We will show a number of case studies of effective CNN

Thats see examples of how convolution layer pooling layer and fully conected layer are put together

we study these case studies to understand how the building blocks learned previously are used to form effective CNN

deeplearning.ai

Case Studies

You can use someone else CNN architecture for your problem which can be different, but that architecture can be useful

Why look at case studies?

Outline

Classic networks:

- LeNet-5
- AlexNet
- VGG

these are effective CNN

these layed the fondation for the current NN

They will be usefull for our own work

ResNet (152) very deep NN 152 layer, has interesting ideas

Inception we will see also this NN

After seeing these architectures of these NN we will get a better intuition

Andrew Ng

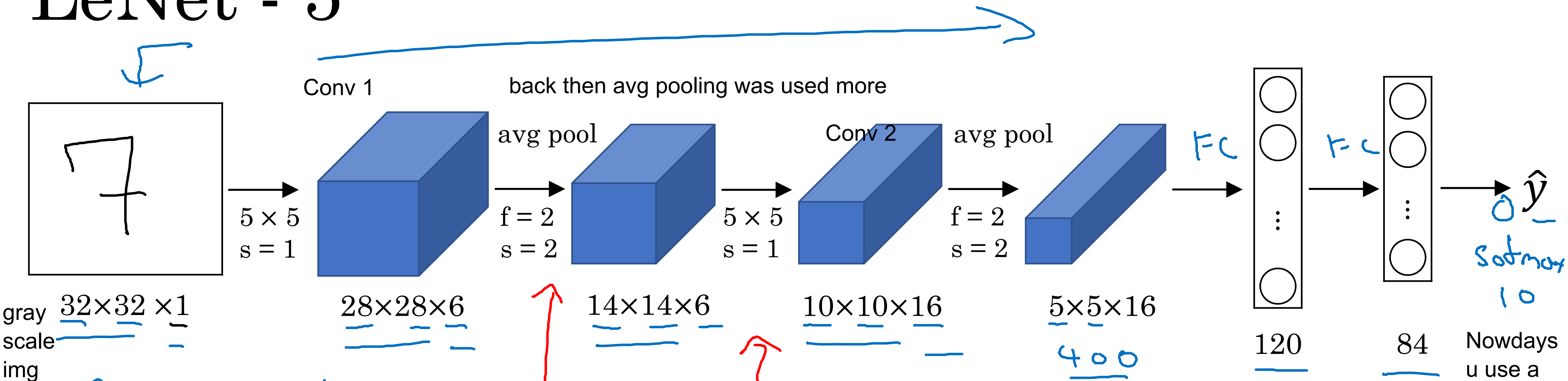
deeplearning.ai

Case Studies

Classic networks

The goal of LeNet .5 was to recognise handwritten digits

LeNet - 5



Nowdays u use a softmax classifier to classify output from 0 to 9
Back then they used different classifier which is usles today

60K parameters.

non-linearity after pooling

$n_H \times n_W \times n_C$ $f \times f \times n_C$

$n_H, n_W \downarrow$ $n_C \uparrow$

hight and width tend to go down, nr of channels increase

conv pool conv pool fc fc output

Advanced: sigmoid/tanh ReLU

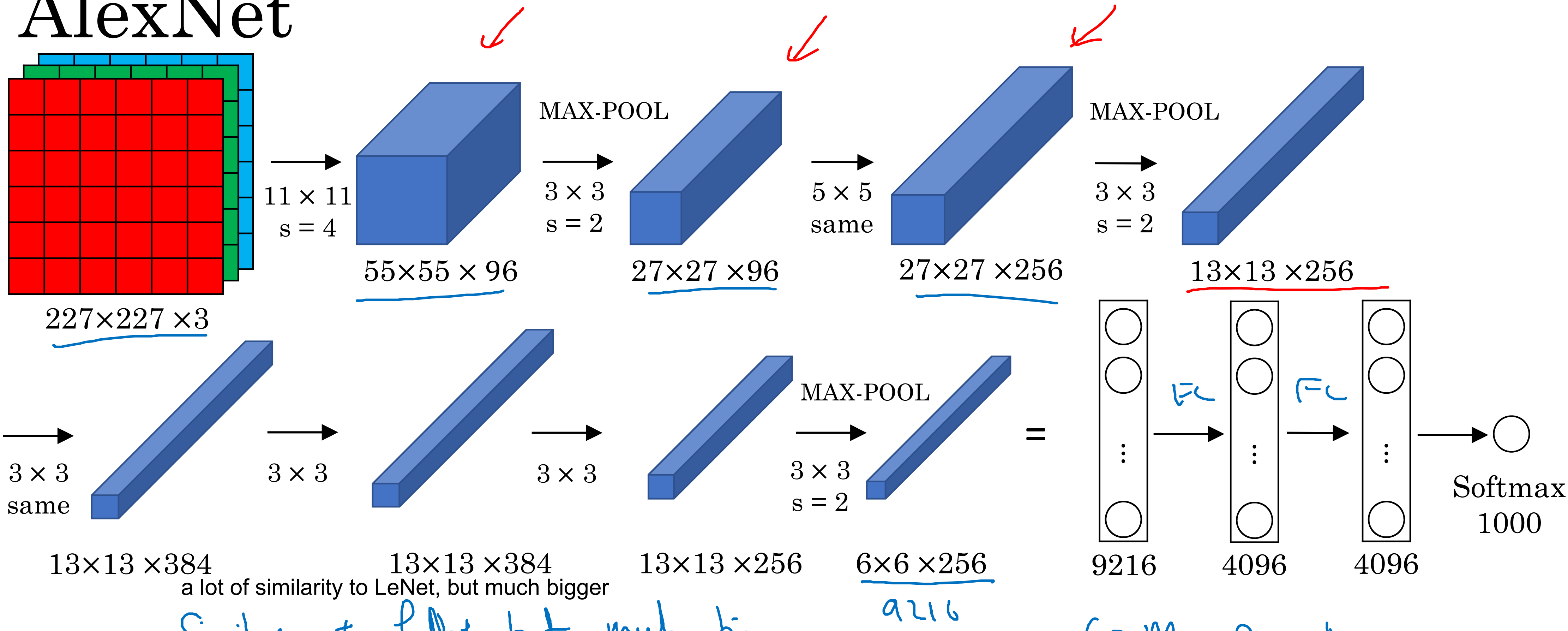
Part in red advanced and can skip. Part for read paper

$\boxed{\text{II}}$ III.

