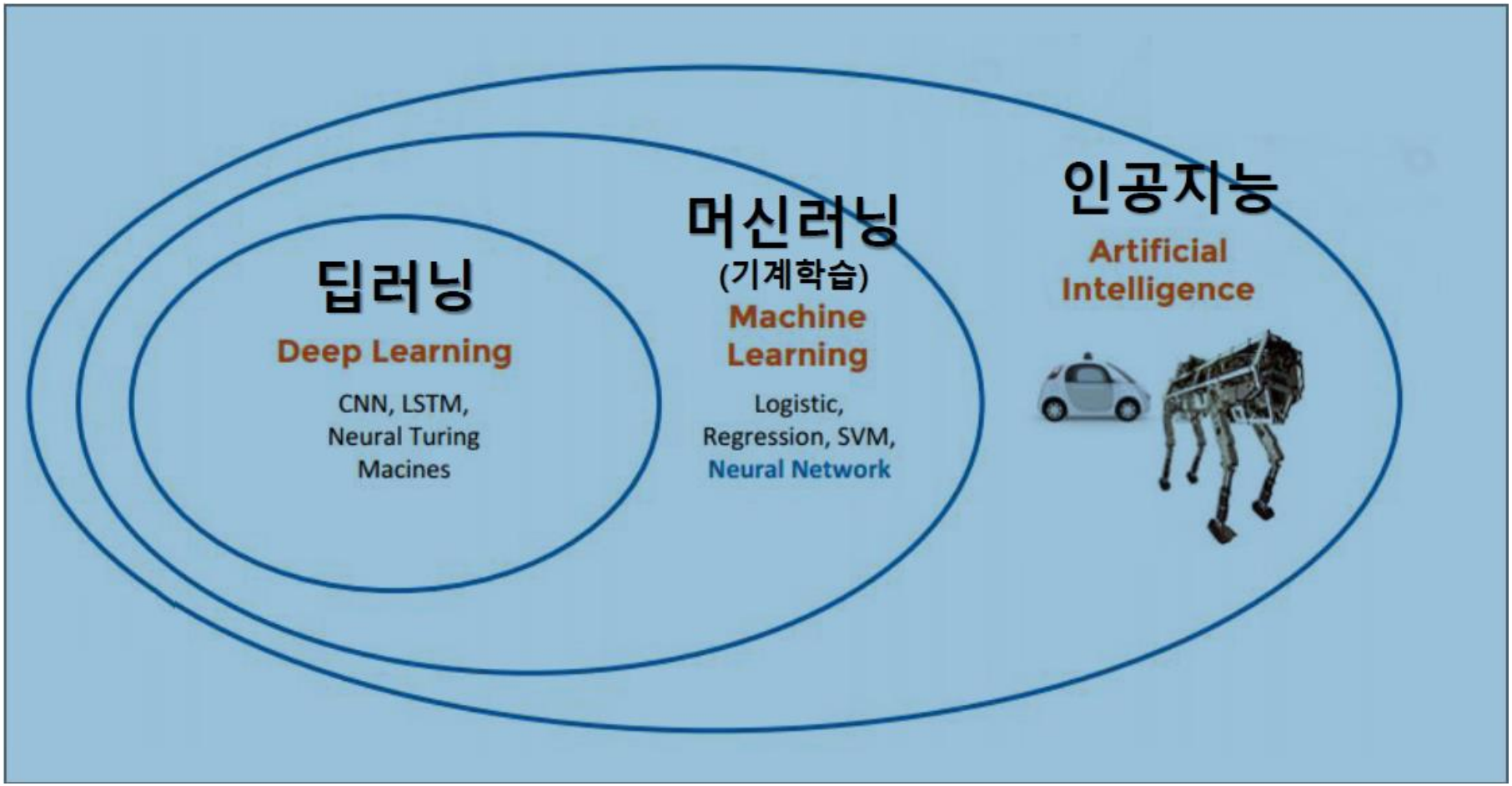
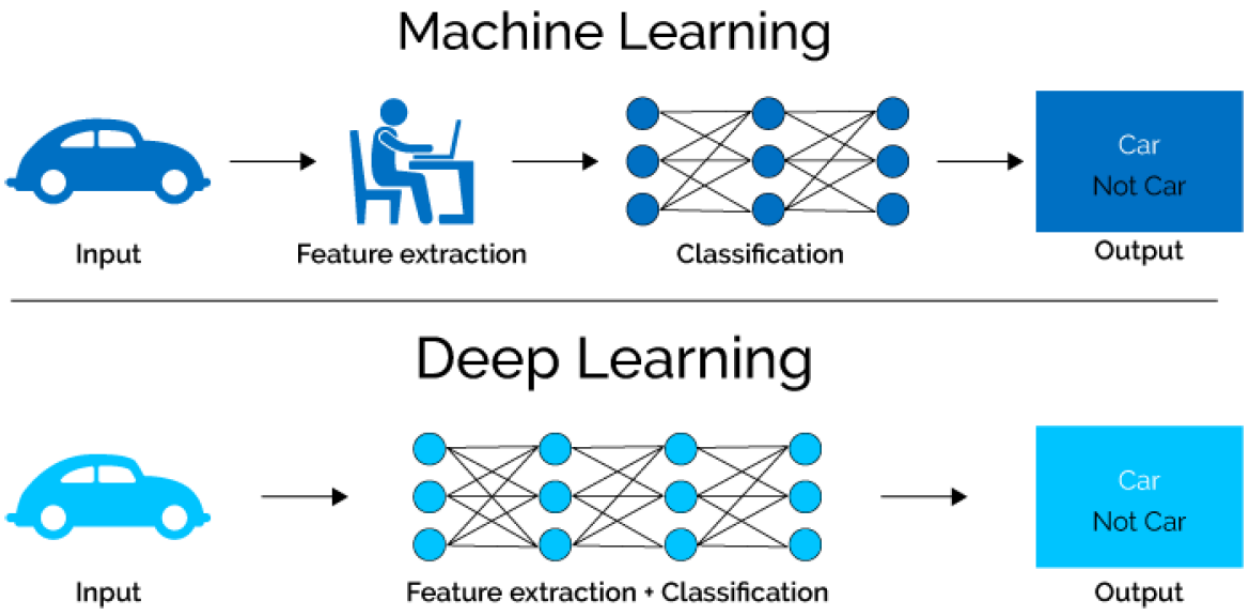
Deep learning:
an introduction

