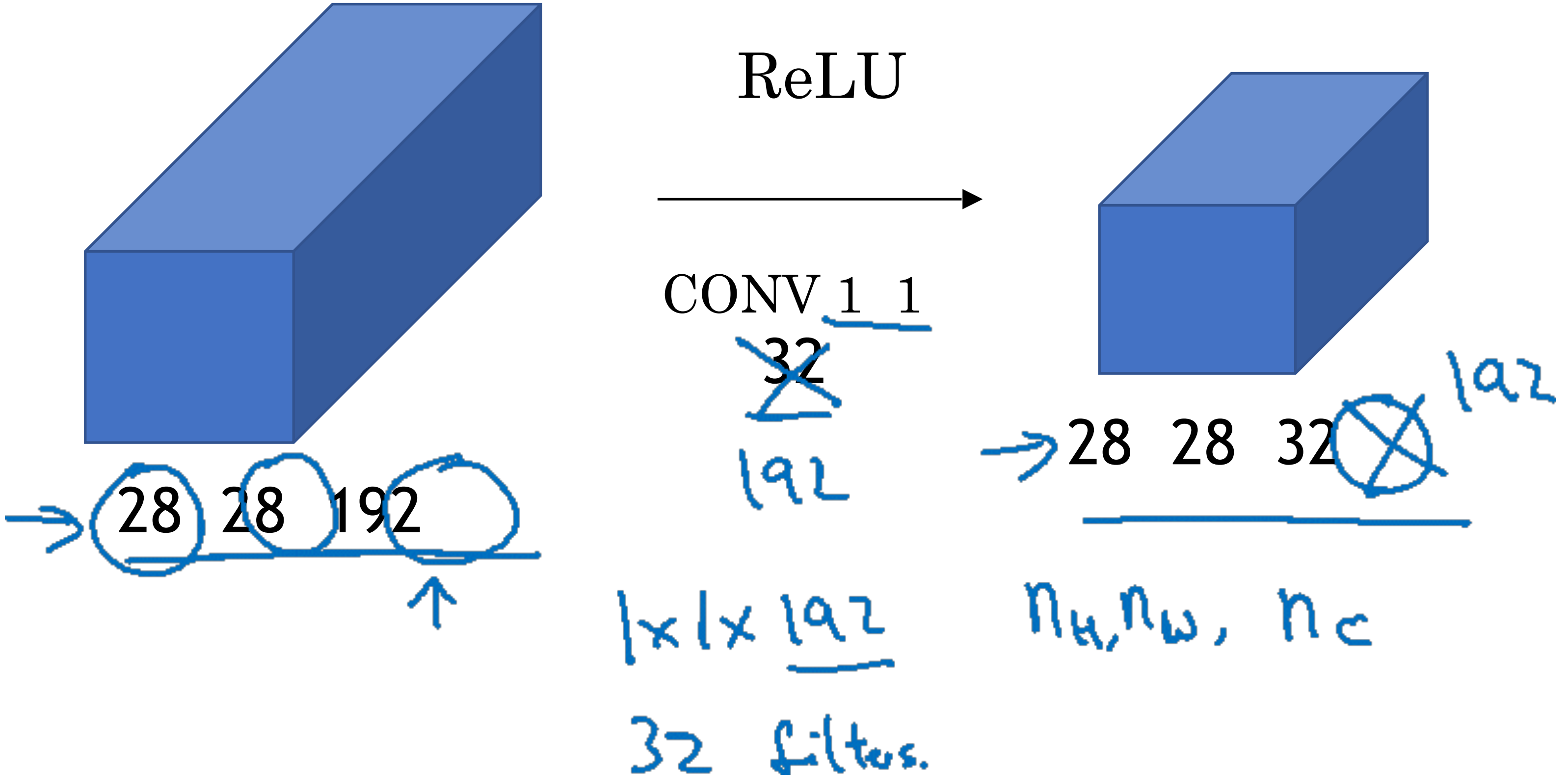
$\underline{\underline{1\ \ 1\ \ 32}}$

$\boxed{6\ \ 6\ \ \#\ \text{filters}}$

[Lin et al., 2013. Network in network]

