# Case Study: AlexNet

# AlexNet
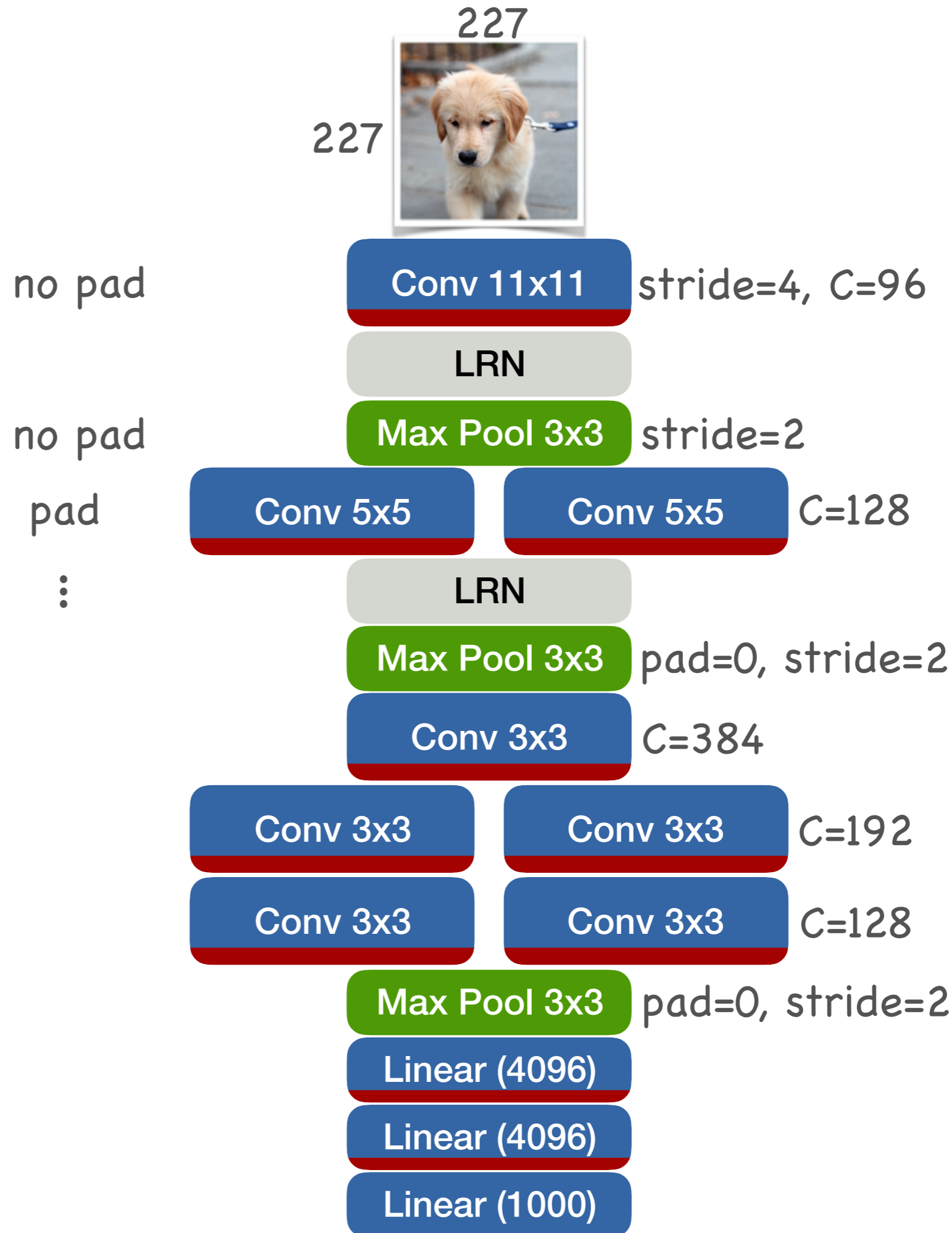
ImageNet challenge

- Won the ImageNet competition 2012

- Start of deep learning revolution in computer vision

Top-5 error