Deep learning in AI



Deep learning in AI

고전적인 신경망과 딥러닝

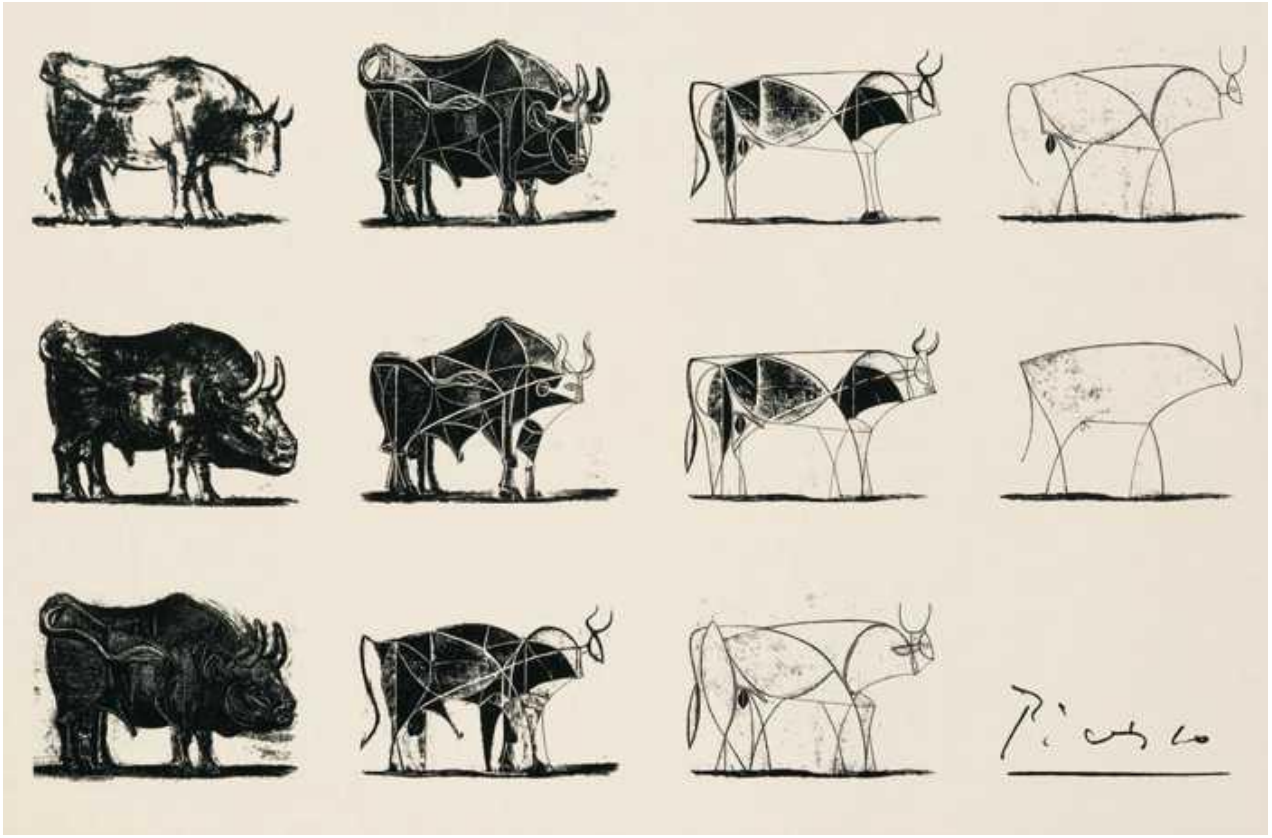


Contents

- Deep learning in a nutshell
- The boom of deep learning: famous achievements!
- Neural networks
- Deep neural networks
- Deep learning variants
- Extensions
- Final remarks

