

# Redes Neurais e Aprendizagem Profunda

## REDES NEURAIS CONVOLUCIONAIS GOOGLENET / RESNET / ANÁLISE

---

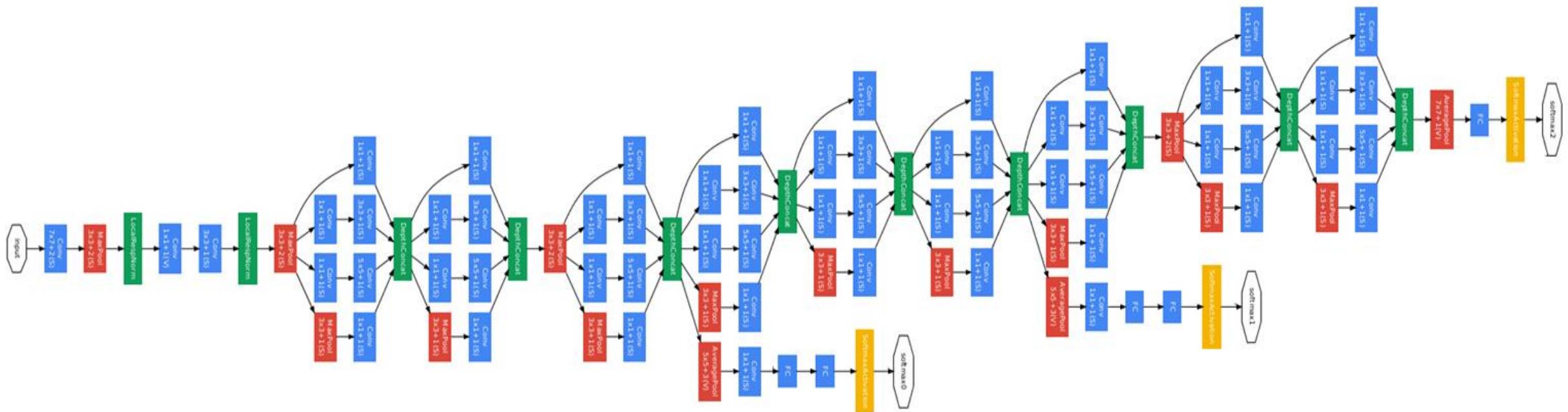
Zenilton K. G. Patrocínio Jr

[zenilton@pucminas.br](mailto:zenilton@pucminas.br)

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

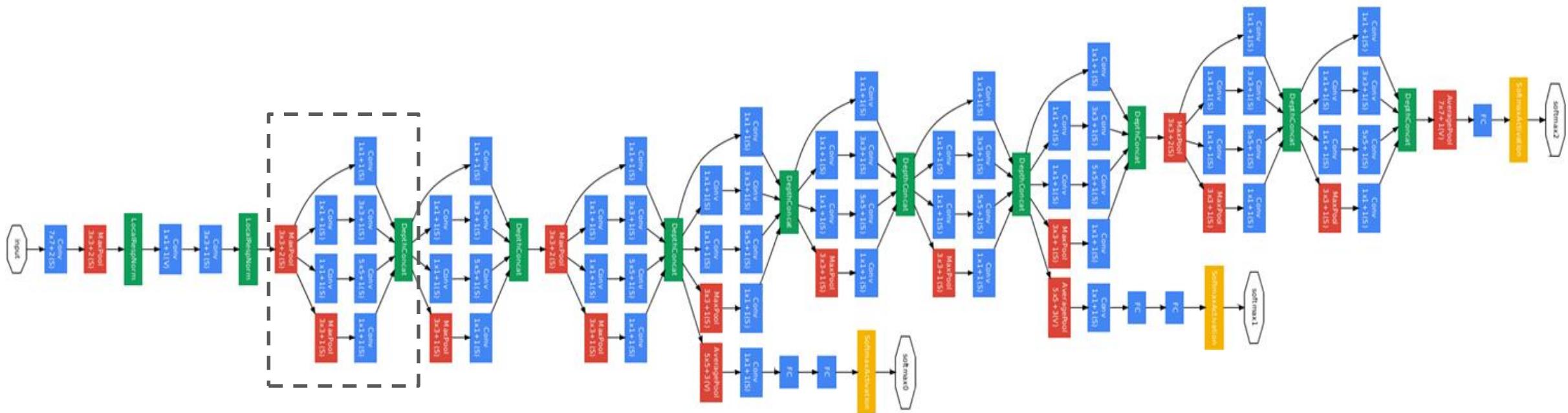
## Uso de vários filtros em paralelo



# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

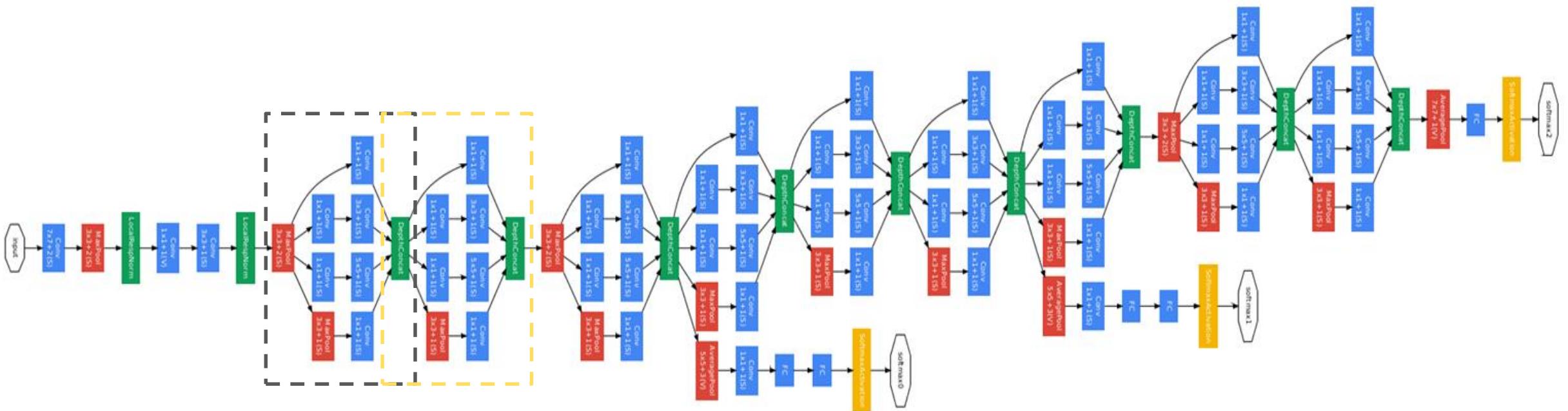


# Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

Uso de vários filtros em paralelo

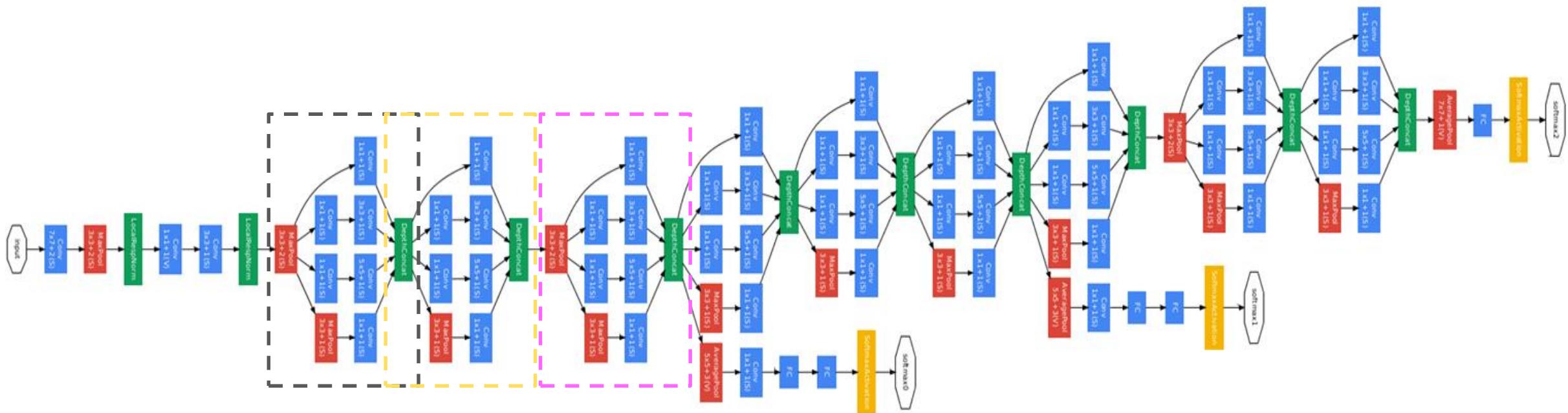


Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

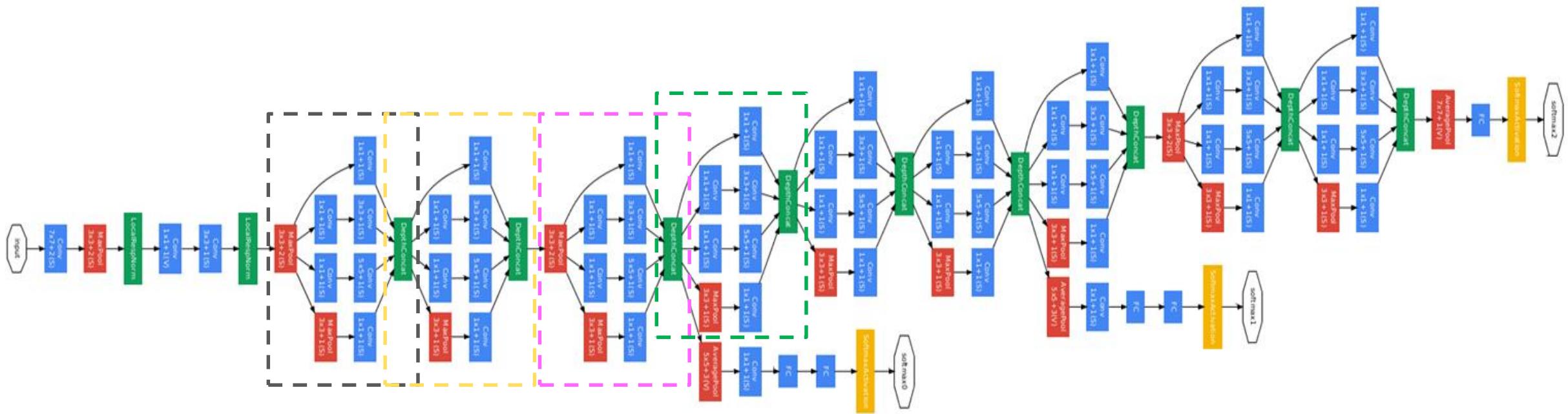


# Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

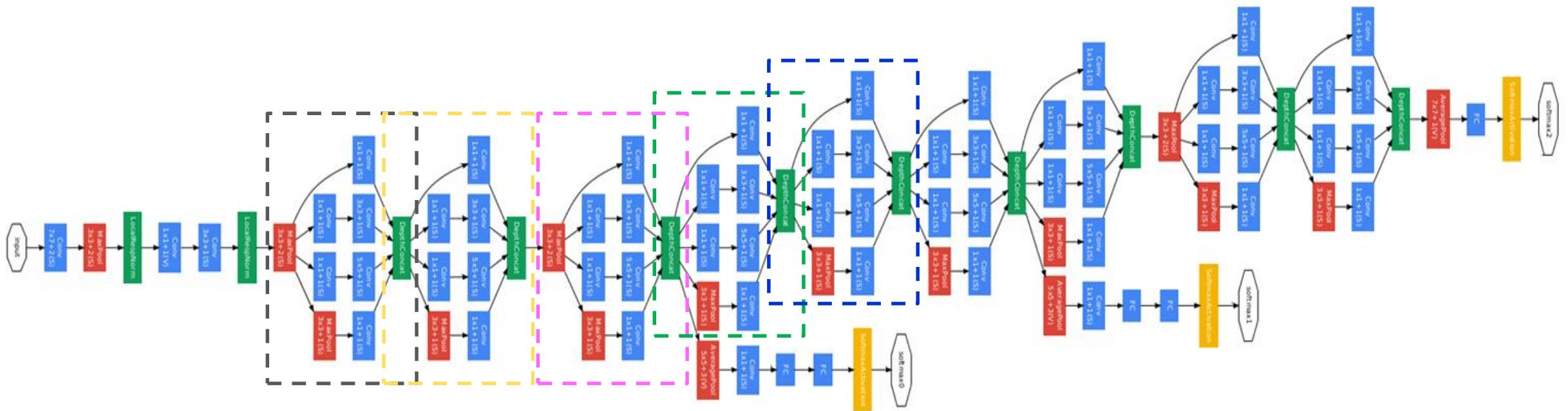


## Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

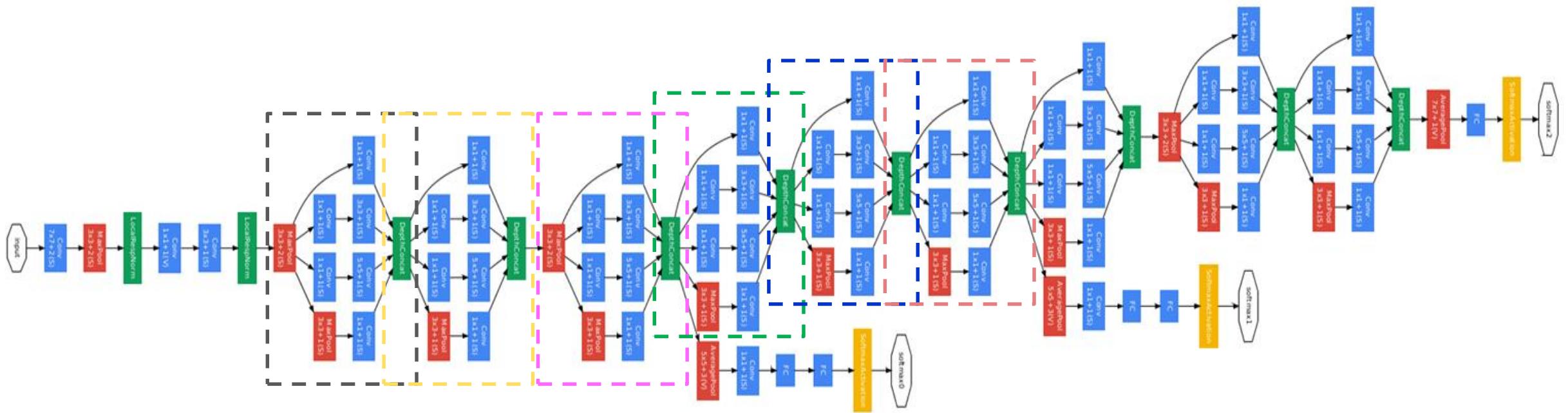


# Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

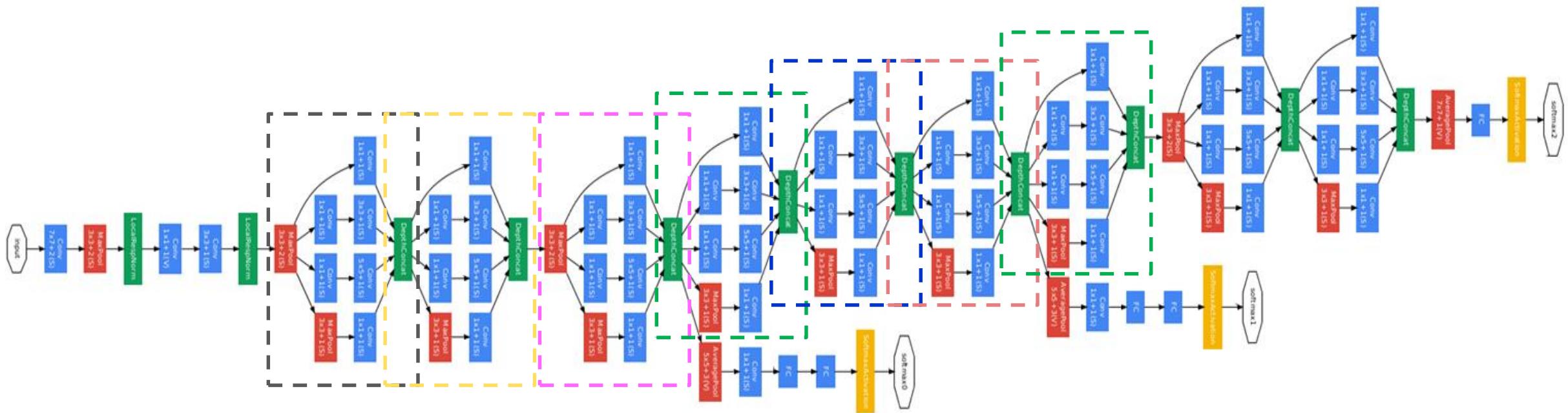


# Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

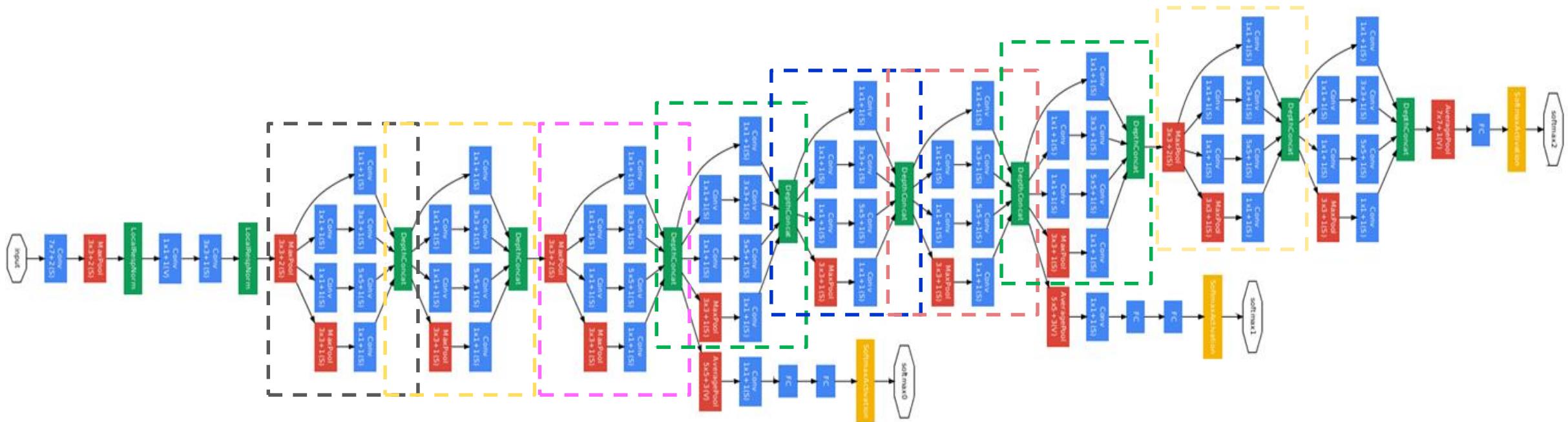


# Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

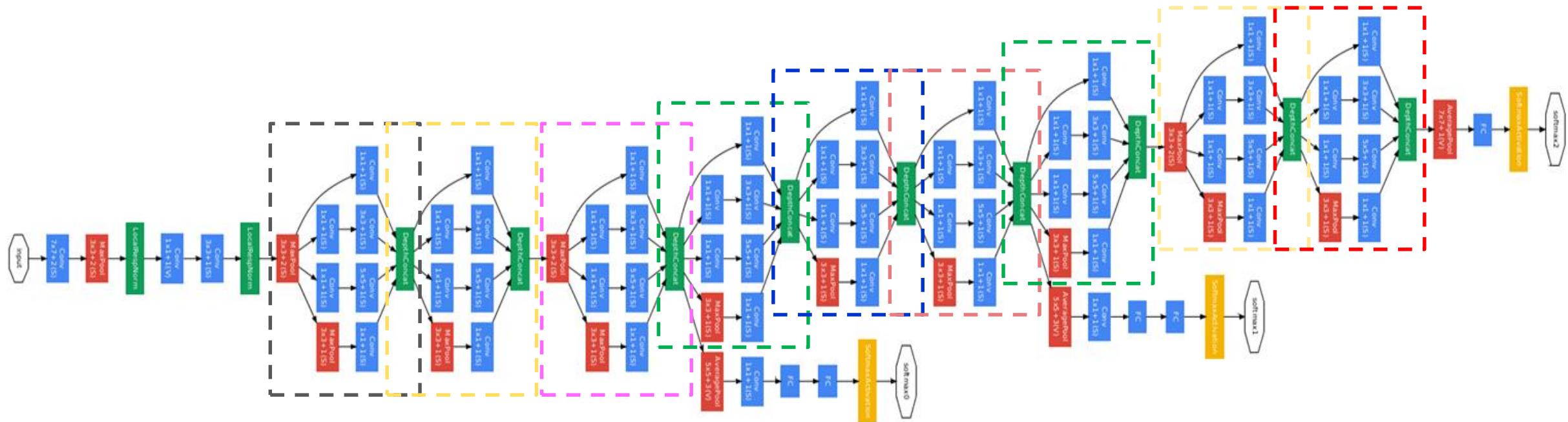


# Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Uso de vários filtros em paralelo