30

22.5

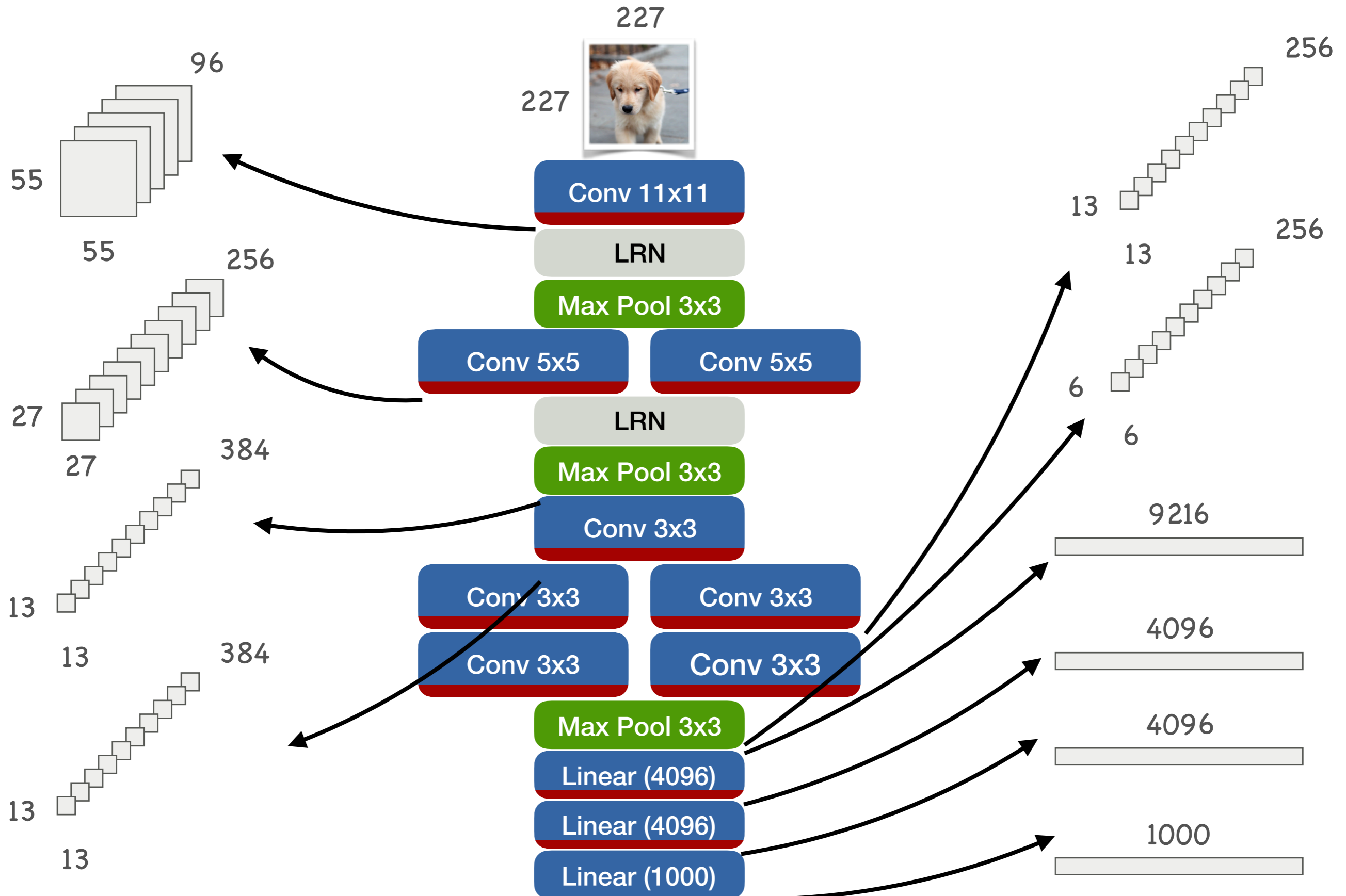15

7.5

0

**Non-deep**

**8 layers**

**19-22 layers**

**152 layers**

2010

2011

2012 (AlexNet)

2013 (ZFNet)

2014 (VGG/GoogLeNet)