Deep learning in a nutshell

- Deep learning is a machine learning methodology that aims at solving (modeling) problems by building layer-wise models with several levels of increasing abstraction
- Layers of these models capture discriminative/ descriptive information from raw data
- Can be used for: supervised/unsupervised learning, reinforcement learning, feature extraction, …
- Examples: multi-layer perceptrons, deep neural networks, convolutional neural networks, deep belief nets, auto encoders, etc.

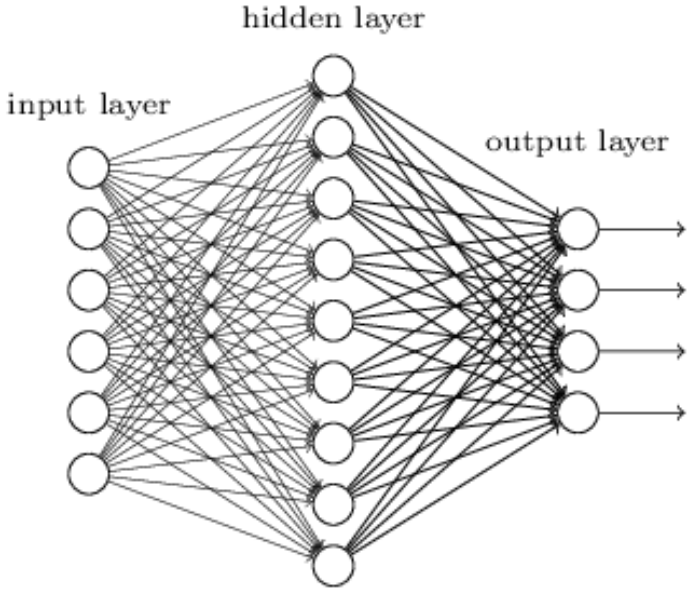


Deep learning in a nutshell

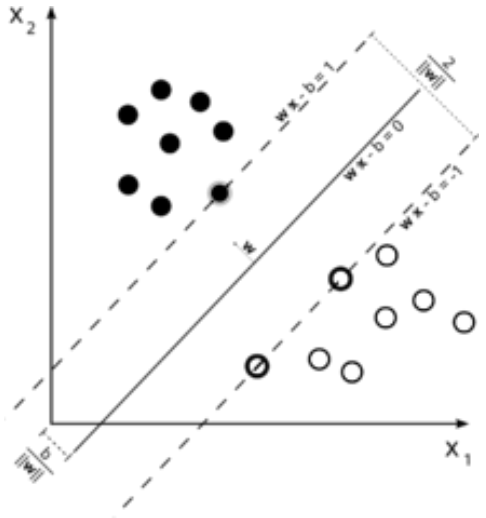
- Features of deep learning methods:
  - Large number of parameters (on the ranges of millions)
  - Require large amounts of data to be trained
  - Can extract features automatically
  - Can leverage unlabeled data
  - Extremely complex models
  - Require of specialized hardware for training them efficiently
  - Dominate the arenas of machine learning applications (e.g., computer vision, NLP)

Deep learning in a nutshell

- How does a non deep model looks like?



