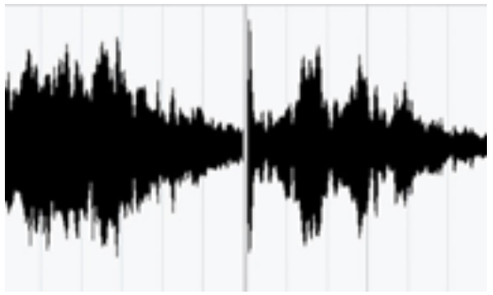
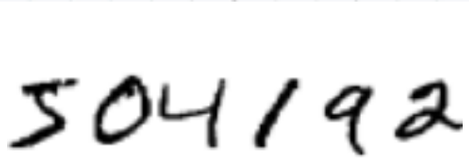
CNN Topologies

Neural Networks

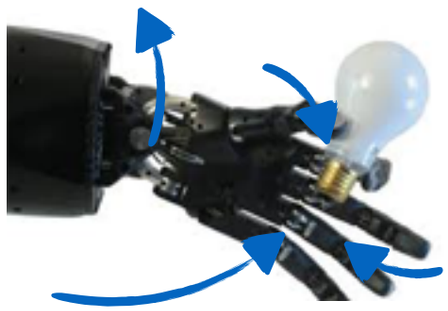
- Learn Mapping Input → Expected Output



→ “Hello World”