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

Andrew Ng

AlexNet



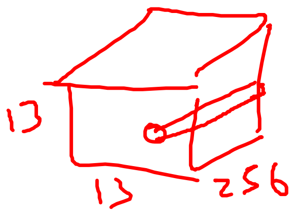
a lot of similarity to LeNet, but much bigger

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

~60M parameters

much bigger as it has around 60M parameters

it had a strange way to train into two GPU



normalize the data in that cilinder, actually researchers have found this does not help that much

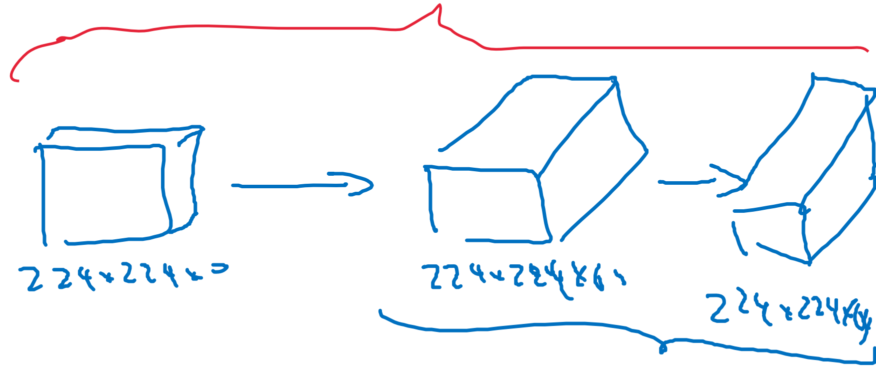
[Krizhevsky et al., 2012. ImageNet classification with deep convolutional neural networks]

Andrew Ng

Had also this layer called the local response Normalization which is not used and we did not talk about it.

Instead of having so many hyperparameter they say thats have a much simpler network
VGG it simplefied these NN

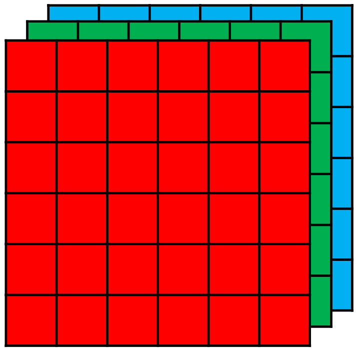
this is what that CONV 64 X2 MEANS
ITS JUST TWO LAYER OF CONV



VGG - 16

VGG-19

CONV = 3×3 filter, s = 1, same   MAX-POOL = 2×2 , s = 2



224×224 ×3

$\xrightarrow[\times 2]{\text{[CONV 64]}}$ 224×224×64 $\xrightarrow{\text{POOL}}$ 112×112 ×64 $\xrightarrow[\times 2]{\text{[CONV 128]}}$ 112×112 ×128 $\xrightarrow{\text{POOL}}$ 56×56 ×128

start with img and the first two layers are convolutions, use 64 filters

$\xrightarrow[\times 3]{\text{[CONV 256]}}$ 56×56 ×256 $\xrightarrow{\text{POOL}}$ 28×28 ×256 $\xrightarrow[\times 3]{\text{[CONV 512]}}$ 28×28 ×512 $\xrightarrow{\text{POOL}}$ 14×14×512

$\xrightarrow[\times 3]{\text{[CONV 512]}}$ 14×14 ×512 $\xrightarrow{\text{POOL}}$ 7×7×512 $\longrightarrow$ FC 4096 $\longrightarrow$ FC 4096 $\longrightarrow$ Softmax 1000 (OUTPUTS ONE OF 1000 CLASSES)

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

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

Andrew Ng

This is a very large network with 138 M parameters, but its architecture is pretty simpler, quite uniform

Very very deep NN are hard to design because of vanishing and exploding gradient type of problems. In this lecture we learn about skip connection which allows to take activation from one layer and suddenly feed it to another layer even much deeper in a NN and using that we will be able to build ResNet which anables us to build very deep NN, sometimes even networks of over 100 layers.

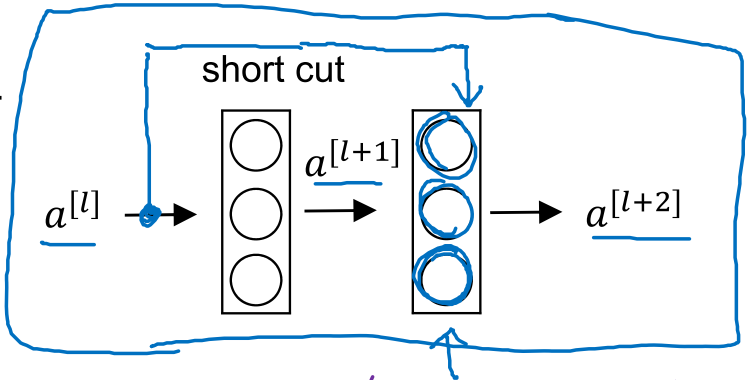
deeplearning.ai

Case Studies

Residual Networks (ResNets)

ResNets are build out of somethink called the residual block

Residual block



Here are two layer of a NN, u start with some activation in layer l and then go to l+1 and then the activation two layers later is a[l +2]

Thats go through the steps of this process "short cut" / skip connection (sometimes is also called skip connection)

apply this linear operator

$a^{[l]} \to$ Linear $\to$ ReLU $\xrightarrow{a^{[l+1]}}$ Linear $\to \oplus \to$ ReLU $\to a^{[l+2]}$

after that u apply the relu nonlinearity "main path"

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

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

So for information to go from a[l] to a[l+2] it has to go through all of these steps, call it the main path. In ResNet we gone make a change to this. We gone take a[l] and take it ahear as shown with purple line, just add a[l]. We gone call this the short cut.

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

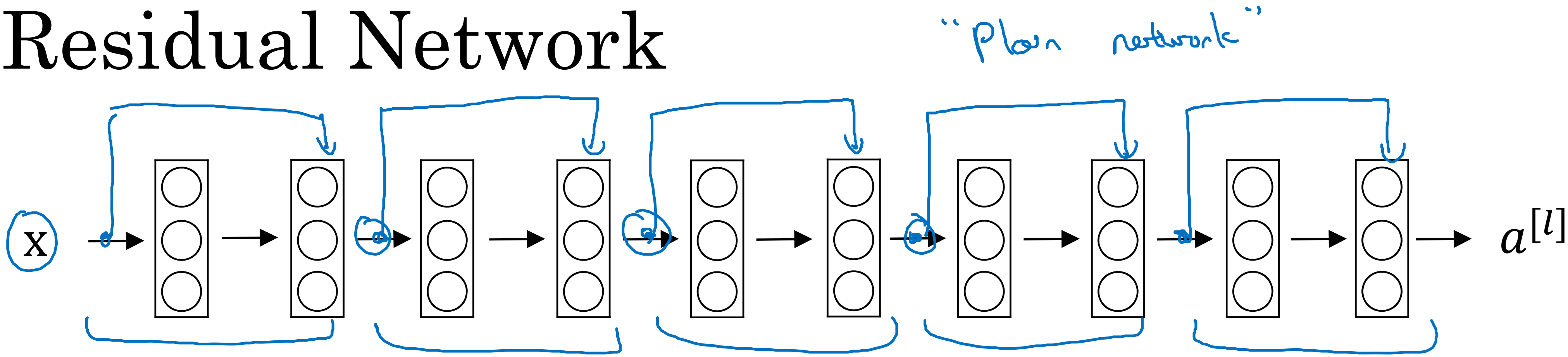
Andrew Ng

what happens now is that we do not do the last step, we make change we add that a[l], so this addition makes it a residual block