Andrew Ng

Using 11 convolutions



ReLU

CONV 1 1
~~32~~
192

28 28 192

28 28 32
192

1x1x192

32 filters.

n_H, n_W, n_C

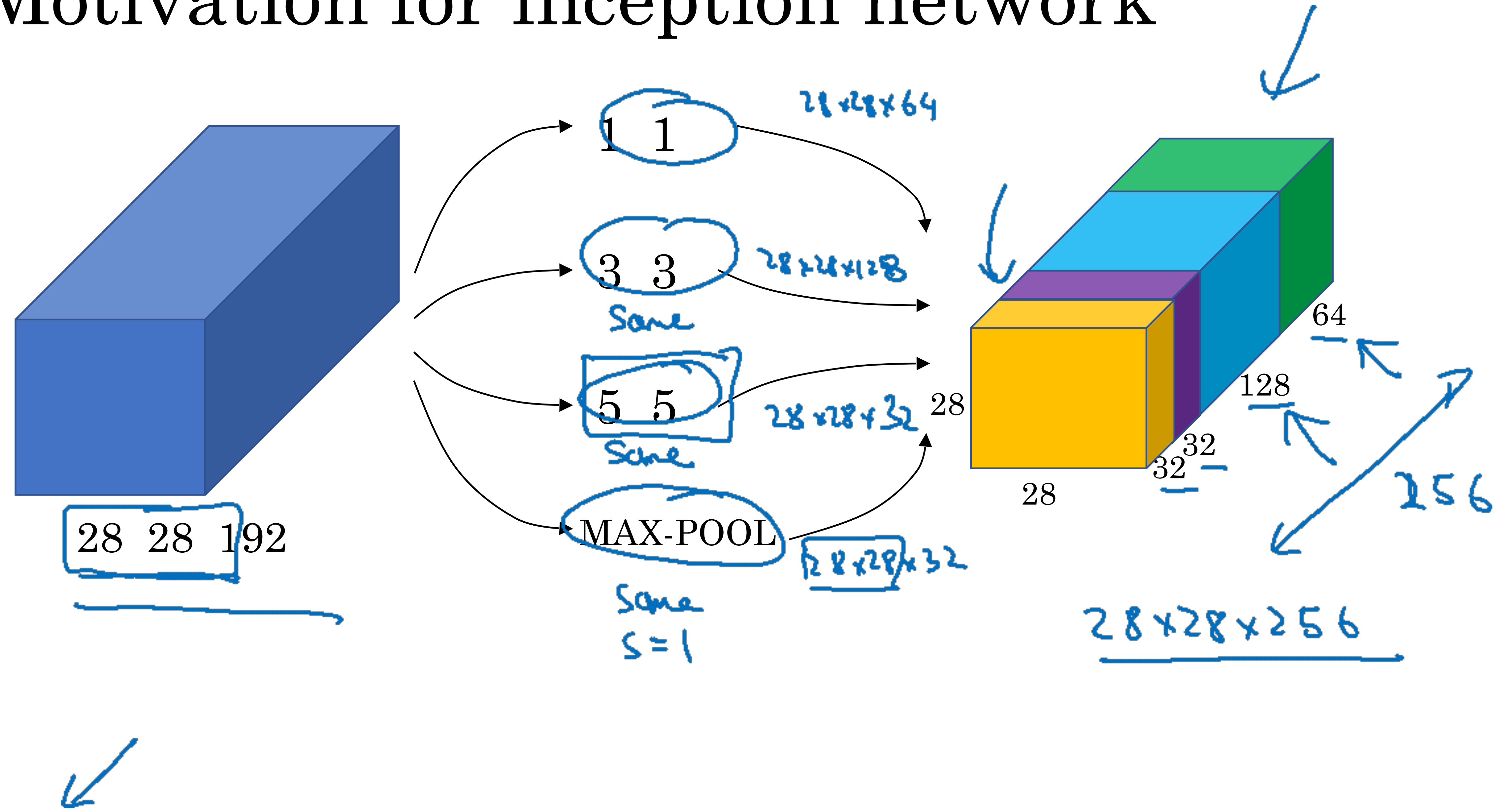
[Lin et al., 2013. Network in network]