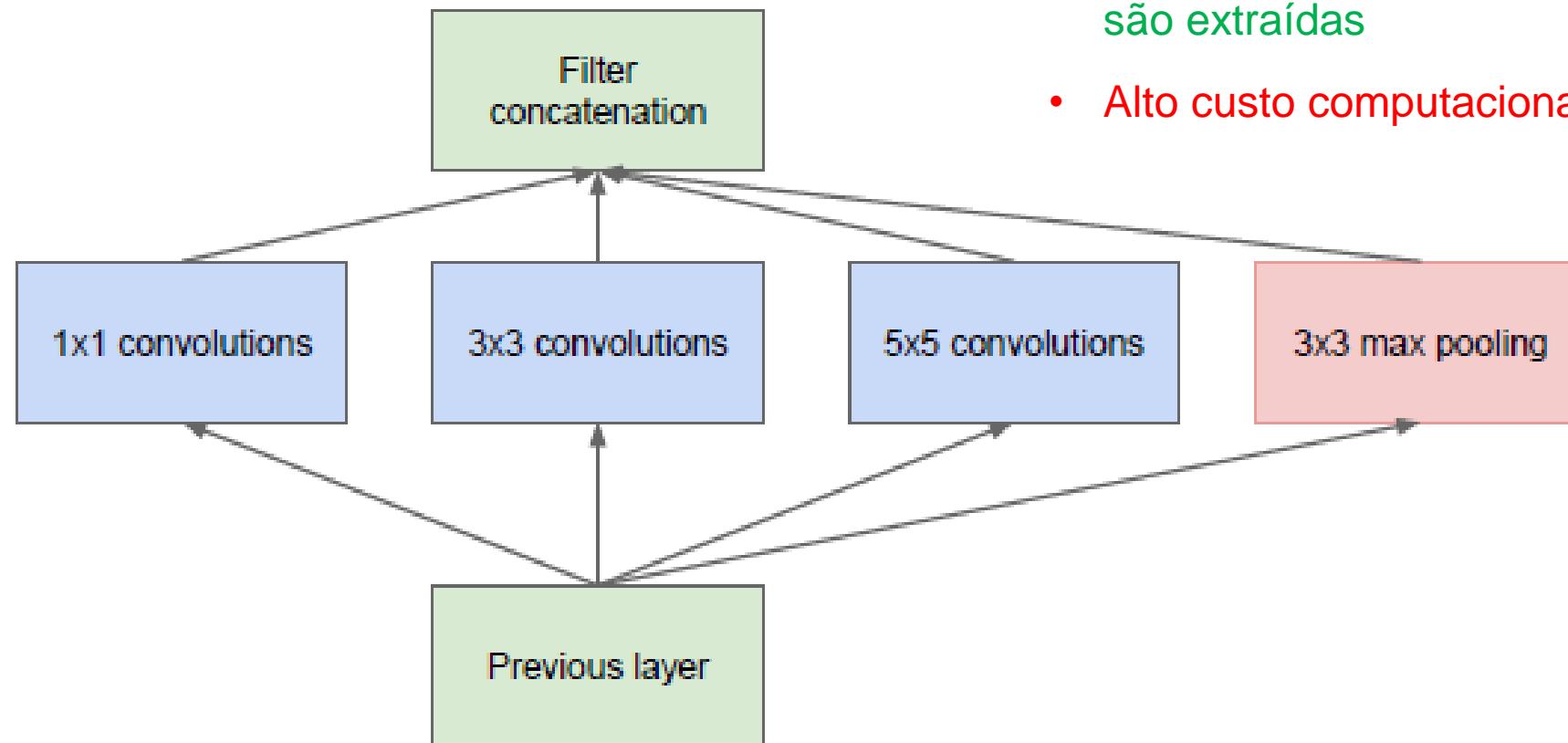


## Módulo *Inception*

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

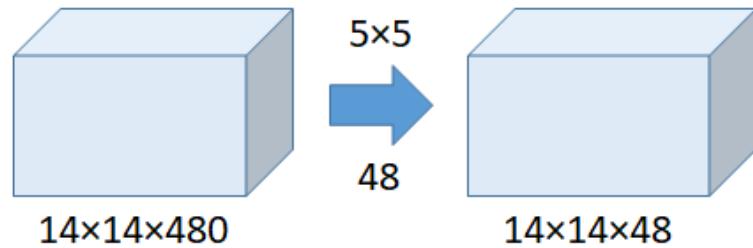
## Módulo *Inception* – Versão básica



# ConvNets – GoogLeNet

[Szegedy et al., 2014]

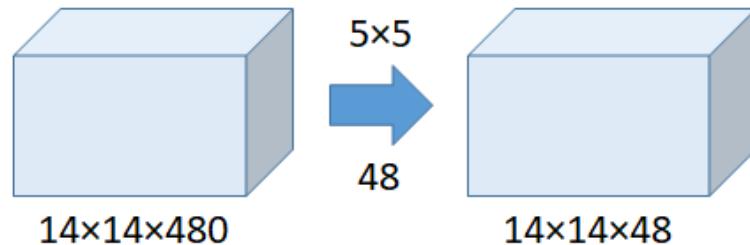
## Módulo *Inception* – Versão básica



# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Módulo *Inception* – Versão básica

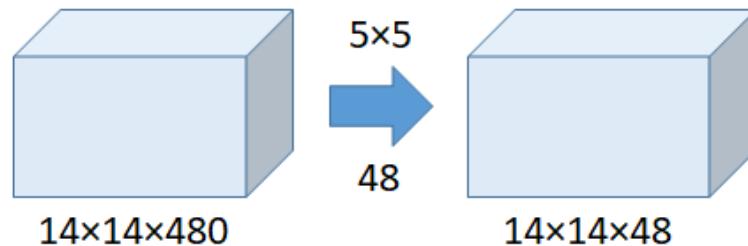


$$\begin{aligned}\text{Número de operações} &= (14 \times 14 \times 48) \times (5 \times 5 \times 480) = \\ &= 112,9\text{M}\end{aligned}$$

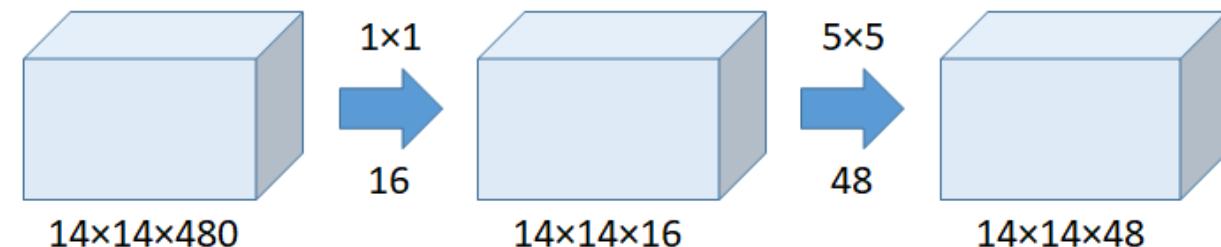
# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Módulo *Inception* – Versão básica



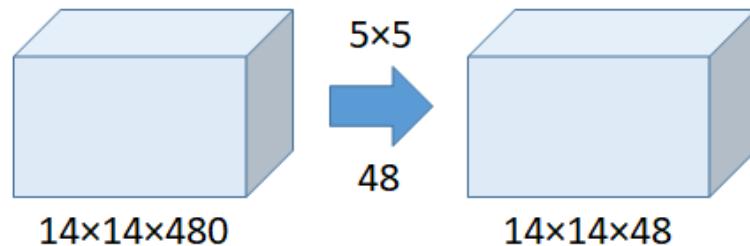
$$\begin{aligned}\text{Número de operações} &= (14 \times 14 \times 48) \times (5 \times 5 \times 480) = \\ &= 112,9\text{M}\end{aligned}$$



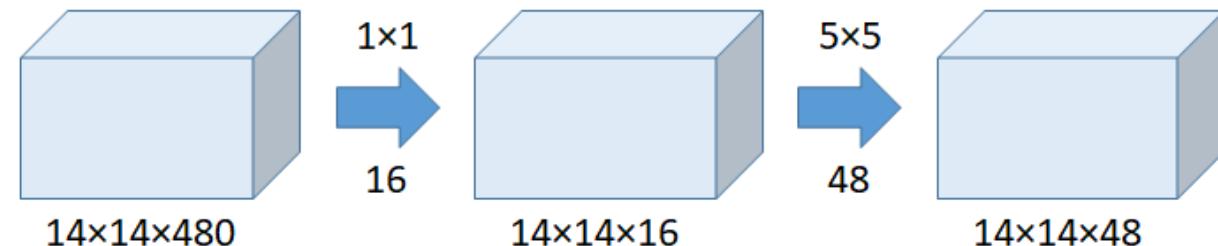
# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Módulo *Inception* – Versão básica



$$\begin{aligned}\text{Número de operações} &= (14 \times 14 \times 48) \times (5 \times 5 \times 480) = \\ &= 112,9\text{M}\end{aligned}$$

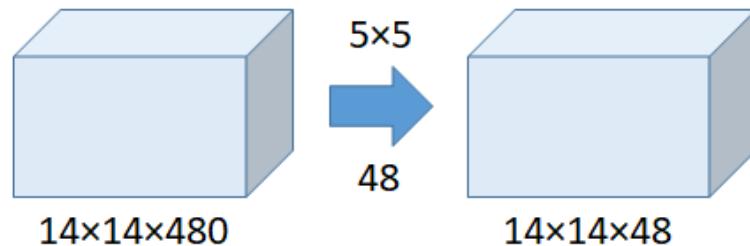


$$\text{Núm. operações para } 1 \times 1 = (14 \times 14 \times 16) \times (1 \times 1 \times 480) = 1,5\text{M}$$

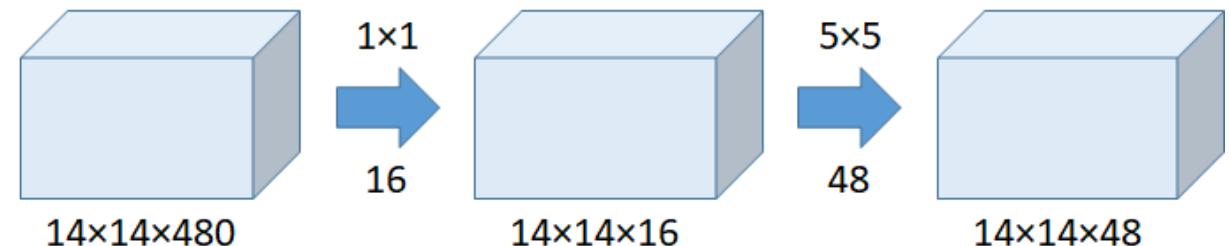
# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Módulo *Inception* – Versão básica



$$\begin{aligned}\text{Número de operações} &= (14 \times 14 \times 48) \times (5 \times 5 \times 480) = \\ &= 112,9\text{M}\end{aligned}$$

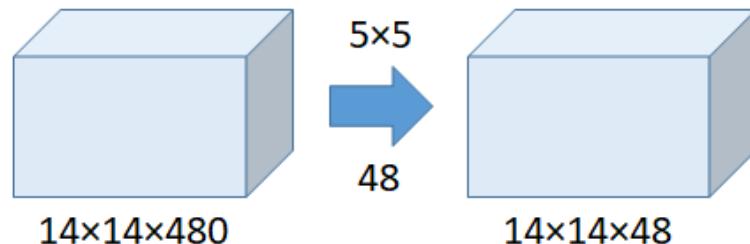


$$\begin{aligned}\text{Núm. operações para } 1 \times 1 &= (14 \times 14 \times 16) \times (1 \times 1 \times 480) = 1,5\text{M} \\ \text{Núm. operações para } 5 \times 5 &= (14 \times 14 \times 48) \times (5 \times 5 \times 16) = 3,8\text{M}\end{aligned}$$

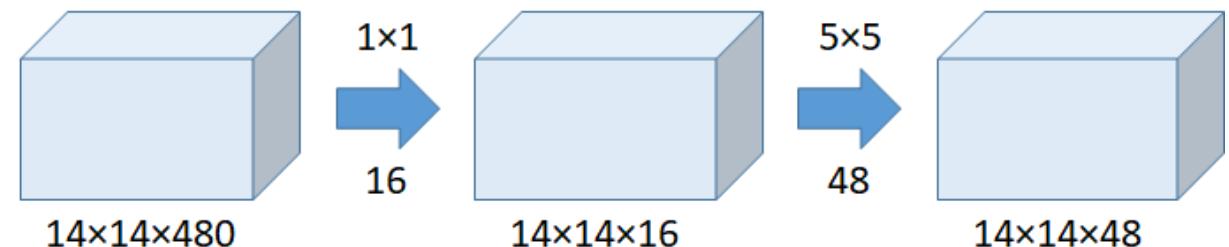
# ConvNets – GoogLeNet

[Szegedy et al., 2014]

## Módulo *Inception* – Versão básica



$$\begin{aligned}\text{Número de operações} &= (14 \times 14 \times 48) \times (5 \times 5 \times 480) = \\ &= \mathbf{112,9M}\end{aligned}$$

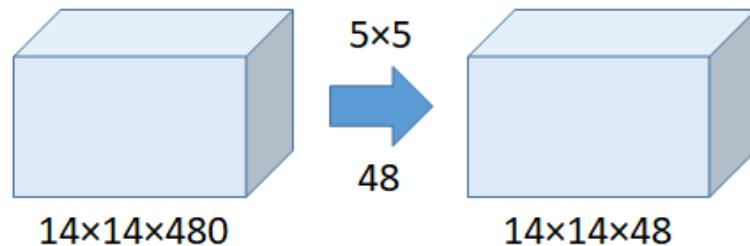


$$\begin{aligned}\text{Núm. operações para } 1 \times 1 &= (14 \times 14 \times 16) \times (1 \times 1 \times 480) = 1,5M \\ \text{Núm. operações para } 5 \times 5 &= (14 \times 14 \times 48) \times (5 \times 5 \times 16) = 3,8M \\ \text{Núm. total de operações} &= 1,5M + 3,8M = \mathbf{5,3M}\end{aligned}$$

# ConvNets – GoogLeNet

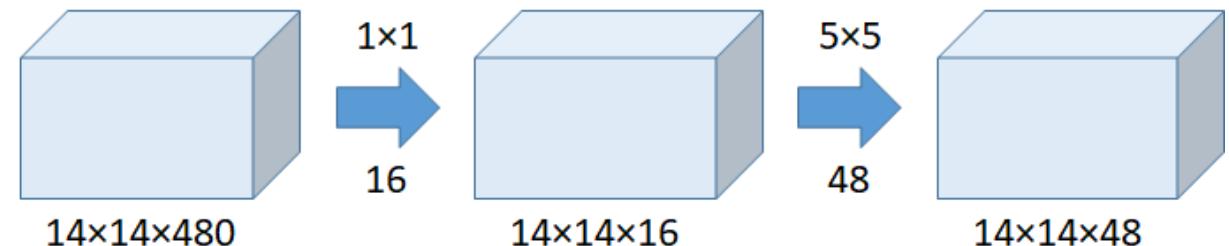
[Szegedy et al., 2014]

## Módulo *Inception* – Versão básica



- CONV  $1 \times 1$  redução de dimensionalidade
- Redução do custo computacional
- Aumento do número de parâmetros

$$\begin{aligned}\text{Número de operações} &= (14 \times 14 \times 48) \times (5 \times 5 \times 480) = \\ &= \mathbf{112,9M}\end{aligned}$$



$$\text{Núm. operações para } 1 \times 1 = (14 \times 14 \times 16) \times (1 \times 1 \times 480) = 1,5M$$