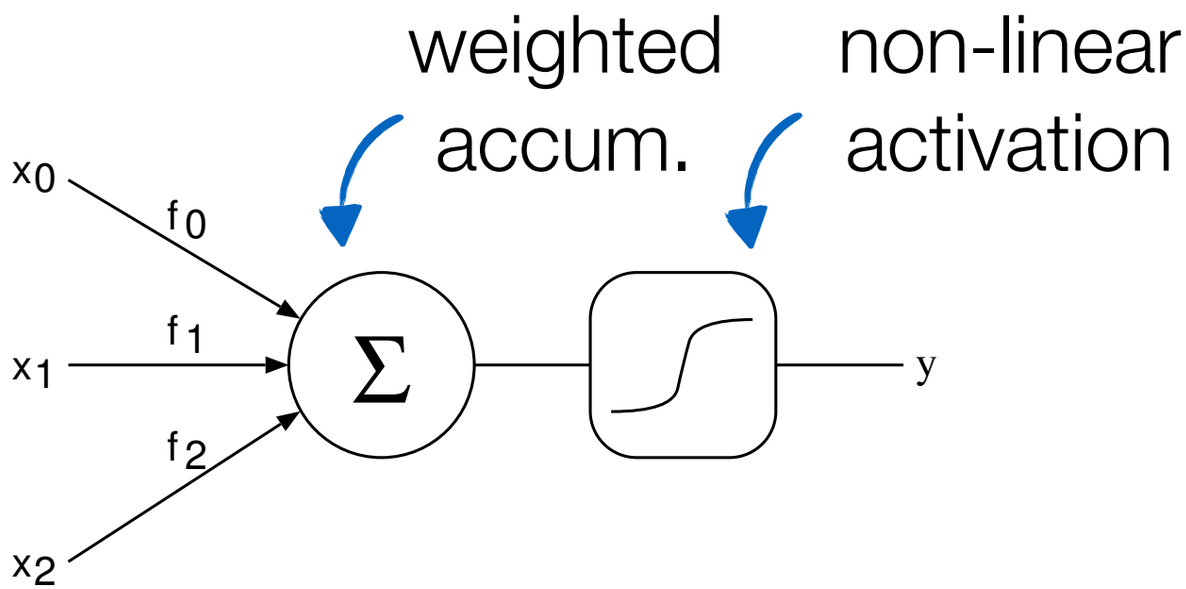
→ 504192

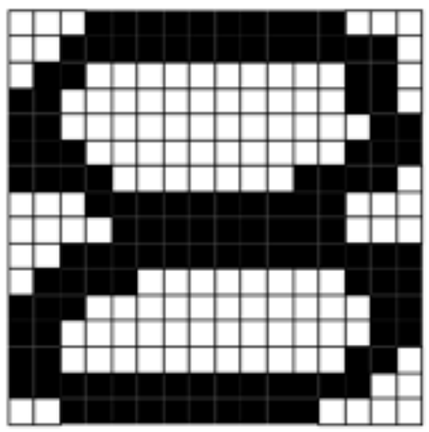


→ *successful gripping trajectory*

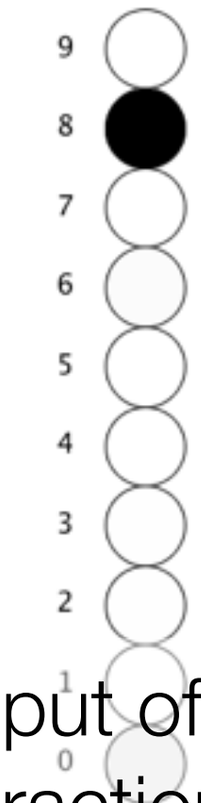
→ *successful game draws*

Neural Networks