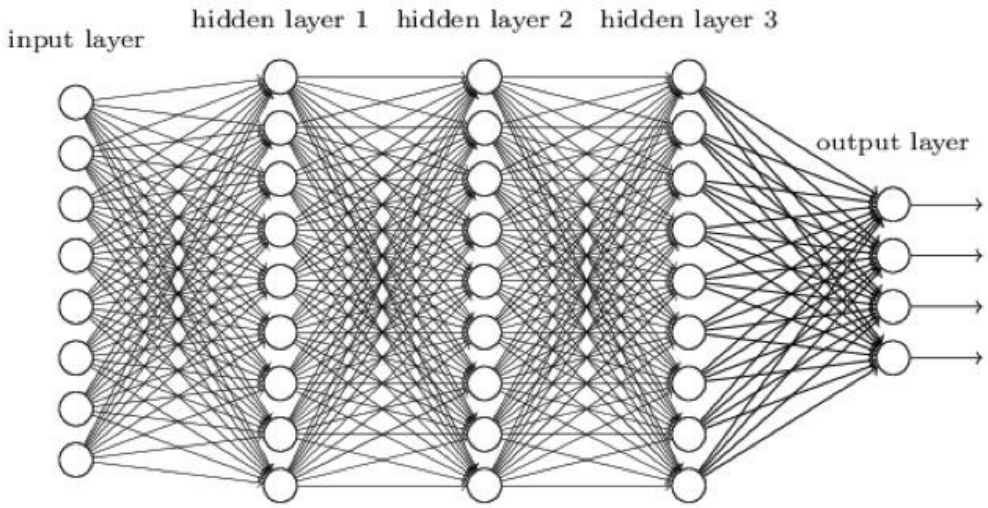
$$f(x) = w\emptyset(x) + b$$



$$f(\mathbf{x}) = \sum_{i}^{N} \alpha_i y_i k(\mathbf{x}_i, \mathbf{x}) + b$$

Deep learning in a nutshell

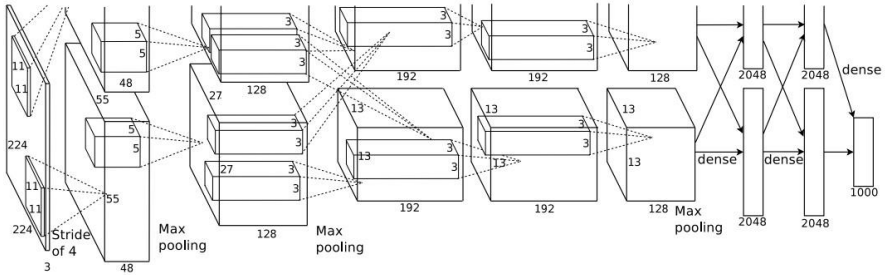
- How does a (not too deep) DL model looks like?

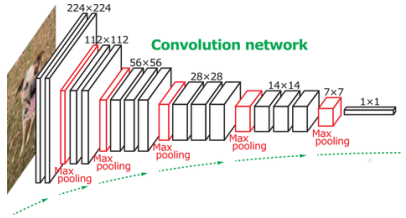


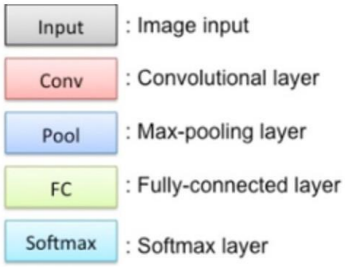
$$f(\mathrm{x}) = W_3\phi_3(W_2\phi_2(W_1\phi_1(X) + b) + b_2) + b_3$$

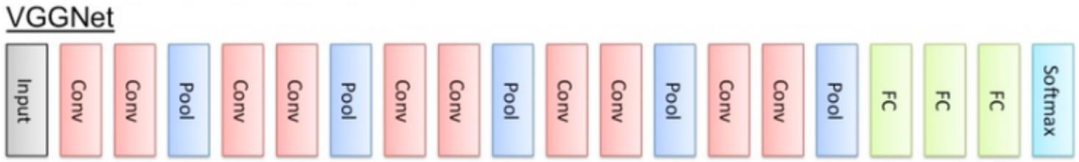
Deep learning in a nutshell

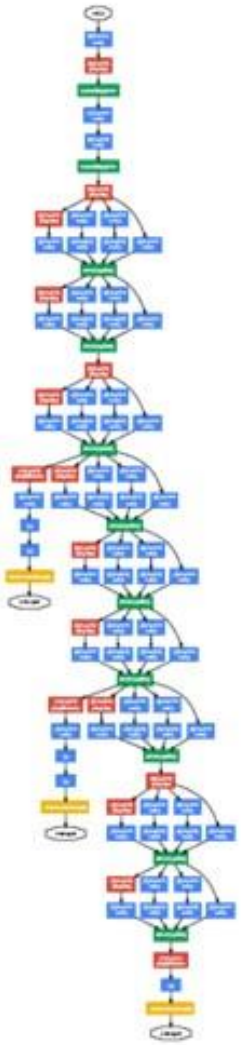
- Going deeper (CNNs)



