Budapest University of Technology and Economics
Department of Artificial Intelligence and Systems Engineering

Artificial intelligence – VIMIAC16-EN, VIMIAC10

2024 Fall Semester

Dr. Gábor Hullám

Deep neural networks

- In the mid-2000s, there were hardly any articles on neural networks that were accepted
- 2004 – G. Hinton - CIFAR (Canadian Institute for Advanced Research)
  - New packaging for neural networks: deep learning
- 2006 - Hinton, Osindero, Yee-Whye Teh: A fast learning algorithm for deep belief nets
  - New breakthrough: teaching deep belief nets layer by layer,
  - Restarts research

Deep learning – ImageNet

2009 - J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei:

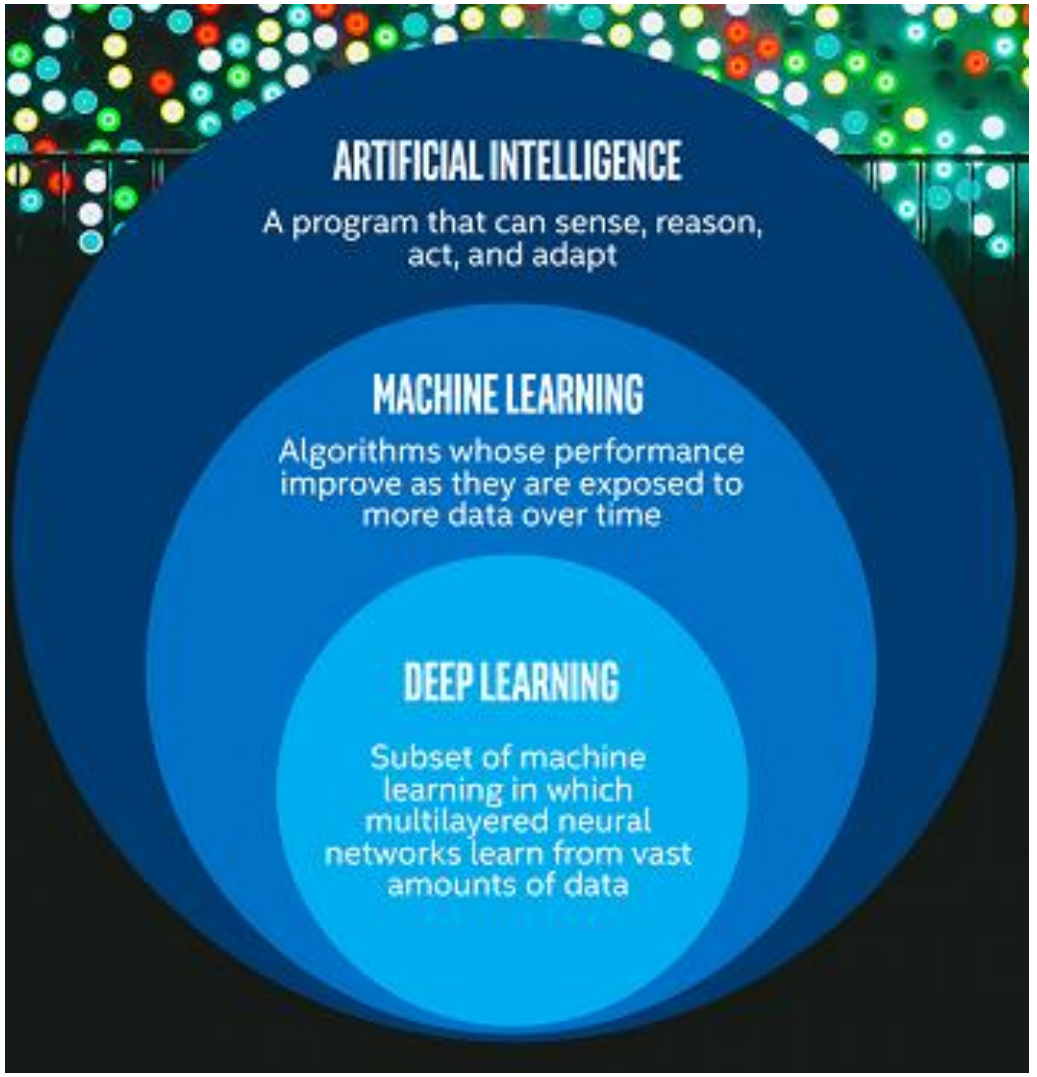
ImageNet: A Large-Scale Hierarchical Image Database

- high-resolution images, 14M images, 10k+ categories
  http://www.image-net.org/about-stats

