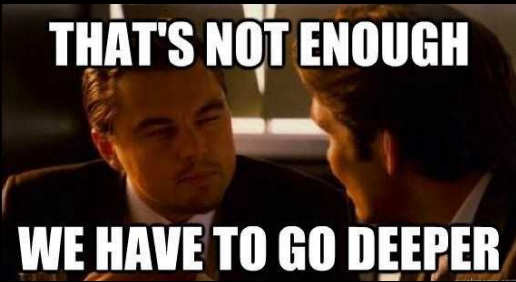
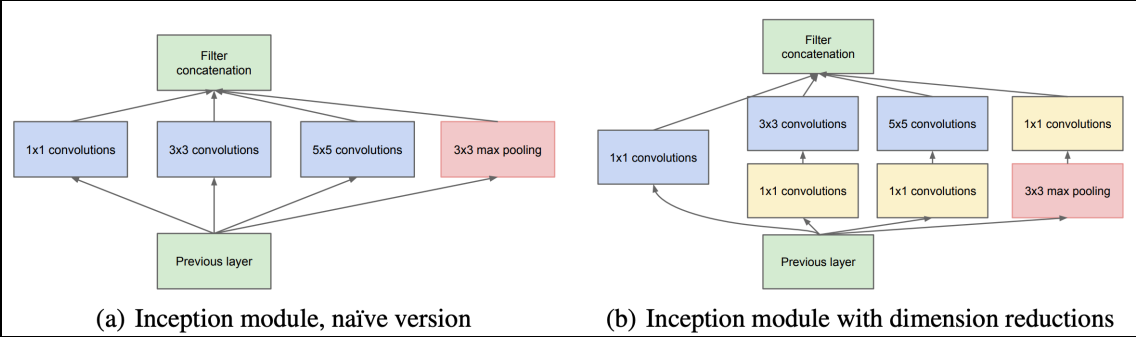
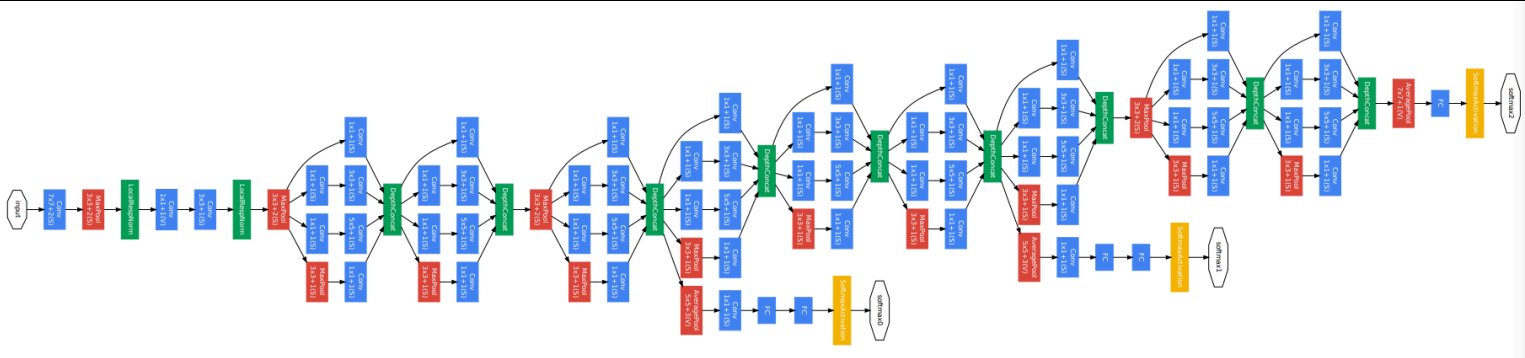
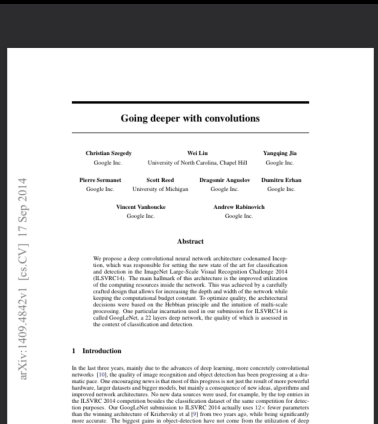


GoogleNet, also known as Inception v1, was introduced by Szegedy et al. in the 2014 ILSVRC (ImageNet Large Scale Visual Recognition Challenge). It was developed by researchers at Google and significantly outperformed previous architectures like AlexNet and VGGNet by reducing computational cost while improving accuracy.



Agenda

- 1. Network Architecture
- 2. Inception Module
- 3. 1x1 Convolutions
- 4. Global Average Pooling (GAP)
- 5. Auxiliary Classifiers
- 6. Variants of Inception



$3 \times 3 \times 64$
 $3 \times 3 \times 128$
 256

$3 \times 3 \times 64$
 $1 \times 1 \times 64$
 $3 \times 3 \times 32$

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								

Structure of Inception

- 1) 1×1 Conv
- 2) 3×3 Conv
- 3) 5×5 Conv
- 4) 3×3 MP

Auxiliary Classifier

2 Aux Class

→ Improves the gradient

→ Regularization

Only Training (No Inference)

Variants of Google Net

- 1) 2014 Inception V1
- 2) 2015 Inception V2
- 3) 2015 Inception V3

$1 \times 1, 3 \times 3, 5 \times 5$

Batch Norm 5×5
 $= 2(3 \times 3)$

Factorized
Convolutions

$7 \times 7 \rightarrow 3(3 \times 3)$

$7 \times 7 \rightarrow 1 \times 7$
 7×1

RMS Prop
(299×299)