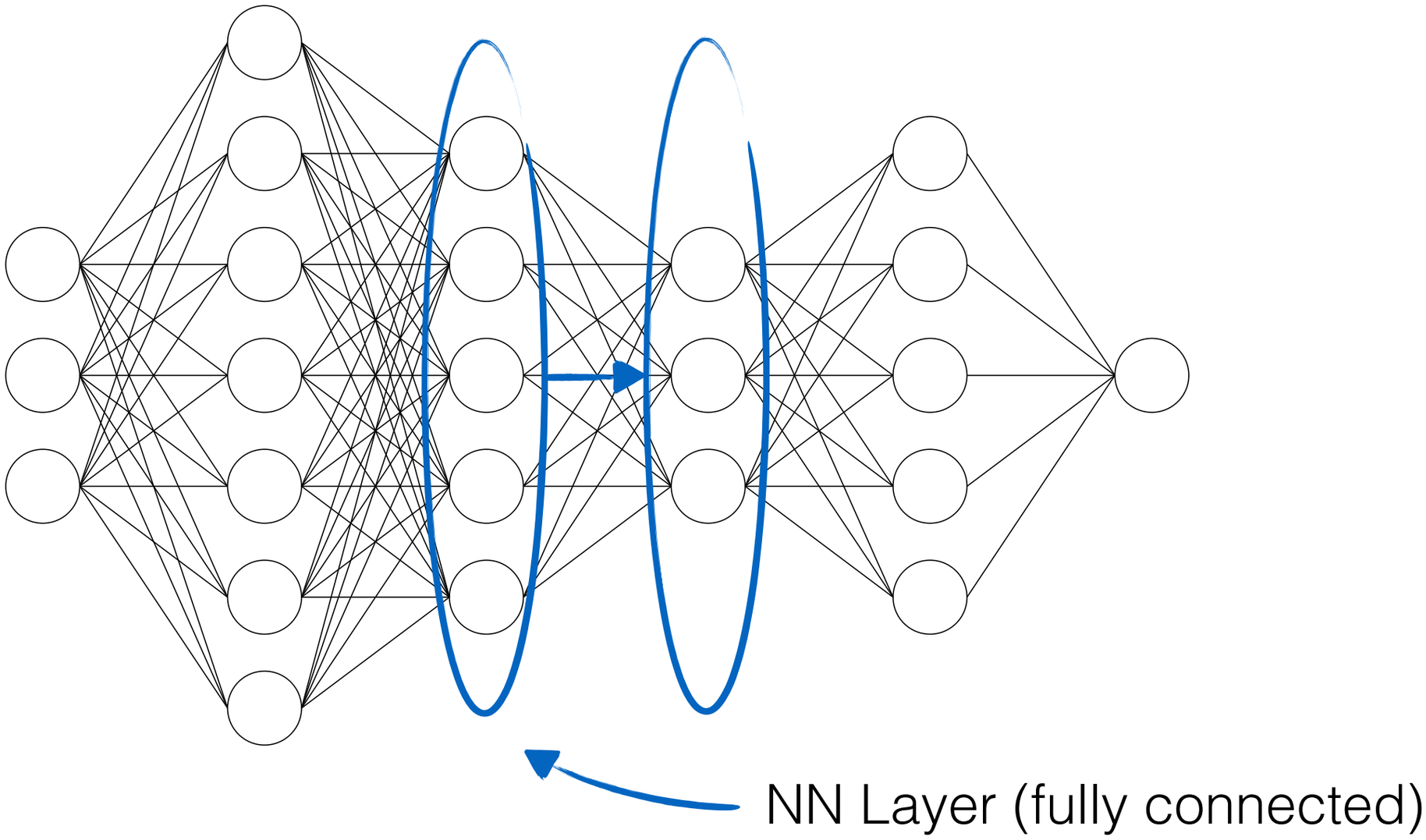
many low-complexity inputs



1 output of higher abstraction level

Neural Networks

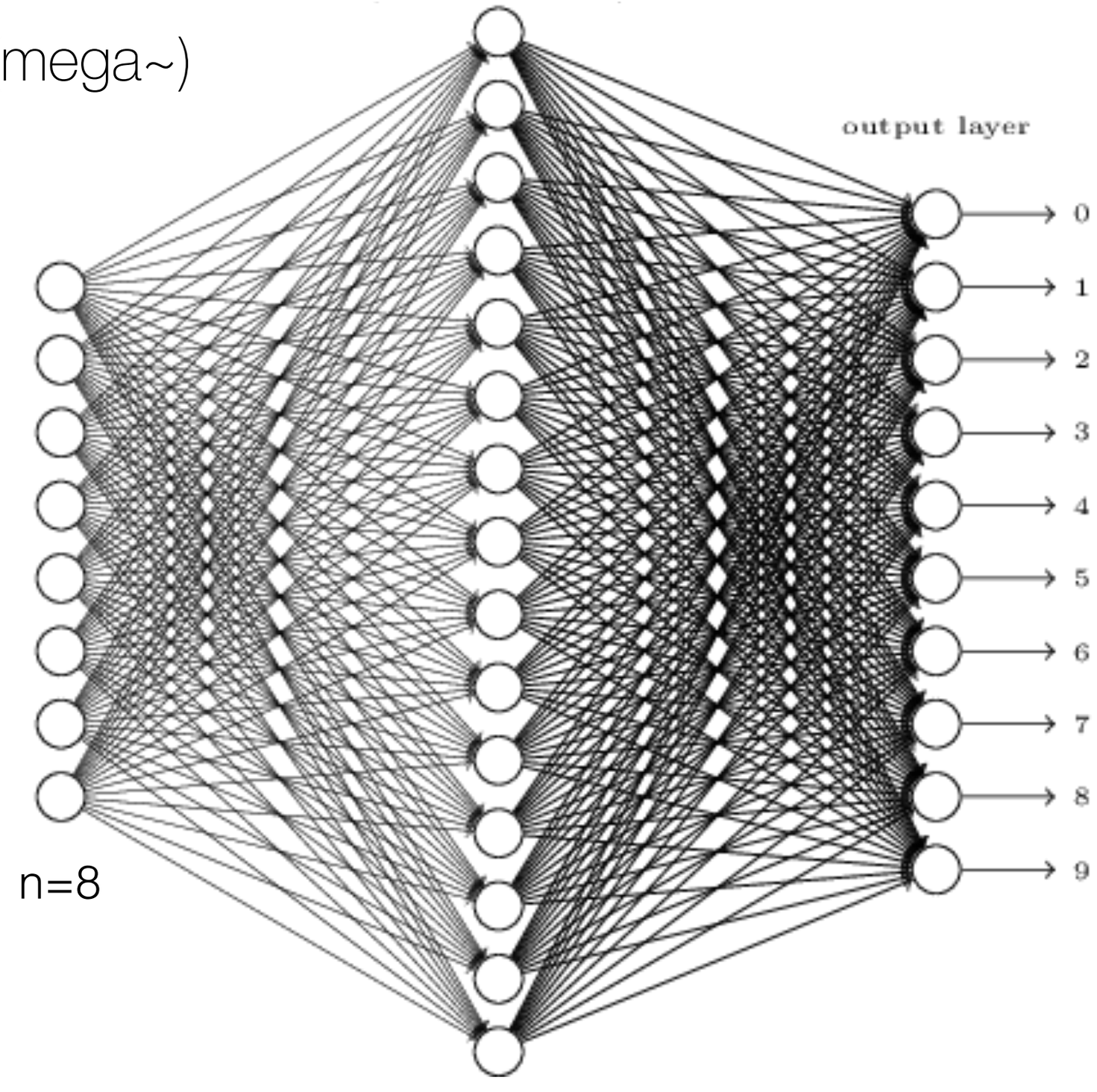
- Neural Network



- Learning = Optimization of Coefficients