deeplearning.ai

Case Studies

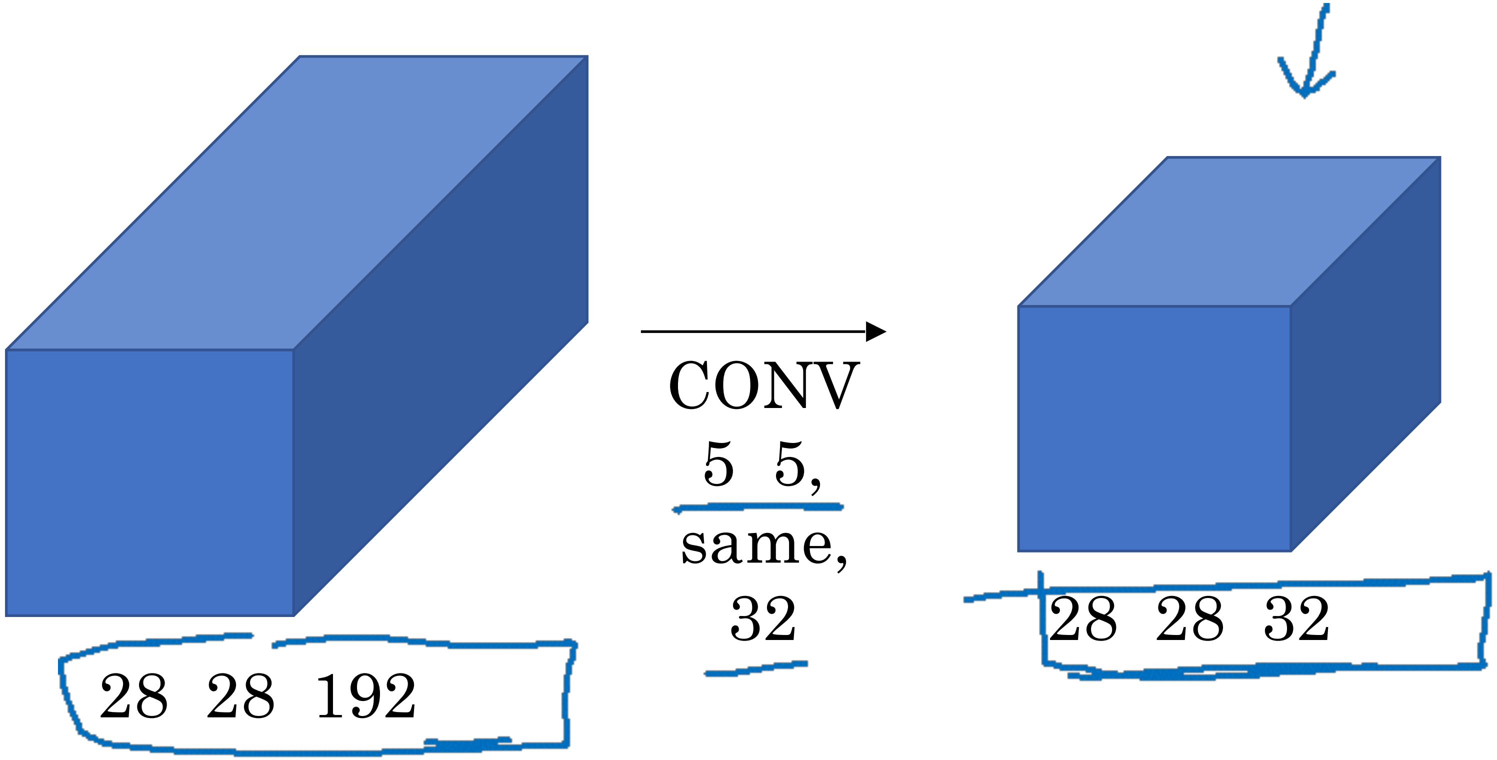
Inception network motivation

Motivation for inception network



[Szegedy et al. 2014. Going deeper with convolutions]