2015 (ResNet)
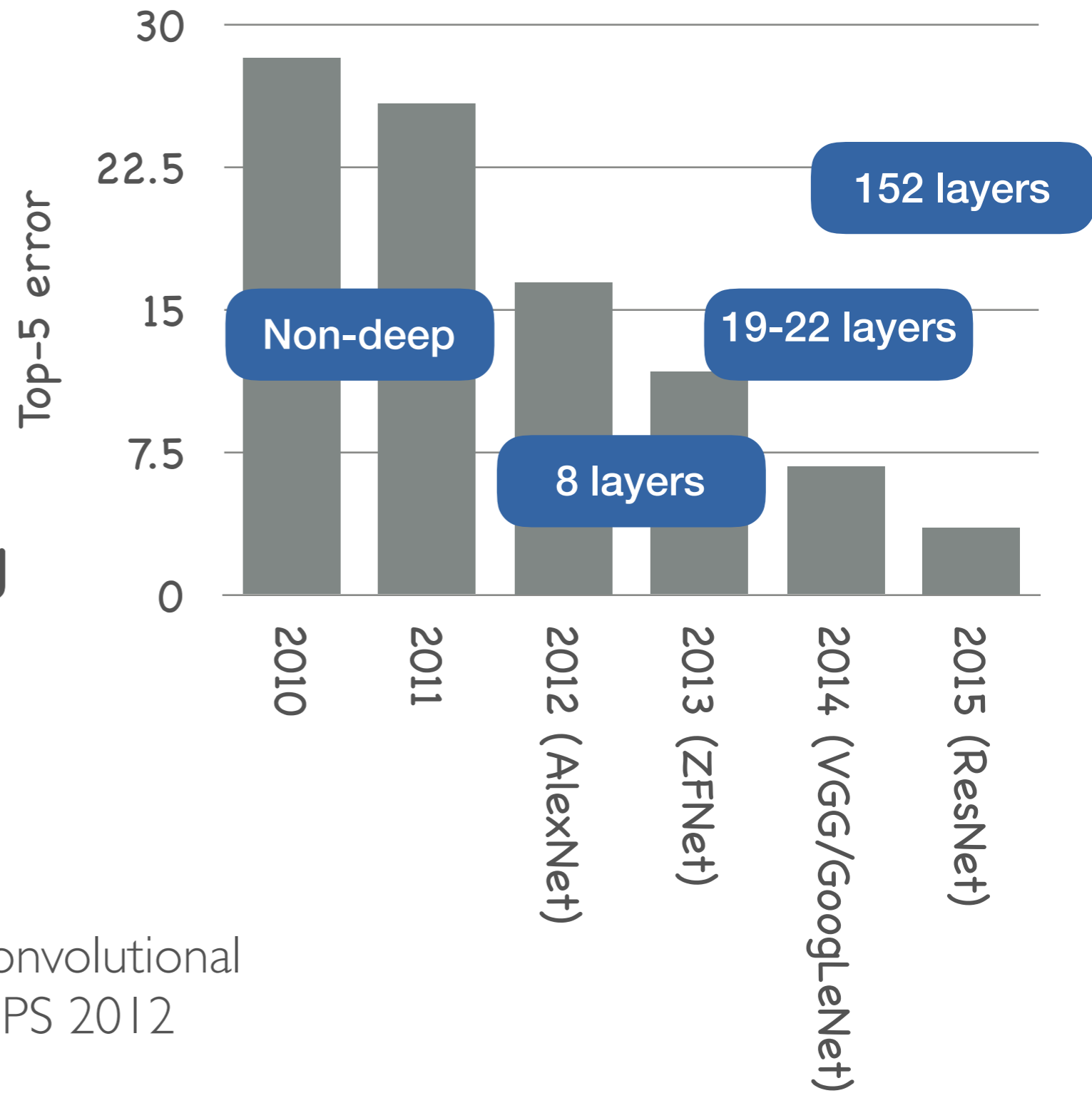
ImageNet Classification with Deep Convolutional Neural Networks, Krizhevsky et al., NIPS 2012