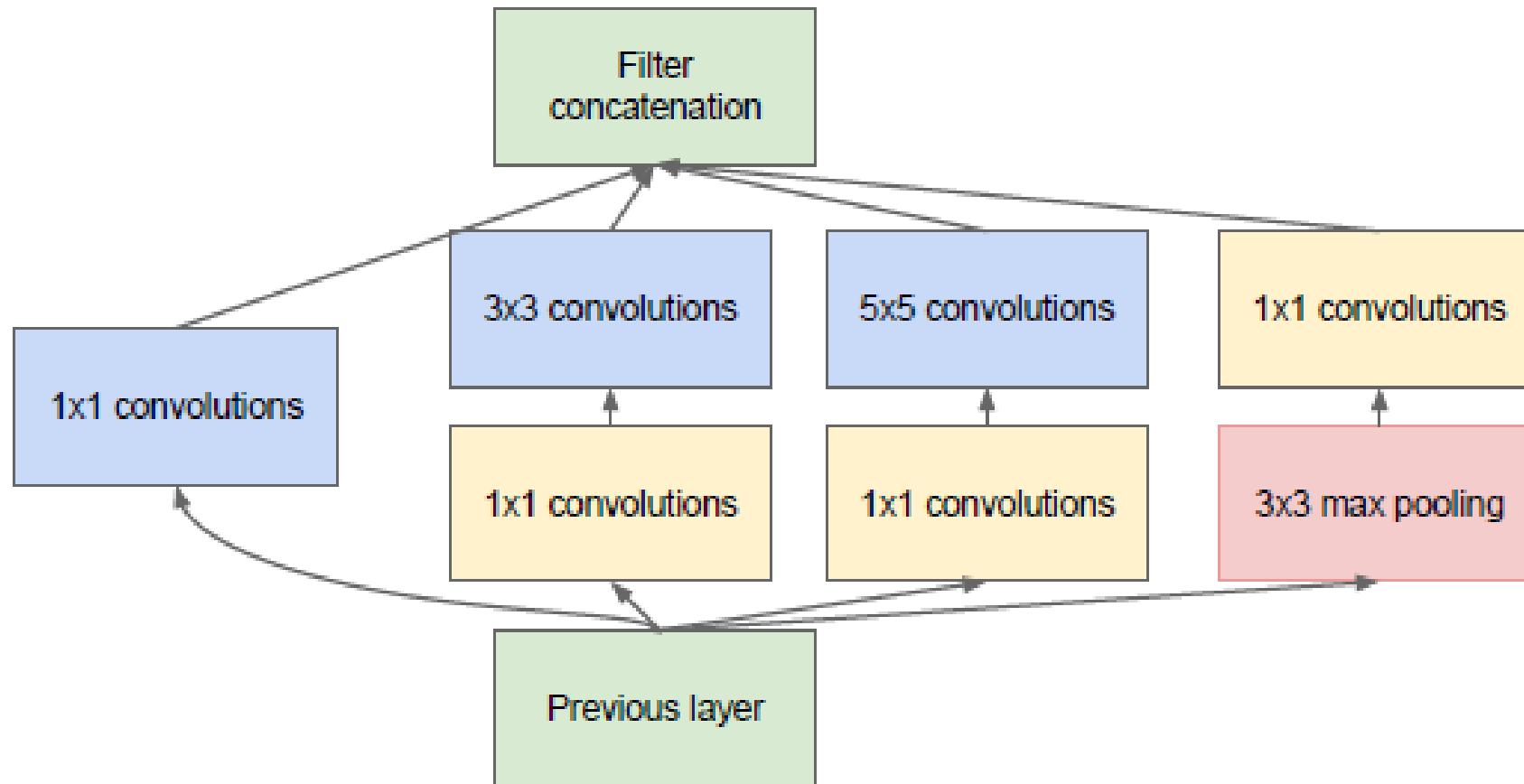
$$\text{Núm. operações para } 5 \times 5 = (14 \times 14 \times 48) \times (5 \times 5 \times 16) = 3,8M$$

$$\text{Núm. total de operações} = 1,5M + 3,8M = \mathbf{5,3M}$$

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

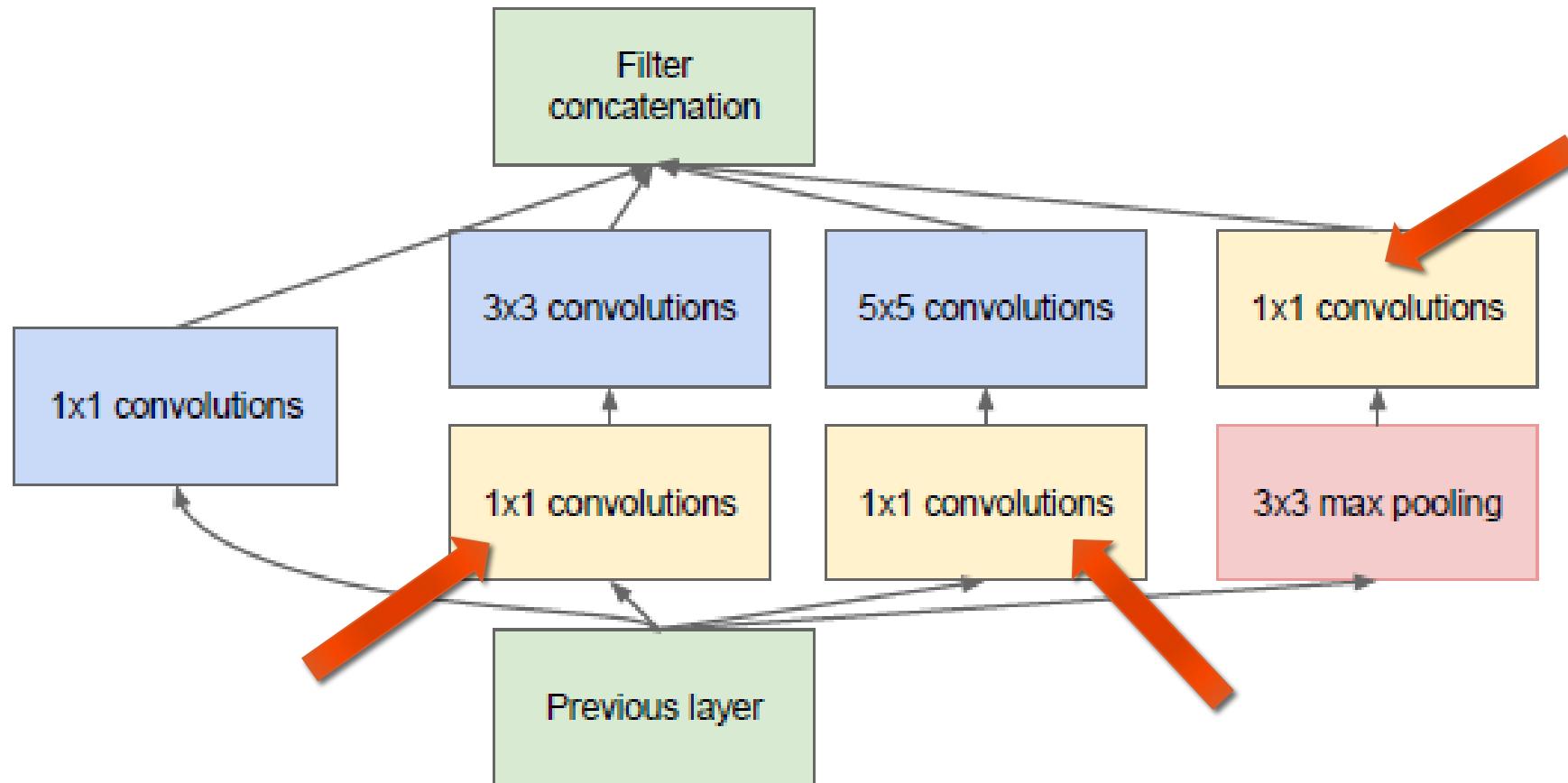
## Módulo *Inception* – Versão completa



# ConvNets – GoogLeNet

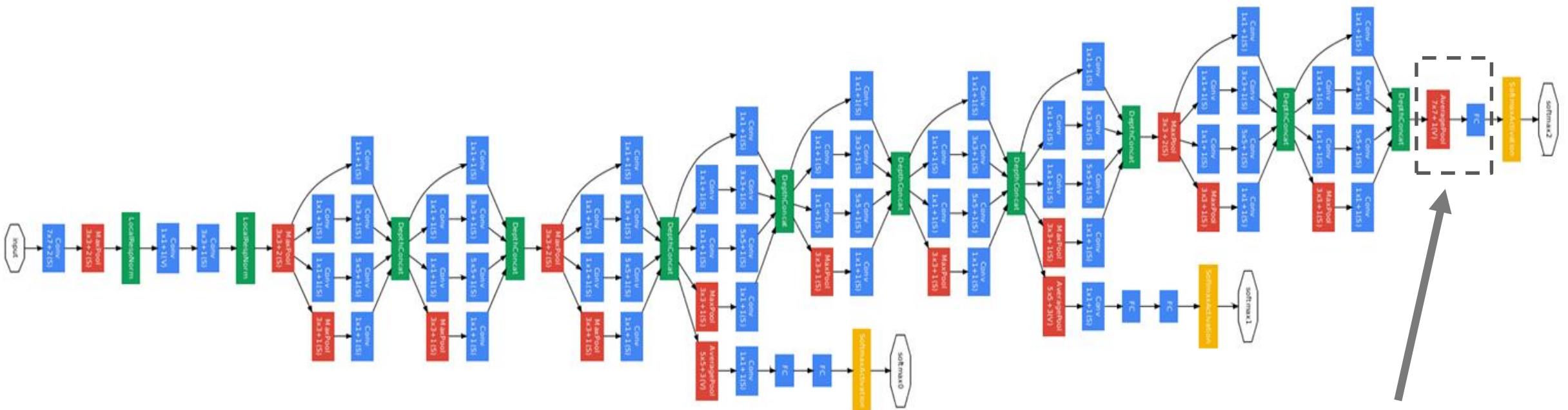
[Szegedy et al., 2014]

## Módulo *Inception* – Versão completa



# ConvNets – GoogLeNet

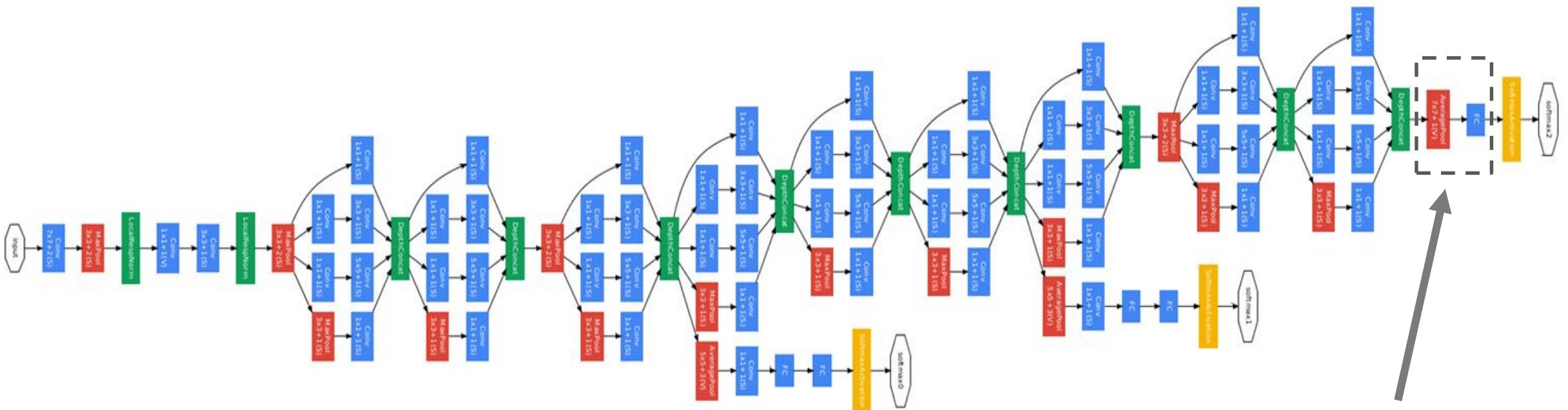
[Szegedy et al., 2014]



Agrupamento  
pela Média  
(average pooling)

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

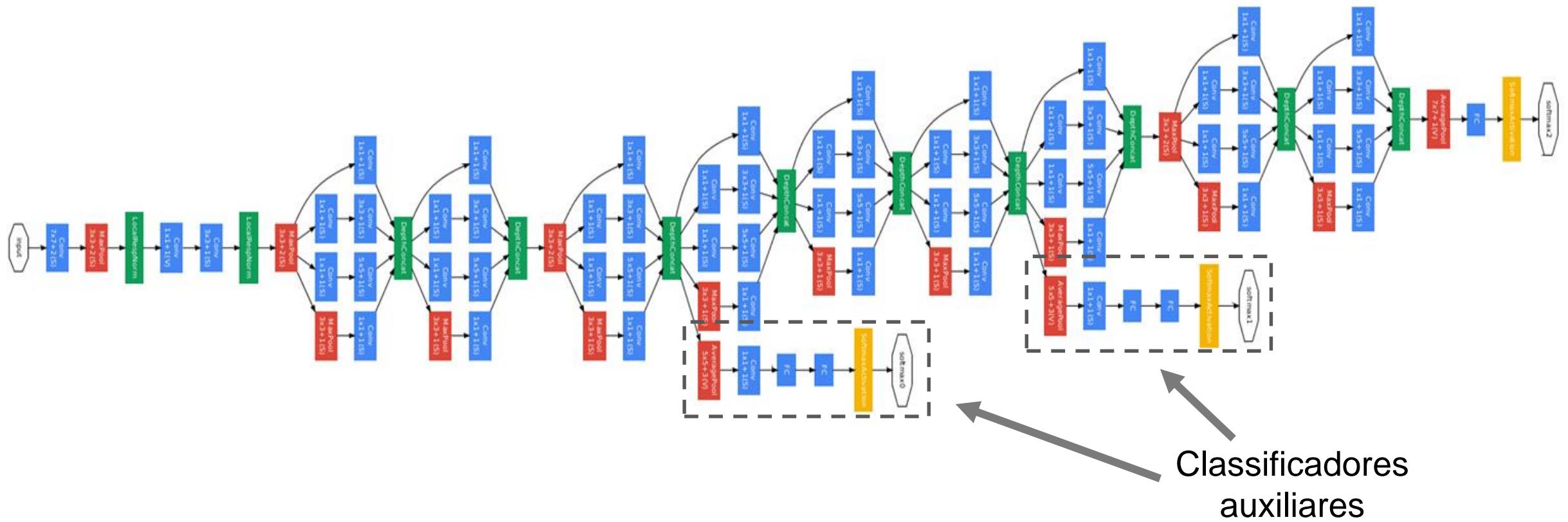


Agrupamento  
pela Média  
(average pooling)

Redução params

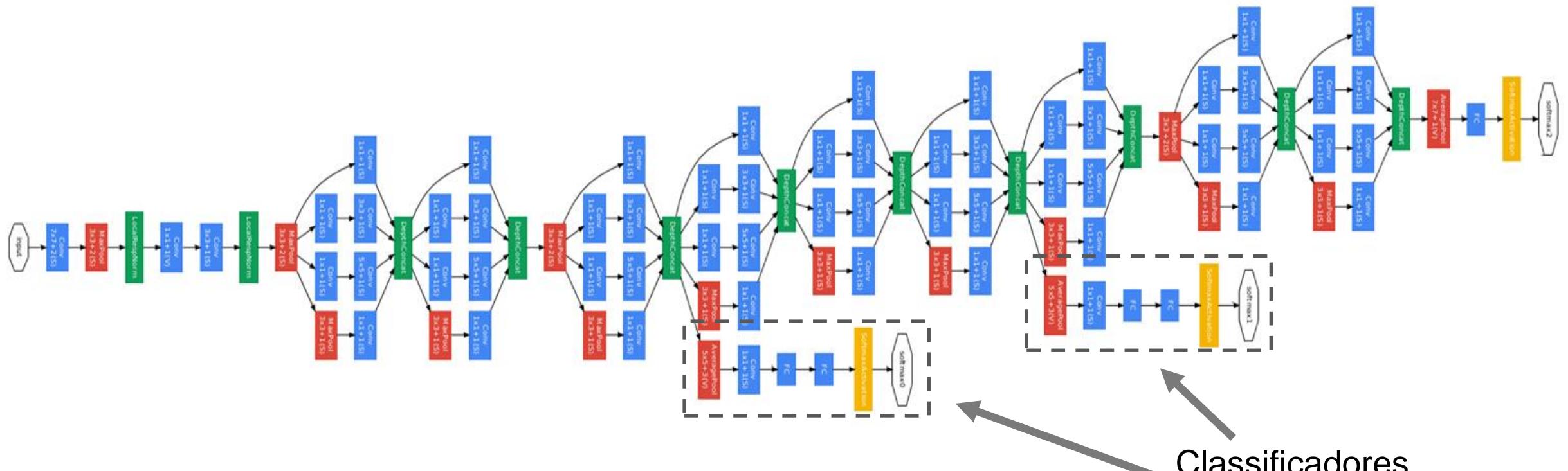
# ConvNets – GoogLeNet

[Szegedy et al., 2014]



# ConvNets – GoogLeNet

[Szegedy et al., 2014]



## **auxiliares**

## Facilita treinamento

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

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								

Vencedora ILSVRC 2014 – 6,7% de erro (top 5)

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

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								

## Detalhes:

- Apenas  $\approx 5M$  params!  
(Remoção completa das camadas FC)

Vencedora ILSVRC 2014 – 6,7% de erro (top 5)

# ConvNets – GoogLeNet

[Szegedy et al., 2014]

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								

Vencedora ILSVRC 2014 – 6,7% de erro (top 5)

## Detalhes:

- Apenas  $\approx 5$ M params!  
(Remoção completa das camadas FC)

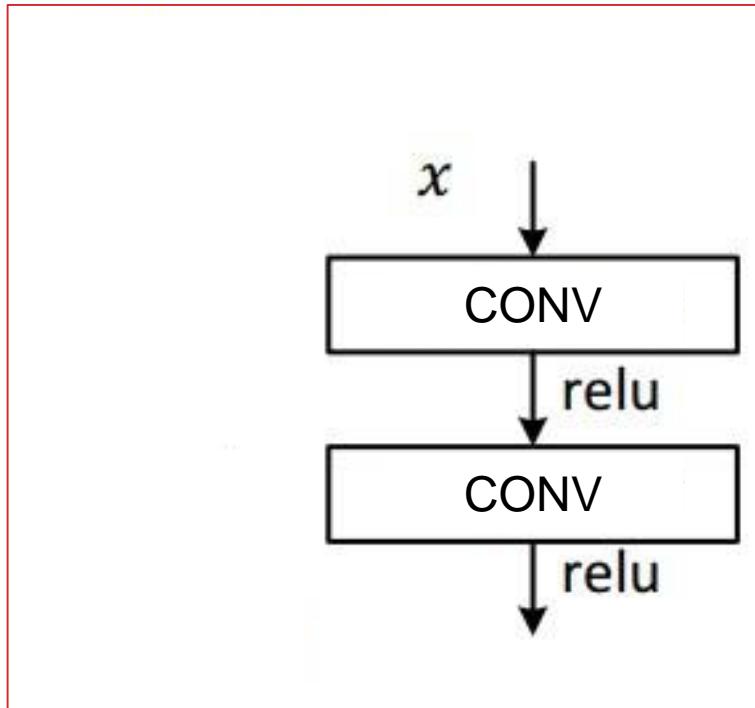
## Comparada a AlexNet:

- 12× menos params
- 2× mais computação
- 6,67% de erro (vs. 16,4%)

# ConvNets – ResNet

[He et al., 2015]