The problem of computational cost



28 28 192

CONV
5 5,
same,
32

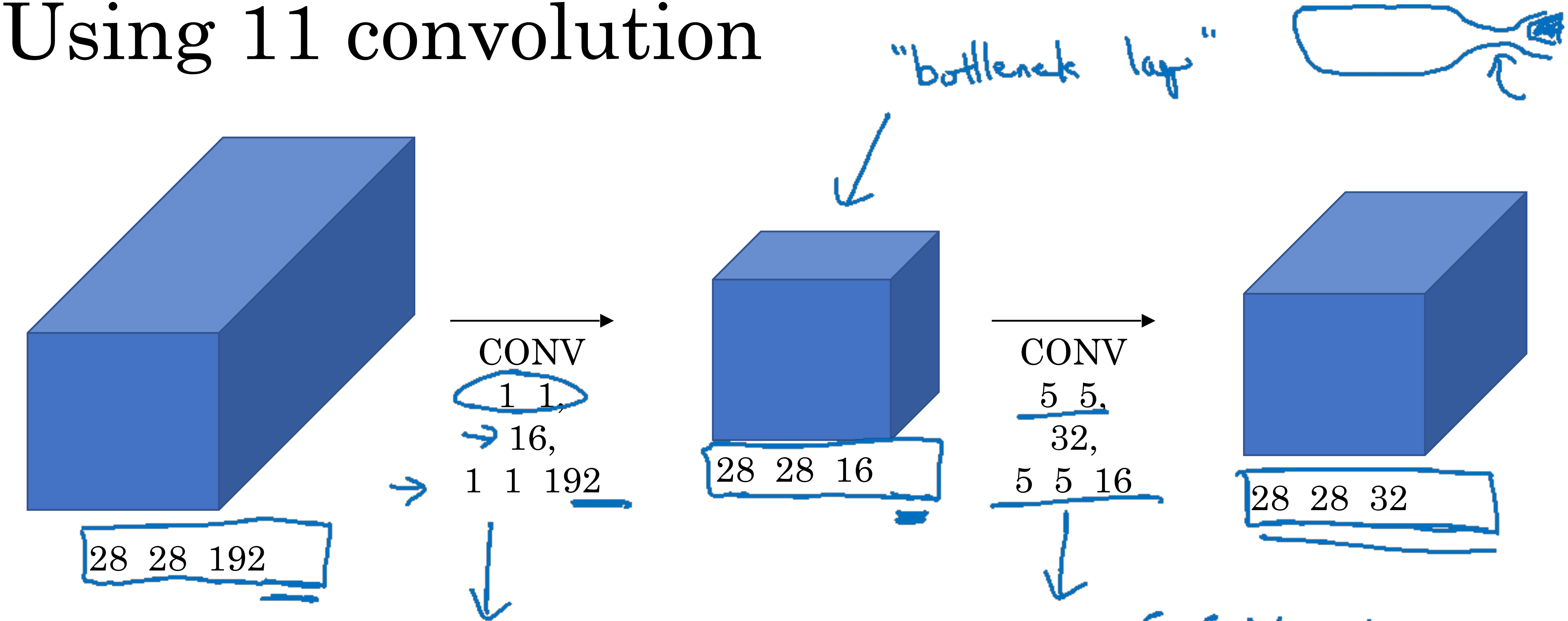
28 28 32

32 filters.

filters are $\underline{5 \times 5 \times 192}$.

$\underline{28 \times 28 \times 32} \times \underline{5 \times 5 \times 192} = \underline{120M}$.