227

227

Conv 11x11 — stride=4, C=96

no pad

LRN

Max Pool 3x3 — stride=2

no pad

Conv 5x5    Conv 5x5 — C=128

pad

⋮

LRN

Max Pool 3x3 — pad=0, stride=2

Conv 3x3 — C=384

Conv 3x3    Conv 3x3 — C=192

Conv 3x3    Conv 3x3 — C=128

Max Pool 3x3 — pad=0, stride=2
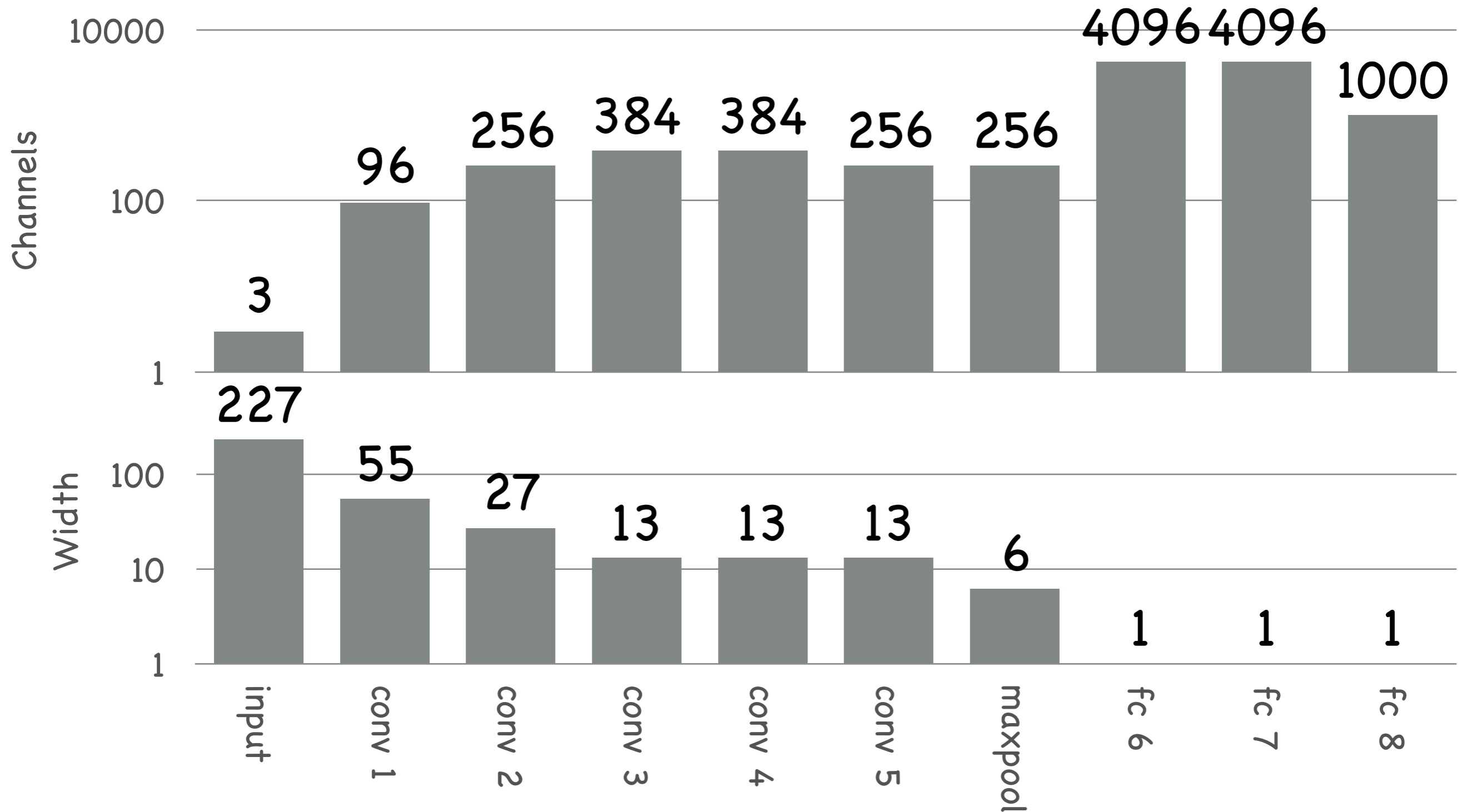
Linear (4096)