WHAT THE INVENTORS OF THIS FOUT IS THAT USING RESIDUAL BLOCKS ALLOWS TO TRAIN MUCH DEEP NN

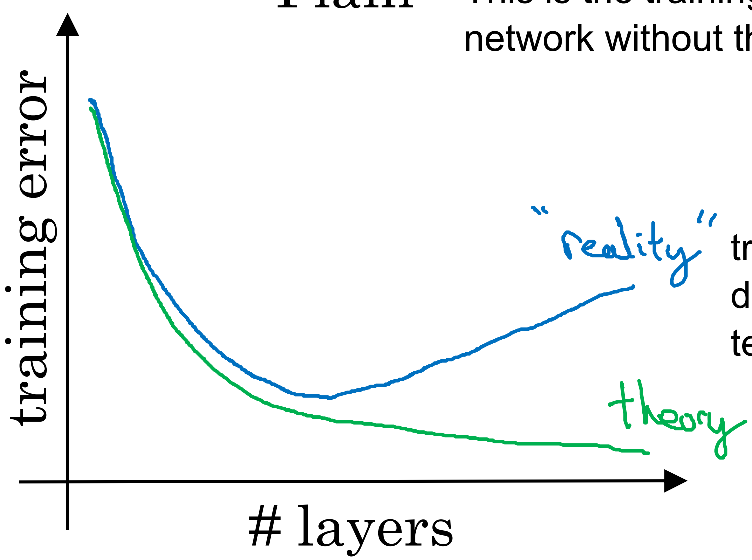
So we have a plain net and to turn this to resNet what u do is to add all those skip connections

Residual Network



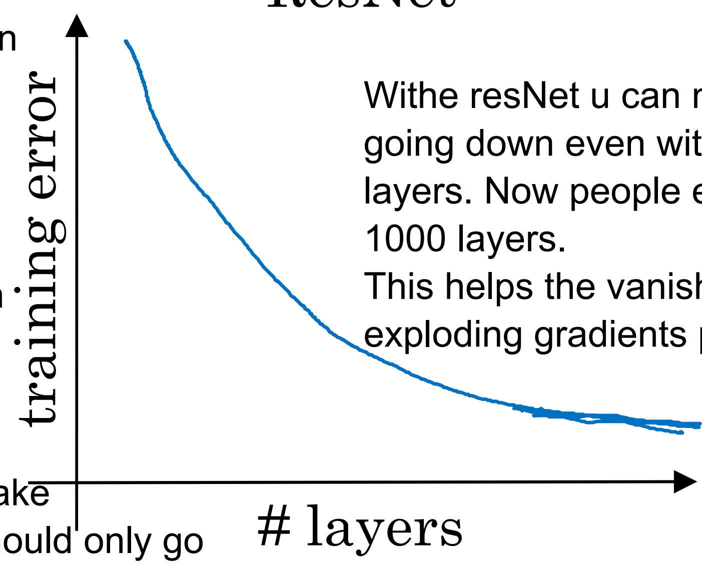
Plain

This is the training of a plain network without the skip connection



training error will decrease and then tend go up

ResNet



Withe resNet u can make it keep going down even with more then 100 layers. Now people experiment with 1000 layers.
This helps the vanishing and exploding gradients problem.

in theory as u make a nn deeper it should only go down the error
In theory havin a deeper nn should help, in reality that does not what happens.

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

We will see an intuition why resNet works so well. thats see an example.