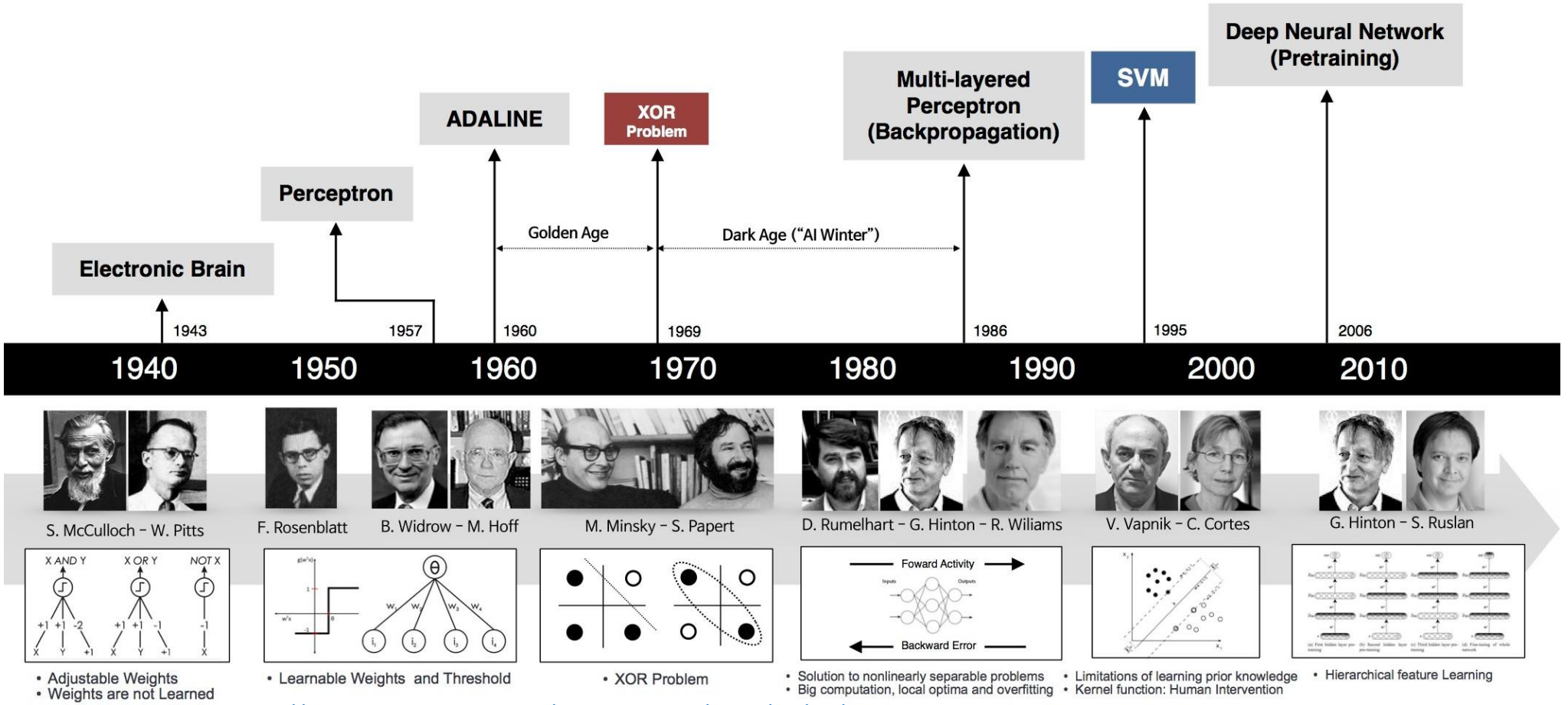








The boom of DL: brief history



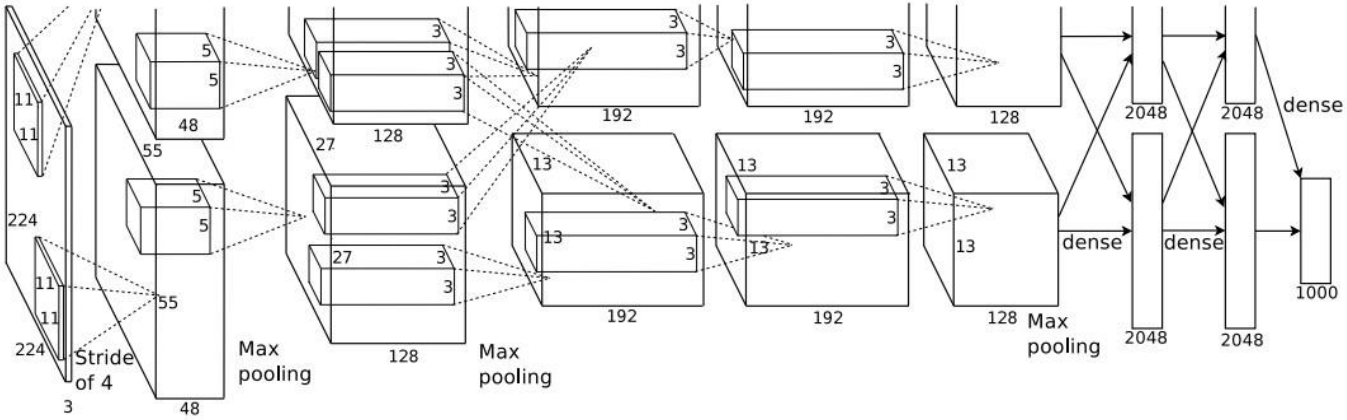
http://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html