Deep learning - ILSVRC

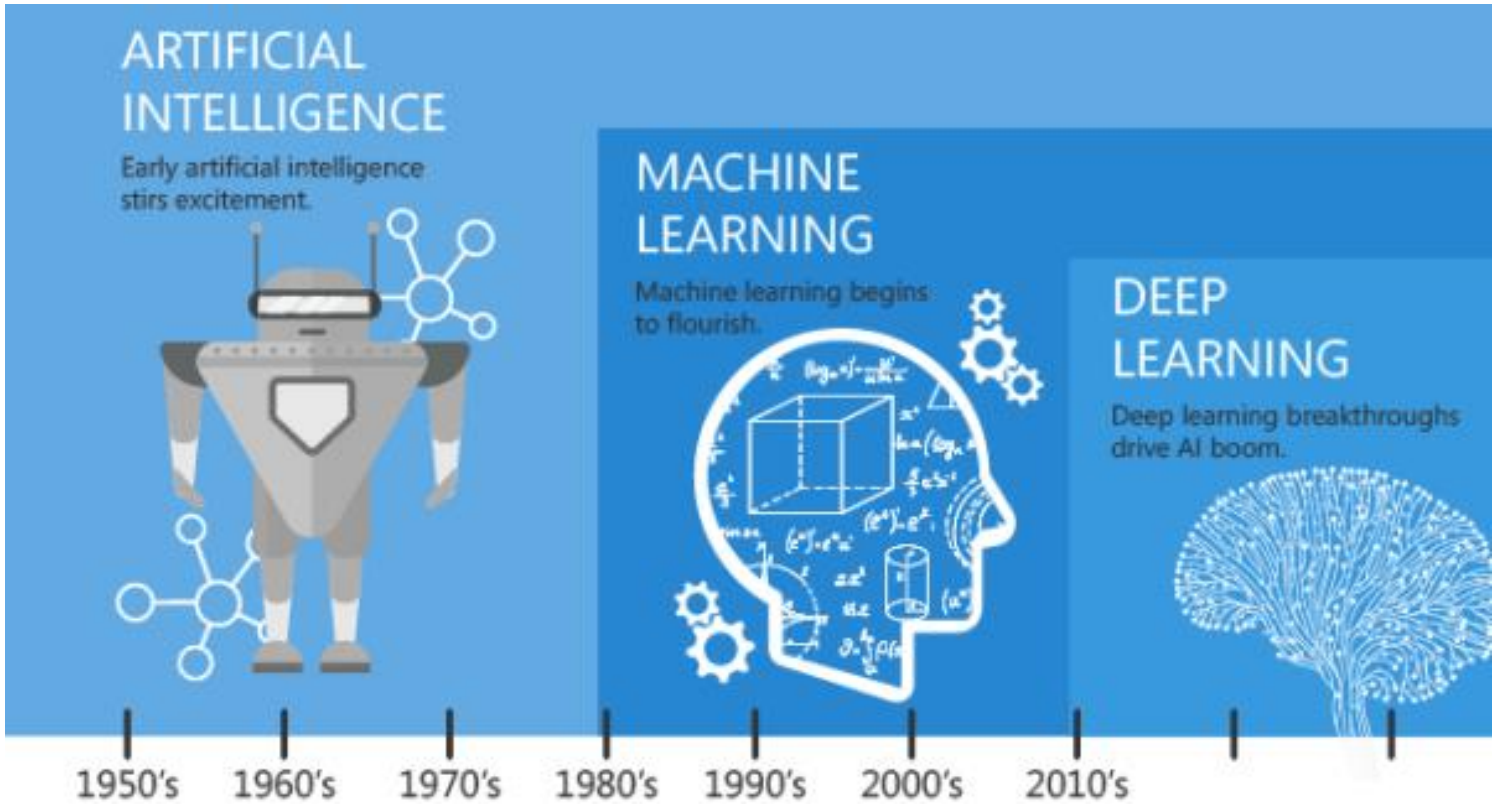
- 2012 - Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton: Imagenet classification with deep convolutional neural networks
  - This is the first purely neural network solution to achieve the best result of the year (SVM until then)
  - From there, deep learning becomes mainstream

AI – ML - DL



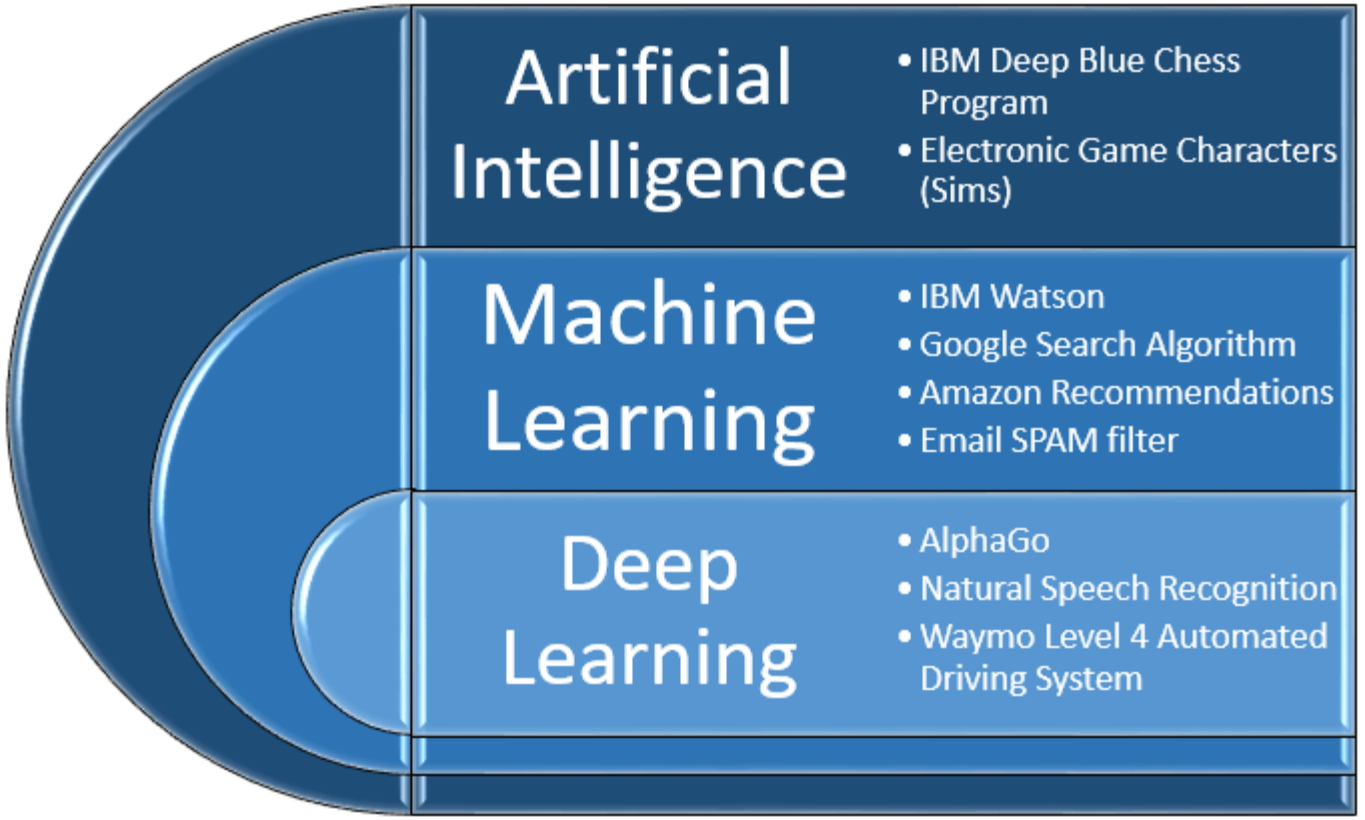
https://steemkr.com/technology/@neurallearner/ai-vs-machine-learning-vs-deep-learning