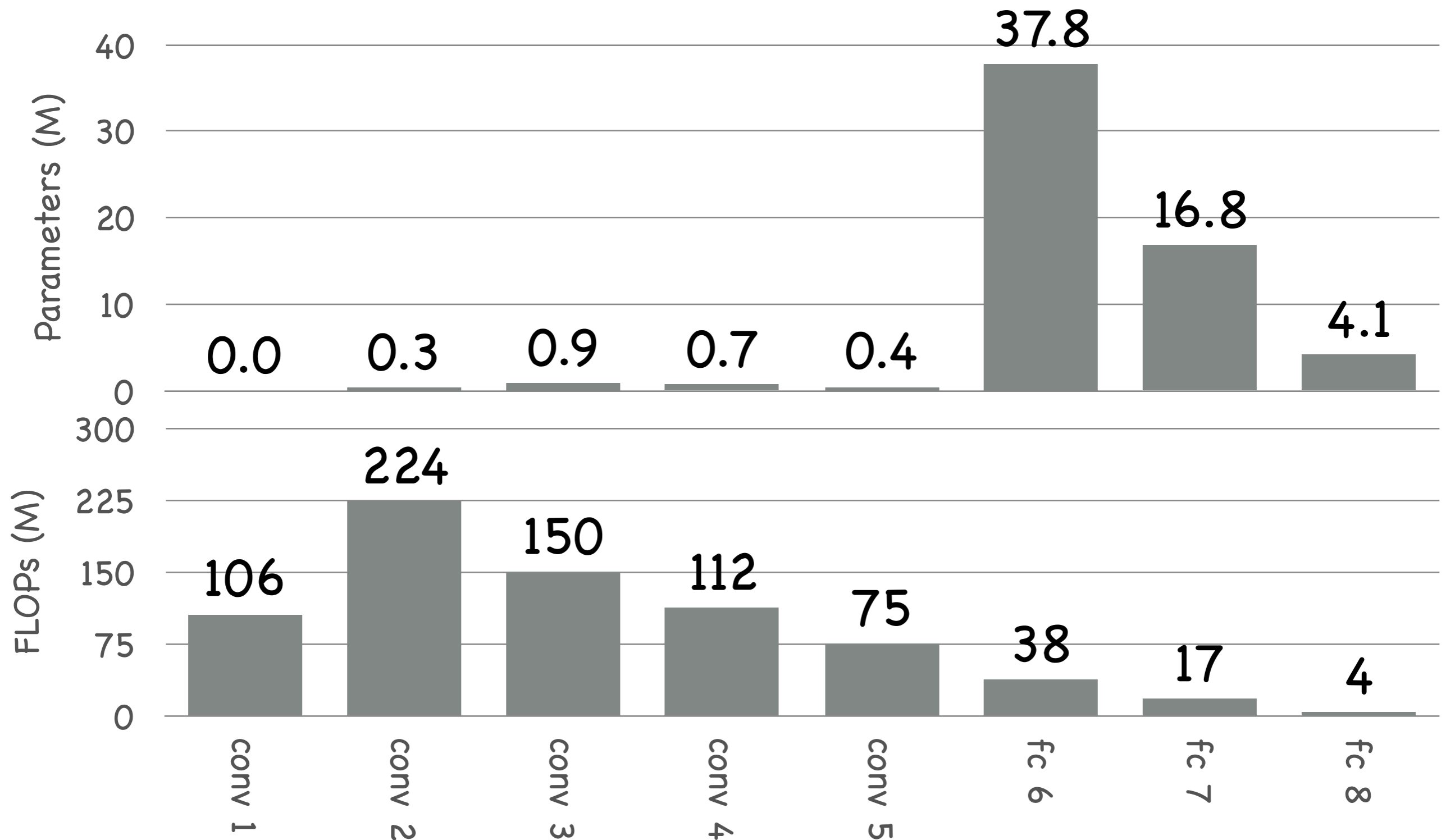Linear (4096)

Linear (1000)