deeplearning.ai

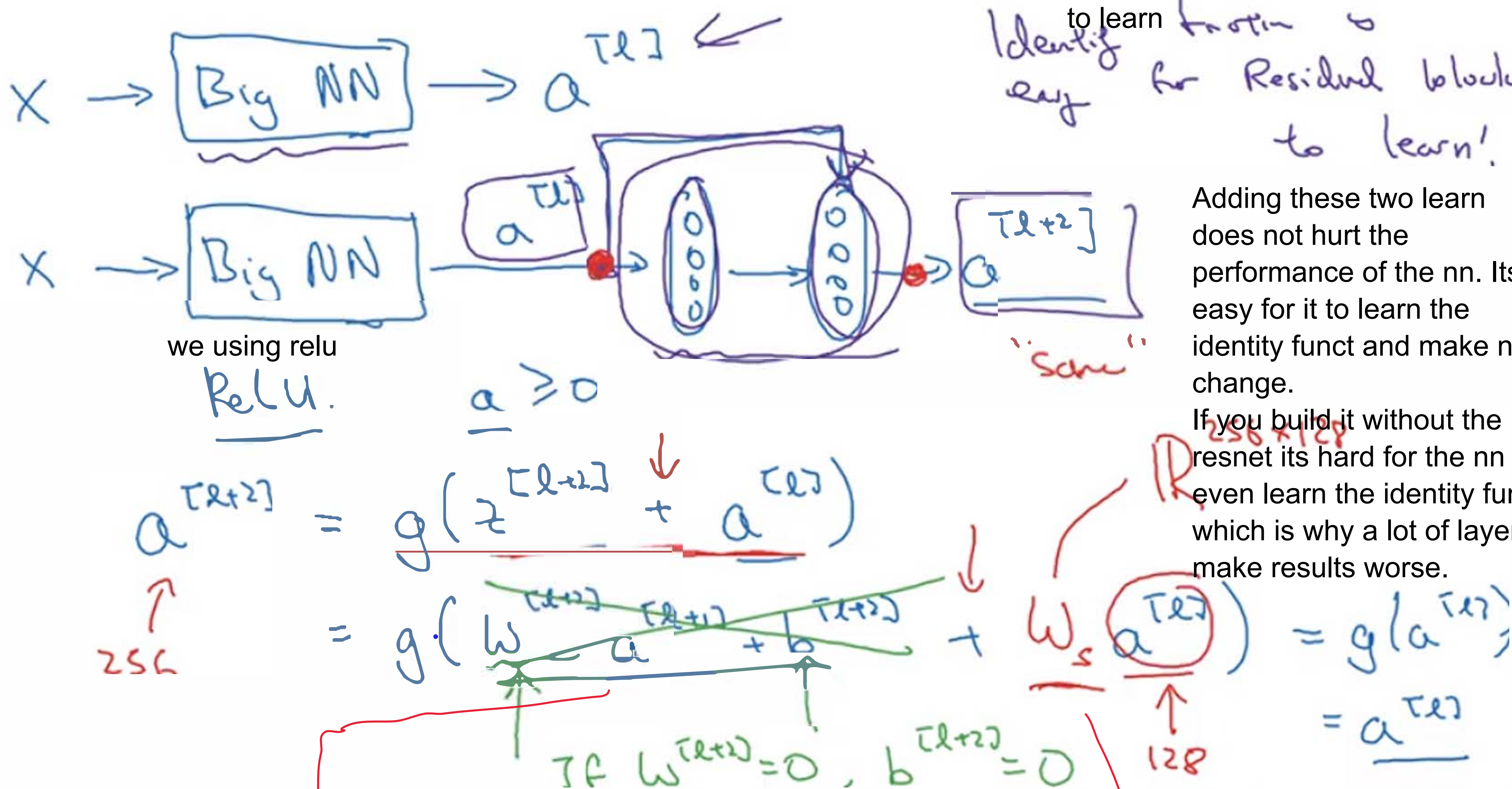
Case Studies

Why ResNets work

So making a nn can hurt its performance, but this is not true when u train a resnet

Why do residual networks work?

identiy funct is easy for residual block to learn



Adding these two learn does not hurt the performance of the nn. Its easy for it to learn the identity funct and make not change.

If you build it without the resnet its hard for the nn to even learn the identity funct which is why a lot of layer make results worse.

we using relu

so these terms if they are not 0, then they are like an

if the input and output have different dimentions

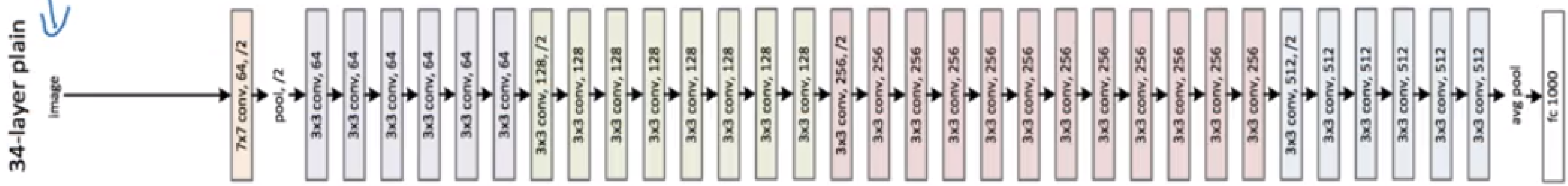
ResNet

you input an img

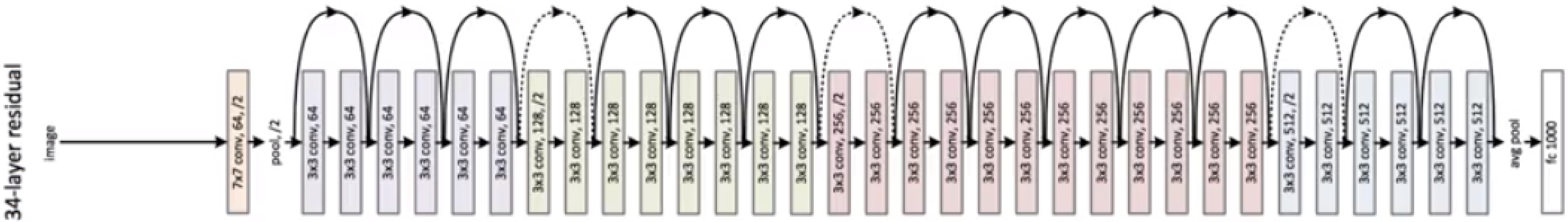
this is an example of a plain nn

softmax output

Plain



ResNet



3x3 same

In terms of designing conv net architectures one of the ideas that realy helps is using a one by one convolution. what does it do ????

deeplearning.ai

Case Studies

Network in Network and 1×1 convolutions

Why does a 1 × 1 convolution do?

1	2	3	6	5	8
3	5	5	1	3	4
2	1	3	4	9	3
4	7	8	5	7	9
1	5	3	7	4	8
5	4	9	8	3	5

6×6 $\times 1$

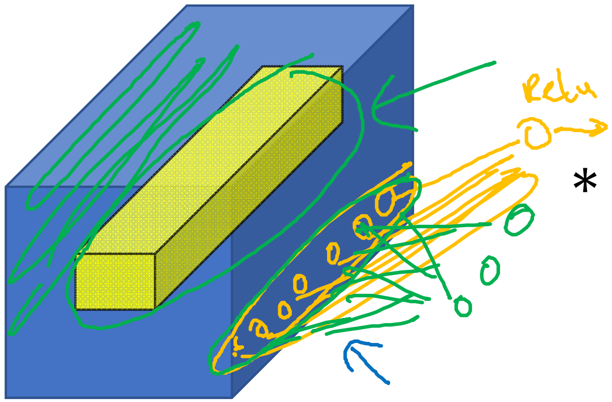
if u take that 6x6 img and convolve it with this 1x1 filter u end up taking img and multiply it by two

here is a 1x1 filter

$* \quad \boxed{2} \quad =$

2	4	6	...		

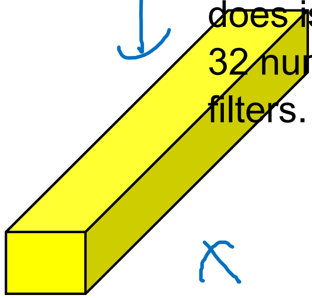
it does not seem particularly useful to just mult by 2



Relu

$6 \times 6 \times 32$

$*$



$1 \times 1 \times 32$

$32 \rightarrow$ # filters.