AI – ML - DL



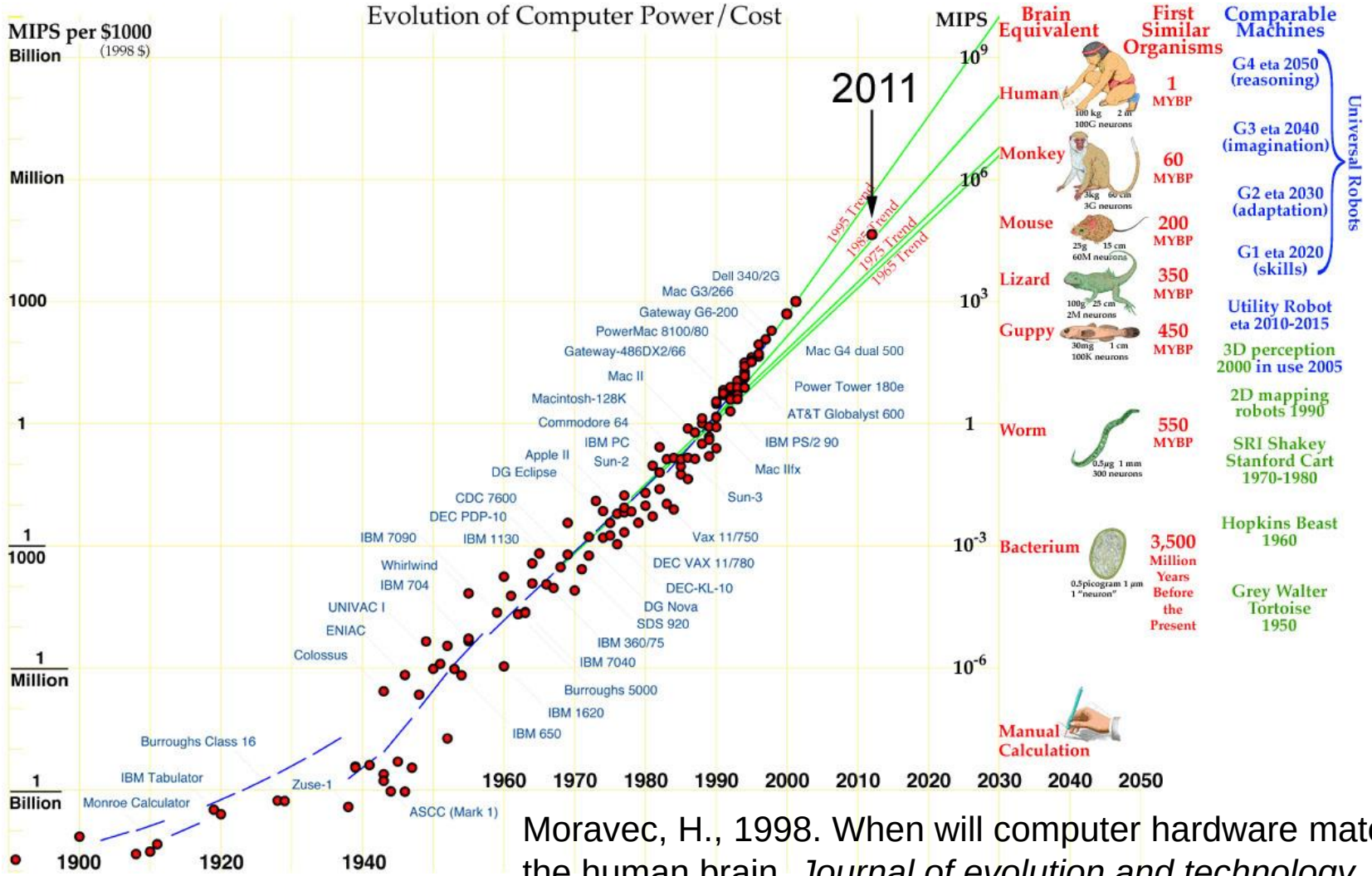
https://shawnennis.com/the-technology-of-machine-learning-with-ai

AI – ML - DL



https://www.cargroup.org/behind-headlines-artificial-intelligence-challenges-using-ai-automotive-industry/

Background requirement: compute power



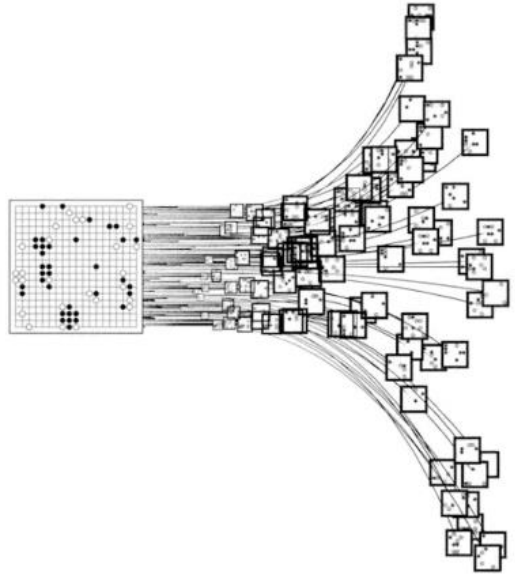
Moravec, H., 1998. When will computer hardware match the human brain. *Journal of evolution and technology*, *1*(1), p.10.

