



Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar
Mesterséges Intelligencia és Rendszertervezés Tanszék



VIMIAC16 2025/26/I.

Mély neurális hálók alapjai - CNN

Előadó: Dr. Hullám Gábor



Mély neurális hálózatok

- 2000-es évek közepén a neurális hálózatok témakörben alig lehetett cikket elfogadtatni
- 2004 – G. Hinton - CIFAR (Canadian Institute for Advanced Research)
 - Új csomagolás a neurális hálózatoknak: **deep learning**
- 2006 - Hinton, Osindero, Yee-Whye Teh: A fast learning algorithm for deep belief nets
 - Új áttörés: mély hiedelem hálók rétegenkénti tanítása,
 - Újra beindítja a kutatást

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

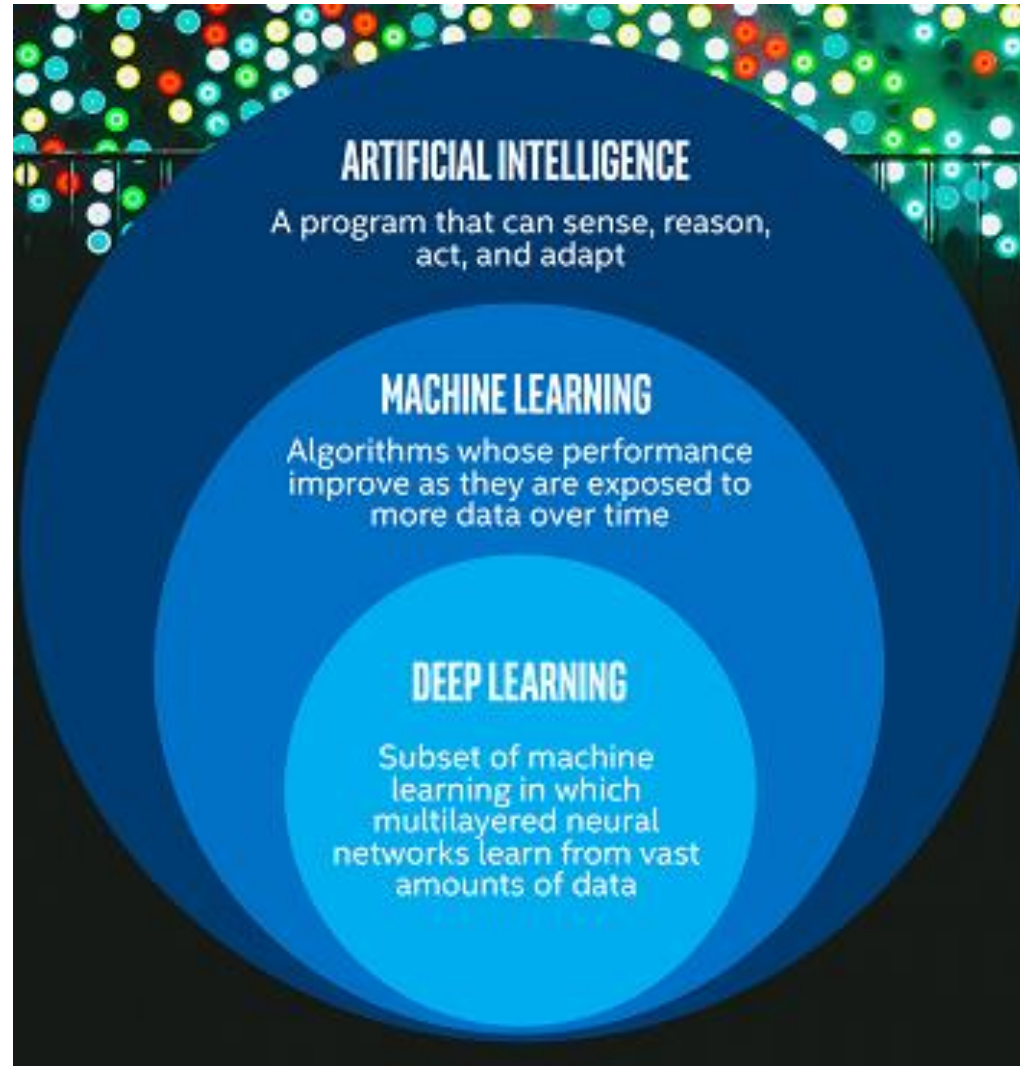
- Adatkészlet, nagy felbontású képek, 14M kép, 10k+ kategória <http://www.image-net.org/about-stats>