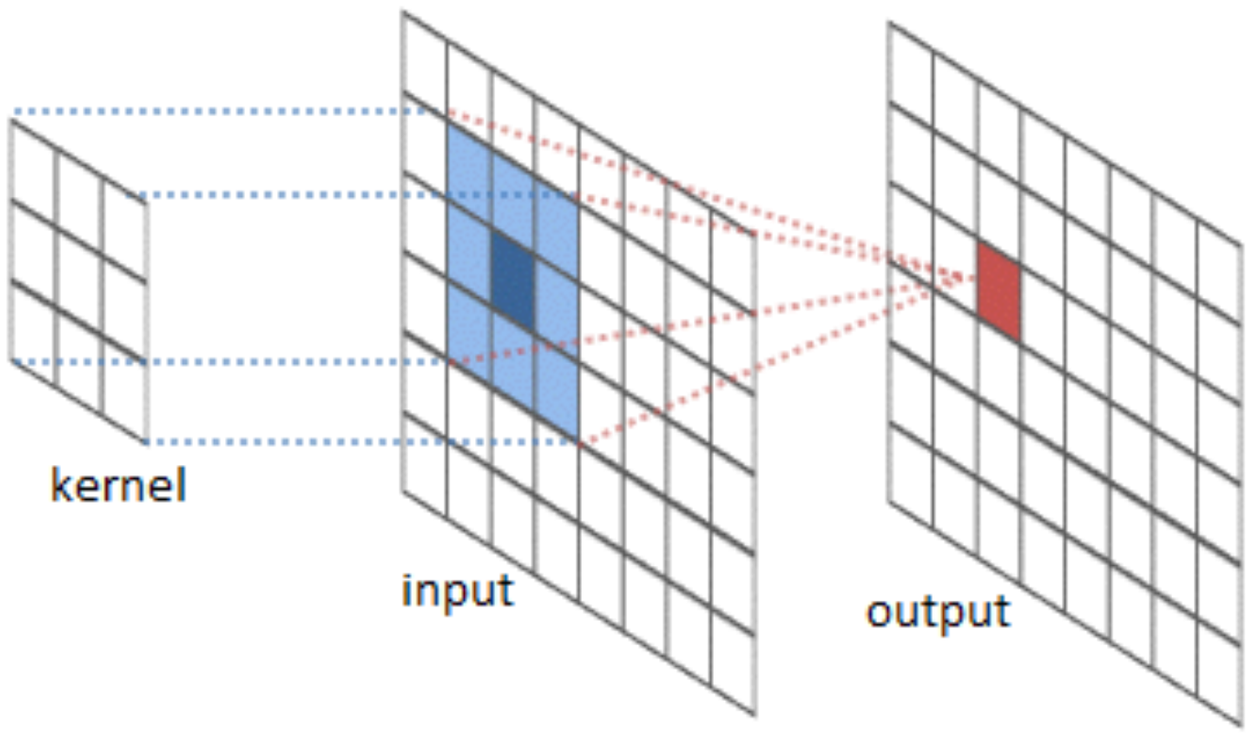
Neural Networks

- images: input = pixels (mega~)
- #coefficients explodes

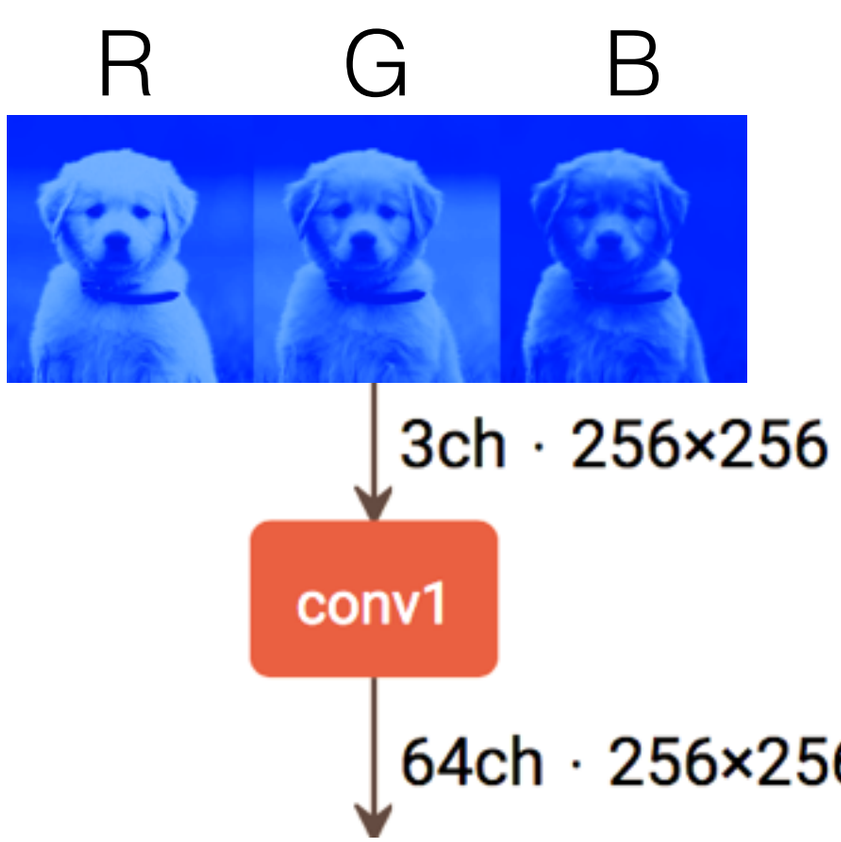


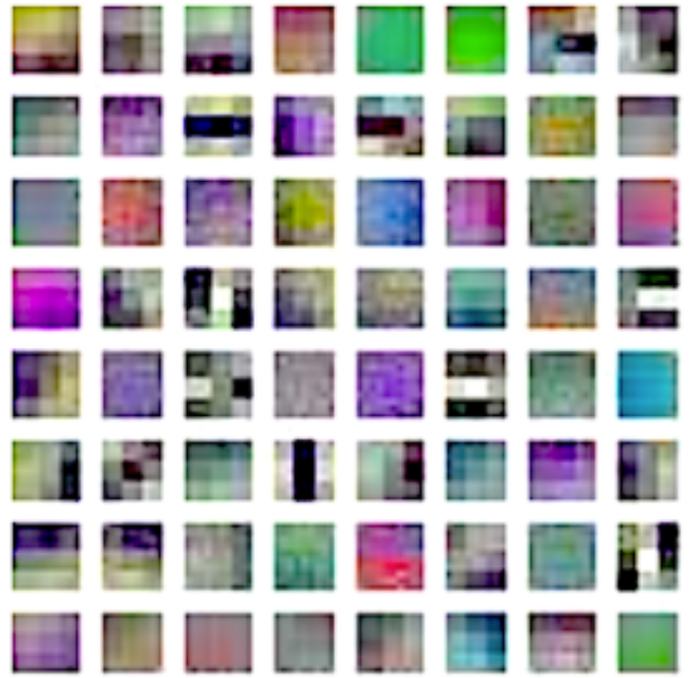
Convolutional Neural Networks