Deep reinforcement learning

- 2013 – DeepMind: Playing Atari with Deep Reinforcement Learning
- 2016 – DeepMind: AlphaGo beats Go World Champion Lee Sedol

Go

- Google DeepMind
- Monte Carlo tree search
- 2016: 9 dan
- 2017: wins against human champion



AlphaGo

ARTICLE

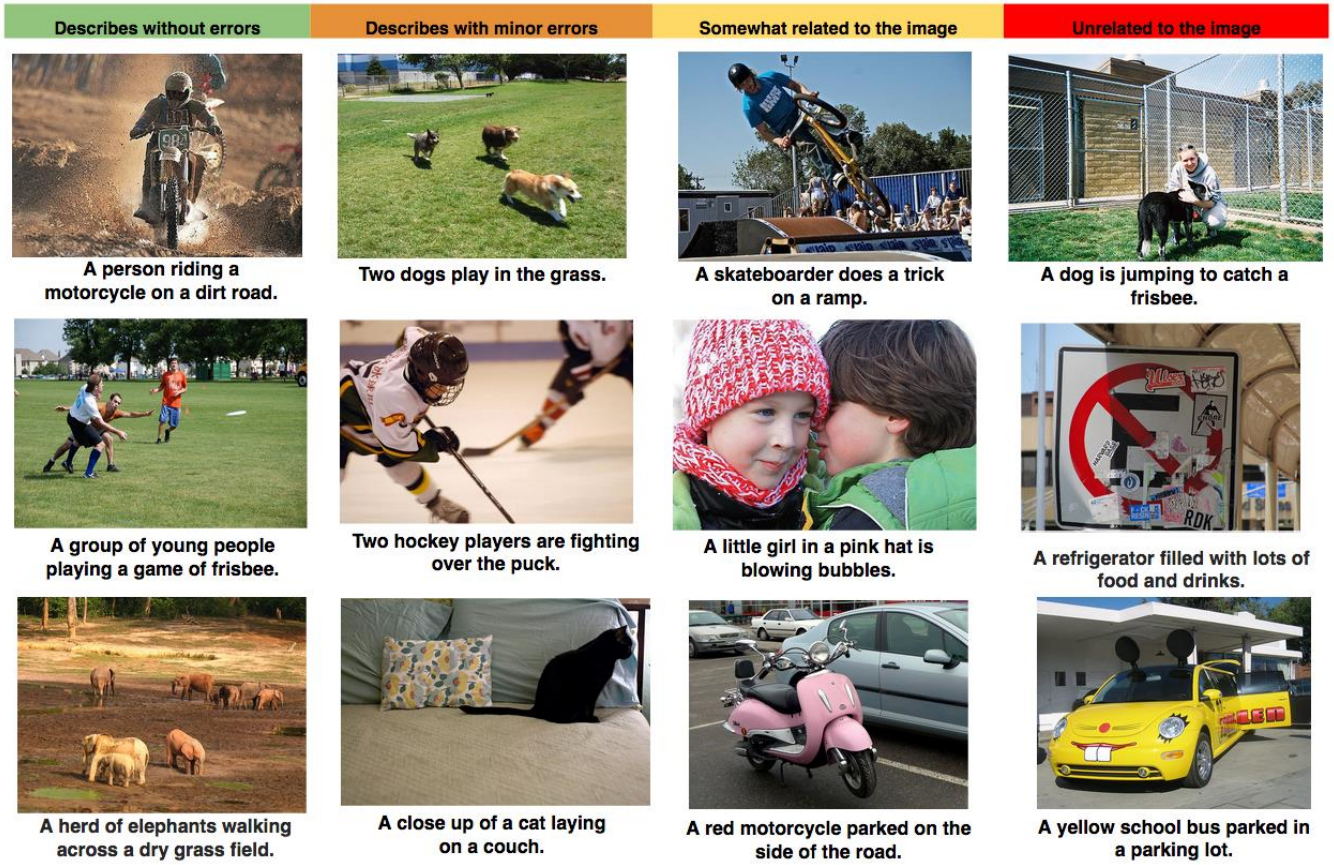
doi:10.1038/nature16961

Mastering the game of Go with deep neural networks and tree search

David Silver[1]*, Aja Huang[1]*, Chris J. Maddison[1], Arthur Guez[1], Laurent Sifre[1], George van den Driessche[1], Julian Schrittwieser[1], Ioannis Antonoglou[1], Veda Panneershelvam[1], Marc Lanctot[1], Sander Dieleman[1], Dominik Grewe[1], John Nham[2], Nal Kalchbrenner[1], Ilya Sutskever[2], Timothy Lillicrap[1], Madeleine Leach[1], Koray Kavukcuoglu[1], Thore Graepel[1] & Demis Hassabis[1]