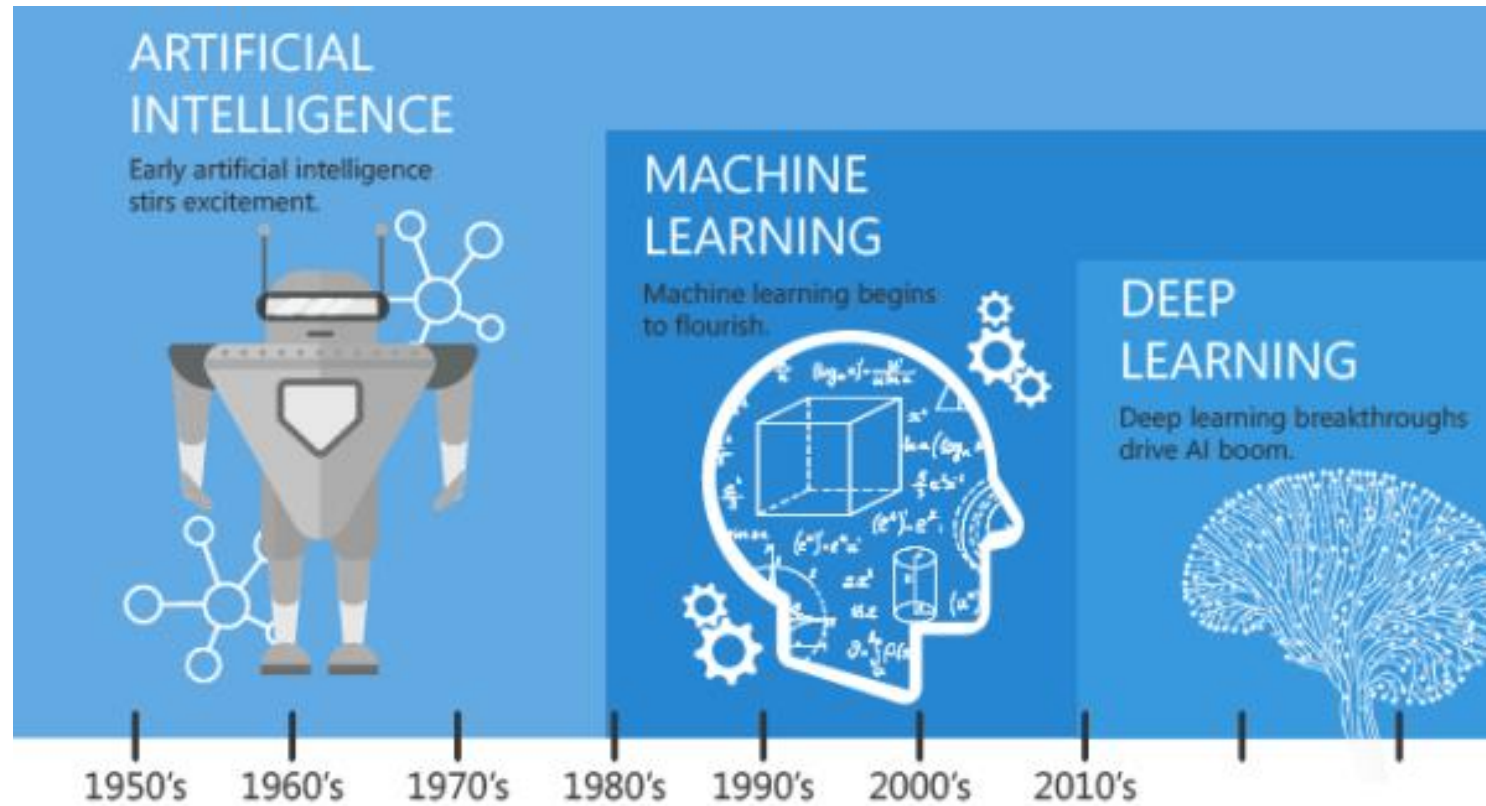
Deep learning - ILSVRC

- 2012 - Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton: Imagenet classification with deep convolutional neural networks
 - Ez az első tisztán neurális hálózatos megoldás, ami az azévi legjobb eredményt éri el (addig SVM)
 - Innentől kezdve lesz mainstream a deep learning