Using 1 1 convolution



"bottleneck layer"

28 28 192

CONV
1 1,
16,
1 1 192

28 28 16

CONV
5 5,
32,
5 5 16

28 28 32

28x28x16 × 192 = 2.4M

28x28x32 × 5x5x16 = 10.0M

12.4M

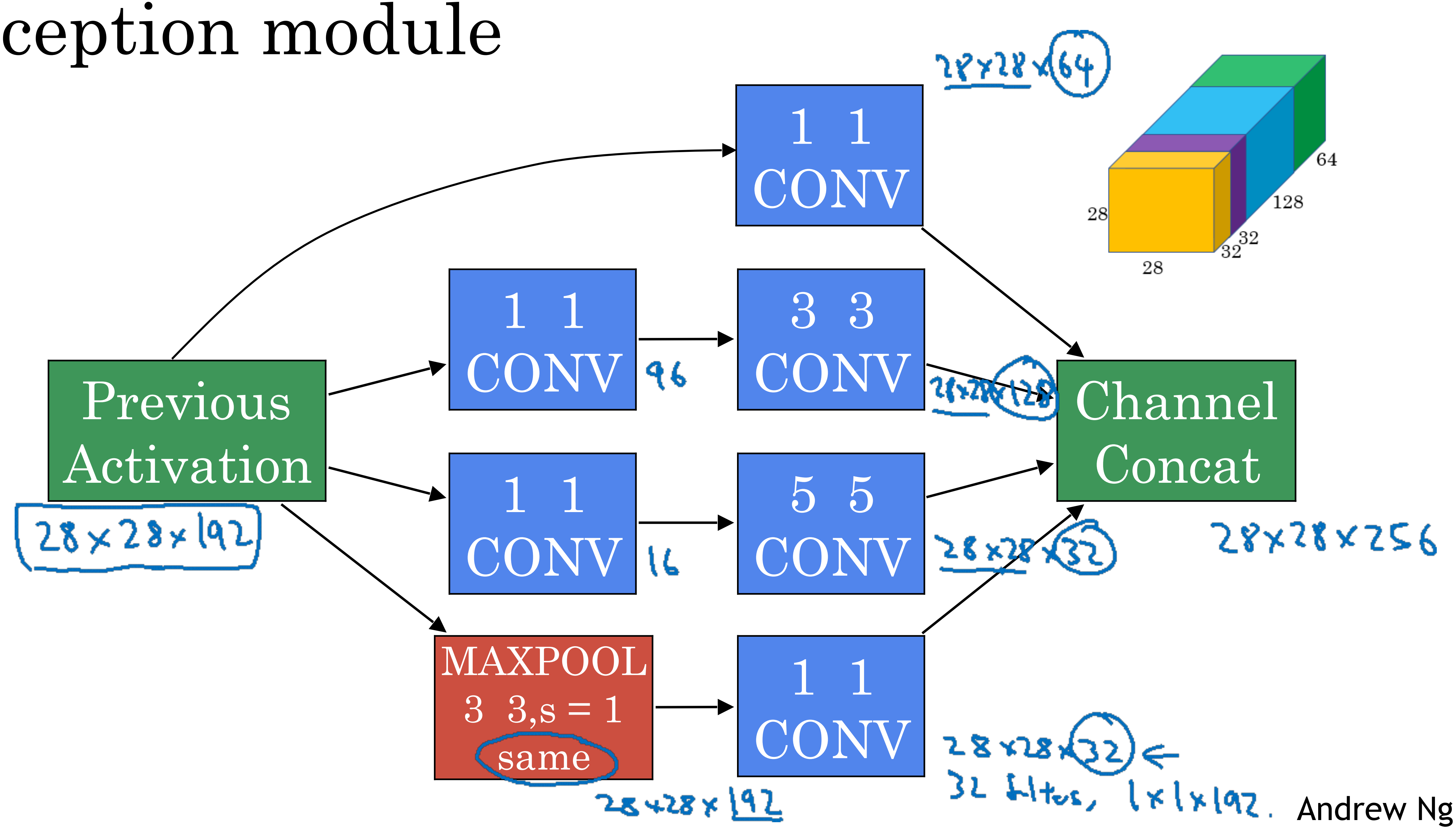
120M → 12.4M

deeplearning.ai