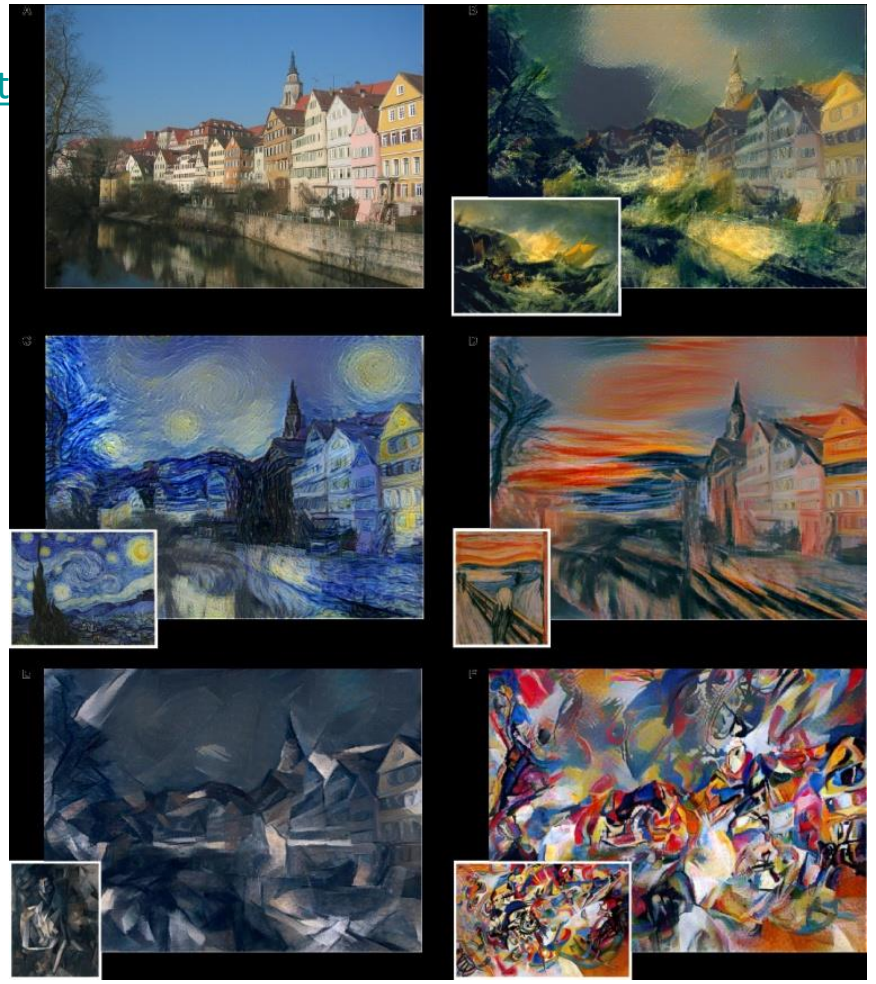
Deep learning - outcomes

- Caption generation for images
  - A. Karpathy, L. Fei-Fei: „Deep Visual-Semantic Alignments for Generating Image Descriptions”
  - https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html



Deep learning - outcomes

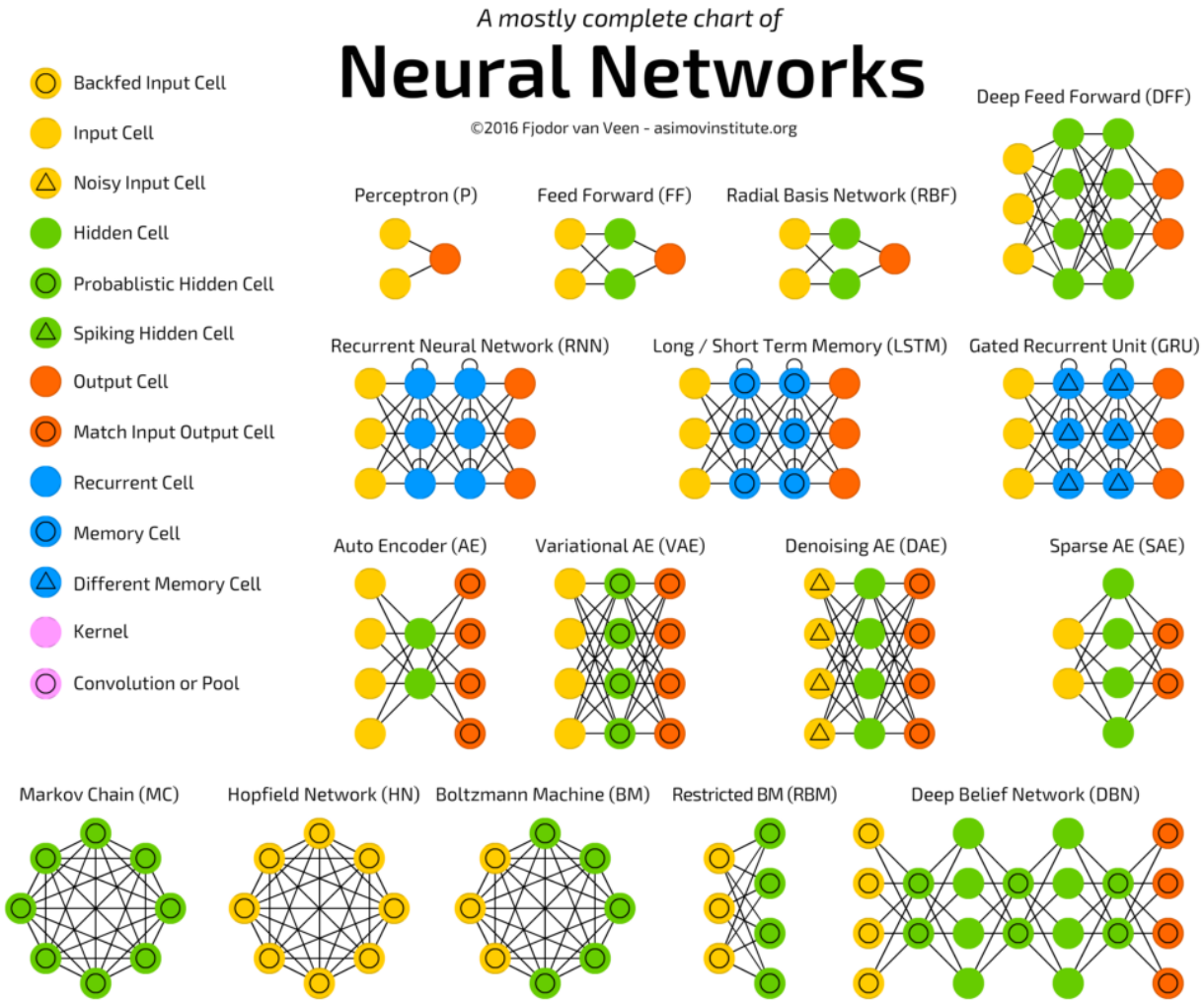
- Style ‘learning’
  - https://github.com/jcjohnson/neural-st
    https://imgur.com/gallery/4LTaQ