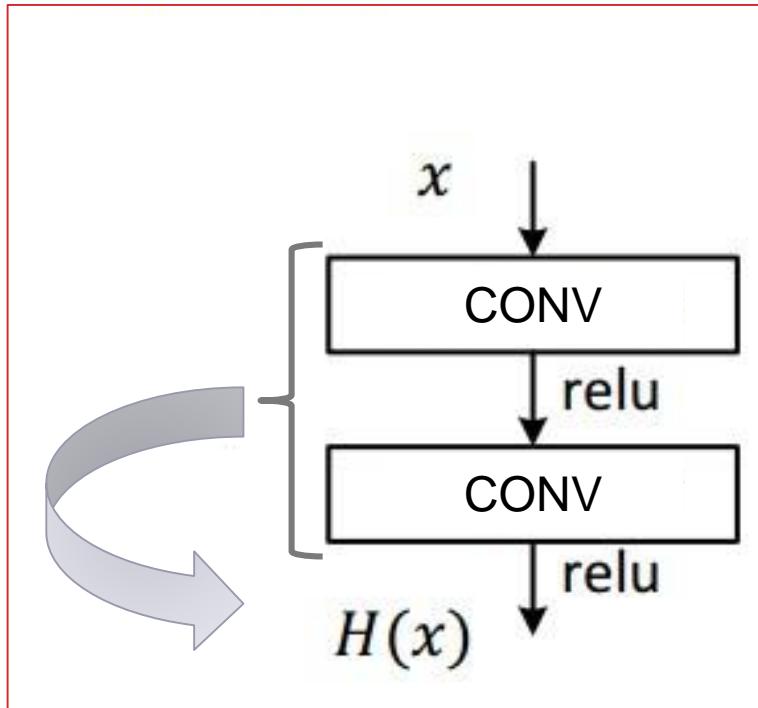
**Rede Tradicional**



# ConvNets – ResNet

[He et al., 2015]

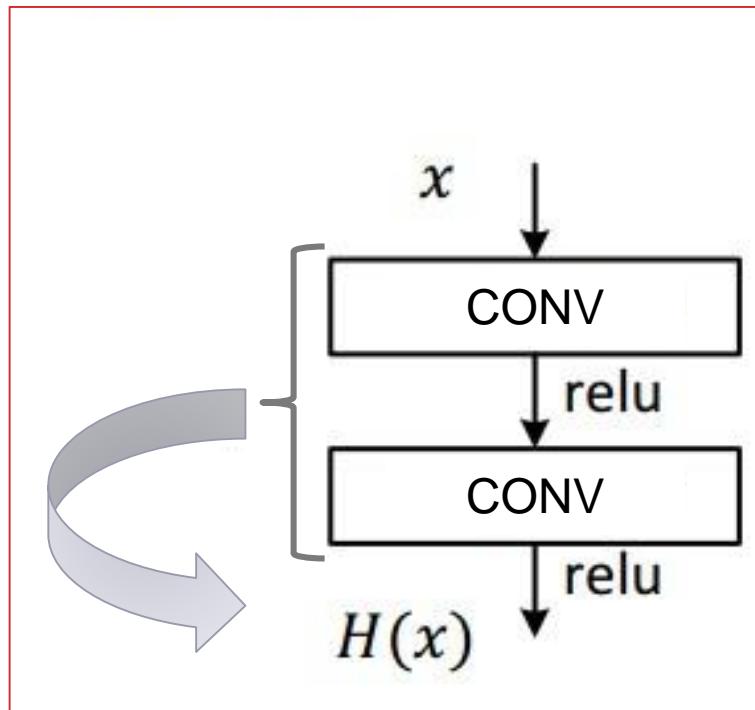
Rede Tradicional



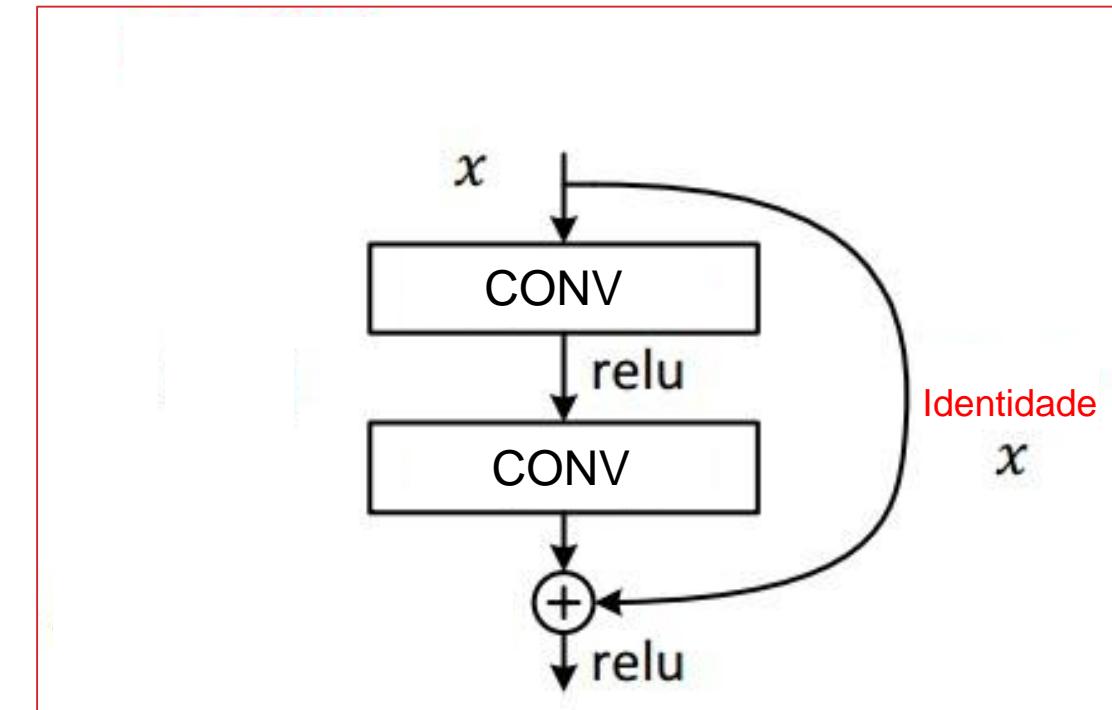
# ConvNets – ResNet

[He et al., 2015]

Rede Tradicional



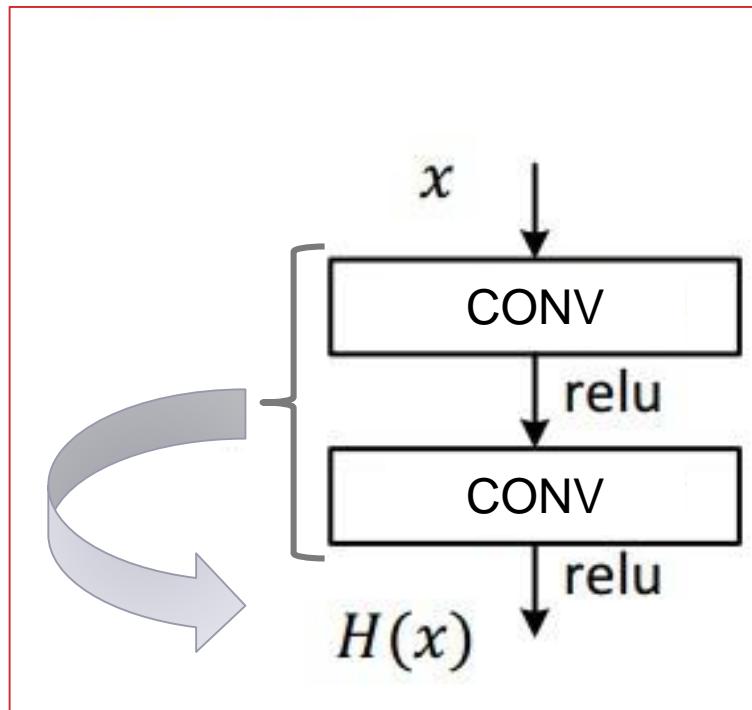
Rede Residual



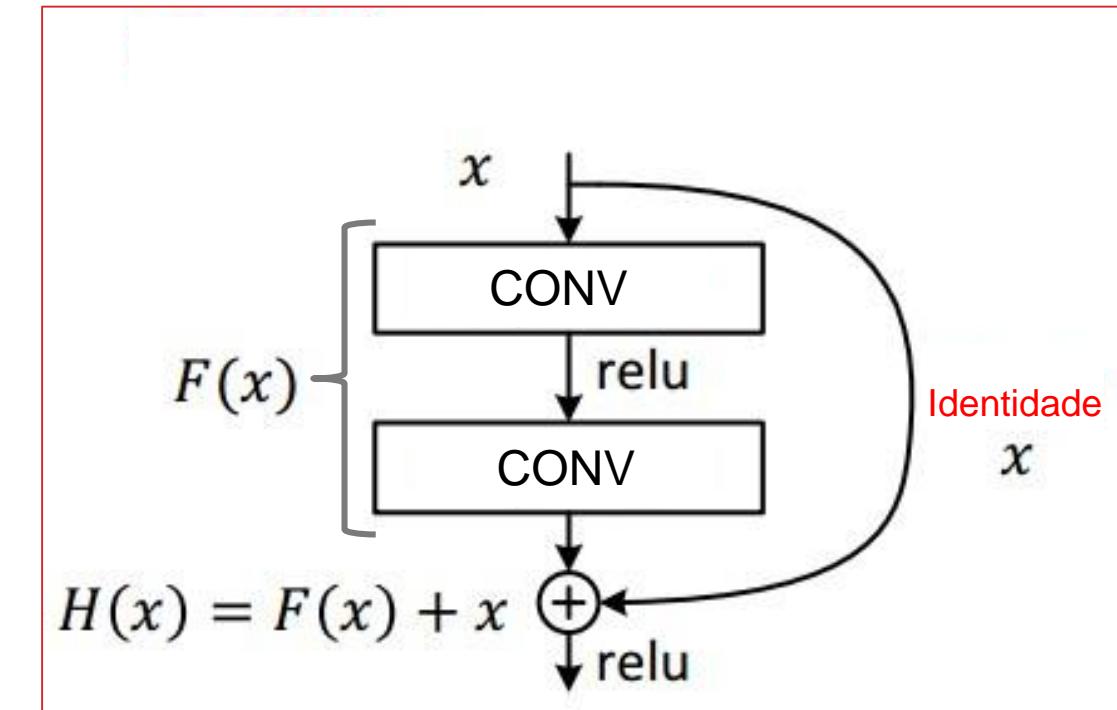
# ConvNets – ResNet

[He et al., 2015]

Rede Tradicional



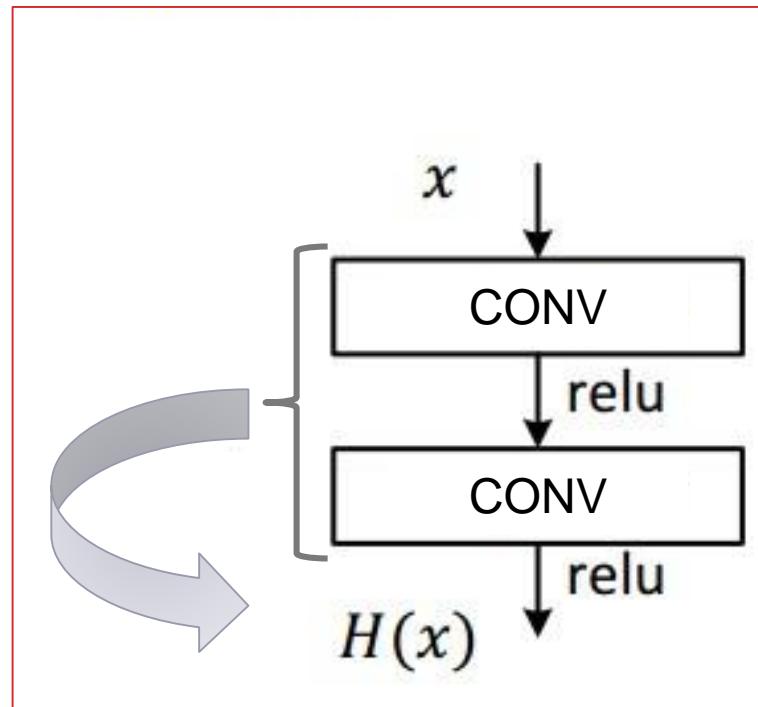
Rede Residual



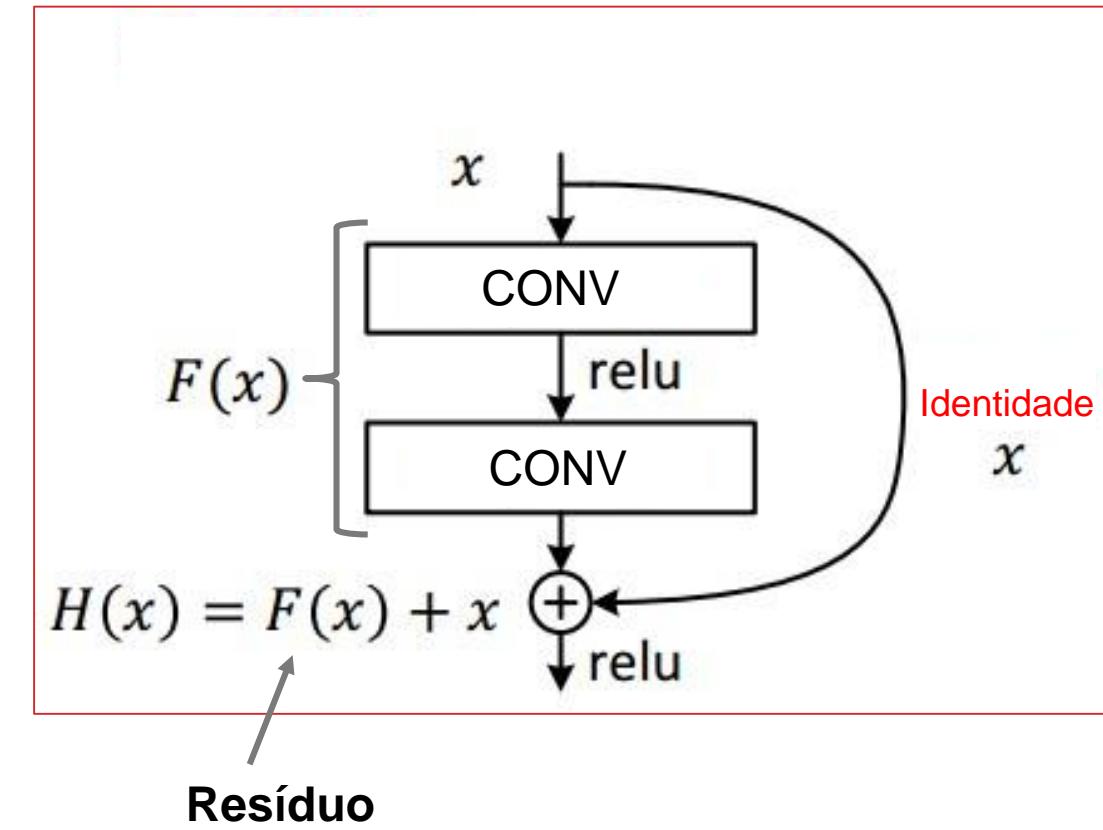
# ConvNets – ResNet

[He et al., 2015]

Rede Tradicional



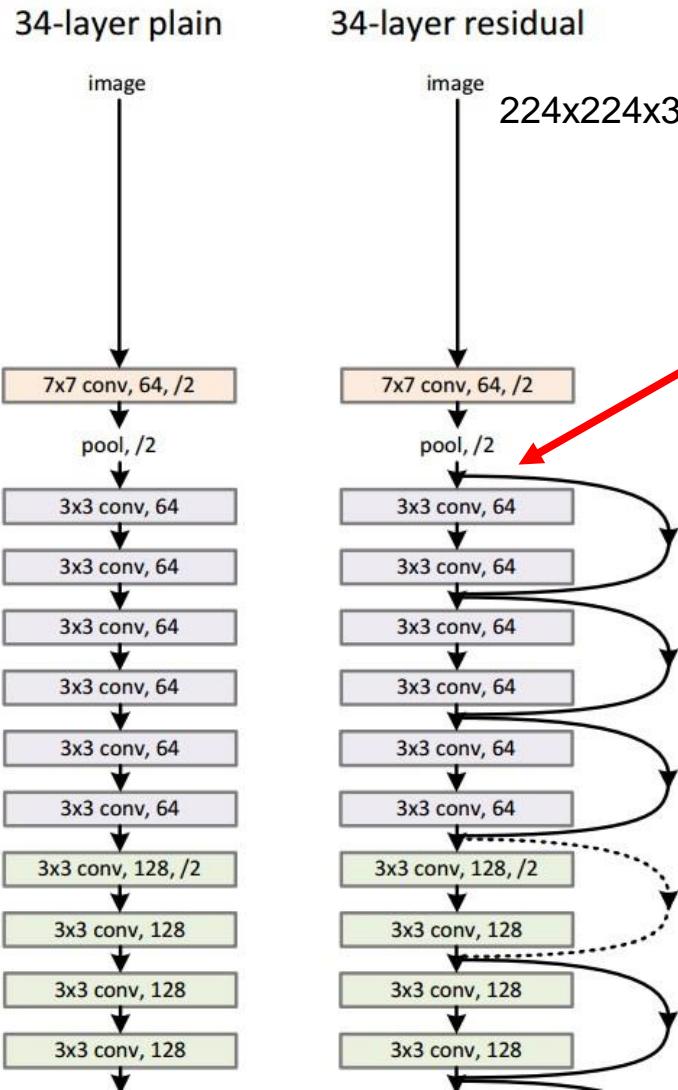
Rede Residual



$$F(x) = H(x) - x$$

# ConvNets – ResNet

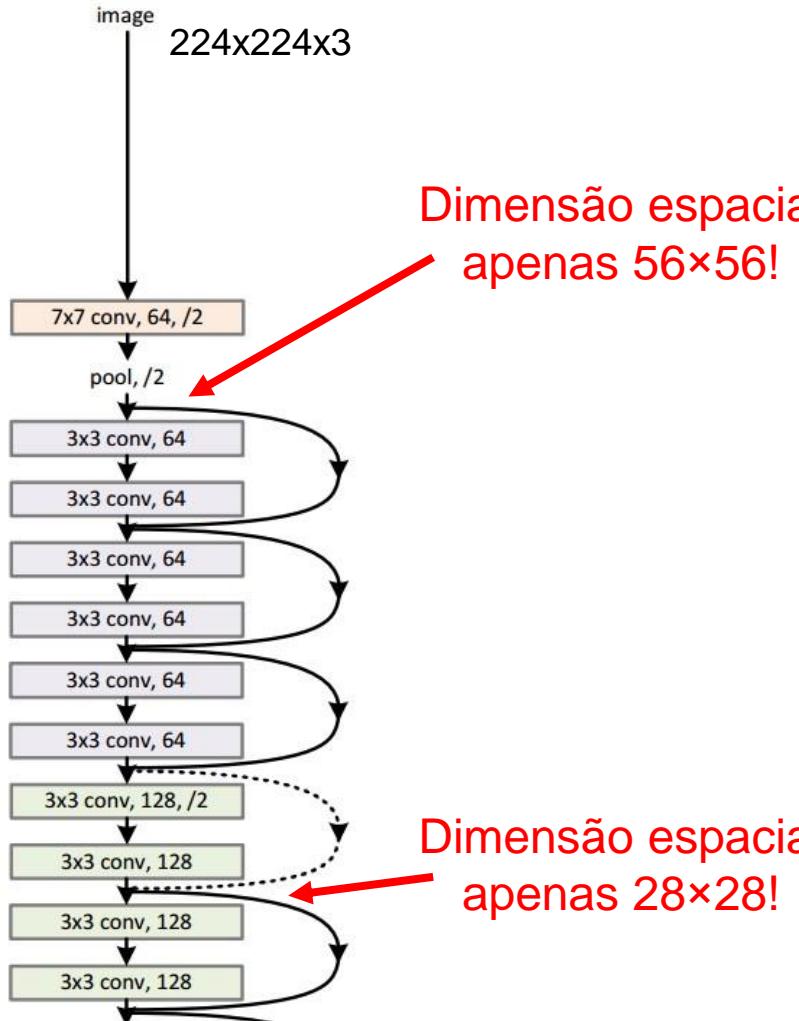
[He et al., 2015]