AI – ML - DL

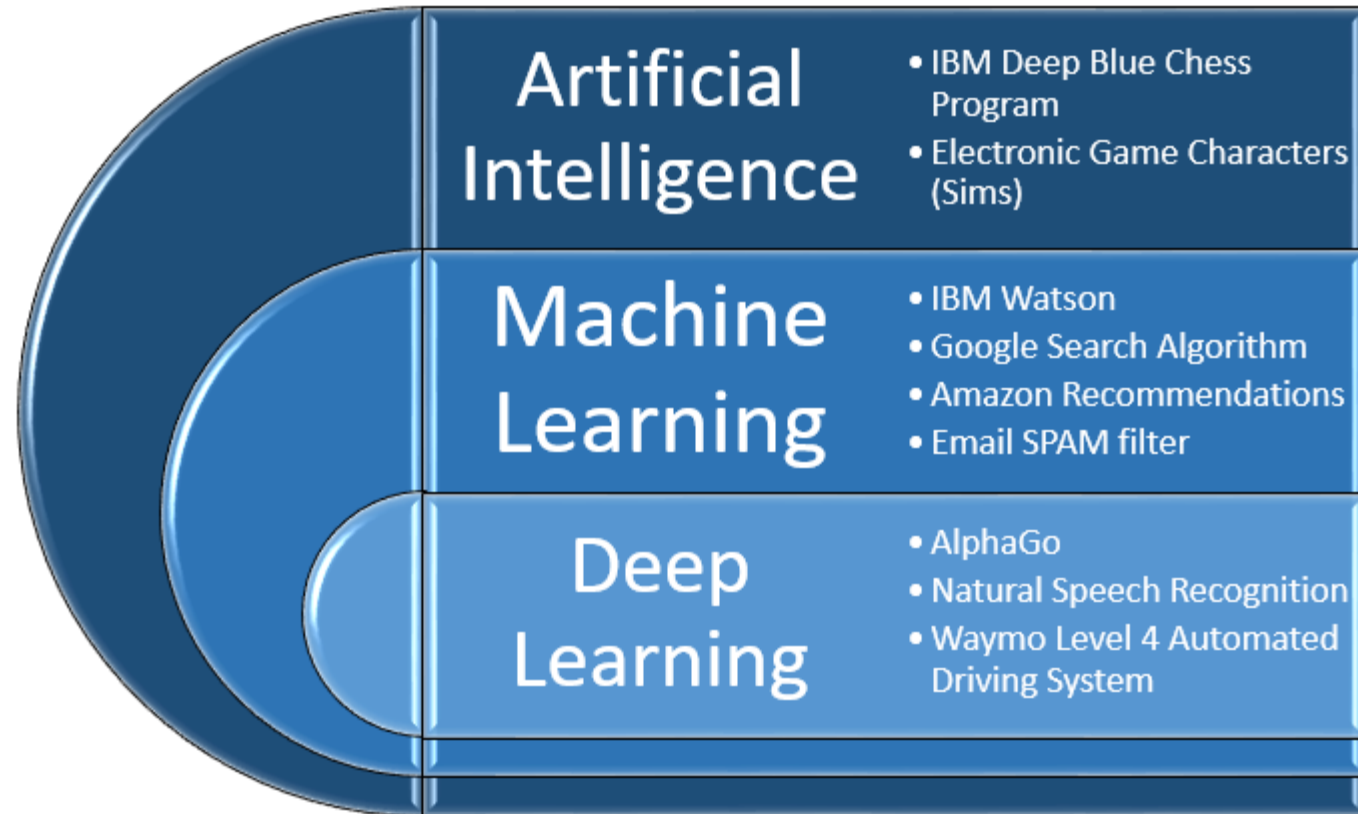


AI – ML - DL



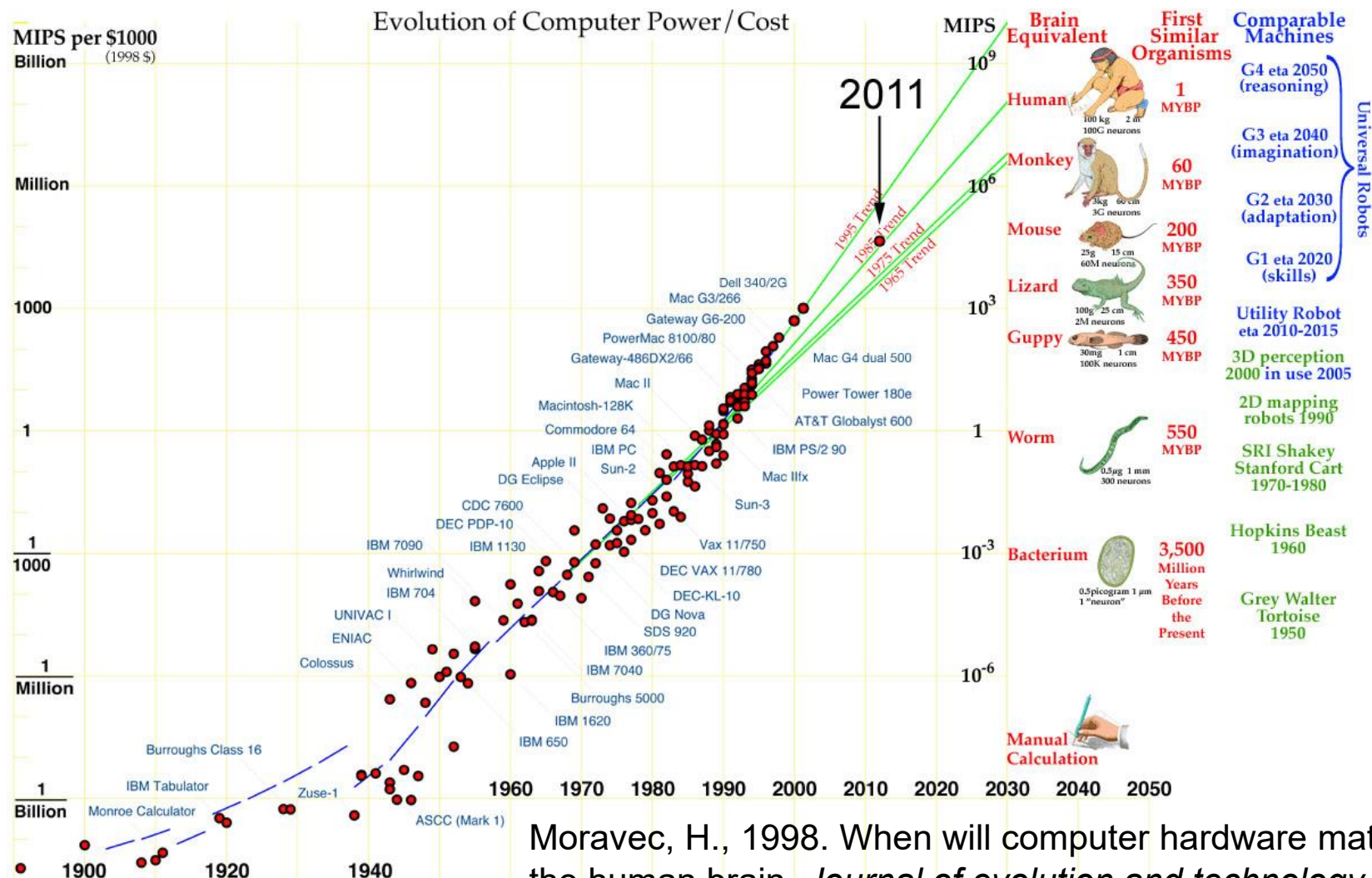
<https://shawnennis.com/the-technology-of-machine-learning-with-ai>

AI – ML - DL



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

Háttérkövetelmény: számítási teljesítmény

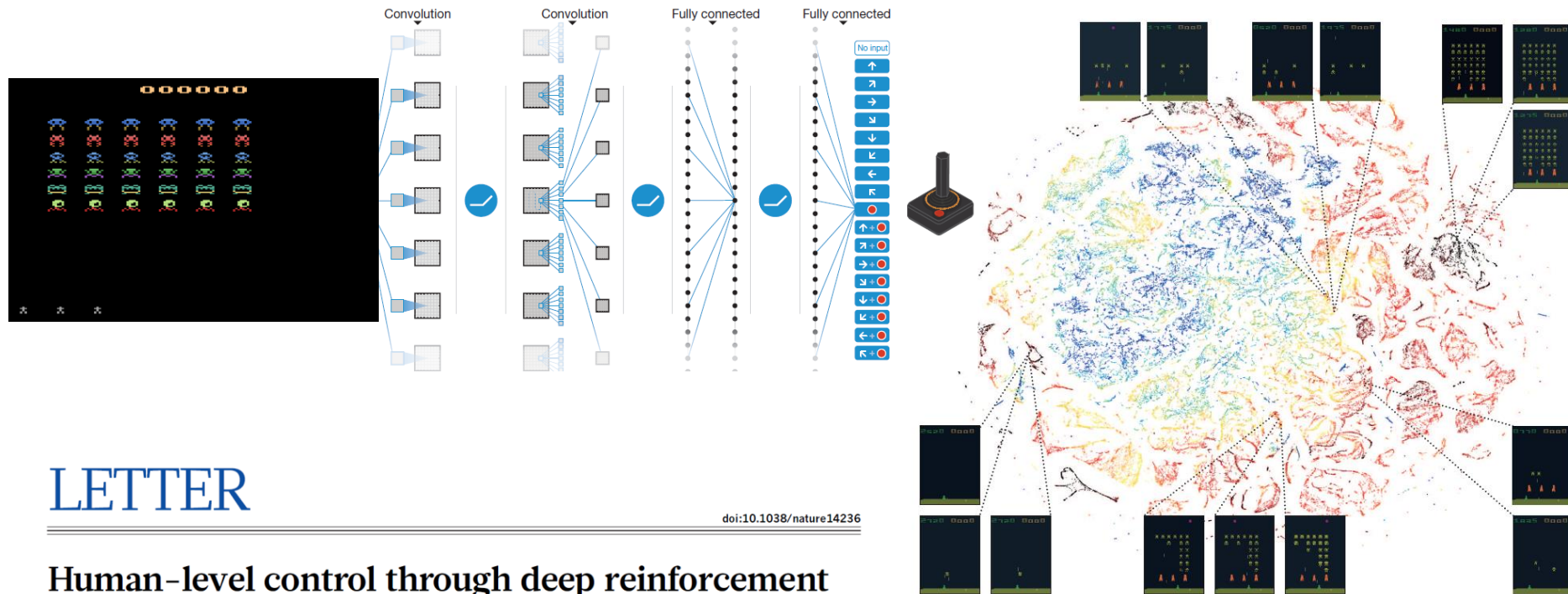


Moravec, H., 1998. When will computer hardware match the human brain. *Journal of evolution and technology*, 1(1), p.10.