Deep learning - outcomes

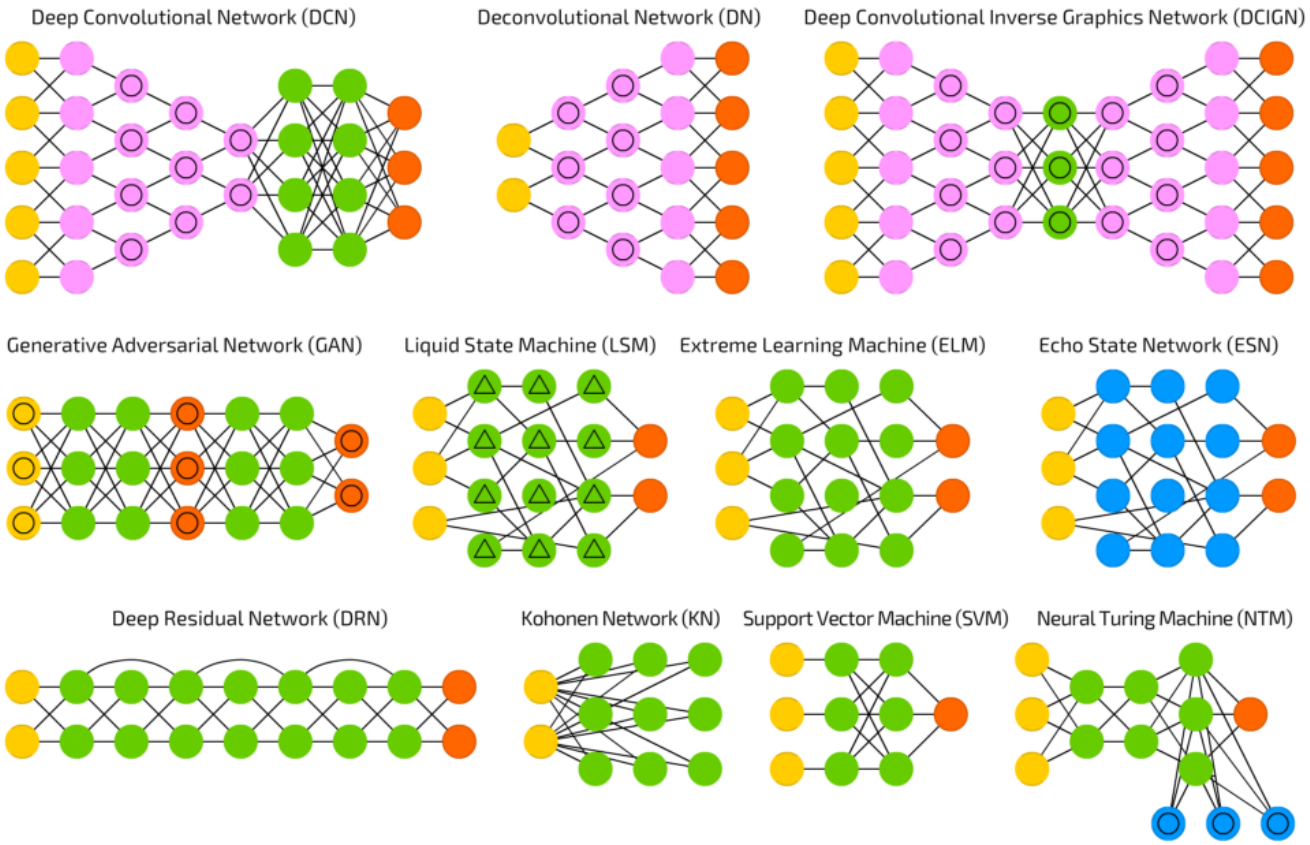
- Comprehension
  - Xiong, W. et al. (2016). Achieving human parity in conversational speech recognition
- Speech generation
  - Oord, Aaron van den, et al. "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499 (2016).
- Image classification
  - C. Szegedy et al. (2014): GoogLeNet – „Going Deeper with Convolutions”

Types of neural networks



https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464

Types of neural networks 2.

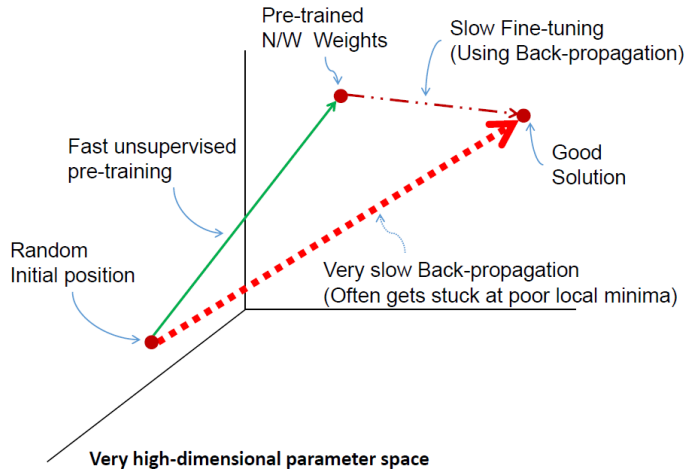


https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464

What makes deep neural networks "better"?