AlexNet

# Activation maps in AlexNet

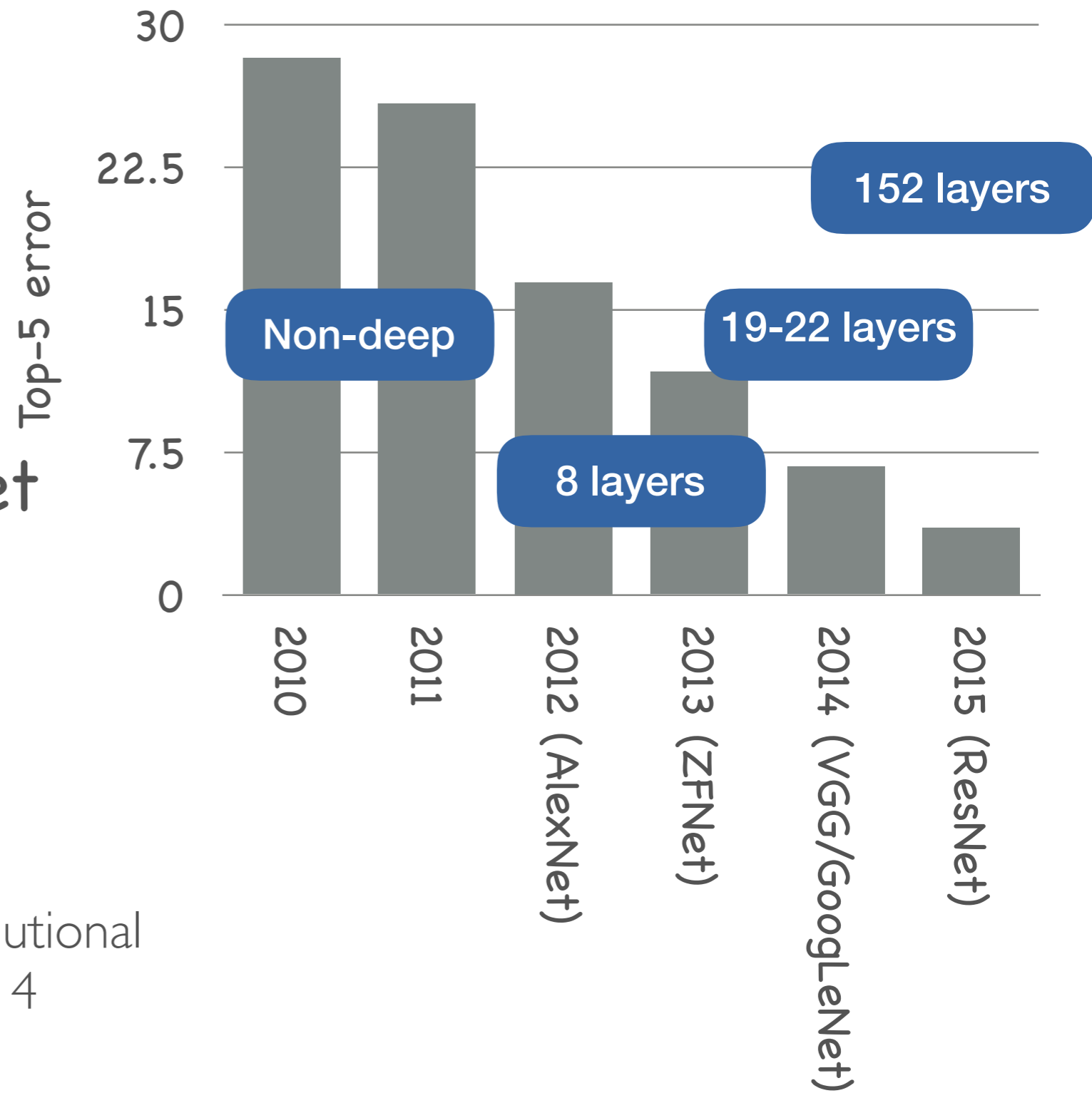# Activation maps in AlexNet

# Parameters and computation

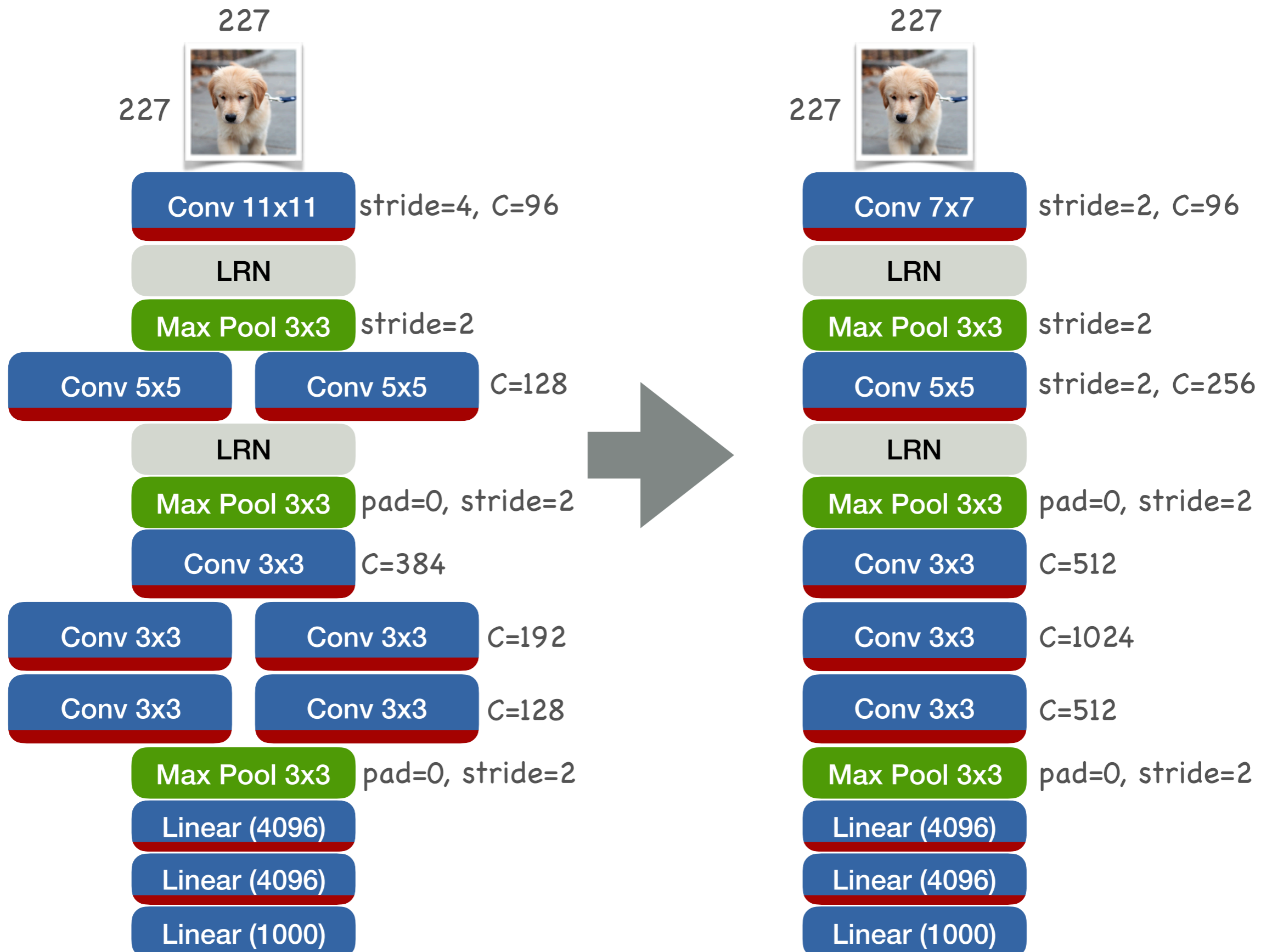# ZFNet

ImageNet challenge

- Won the ImageNet competition 2013

- Nice analysis and visualization of AlexNet
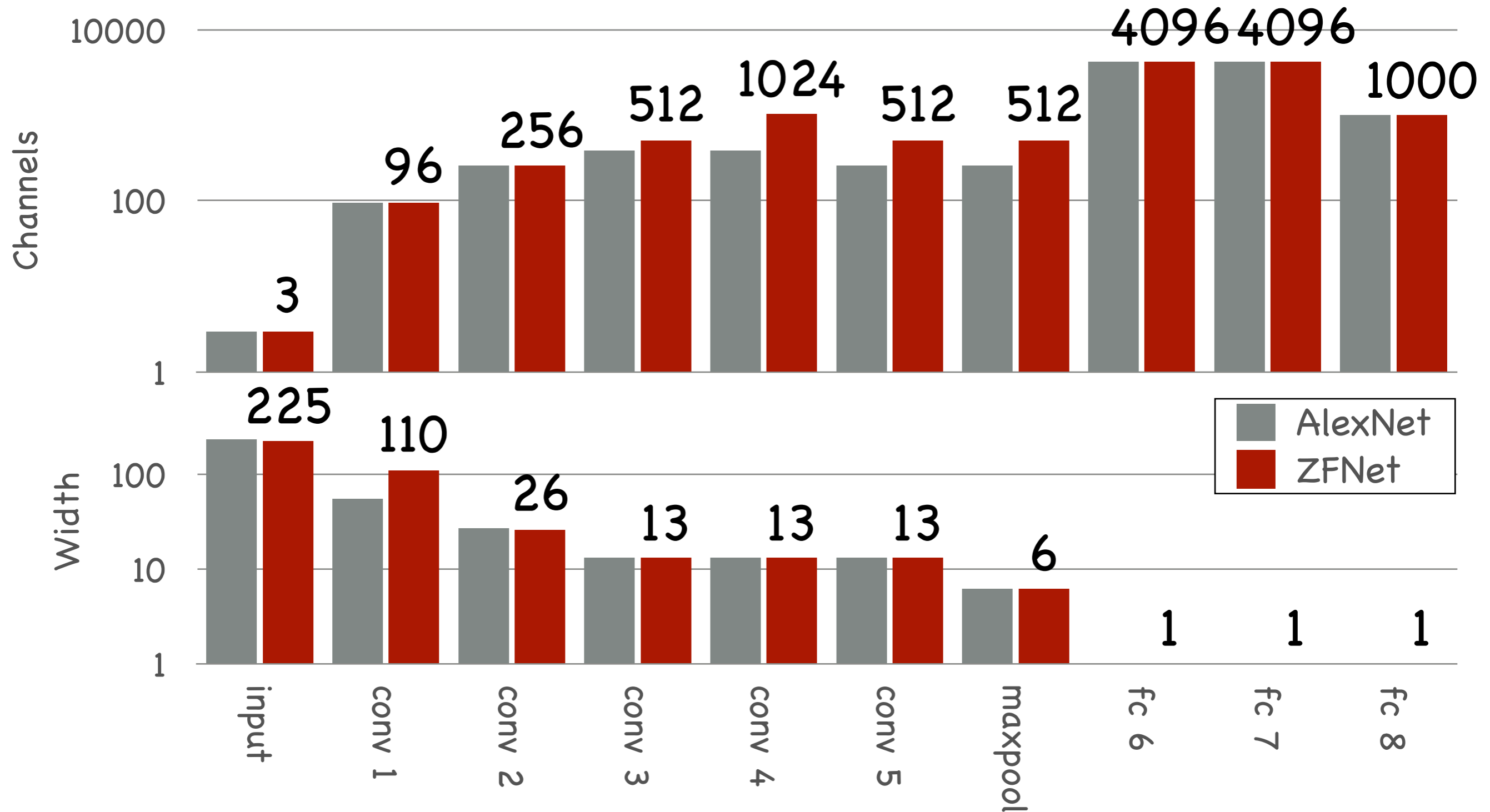
  - Introduced up-conv



Visualizing and Understanding Convolutional Networks, Zeiler & Fergus, ECCV 2014

# AlexNet to ZFNet

227

227

| Conv 11x11 | stride=4, C=96 |
| LRN | |
| Max Pool 3x3 | stride=2 |

| Conv 5x5 | Conv 5x5 | C=128 |

| LRN | |
| Max Pool 3x3 | pad=0, stride=2 |
| Conv 3x3 | C=384 |

| Conv 3x3 | Conv 3x3 | C=192 |
| Conv 3x3 | Conv 3x3 | C=128 |

| Max Pool 3x3 | pad=0, stride=2 |
| Linear (4096) | |
| Linear (4096) | |
| Linear (1000) | |

227

227

| Conv 7x7 | stride=2, C=96 |
| LRN | |
| Max Pool 3x3 | stride=2 |
| Conv 5x5 | stride=2, C=256 |
| LRN | |
| Max Pool 3x3 | pad=0, stride=2 |
| Conv 3x3 | C=512 |
| Conv 3x3 | C=1024 |
| Conv 3x3 | C=512 |
| Max Pool 3x3 | pad=0, stride=2 |
| Linear (4096) | |
| Linear (4096) | |
| Linear (1000) | |

# Parameters and computation