# ConvNets – ResNet

[He et al., 2015]

34-layer plain      34-layer residual

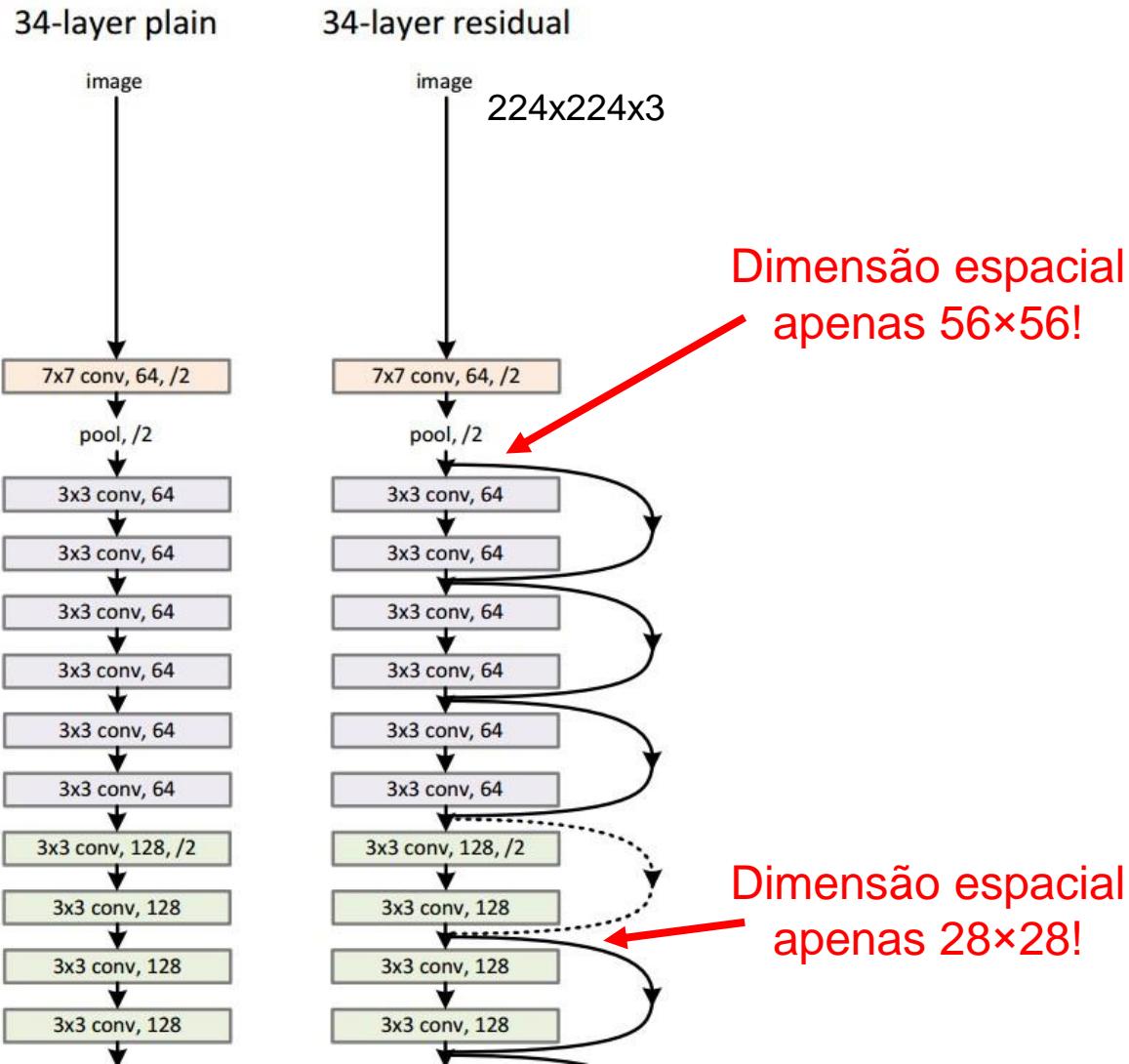


Dimensão espacial  
apenas 56×56!

Dimensão espacial  
apenas 28×28!

# ConvNets – ResNet

[He et al., 2015]



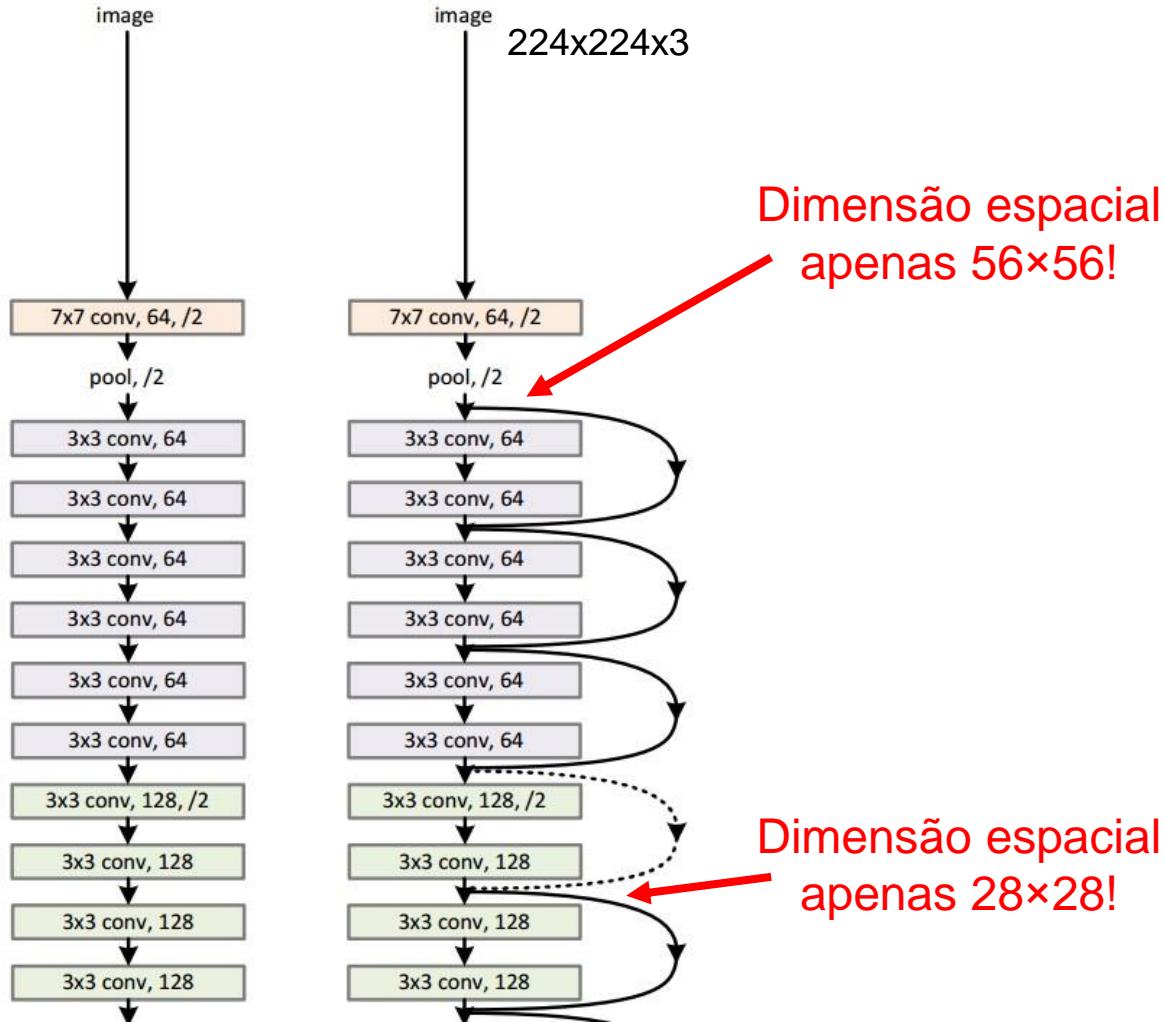
## Detalhes:

- Normalização em lote após cada CONV

# ConvNets – ResNet

[He et al., 2015]

34-layer plain      34-layer residual

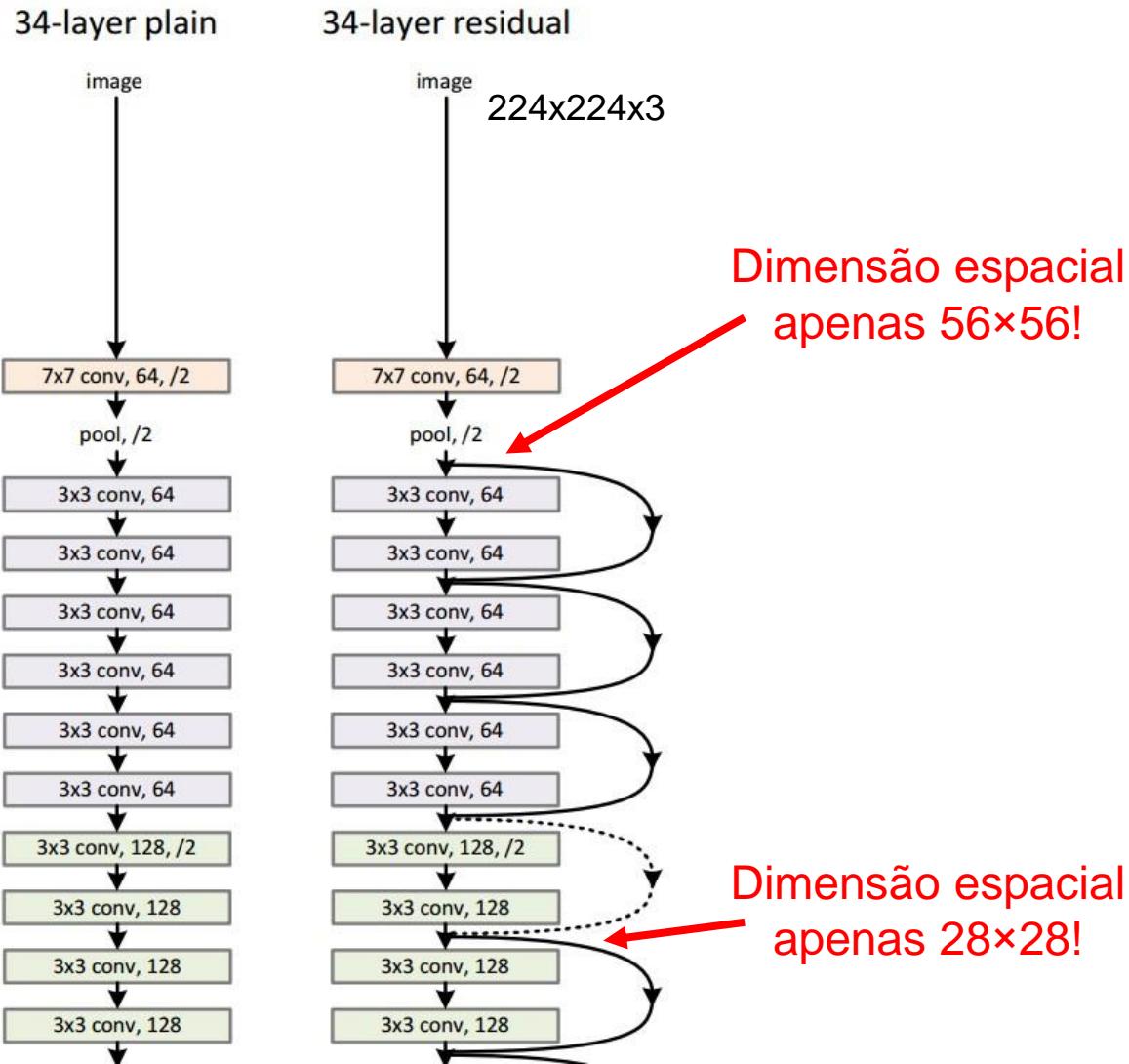


## Detalhes:

- Normalização em lote após cada CONV
- Inicialização modificada (Xavier + /2)

# ConvNets – ResNet

[He et al., 2015]



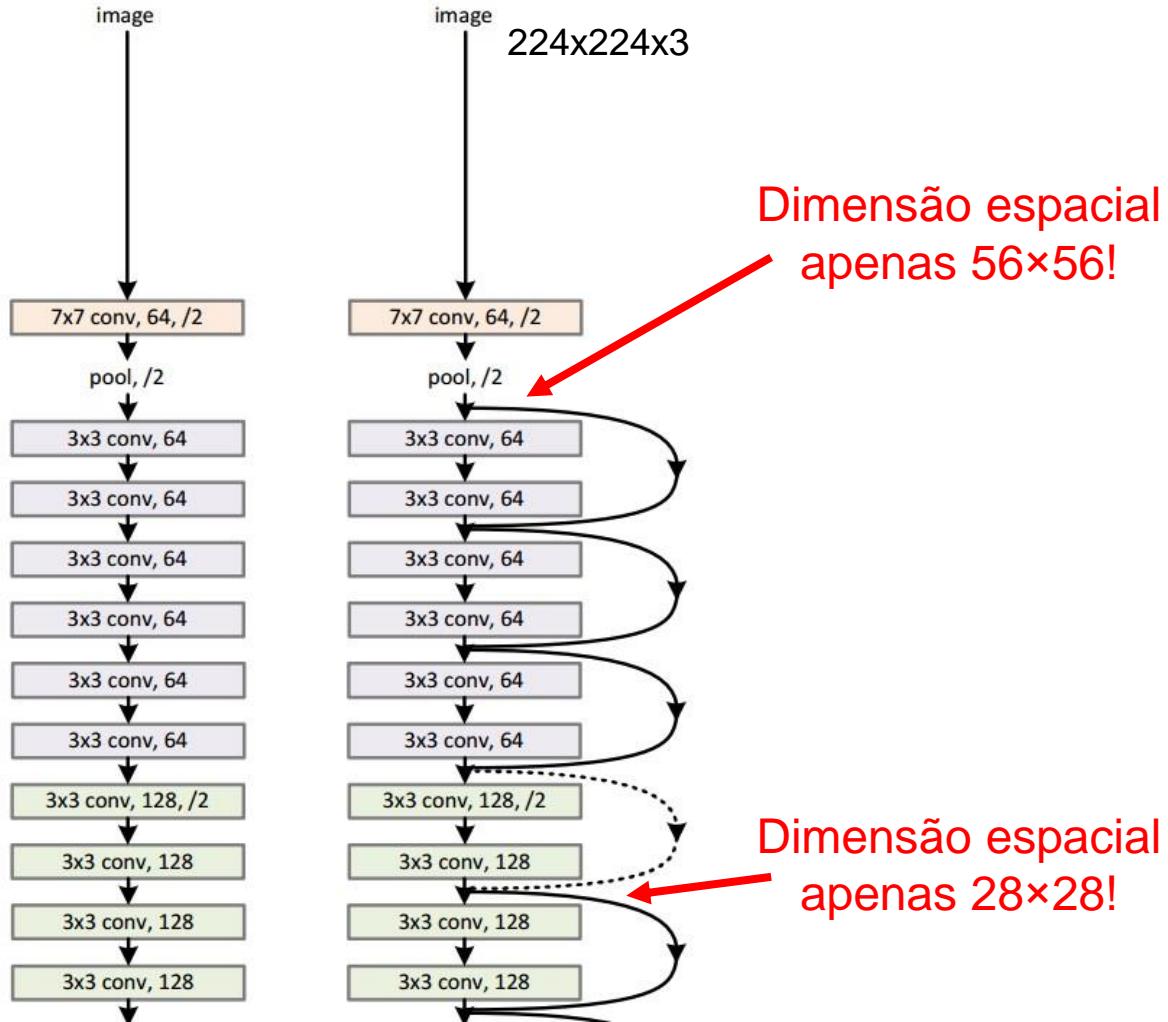
## Detalhes:

- Normalização em lote após cada CONV
- Inicialização modificada (Xavier + /2)
- SGD+Momentum (0,9)

# ConvNets – ResNet

[He et al., 2015]