Deep reinforcement learning

- 2013 – DeepMind: Playing Atari with Deep Reinforcement Learning
- 2016 – DeepMind: AlphaGo legyőzi a Go világbajnok Lee Sedolt

DRL: Arcade játékok tanítása



LETTER

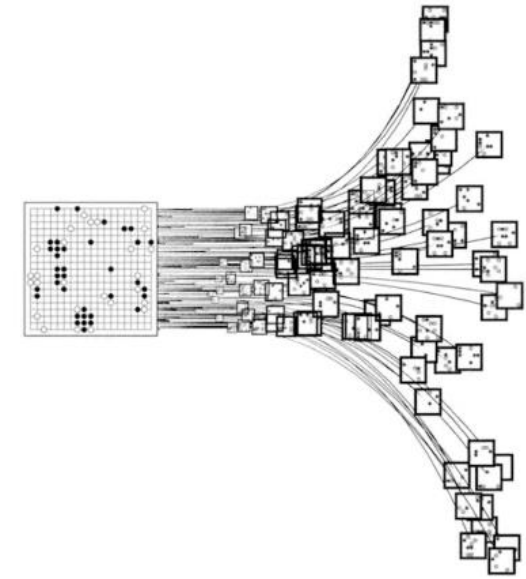
doi:10.1038/nature14236

Human-level control through deep reinforcement learning

Volodymyr Mnih^{1*}, Koray Kavukcuoglu^{1*}, David Silver^{1*}, Andrei A. Rusu¹, Joel Veness¹, Marc G. Bellemare¹, Alex Graves¹, Martin Riedmiller¹, Andreas K. Fiedelnd¹, Georg Ostrovski¹, Stig Petersen¹, Charles Beattie¹, Amir Sadik¹, Ioannis Antonoglou¹, Helen King¹, Dharshan Kumaran¹, Daan Wierstra¹, Shane Legg¹ & Demis Hassabis¹

Go

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



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. Maddison¹, Arthur Guez¹, Laurent Sifre¹, George van den Driessche¹, Julian Schrittwieser¹, Ioannis Antonoglou¹, Veda Panneershelvam¹, Marc Lanctot¹, Sander Dieleman¹, Dominik Grewe¹, John Nham², Nal Kalchbrenner¹, Ilya Sutskever², Timothy Lillicrap¹, Madeleine Leach¹, Koray Kavukcuoglu¹, Thore Graepel¹ & Demis Hassabis¹

Deep learning – első eredmények

■ Felirat generálás képekhez

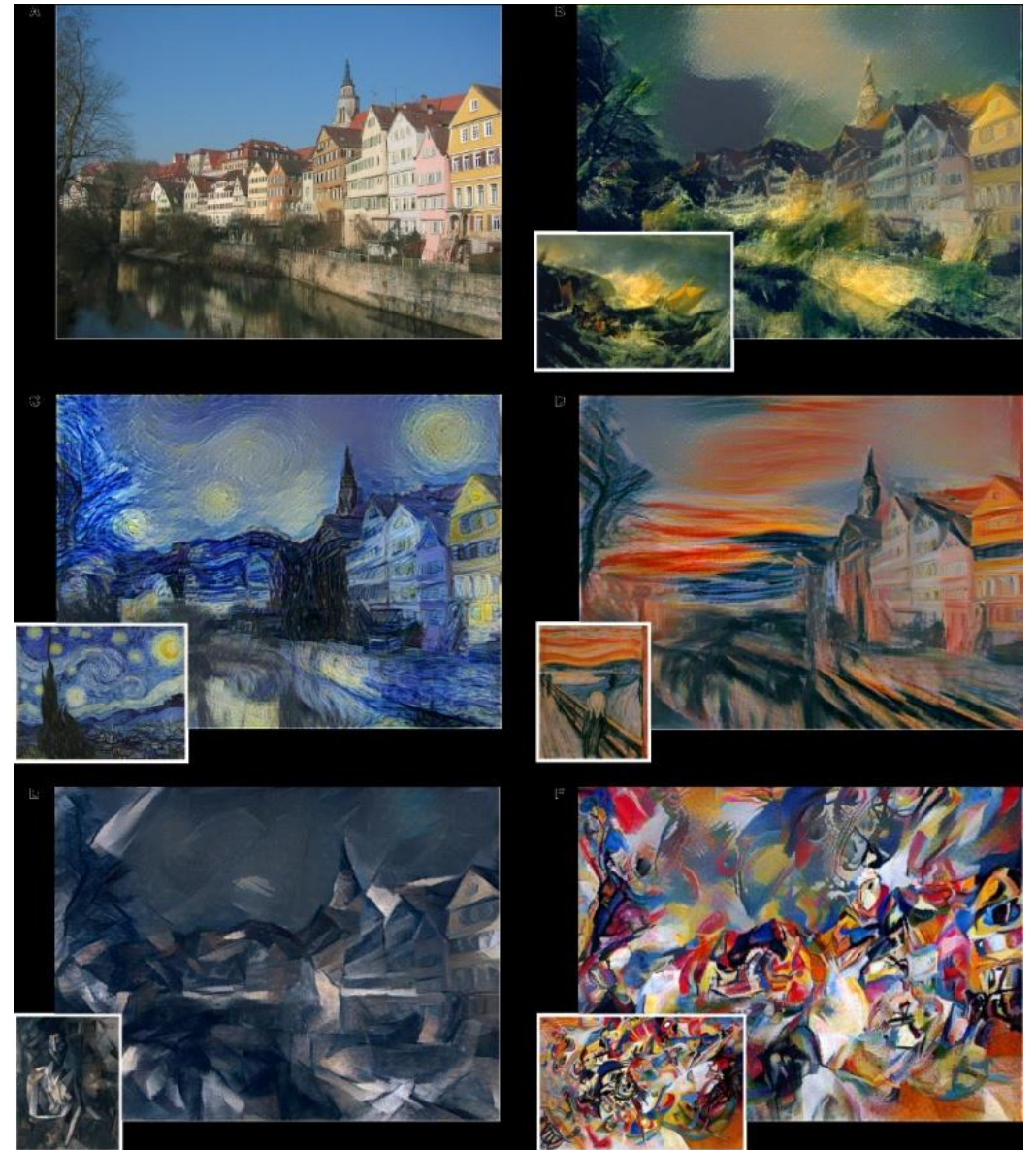
- A. Karpathy, L. Fei-Fei: „Deep Visual-Semantic Alignments for Generating Image Descriptions”
- <https://research.googleblog.com/2014/11/a-picture-is-worth-thousand-coherent.html>