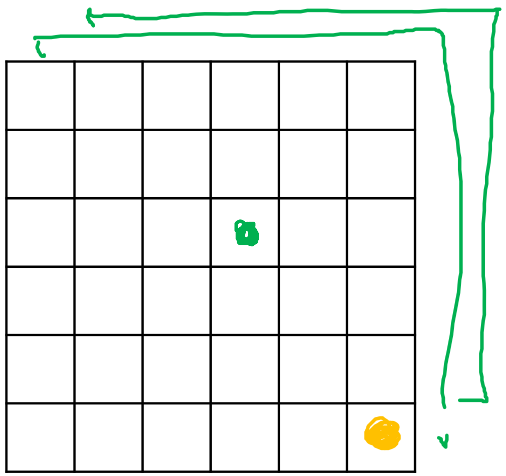
$n_c^{[l+1]}$

what thsi fc nn does is input 32 numbers and outputs nr of filters.

$=$

ReLU

Network in Network



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

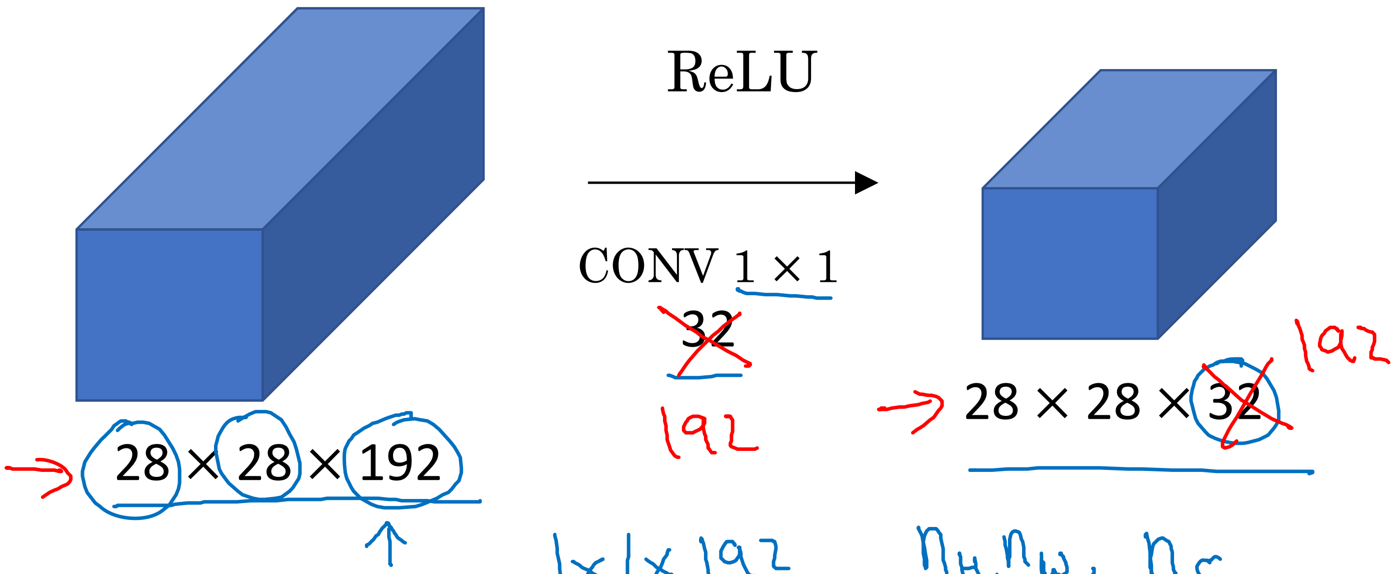
Andrew Ng

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

if u have a 6x6x32 instead of 6x6x1, the a convolution with a one by one x 32 will do a lot what it fsdfasdfasd NOT SO NICE EXPLANATION

sometime this idea is called network in network

Using 1×1 convolutions



$1 \times 1 \times 192$ n_H, n_W, n_c

32 filters.

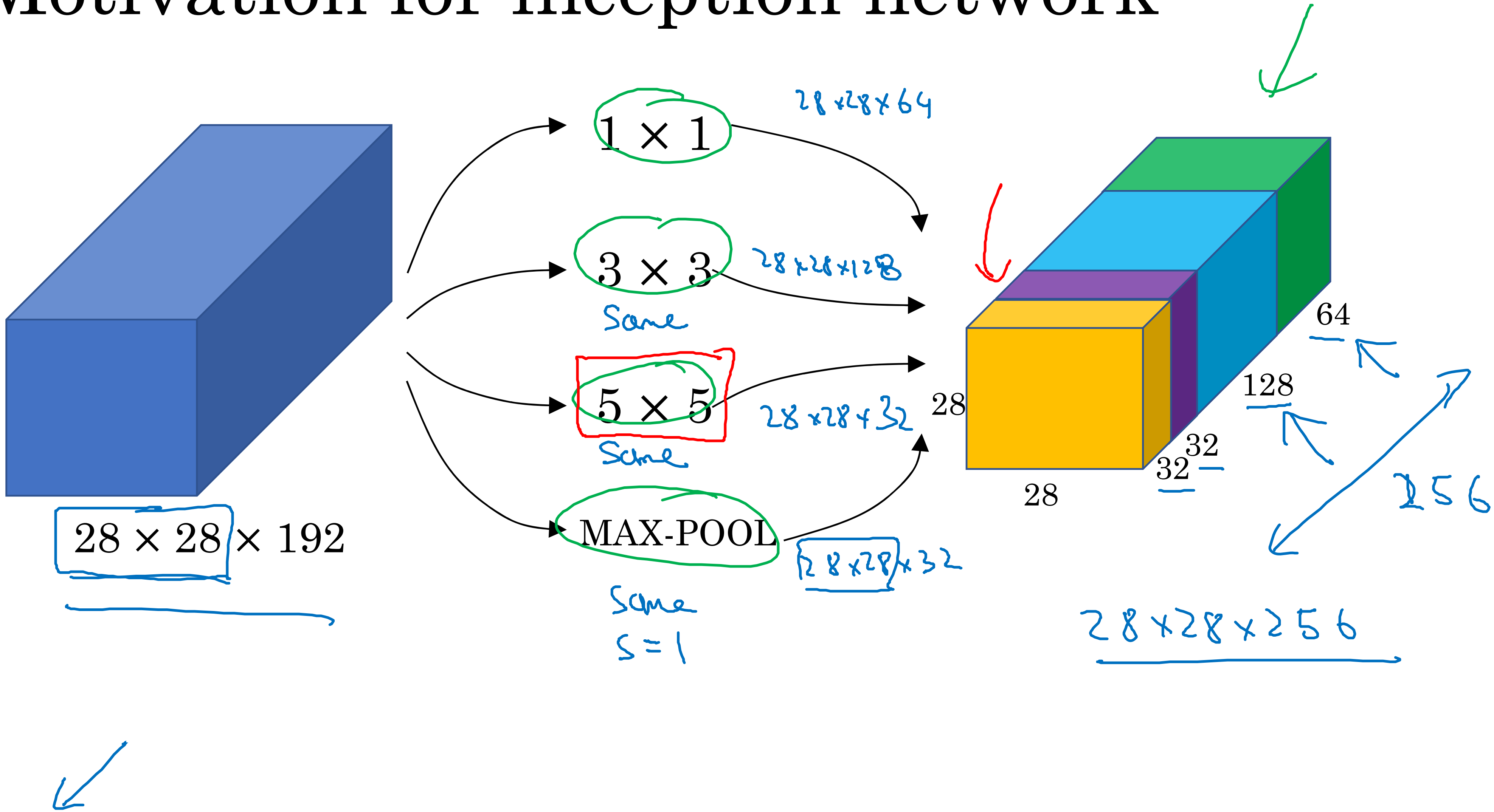
[Lin et al., 2013. Network in network]