34-layer plain      34-layer residual



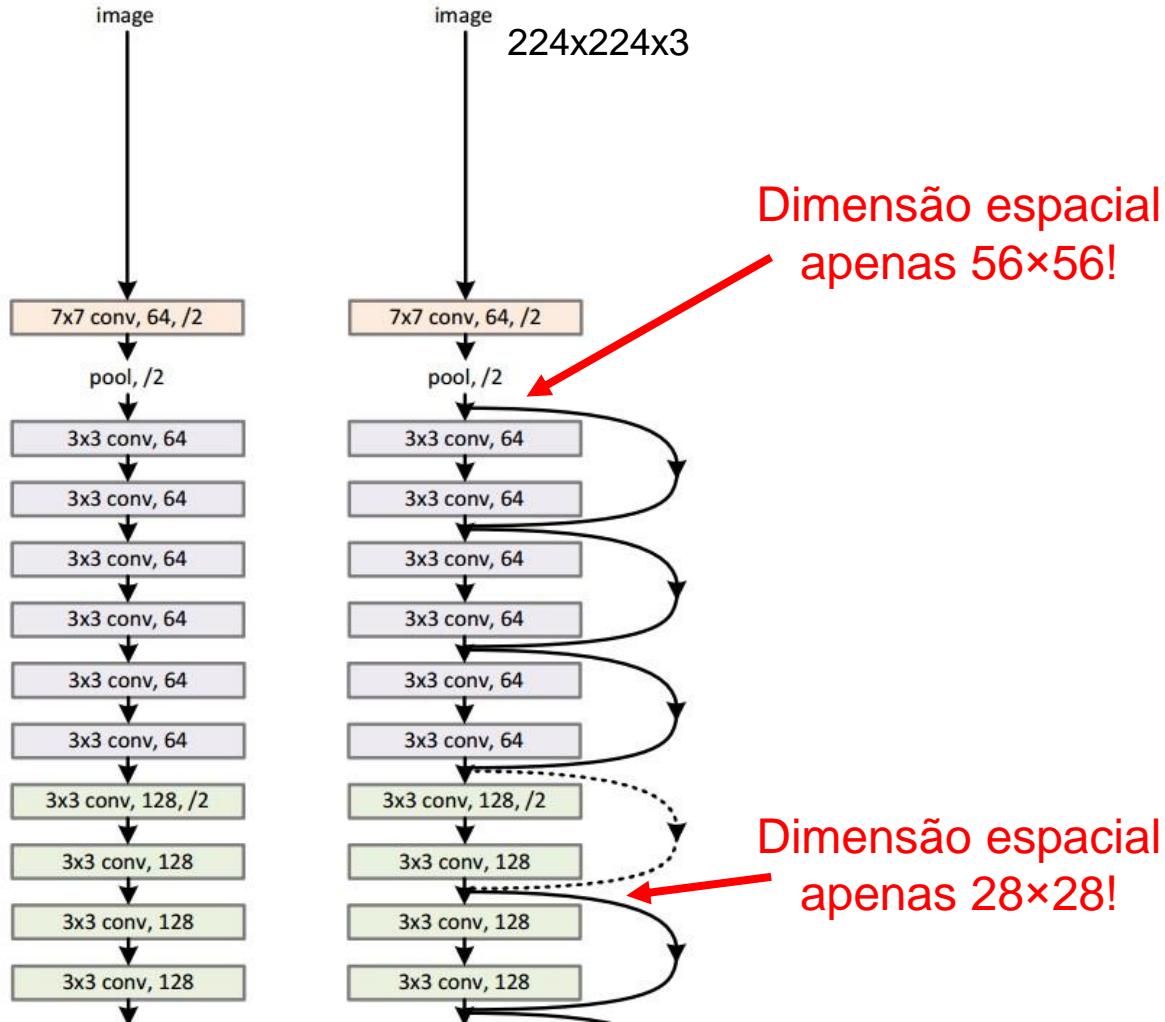
## Detalhes:

- Normalização em lote após cada CONV
- Inicialização modificada (Xavier + /2)
- SGD+Momentum (0,9)
- Tx aprendizado =  $10^{-1}$ , dividida por 10 qdo erro de validação para de reduzir

# ConvNets – ResNet

[He et al., 2015]

34-layer plain      34-layer residual

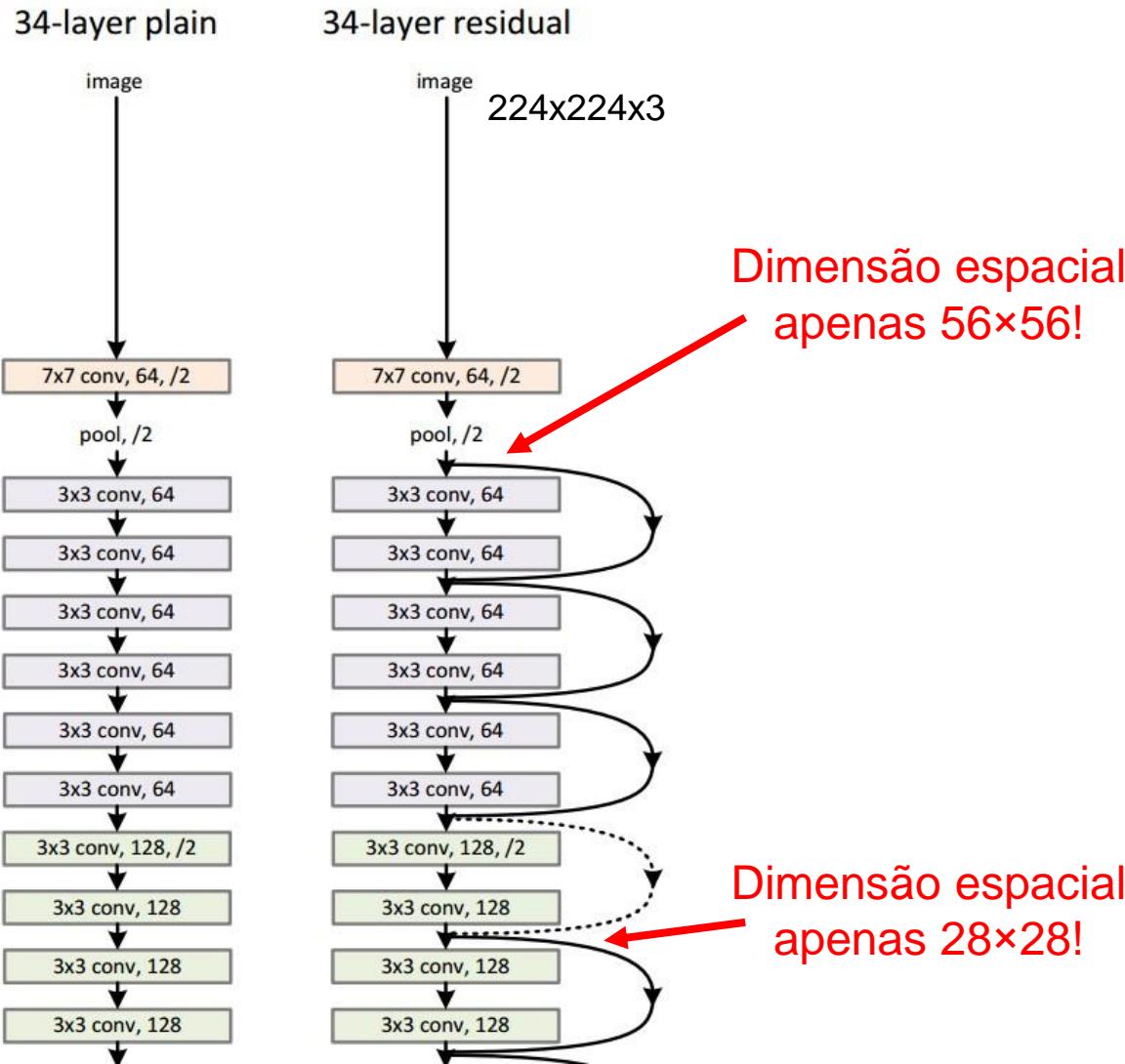


## Detalhes:

- Normalização em lote após cada CONV
- Inicialização modificada (Xavier + /2)
- SGD+*Momentum* (0,9)
- Tx aprendizado =  $10^{-1}$ , dividida por 10 qdo erro de validação para de reduzir
- Tamanho do *minibatch* = 256

# ConvNets – ResNet

[He et al., 2015]

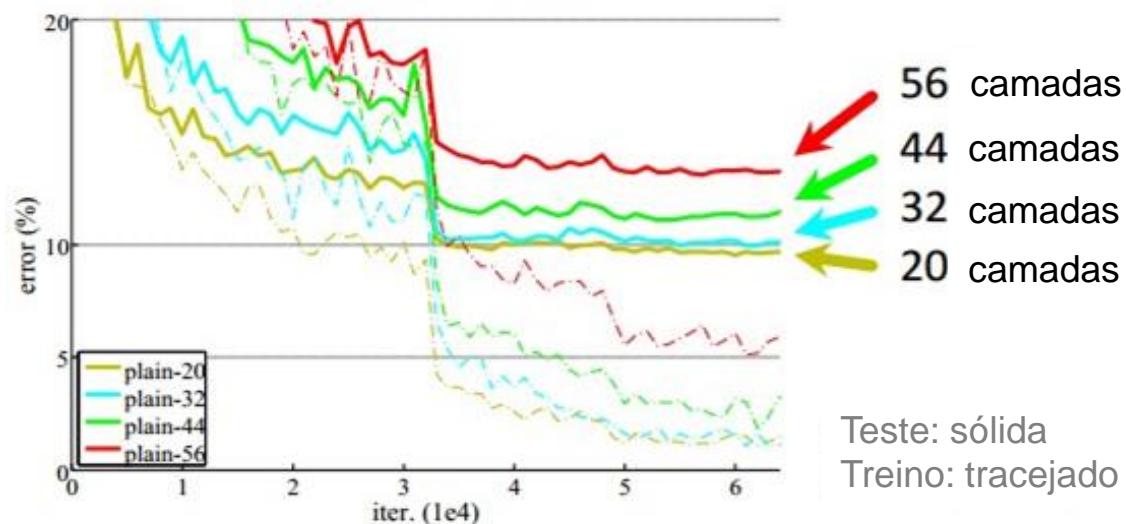


## Detalhes:

- Normalização em lote após cada CONV
- Inicialização modificada (Xavier + /2)
- SGD+*Momentum* (0,9)
- Tx aprendizado =  $10^{-1}$ , dividida por 10 qdo erro de validação para de reduzir
- Tamanho do *minibatch* = 256
- *Dropout* não é usado

# ConvNets – ResNet

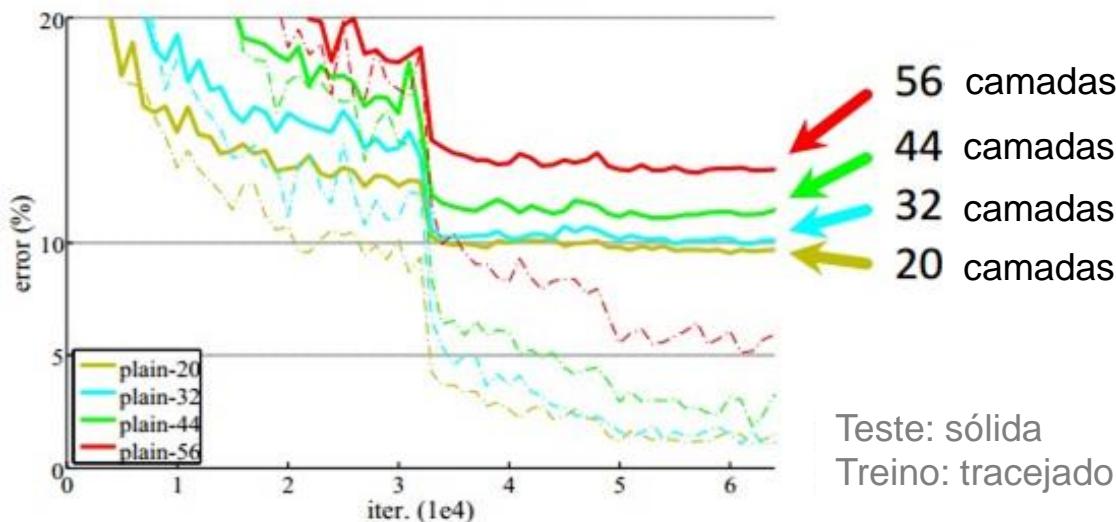
Experimentos com o conjunto de dados CIFAR-10



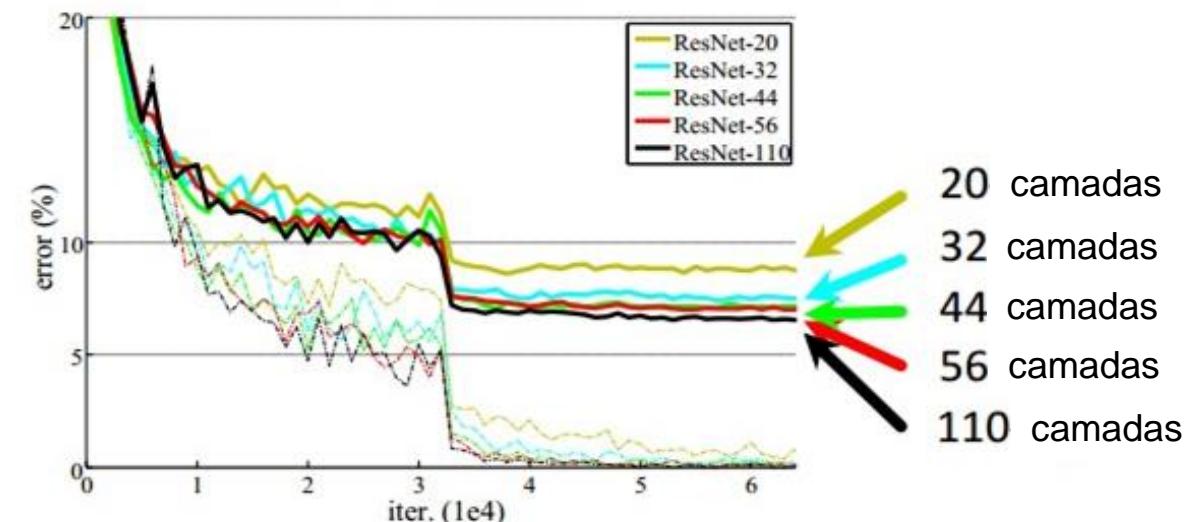
**Rede Tradicional**

# ConvNets – ResNet

Experimentos com o conjunto de dados CIFAR-10



**Rede Tradicional**

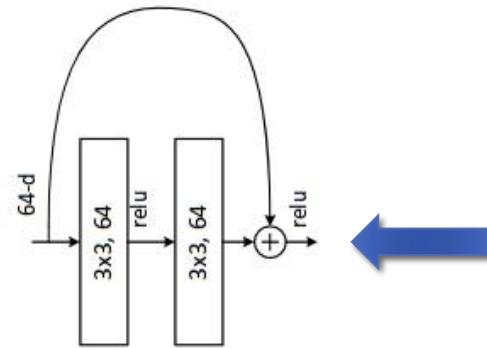


**Rede Residual**

# ConvNets – ResNet

[He et al., 2015]

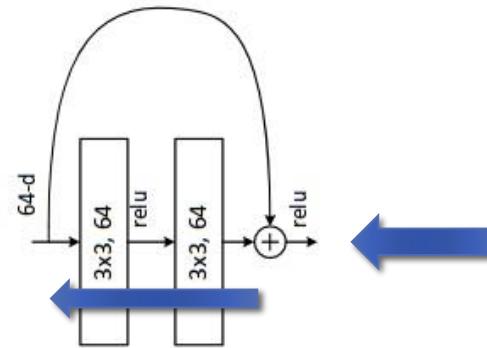
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

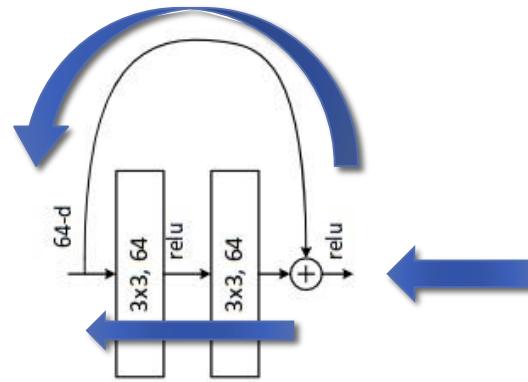
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

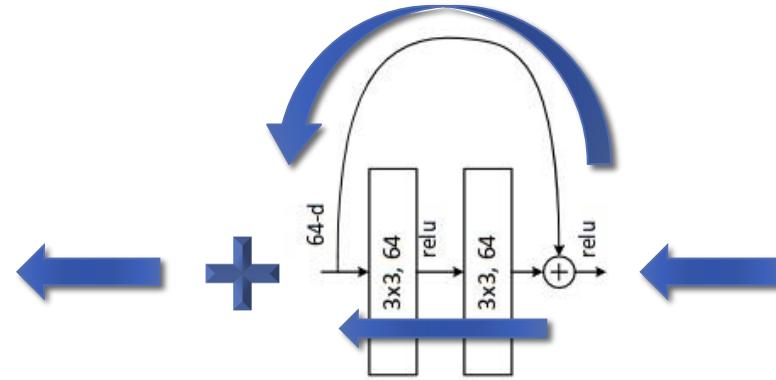
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

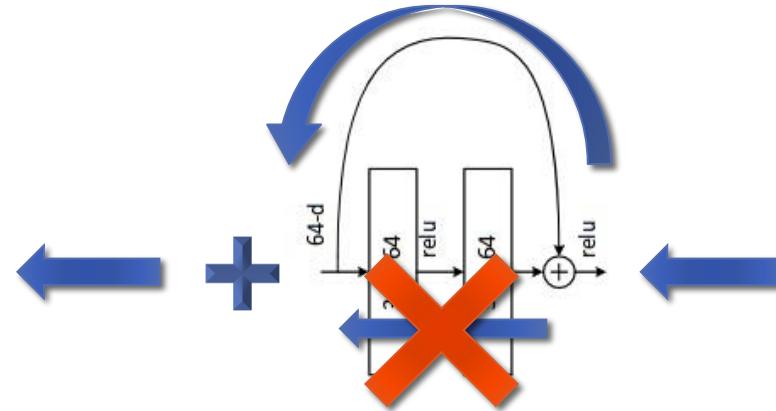
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

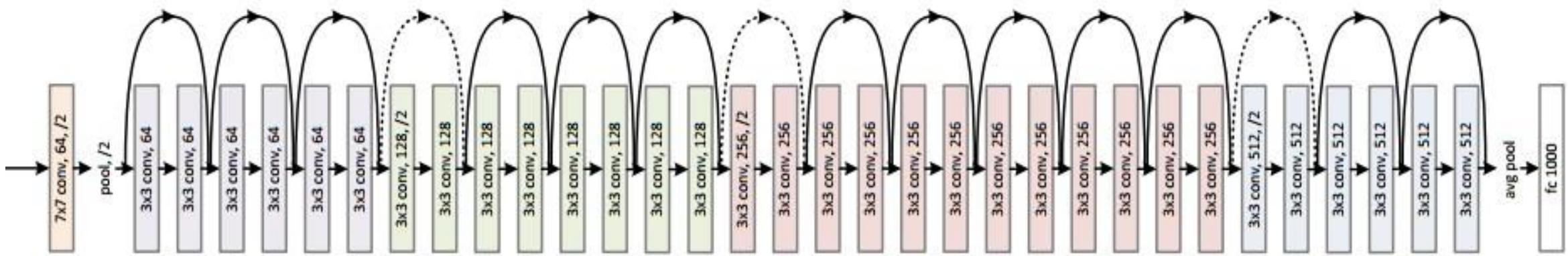
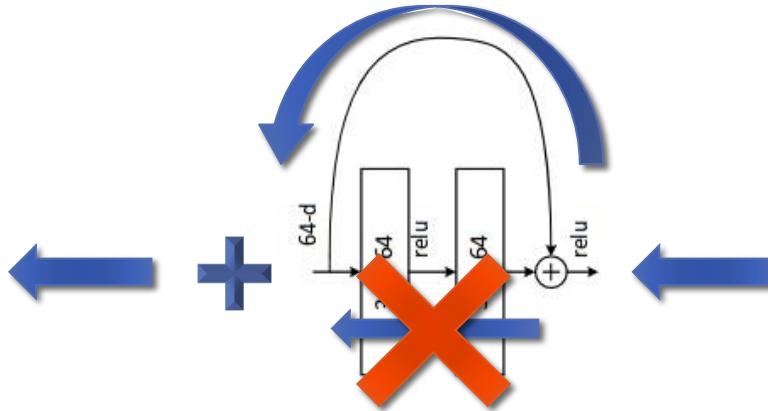
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