Describes without errors	Describes with minor errors	Somewhat related to the image	Unrelated to the image
 <p>A person riding a motorcycle on a dirt road.</p>	 <p>Two dogs play in the grass.</p>	 <p>A skateboarder does a trick on a ramp.</p>	 <p>A dog is jumping to catch a frisbee.</p>
 <p>A group of young people playing a game of frisbee.</p>	 <p>Two hockey players are fighting over the puck.</p>	 <p>A little girl in a pink hat is blowing bubbles.</p>	 <p>A refrigerator filled with lots of food and drinks.</p>
 <p>A herd of elephants walking across a dry grass field.</p>	 <p>A close up of a cat laying on a couch.</p>	 <p>A red motorcycle parked on the side of the road.</p>	 <p>A yellow school bus parked in a parking lot.</p>

Deep learning – első eredmények

- Stílus ‘tanulás’

- <https://github.com/jcjohnson/neural-style>
<https://imgur.com/gallery/4LTaQ>



Deep learning – első eredmények

■ Beszéderítés

- Xiong, W. et al. (2016). Achieving human parity in conversational speech recognition

■ Beszédkeltés

- Oord, Aaron van den, et al. "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499 (2016).

■ Képek osztályozása

- C. Szegedy et al. (2014): GoogLeNet – „Going Deeper with Convolutions”

A mostly complete chart of Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

-  Backfed Input Cell
-  Input Cell
-  Noisy Input Cell
-  Hidden Cell
-  Probabilistic Hidden Cell
-  Spiking Hidden Cell
-  Output Cell
-  Match Input Output Cell
-  Recurrent Cell
-  Memory Cell
-  Different Memory Cell
-  Kernel
-  Convolution or Pool

Perceptron (P)



Feed Forward (FF)



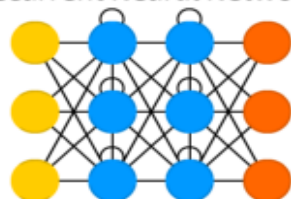
Radial Basis Network (RBF)



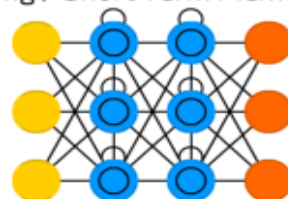
Deep Feed Forward (DFF)



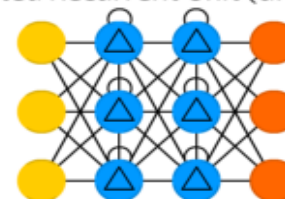
Recurrent Neural Network (RNN)



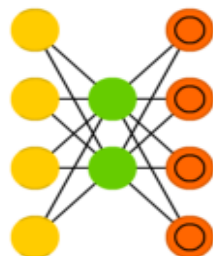
Long / Short Term Memory (LSTM)



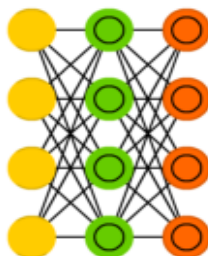
Gated Recurrent Unit (GRU)



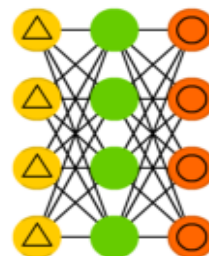
Auto Encoder (AE)



Variational AE (VAE)



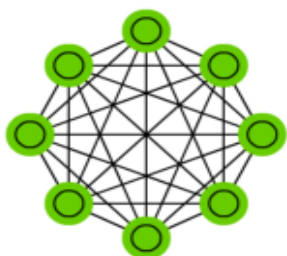
Denoising AE (DAE)



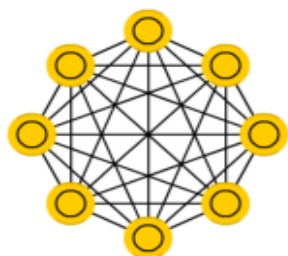
Sparse AE (SAE)



Markov Chain (MC)



Hopfield Network (HN)



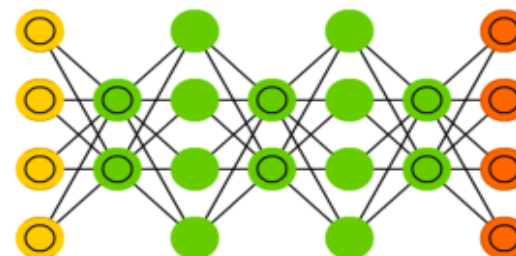
Boltzmann Machine (BM)



Restricted BM (RBM)



Deep Belief Network (DBN)

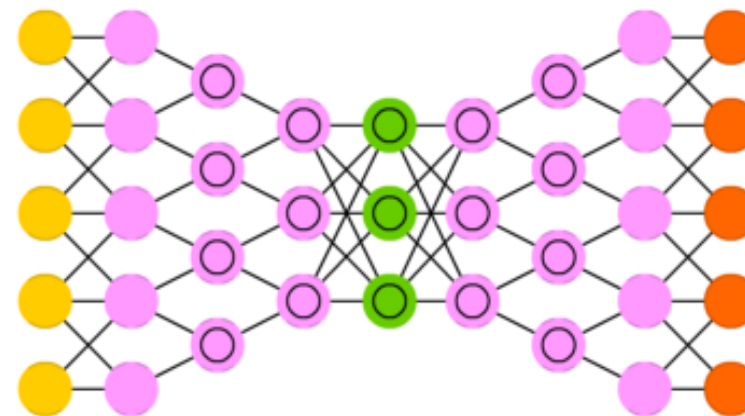


Neurális hálók típusai

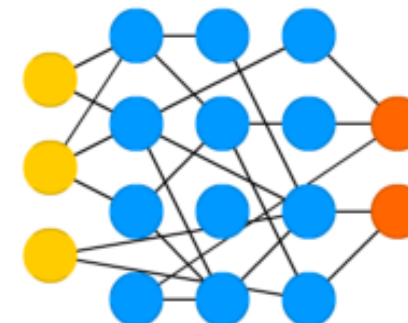
<https://towardsdatascience.com/the-mostly-complete-chart-of-neural-networks-explained-3fb6f2367464>

típusai 2.