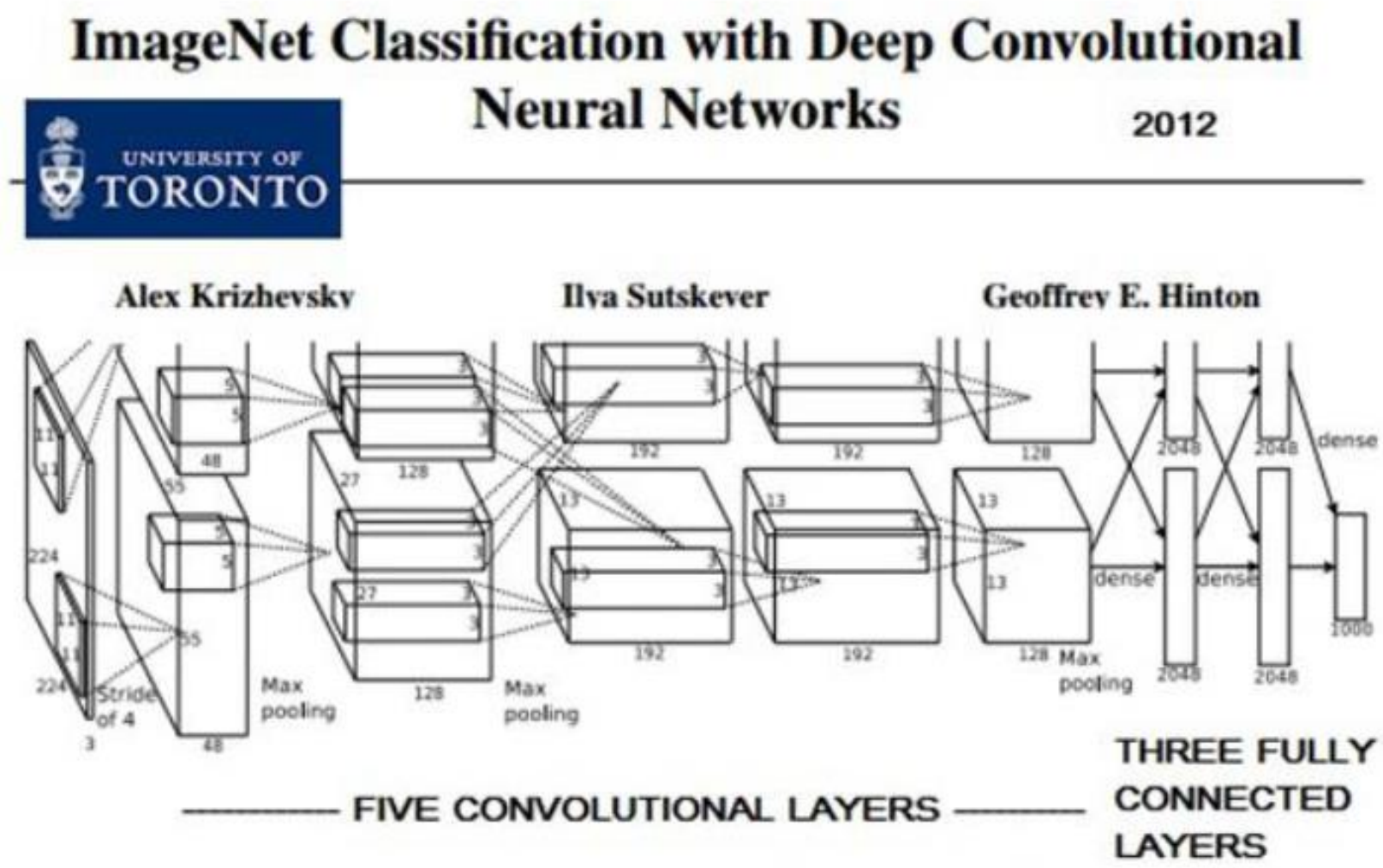
- Humans also detect features of various levels of abstraction

A Solution – Deep Belief Networks
(Hinton et al. 2006)



- These from a hierarchy OR
- There is some form of highlighting = feature subset selection
- These principles should be applied
  - New action for compression / essence highlighting
  - Layer-by-layer weight training
  - Regularization

Alexnet 2012



Convolution

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

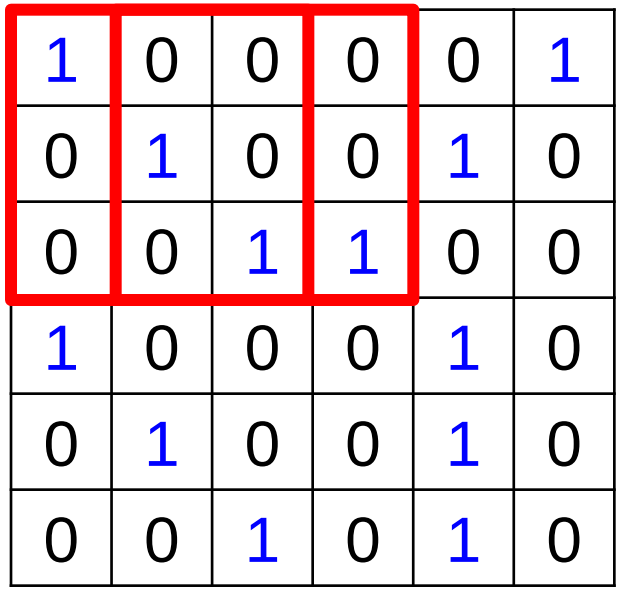
Filter 2

⋮ ⋮

Each filter identifies a simple pattern(3 x 3).