The boom of DL: noticeable achivements

- Large scale image classification
- Speech recognition
- Face recognition
- Deep reinforcement learning
- Other achievements
  - Image captioning
  - Word embeddings
  - Gesture / action recognition
  - Super resolution
  - …

Breakthrough achievements I (ImageNET)

- In 2012, Krizhevsky et al. succeded at training a convolutional neural network with about 1 million images, approaching the ImageNET large scale classification challenge (1000 of classes, millions of images)



IMAGENET

Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. **ImageNet Classification with Deep Convolutional Neural Networks.** Advances in Neural Information Processing Systems 25 (NIPS 2012) – AlexNet (8 Layers)

Breakthrough achievements I (ImageNET)

- Imagenet - 1000개의 카테고리와 1,431,167장의 이미지로 구성된 데이터셋

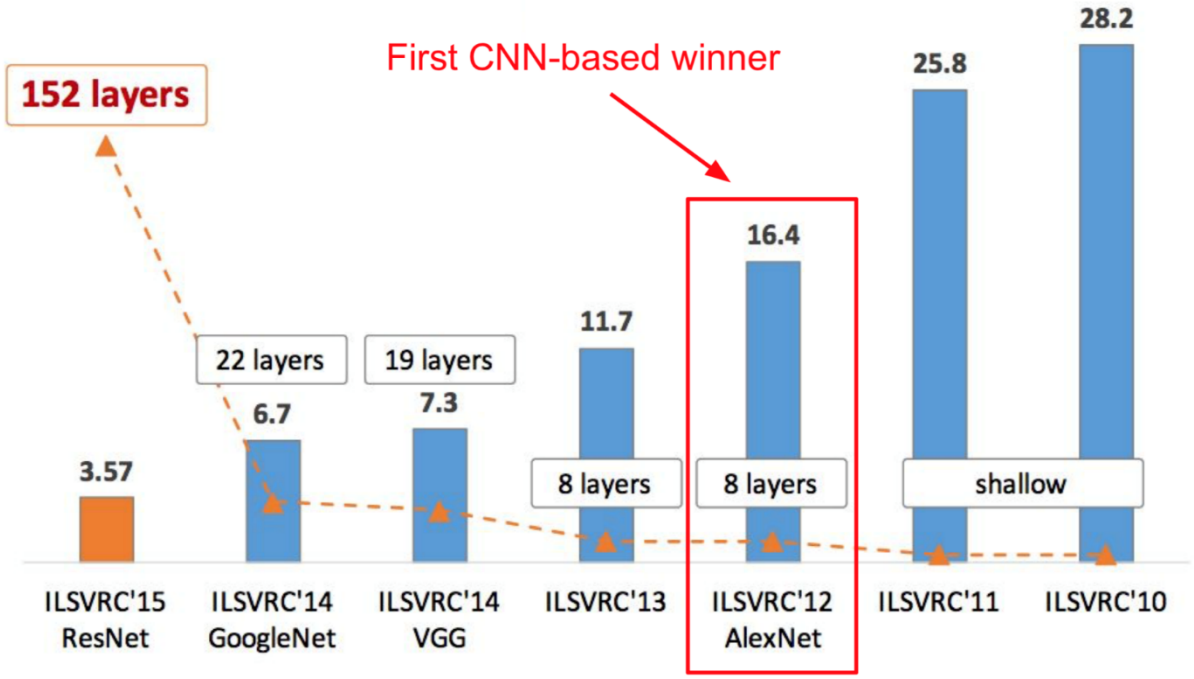
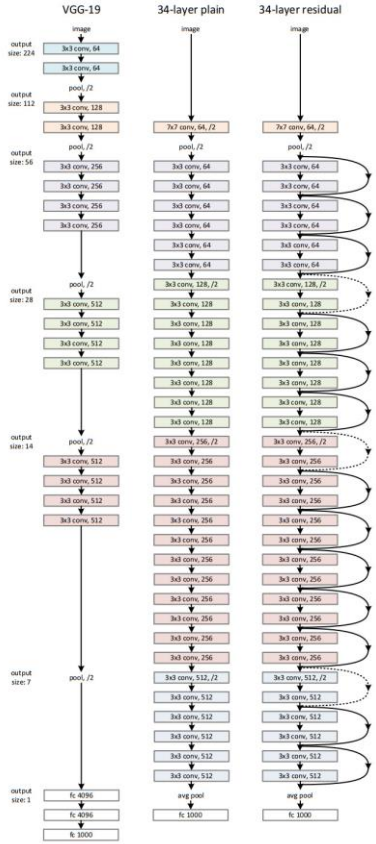