Deep Convolutional Inverse Graphics Network (DCIGN)



Echo State Network (ESN)

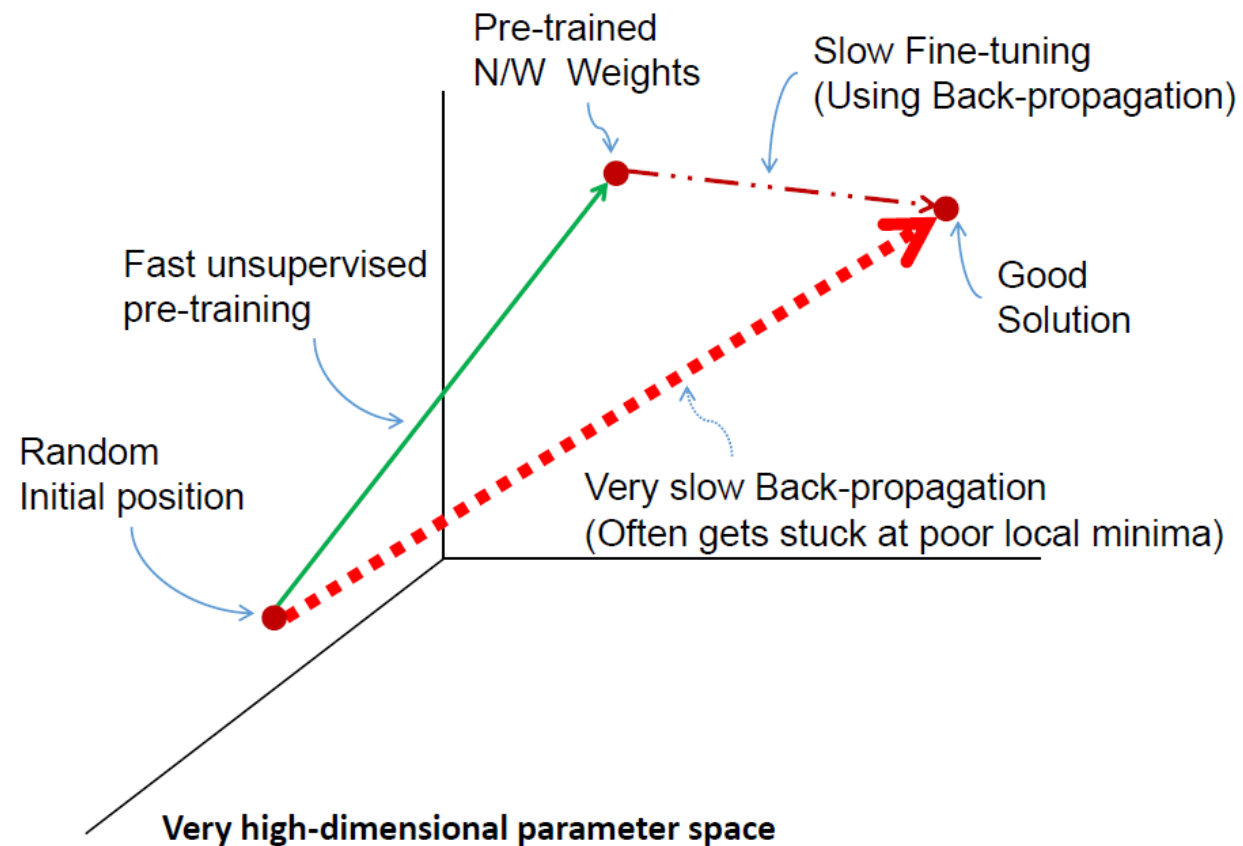


Neural Turing Machine (NTM)



Mitől „jobbak” a mély neurális hálók?

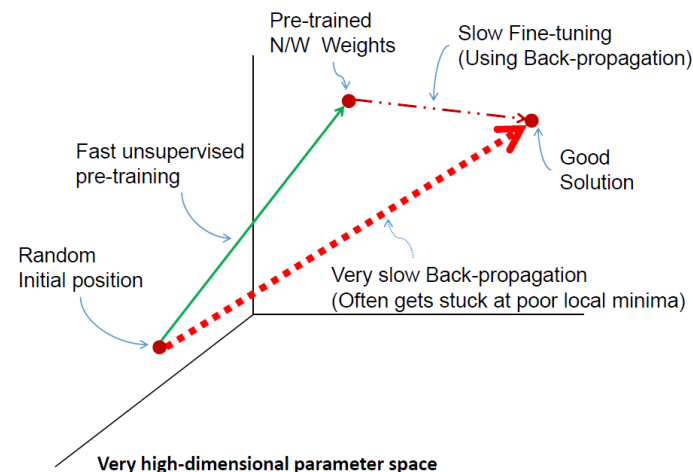
A Solution – Deep Belief Networks
(Hinton et al. 2006)



Mitől „jobbak” a mély neurális hálók?

- Az ember is különböző absztrakciós szintű jegyeket detektál
- Ezek hierarchiába szerveződnek VAGY
- Van valamilyen lényeg-kiemelés
- Ezeket az elveket kéne alkalmazni!
 - Új művelet tömörítésre / lényeg kiemelésre
 - Rétegenkénti súly-előtanítás
 - Regularizáció

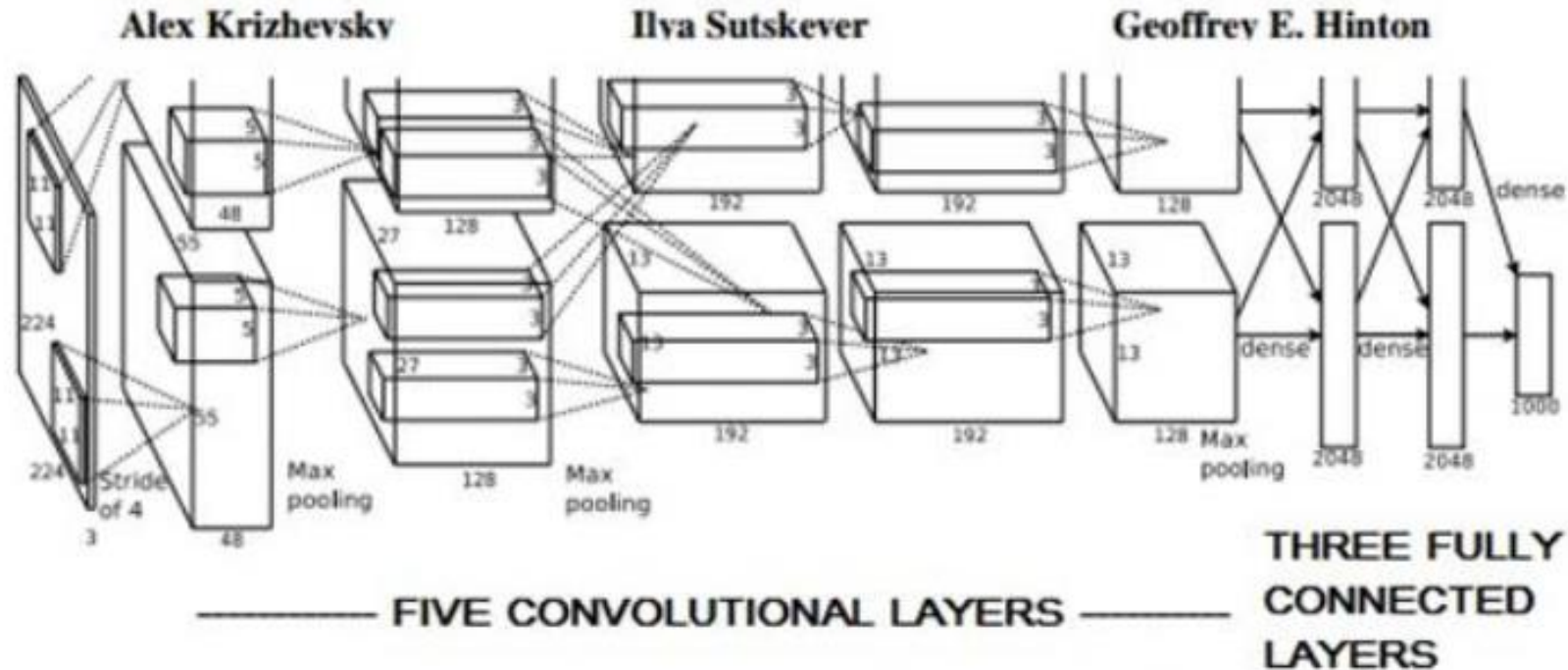
A Solution – Deep Belief Networks
(Hinton et al. 2006)