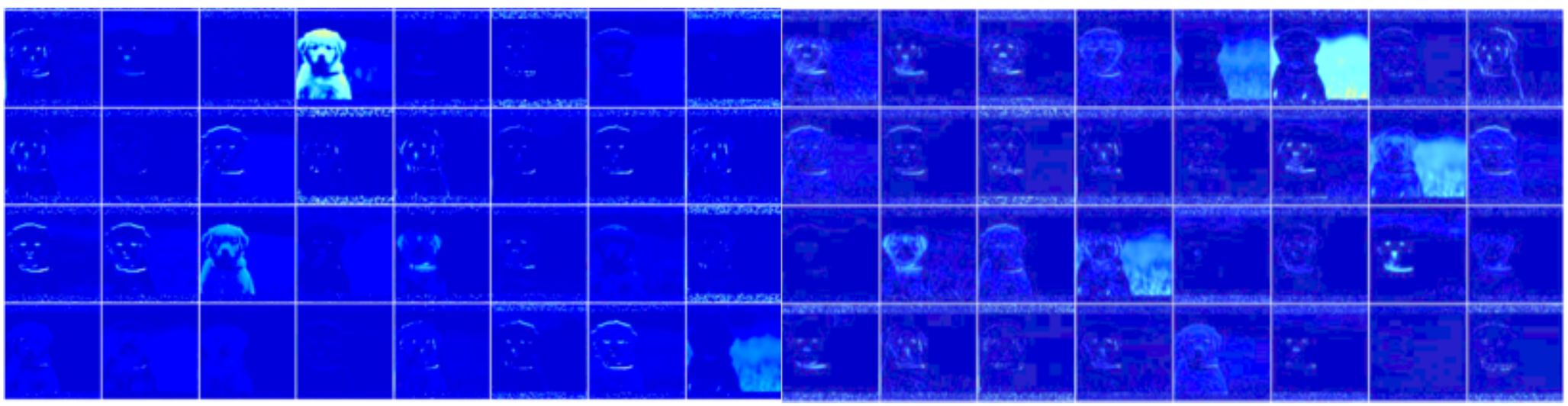
- operation on 2D data (pixels)
- *important*: neighborhood relations (line/edge/circle/eye/…)
- *not important*: color of very distant pixels
- “weighting” by filter kernels → weight sharing