Figure copyright Kaiming He, 2016. Reproduced



The ImageNET challenge

- ImageNET: A huge resource comprising millions of images.
- Images were downloaded from the web using synsets from WordNet (유의어)
- The ImageNET challenge is or ganized since 2011
  - Classification
  - Object detection
  - Object localization

Year	Train images (per class)	Val images (per class)	Test images (per class)
Image classification annotations (1000 object classes)			
ILSVRC2010	1,261,406 (668–3047)	50,000 (50)	150,000 (150)
ILSVRC2011	1,229,413 (384–1300)	50,000 (50)	100,000 (100)
ILSVRC2012-14	1,281,167 (732–1300)	50,000 (50)	100,000 (100)

The numbers in parentheses correspond to (minimum per class–maximum per class). The 1000 classes change from year to year but are consistent between image classification and single-object localization tasks in the same year. All images from the image classification task may be used for single-object localization

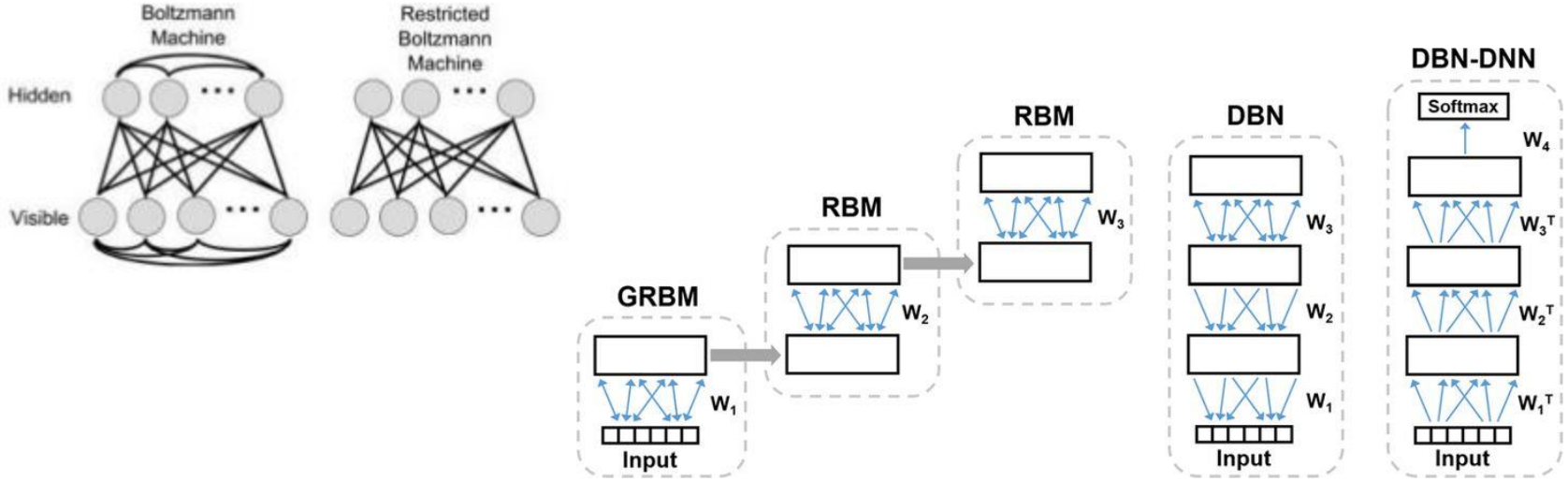
Breakthrough achievements I (ImageNET)

- Performance improvement with solutions from those days was impressive
- Key for success:
  - GPU based training
  - RELU activation functions
  - Dropout regularization
  - Big data / complex model

IMAGENET

Breakthrough achievements II (Speech recognition)

- Around 2012, the most important IT companies converged to the use of Restricted Boltzman Machines for Speech Recognition
- Key idea: RBM-pretraining + fine tunning + HMM