Alexnet 2012

ImageNet Classification with Deep Convolutional Neural Networks

2012



Konvolúció

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 kép

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

Filter 1

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

Filter 2

⋮ ⋮

Minden szűrő egyszerű
mintázatot azonosít (3 x 3).

Konvolúció

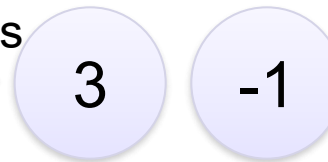
1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

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

Skalár
szorzás



6 x 6 kép

Konvolúció

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

Filter 1

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 kép

3 -3

Konvolúció

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

Filter 1

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 kép

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

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

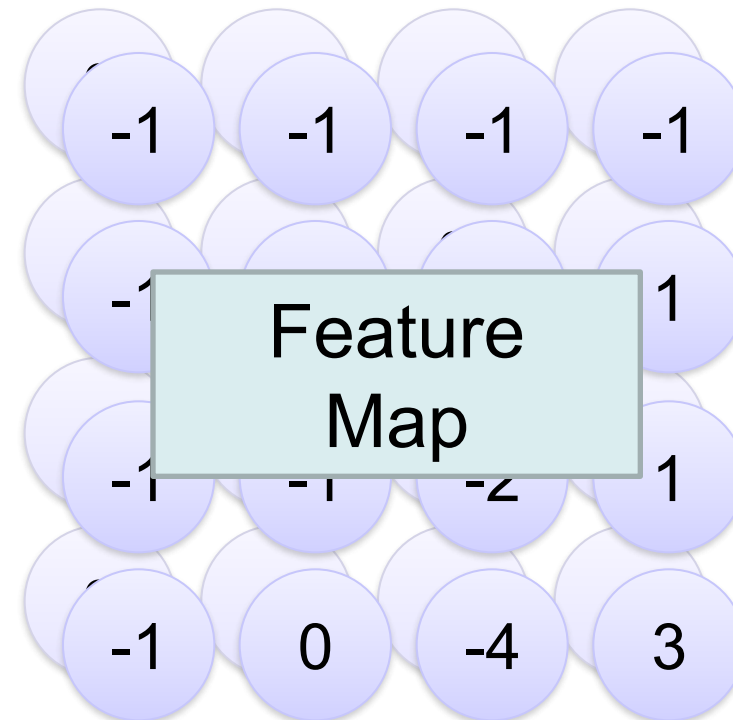
Filter 2

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 kép

Minden filterre



két 4 x 4 kép
2 x 4 x 4 mátrixot alkot

Max Pooling

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

Filter 1

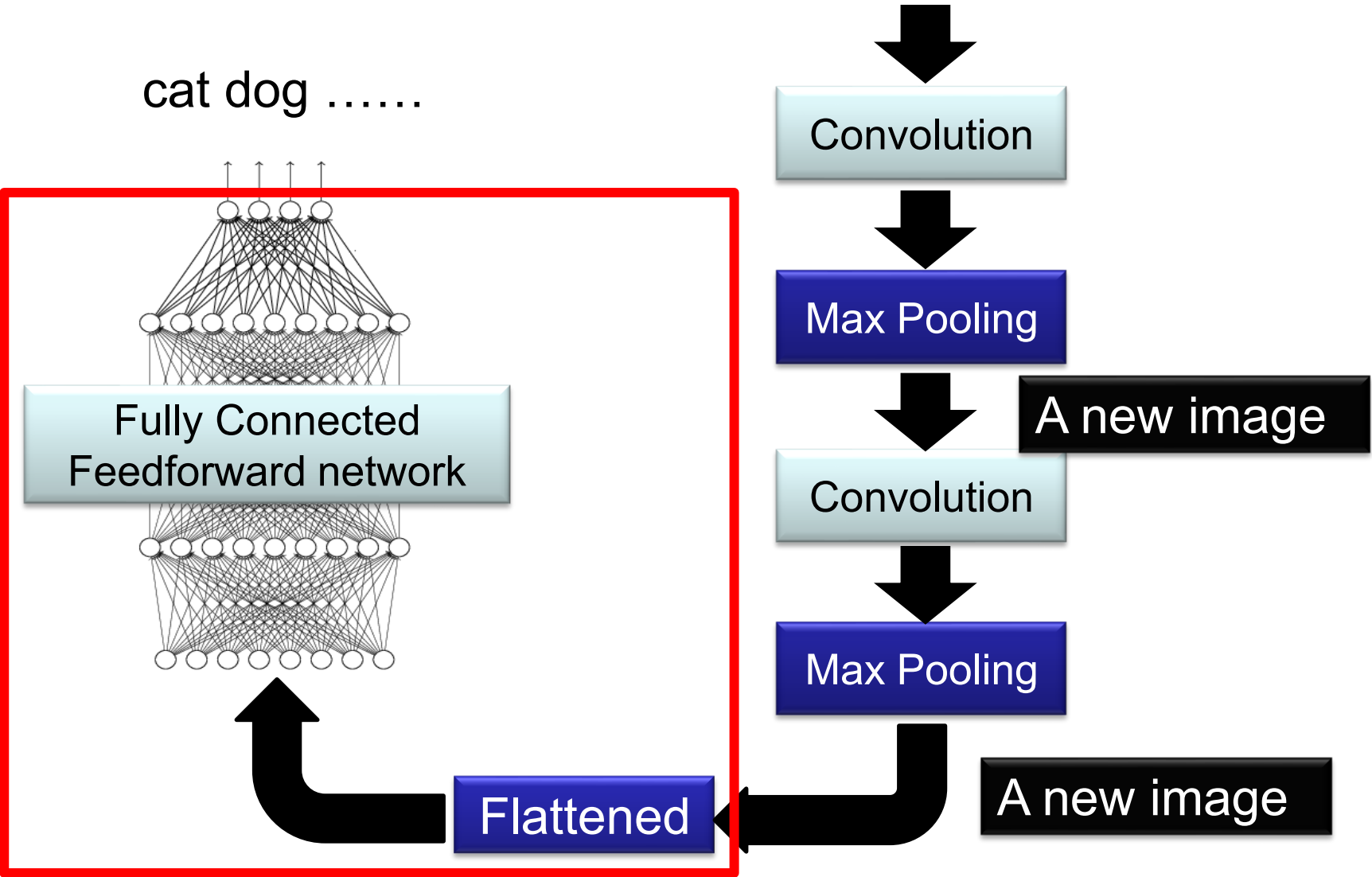
-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