Convolutional Layer







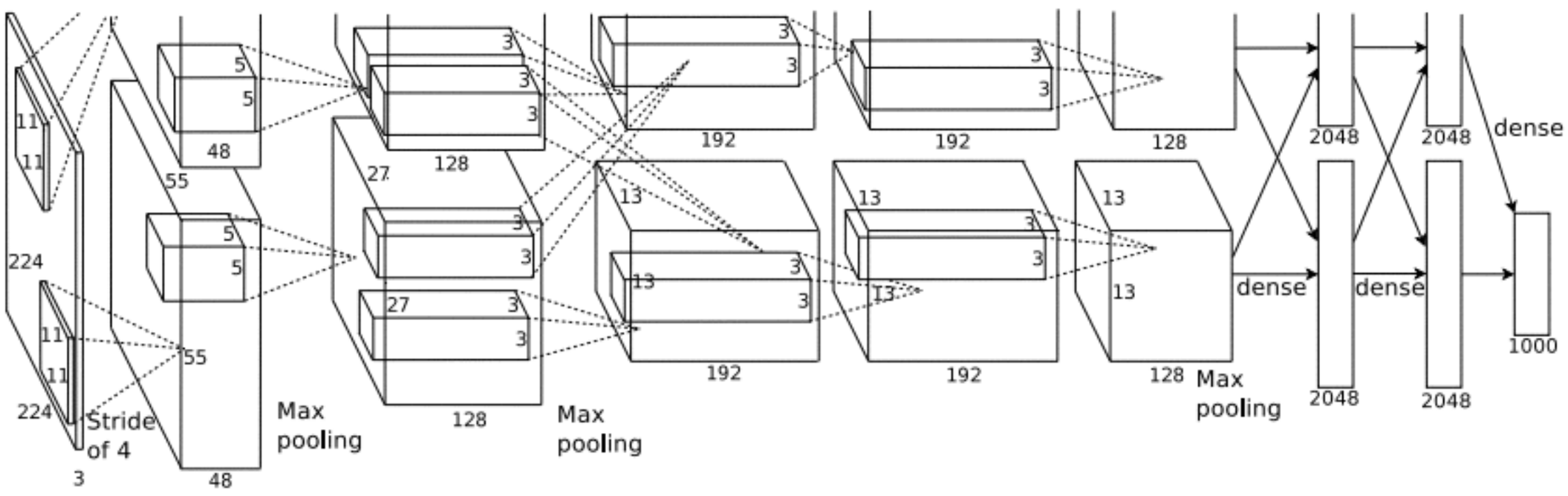
Topologies: Overview

ImageNet Data (1.2mio hi-res images, 1000 classes)

	# layers	# parameters	MACC	top-5 accuracy
AlexNet	5	60 M	1’100 M	80 %
VGG-16	16	140 M	16’000 M	90 %
GoogLeNet	> 22	7 M	1’600 M	89 %
SqueezeNet	18	1.2 M	860 M	79 %
SqueezeNet+	18	2.5 M	570 M	85 %

	# layers	# parameters	MACC	top-5 acc.
AlexNet	**5**	**60 M**	**1’100 M**	**80 %**
VGG-16	16	140 M	16’000 M	90 %
GoogLeNet	> 22	7 M	1’600 M	89 %
SqueezeNet	18	1.2 M	860 M	79 %
SqueezeNet+	18	2.5 M	570 M	85 %