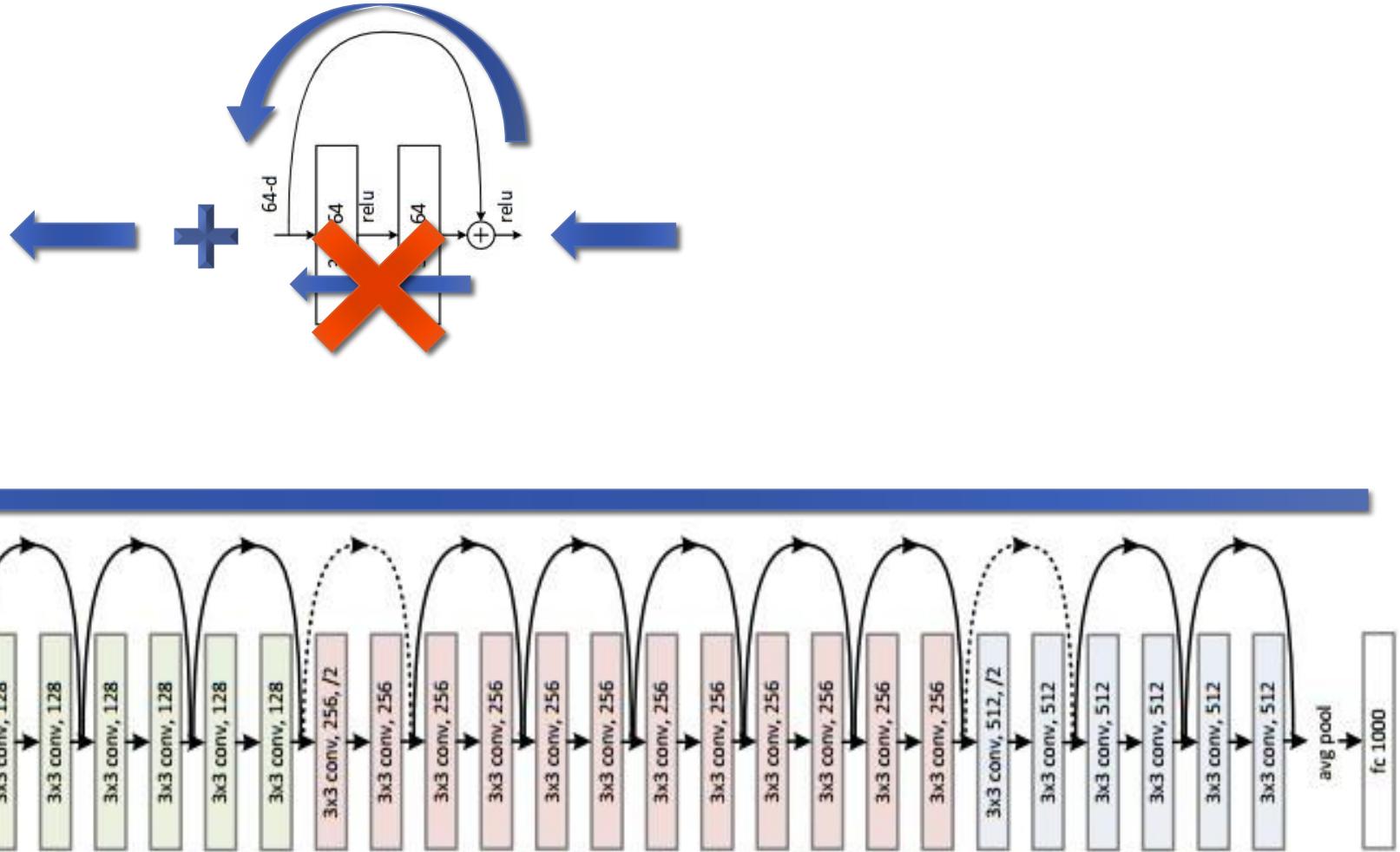
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

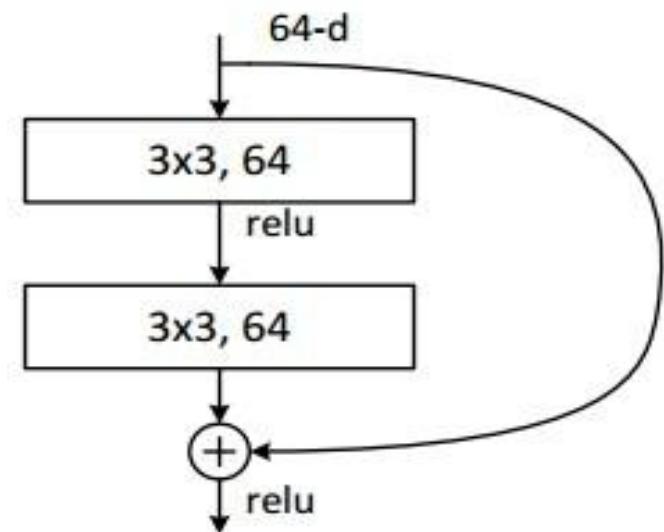
## Fluxo de Gradientes



# ConvNets – ResNet

[He et al., 2015]

Lidando com redes de 50+ camadas

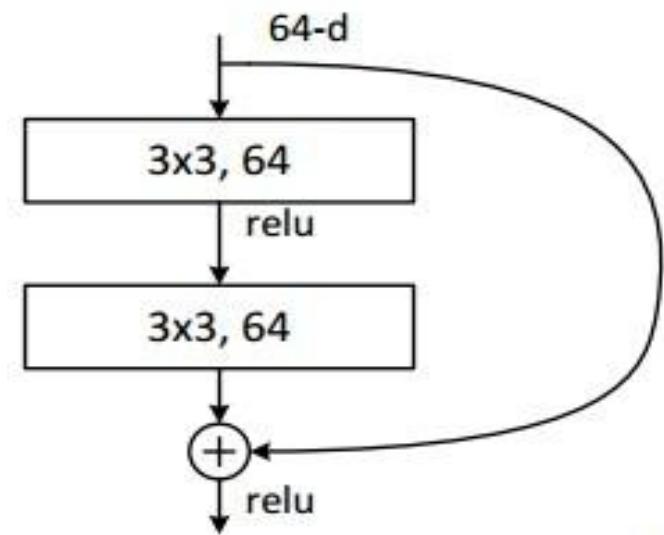


Todas  
CONV 3x3

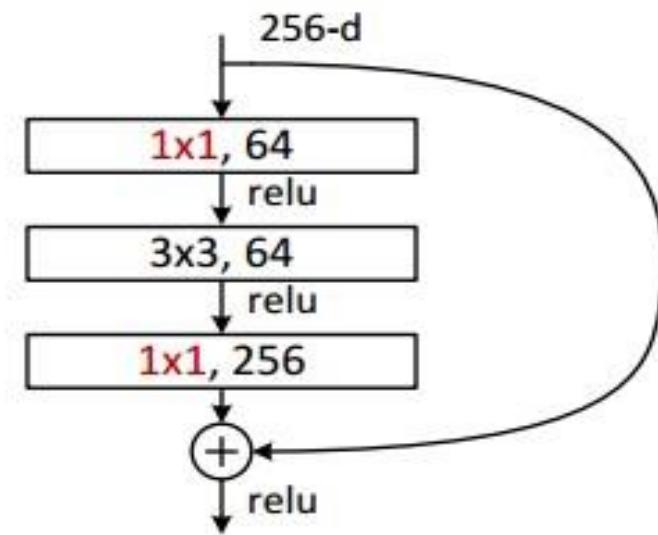
# ConvNets – ResNet

[He et al., 2015]

Lidando com redes de 50+ camadas



Todas  
CONV 3x3



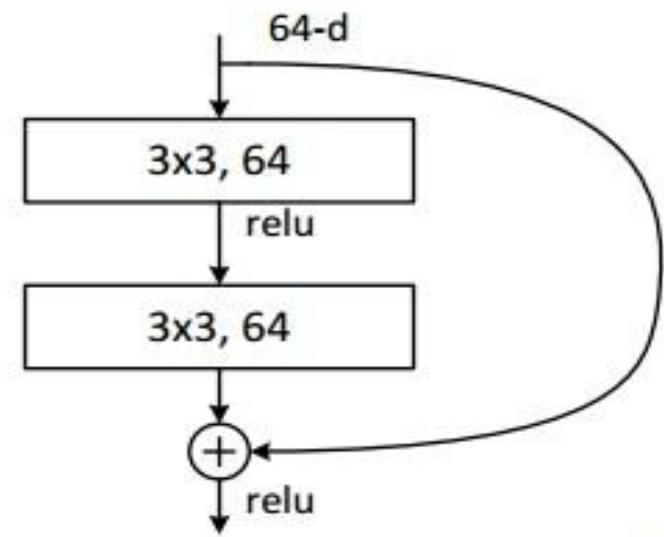
Evitando o “gargalo”  
(para ResNet-50/101/152)



# ConvNets – ResNet

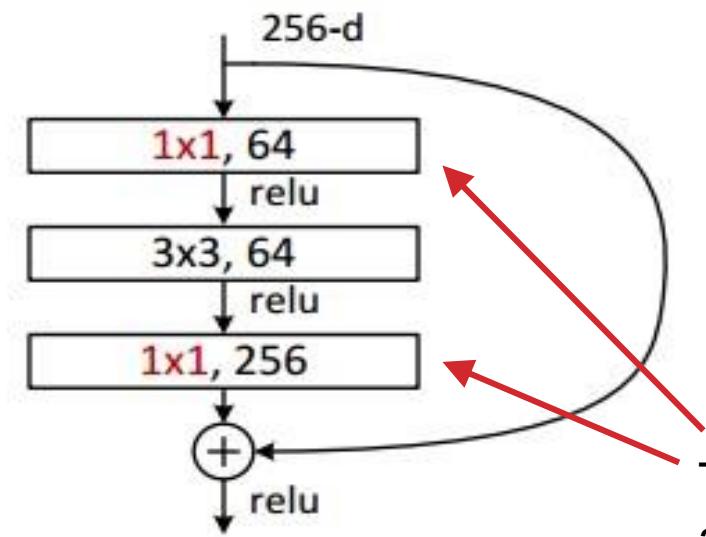
[He et al., 2015]

Lidando com redes de 50+ camadas



Todas  
CONV 3x3

Complexidade  
similar

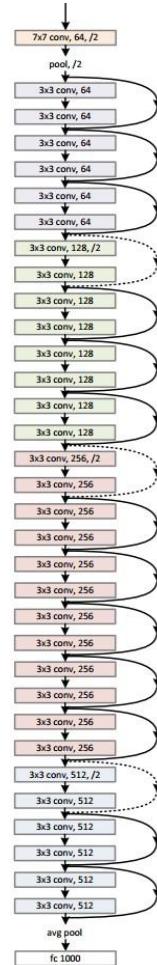


Evitando o “gargalo”  
(para ResNet-50/101/152)

Truque similar  
a GoogLeNet

# ConvNets – ResNet

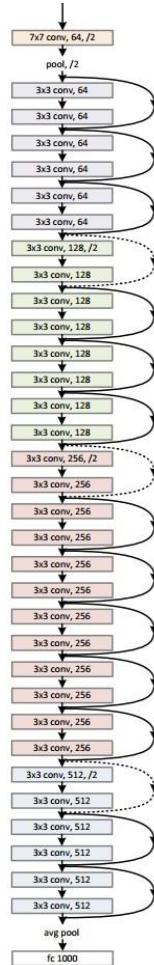
[He et al., 2015]



layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112			7×7, 64, stride 2		
conv2_x	56×56	$\left[ \begin{matrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$
conv3_x	28×28	$\left[ \begin{matrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 8$
conv4_x	14×14	$\left[ \begin{matrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{matrix} \right] \times 6$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 6$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 23$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 36$
conv5_x	7×7	$\left[ \begin{matrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

# ConvNets – ResNet

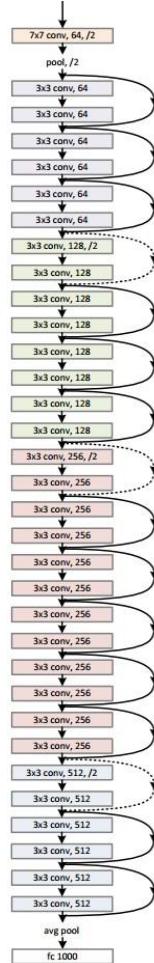
[He et al., 2015]



layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112			7×7, 64, stride 2		
conv2_x	56×56	$\left[ \begin{matrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$
conv3_x	28×28	$\left[ \begin{matrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 8$
conv4_x	14×14	$\left[ \begin{matrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{matrix} \right] \times 6$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 6$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 23$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 36$
conv5_x	7×7	$\left[ \begin{matrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

# ConvNets – ResNet

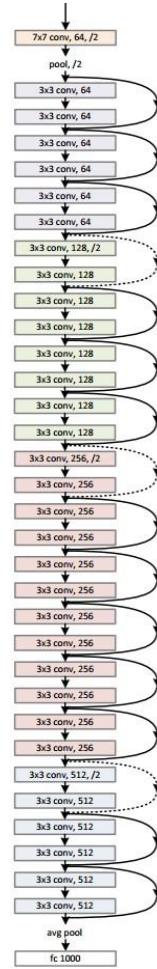
[He et al., 2015]



layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112			7×7, 64, stride 2		
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

# ConvNets – ResNet

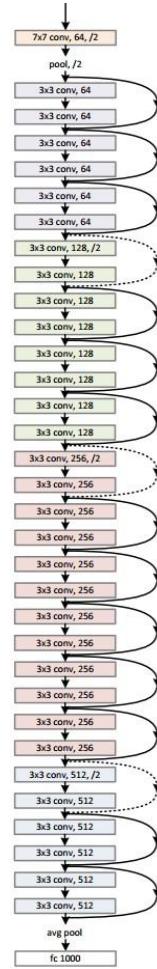
[He et al., 2015]



layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112			7×7, 64, stride 2		
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

# ConvNets – ResNet

[He et al., 2015]

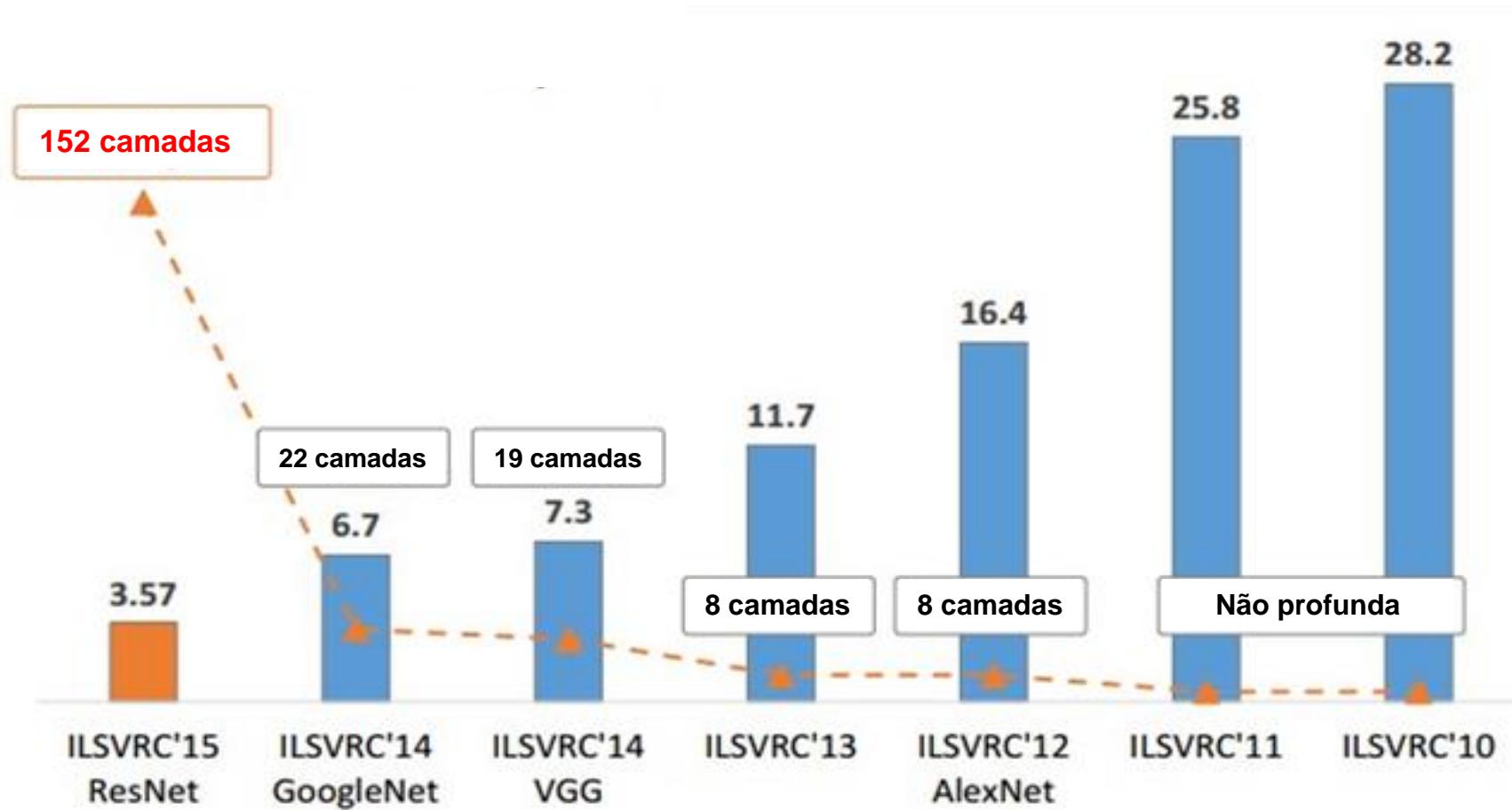


layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112x112			7x7, 64, stride 2		
conv2_x	56x56	$\left[ \begin{matrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{matrix} \right] \times 3$
conv3_x	28x28	$\left[ \begin{matrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 4$	$\left[ \begin{matrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{matrix} \right] \times 8$
conv4_x	14x14	$\left[ \begin{matrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{matrix} \right] \times 6$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 23$	$\left[ \begin{matrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{matrix} \right] \times 36$	
conv5_x	7x7	$\left[ \begin{matrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{matrix} \right] \times 2$	$\left[ \begin{matrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{matrix} \right] \times 3$	$\left[ \begin{matrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{matrix} \right] \times 3$		
	1x1			average pool, 1000-d fc, softmax		
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