deeplearning.ai

Case Studies

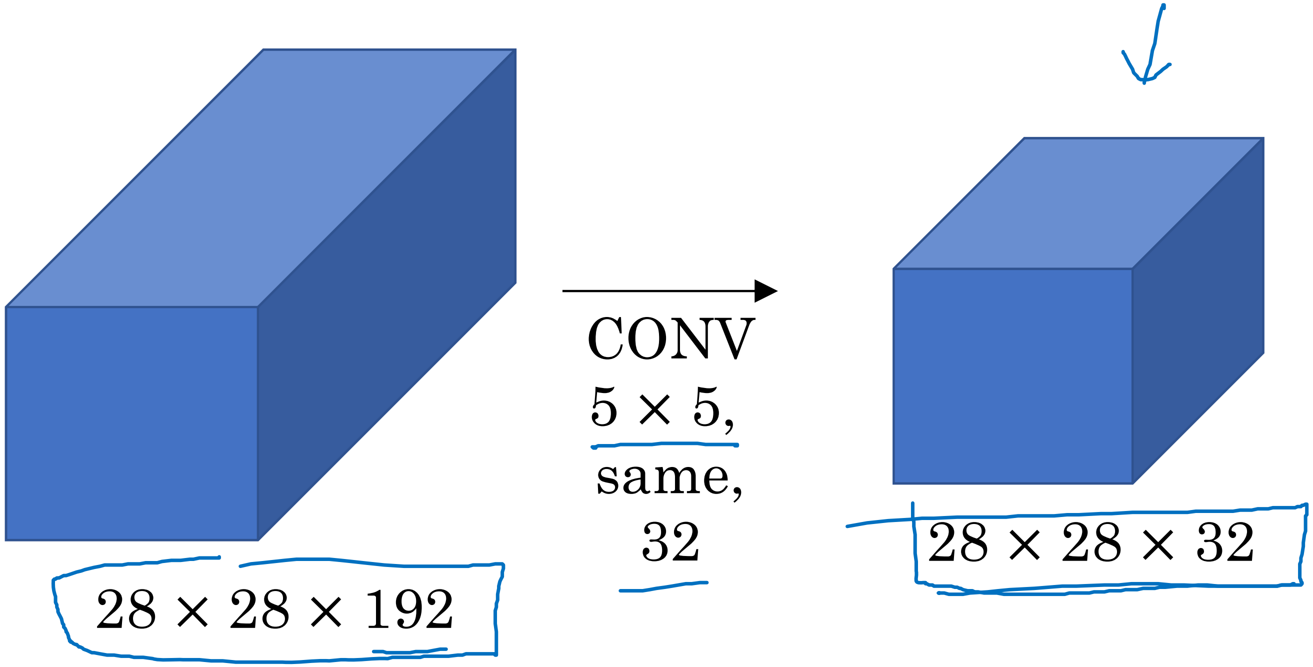
Inception network motivation

Motivation for inception network



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

The problem of computational cost

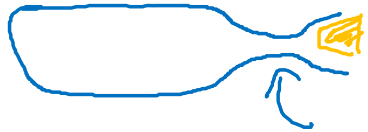


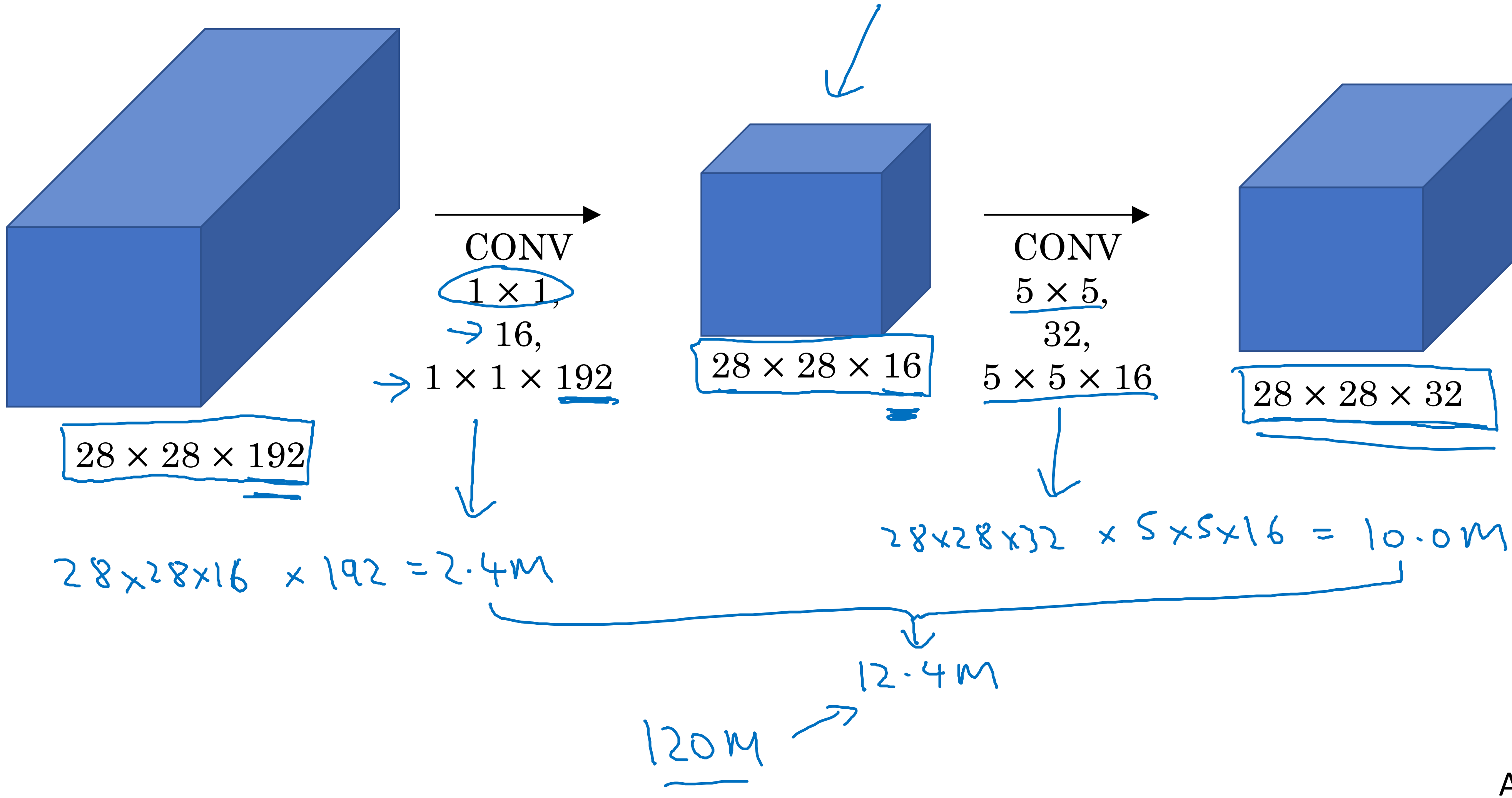
32 filters. filters are 5×5×192.