Alex Krizhevsky et. al., U Toronto, 2012

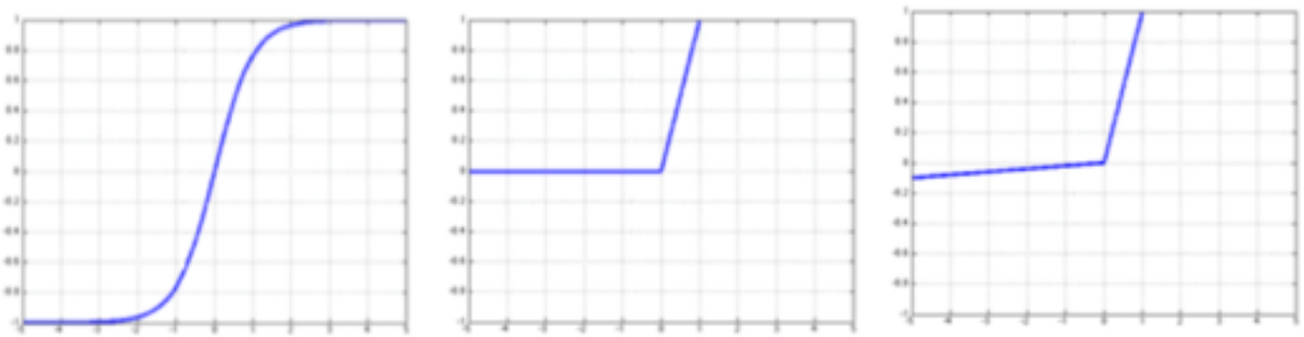


(re-)started the boom; aka. CaffeNet

	# layers	# parameters	MACC	top-5 acc.
AlexNet	**5**	**60 M**	**1’100 M**	**80 %**
VGG-16	16	140 M	16’000 M	90 %
GoogLeNet	> 22	7 M	1’600 M	89 %
SqueezeNet	18	1.2 M	860 M	79 %
SqueezeNet+	18	2.5 M	570 M	85 %

- **CONV**
  - Kernel
  - Stride
  - Pad
- **Non-Linearity**
- **Pooling**
- **LRN**

-> Netscope AlexNet



tanh ReLU pReLU

	# layers	# parameters	MACC	top-5 acc.
AlexNet	5	60 M	1’100 M	80 %
VGG-16	**16**	**140 M**	**16’000 M**	**90 %**
GoogLeNet	> 22	7 M	1’600 M	89 %
SqueezeNet	18	1.2 M	860 M	79 %
SqueezeNet+	18	2.5 M	570 M	85 %

Visual Geometry Group, Oxford, 2014

-> Netscope VGG-16

very deep
regular (3x3 Conv, 2x2 Maxpool)
new record

	# layers	# parameters	MACC	top-5 acc.
AlexNet	5	60 M	1’100 M	80 %
VGG-16	16	140 M	16’000 M	90 %
GoogLeNet	**> 22**	**7 M**	**1’600 M**	**89 %**
SqueezeNet	18	1.2 M	860 M	79 %
SqueezeNet+	18	2.5 M	570 M	85 %

Szegedy et al., Google, 2014

-> Netscope GoogLeNet

Inception Module: Network-in-Network
(more non-linearity, less parameters)

CONV 1x1, 3x3, 5x5 in parallel
reduce-expand architecture

LRN, hard to train

	# layers	# parameters	MACC	top-5 acc.
AlexNet	5	60 M	1’100 M	80 %
VGG-16	16	140 M	16’000 M	90 %
GoogLeNet	> 22	7 M	1’600 M	89 %
SqueezeNet	**18**	**1.2 M**	**860 M**	**79 %**
SqueezeNet+	18	2.5 M	570 M	85 %

Iandola et al., Berkeley, 2016

-> Netscope SqueezeNet

Reduce-Expand Modules

very small model
1/50 of AlexNet @ same accuracy

	# layers	# parameters	MACC	top-5 acc.
AlexNet	5	60 M	1’100 M	80 %
VGG-16	16	140 M	16’000 M	90 %
GoogLeNet	> 22	7 M	1’600 M	89 %
SqueezeNet	18	1.2 M	860 M	79 %
SqueezeNet+	**18**	**2.5 M**	**570 M**	**85 %**

Gschwend et al., SCS, 2016

-> Netscope SqueezeNet+ b2a_ext3

no Pooling
only 1x1, 3x3 convolution
very regular