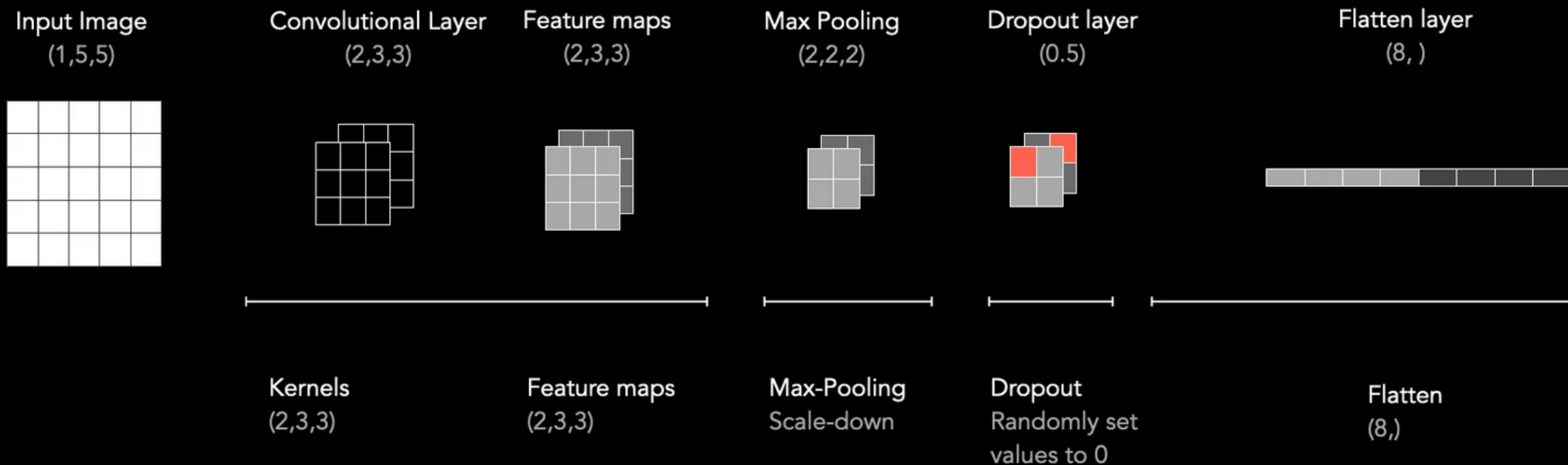
3	-1	-3	-1
-3	1	0	-3
-3	-3	0	1
3	-2	-2	-1

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

Konvolúciós neurális hálózat (CNN)

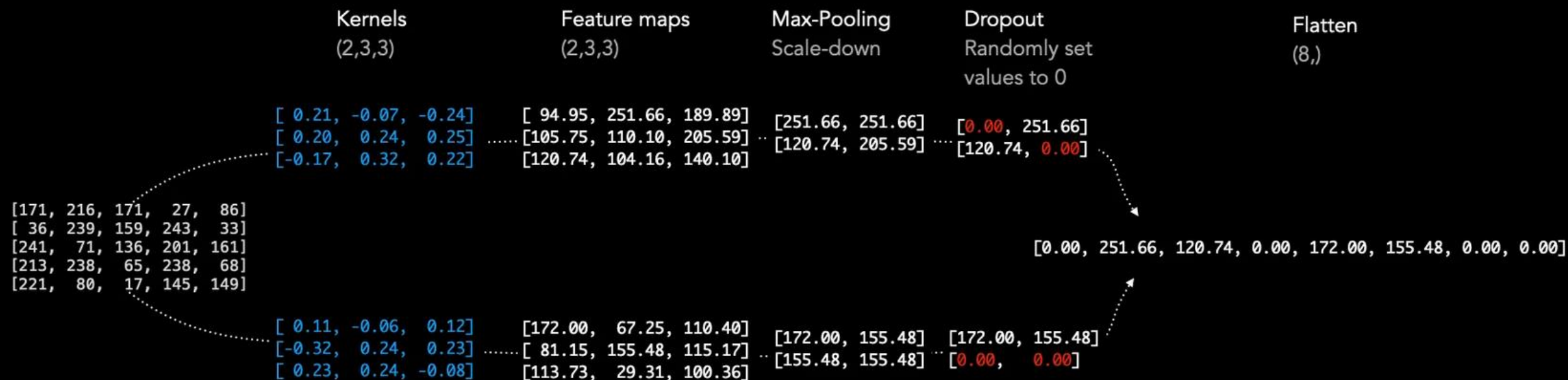


CNN: bemenettől a „fej” részig



<https://medium.com/advanced-deep-learning/cnn-operation-with-2-kernels-resulting-in-2-feature-mapsunderstanding-the-convolutional-filter-c4aad26cf32>

CNN: bemenettől a „fej” részig



<https://medium.com/advanced-deep-learning/cnn-operation-with-2-kernels-resulting-in-2-feature-mapsunderstanding-the-convolutional-filter-c4aad26cf32>