$28\times28\times32 \times 5\times5\times192 = 120M$.

Using 1×1 convolution

"bottleneck layer"





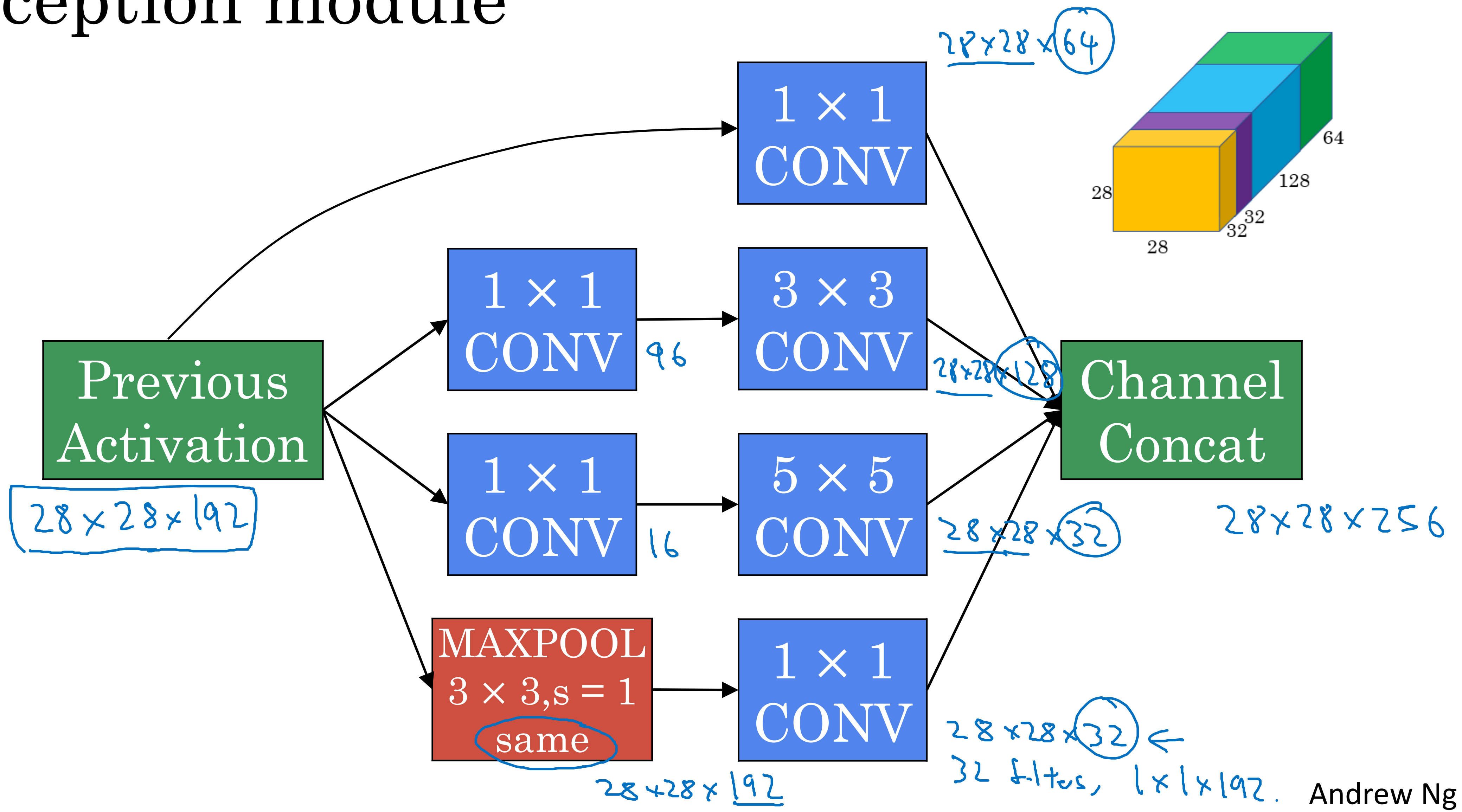
Andrew Ng

deeplearning.ai

Case Studies

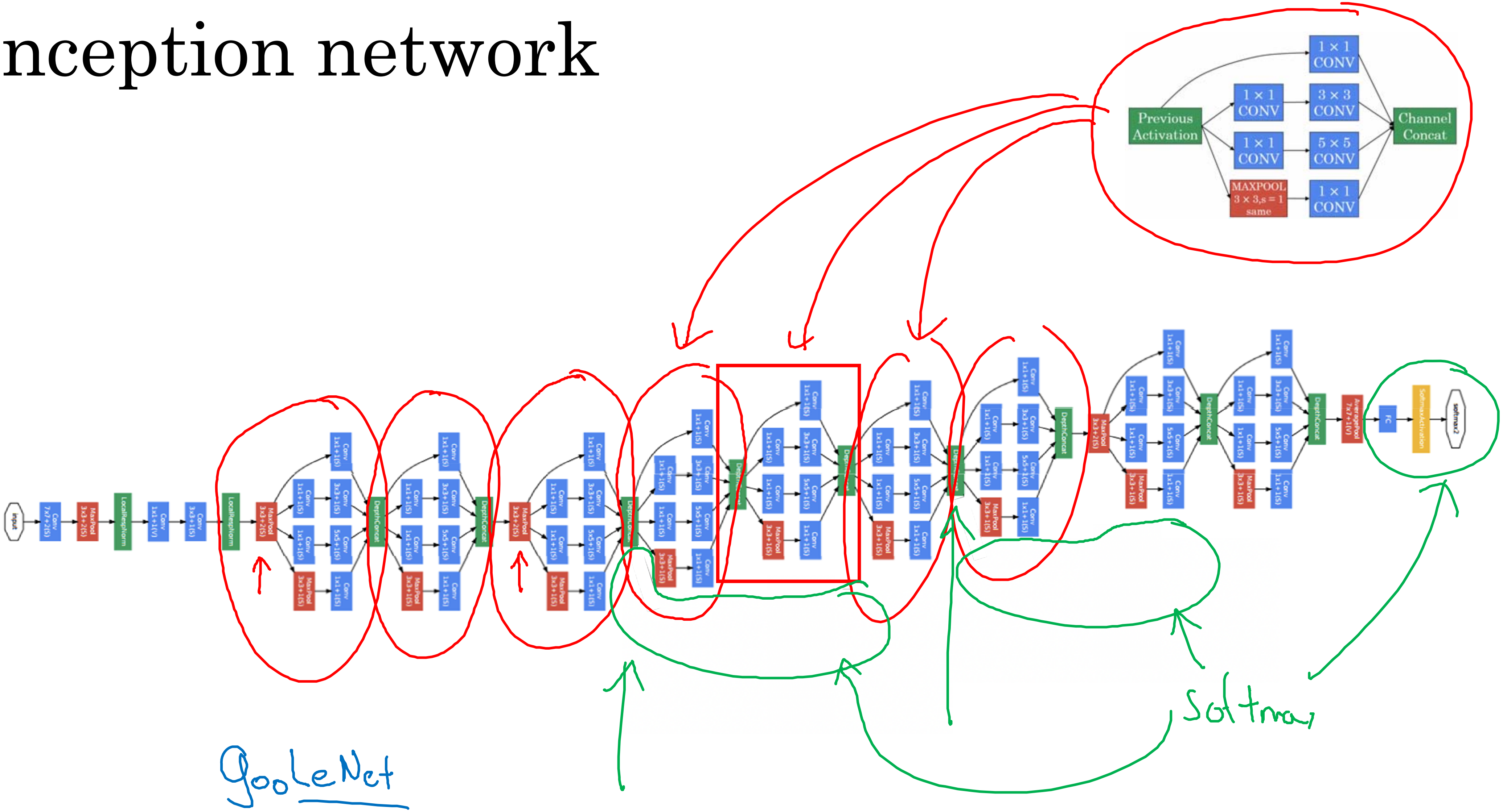