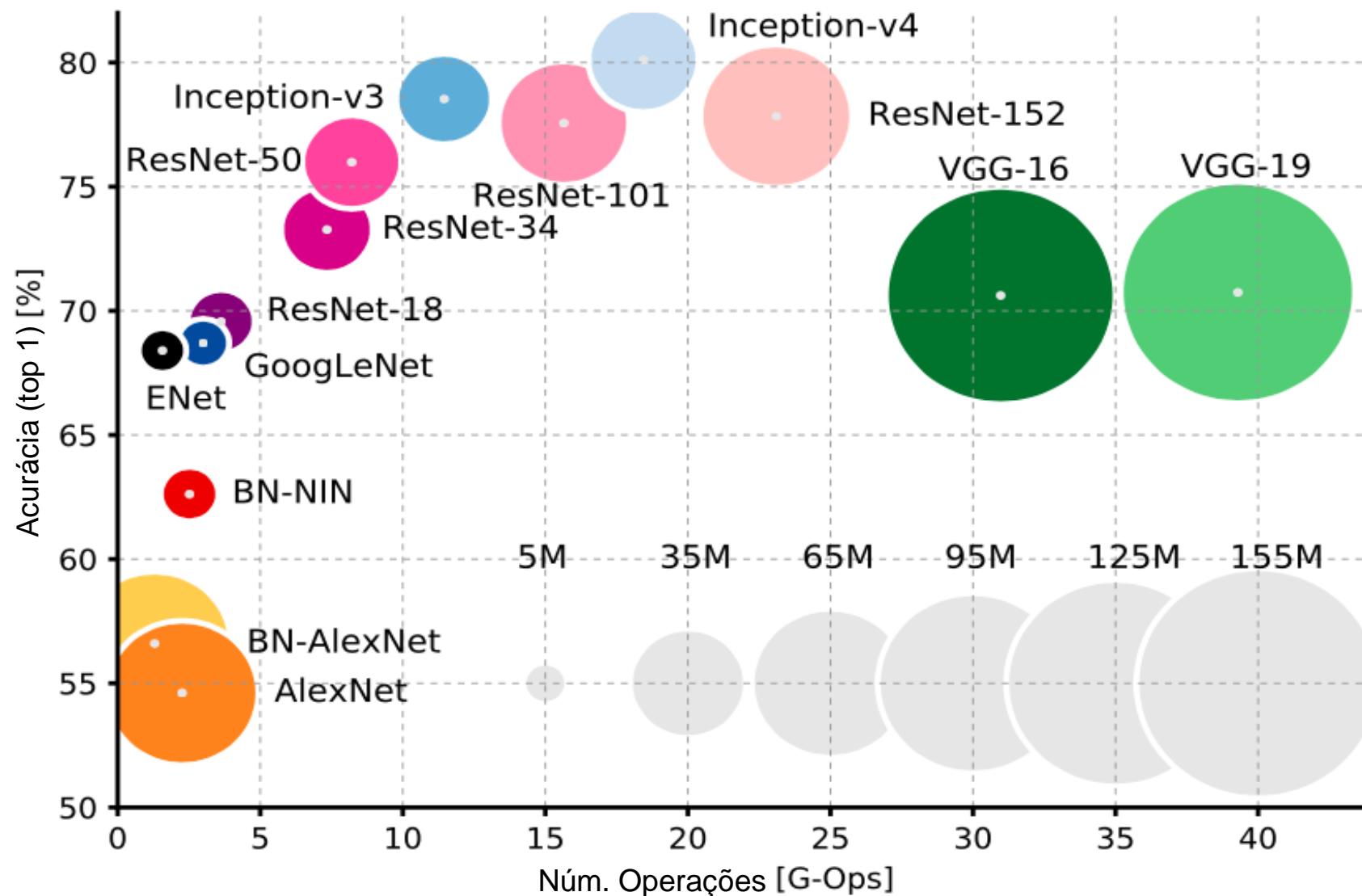
VGG16 –  $15.3 \times 10^9$   
VGG19 –  $19.6 \times 10^9$

Vencedora ILSVRC 2015 – 3,6% de erro (top 5)

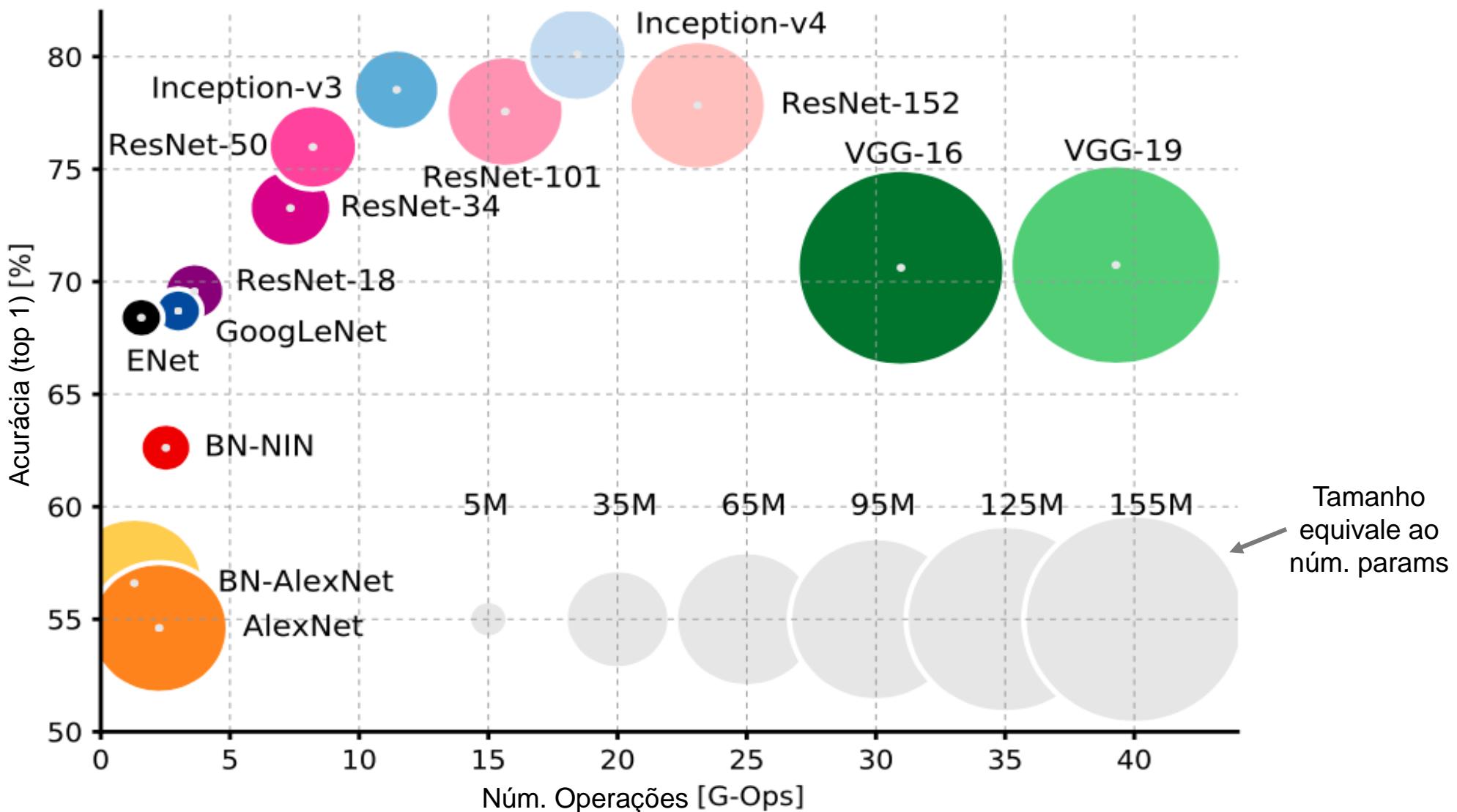
# ConvNets – Análise Comparativa



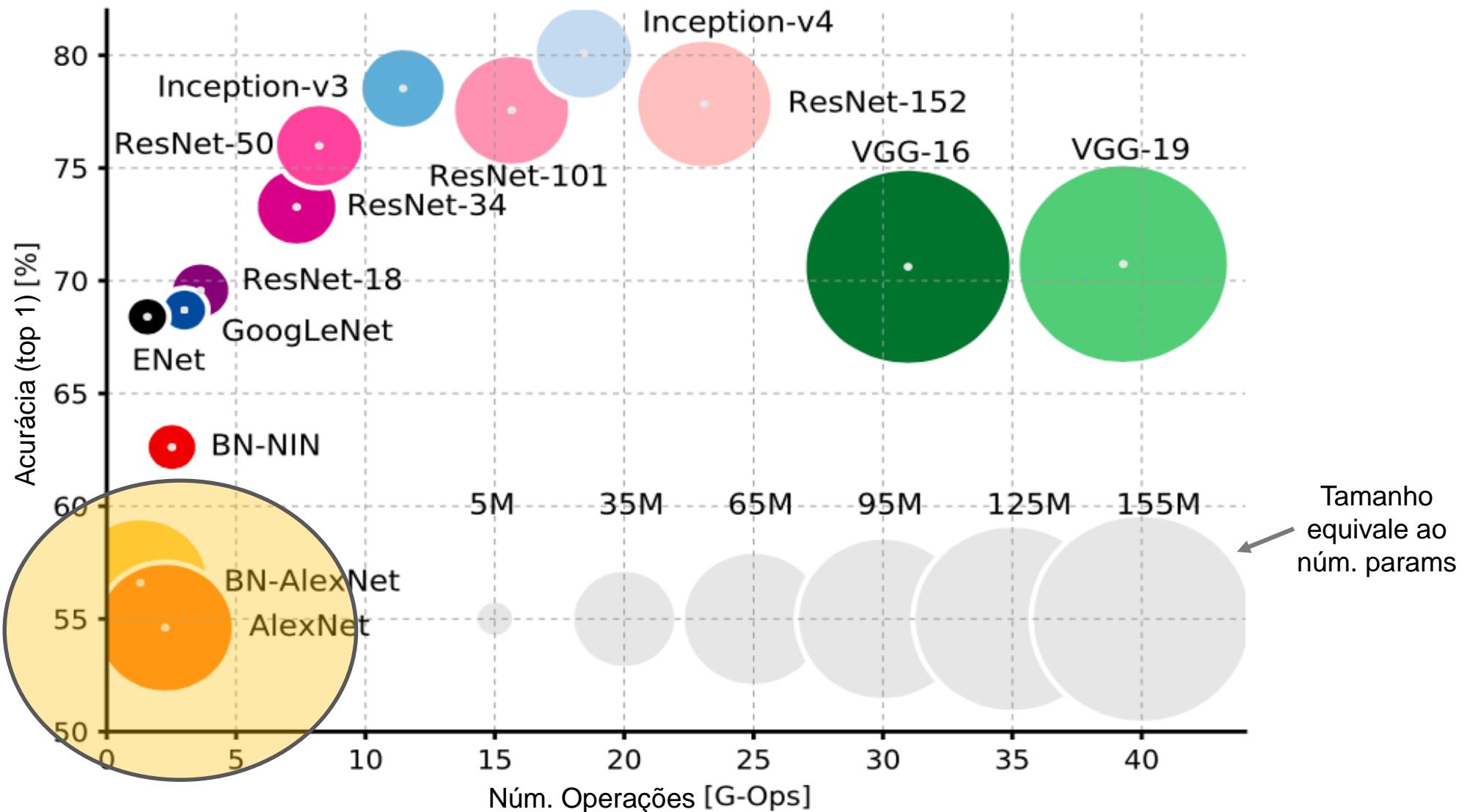
# ConvNets – Análise Comparativa



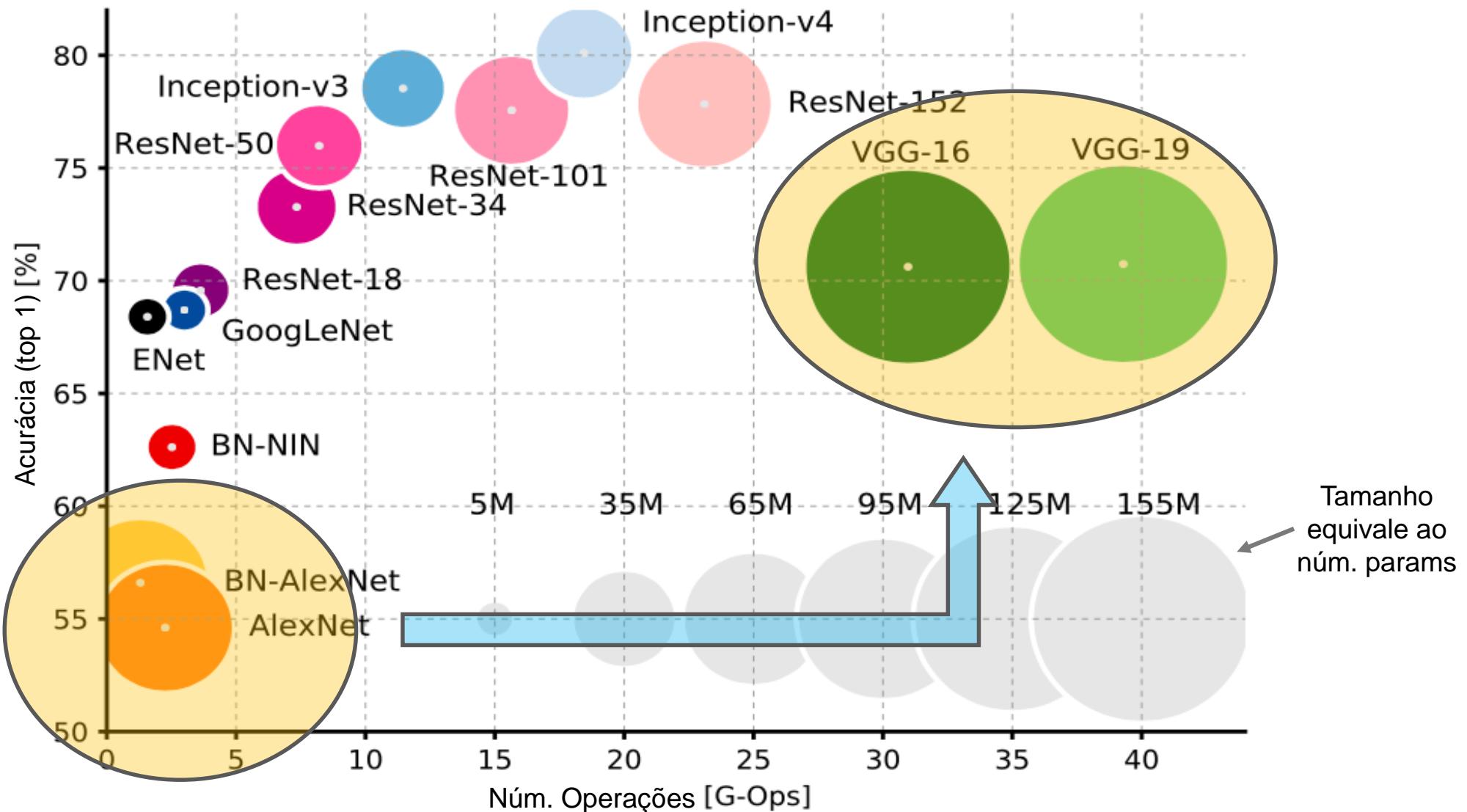
# ConvNets – Análise Comparativa



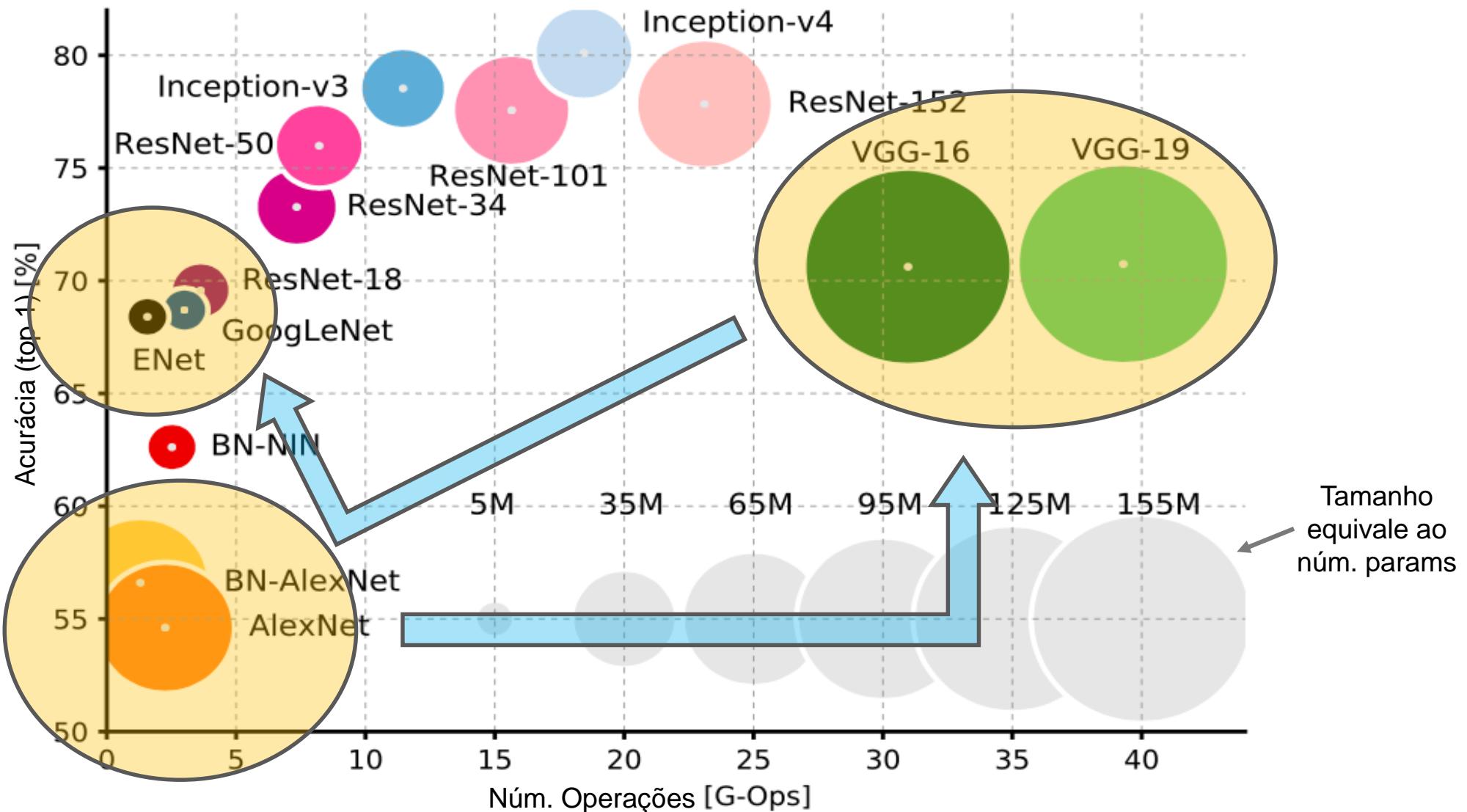
# ConvNets – Análise Comparativa



# ConvNets – Análise Comparativa



# ConvNets – Análise Comparativa



# ConvNets – Análise Comparativa