Case Studies

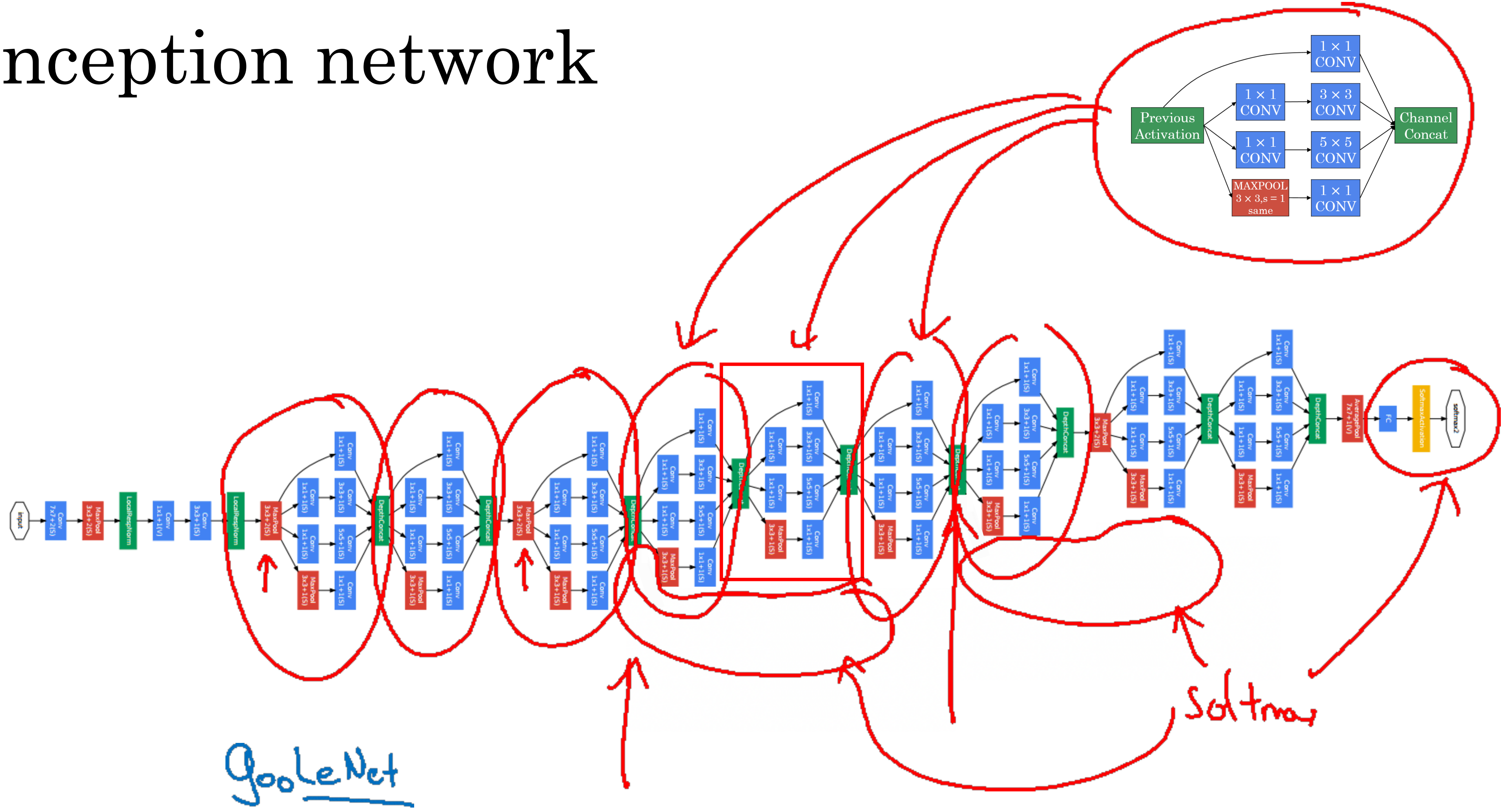
Inception network

Inception module



Andrew Ng

Inception network



[Szegedy et al., 2014, Going Deeper with Convolutions]

Andrew Ng

http://knowyourmeme.com/memes/we-need-to-go-deeper