Inception network

Inception module



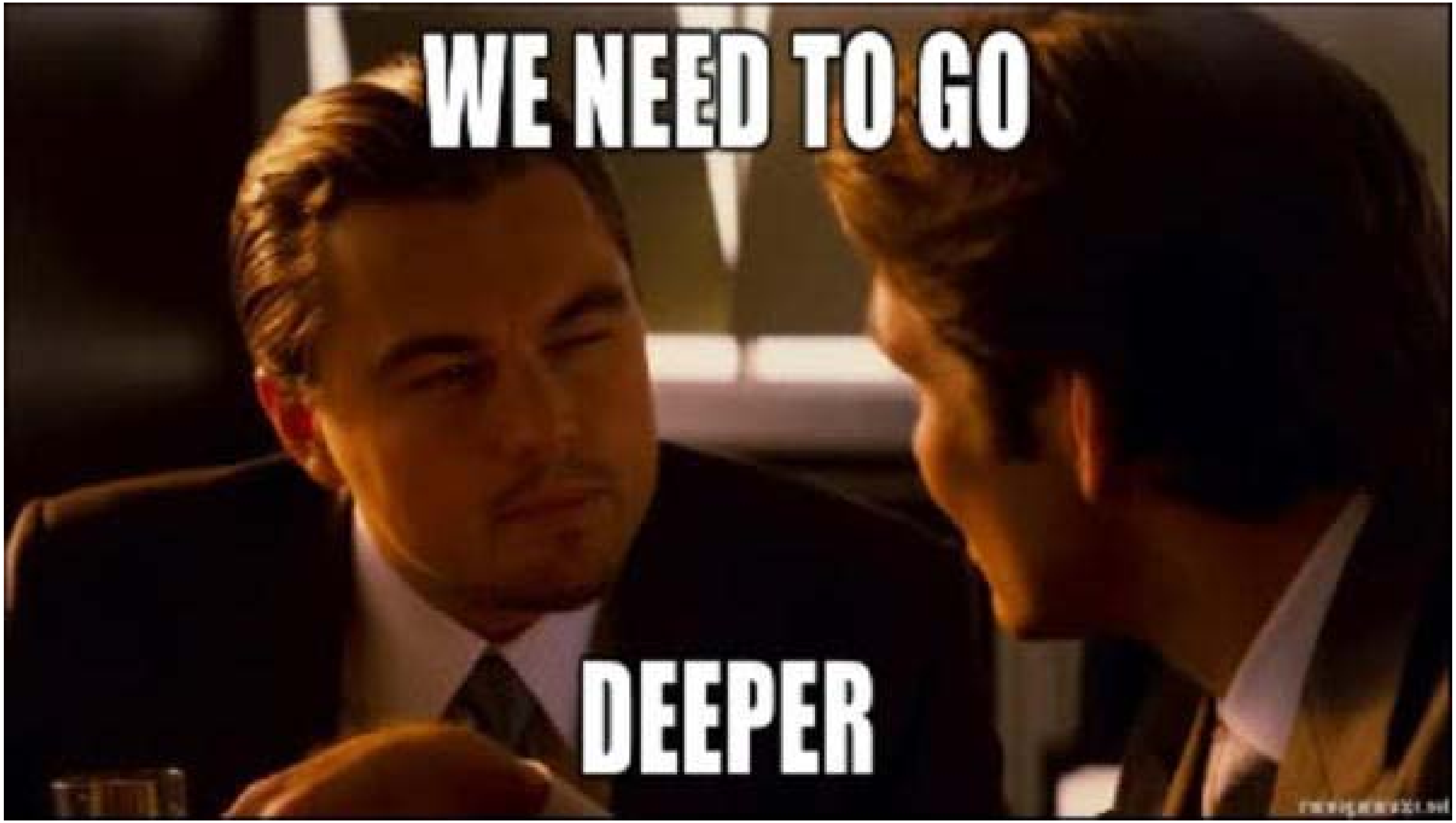
Andrew Ng

Inception network



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

Andrew Ng



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

Andrew Ng