https://cs.uwaterloo.ca/~mli/Deep-Learning-2017-Lecture5CNN.ppt

stride=1



6 x 6 image

Scalar multiplication

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

https://cs.uwaterloo.ca/~mli/Deep-Learning-2017-Lecture5CNN.ppt

Convolution

If stride=2

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

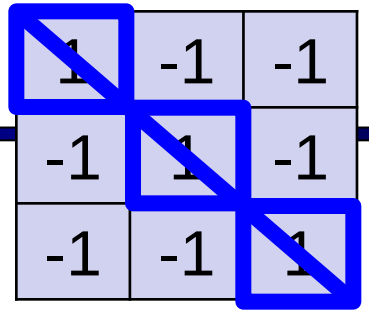
3 -3

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

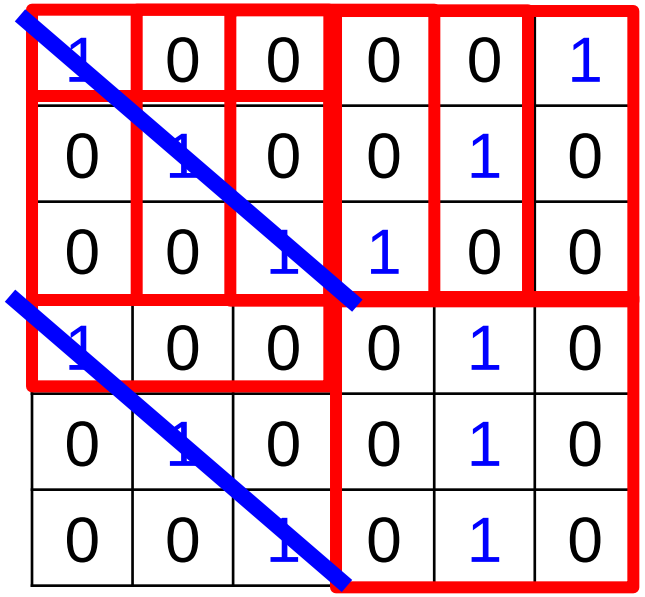
https://cs.uwaterloo.ca/~mli/Deep-Learning-2017-Lecture5CNN.ppt

Convolution

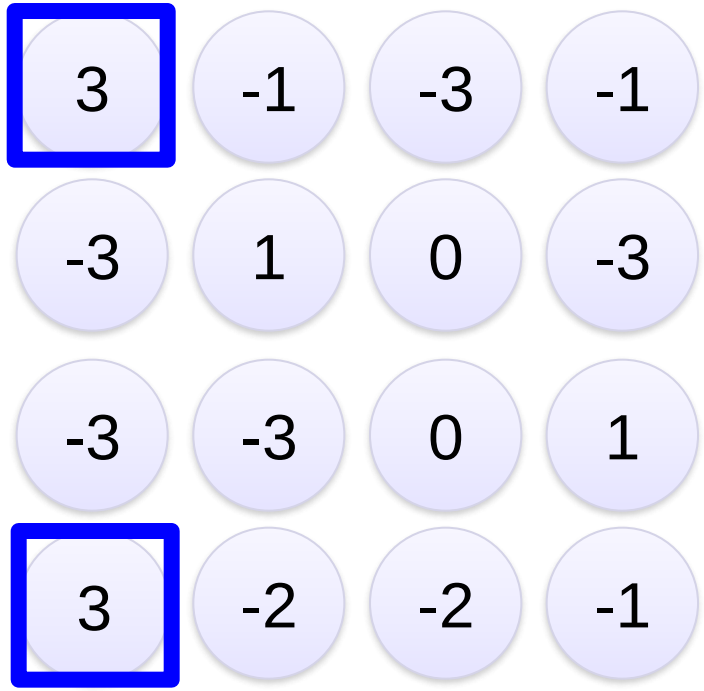


Filter 1

stride=1



6 x 6 image



https://cs.uwaterloo.ca/~mli/Deep-Learning-2017-Lecture5CNN.ppt

-1	1	-1
-1	1	-1
-1	1	-1

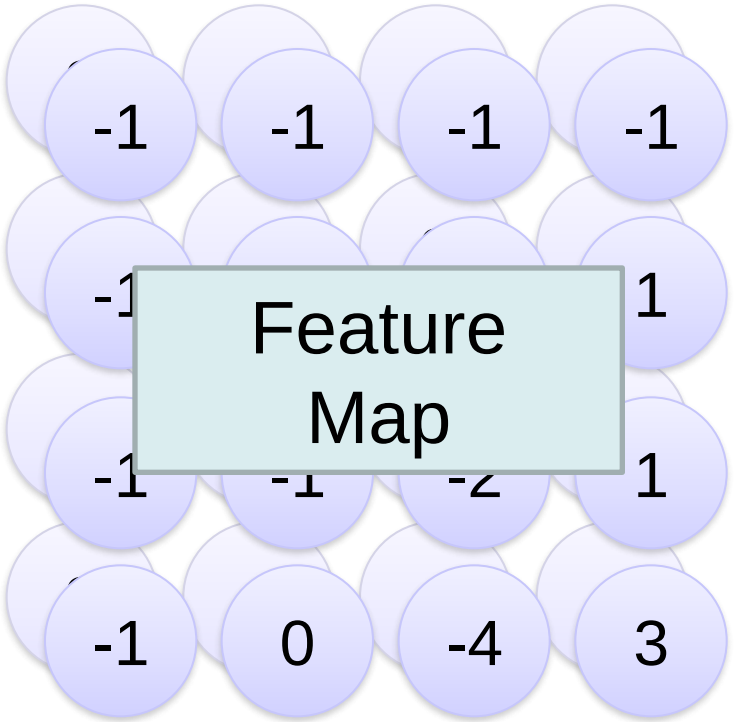
Filter 2

stride=1

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

For all filters



two 4 x 4 images from a
2 x 4 x 4 matrix

Max Pooling

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

-1	-1	-1	-1
-1	-1	-2	1
-1	-1	-2	1
-1	0	-4	3

https://cs.uwaterloo.ca/~mli/Deep-Learning-2017-Lecture5CNN.ppt

CNN

Convolution

Max Pooling

A new image

Convolution

Max Pooling

A new image

Flattened

Fully Connected Feedforward network

cat dog ……