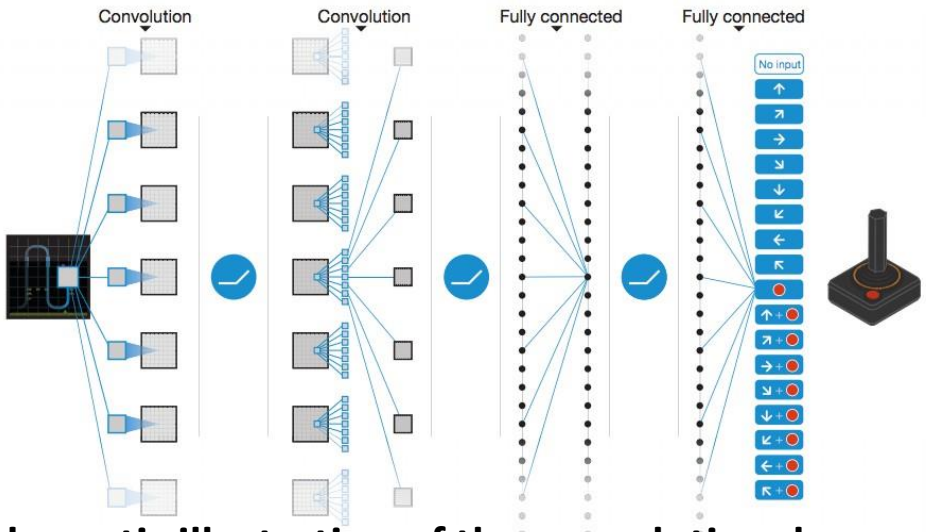
greedy-layer-wise pre-training

supervised fine-tuned as one DNN

Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. **Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups.** IEEE Signal Processing Magazine, Vol 29(6):82 - 97, 2012

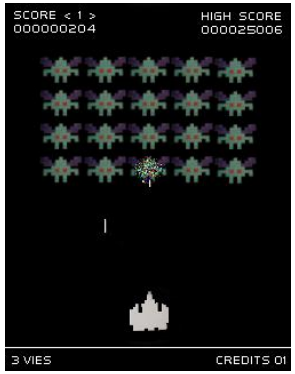
Breakthrough achievements ++: DeepRL

- In 2015, the deepmind team published their Deep-Q network: a DL architecture that by ”looking” at the pixels produced in videogames and using game scores, was able to learn to play Atari



Schematic illustration of the convolutional neural network.

Volodymyr Mnih, et al**. Human-level Control through Deep Reinforcement Lear ning** In Nature, 518: 529–533, 2015.

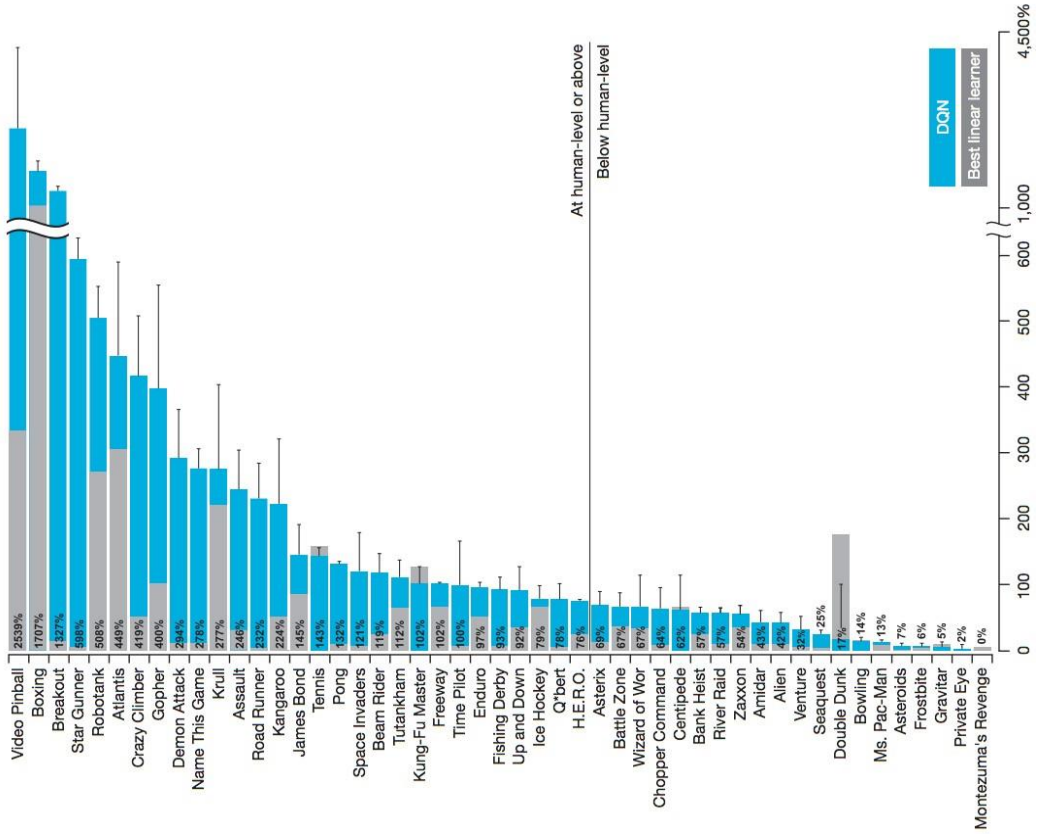


https://youtu.be/W2CAghUiofY

https://youtu.be/TmPfTpjtdgg

Breakthrough achievements ++: DeepRL

- DQ outperformed all previous solutions in a suite of 50 Atari games
- Achieving human (expert) level performance in a large portion of the games

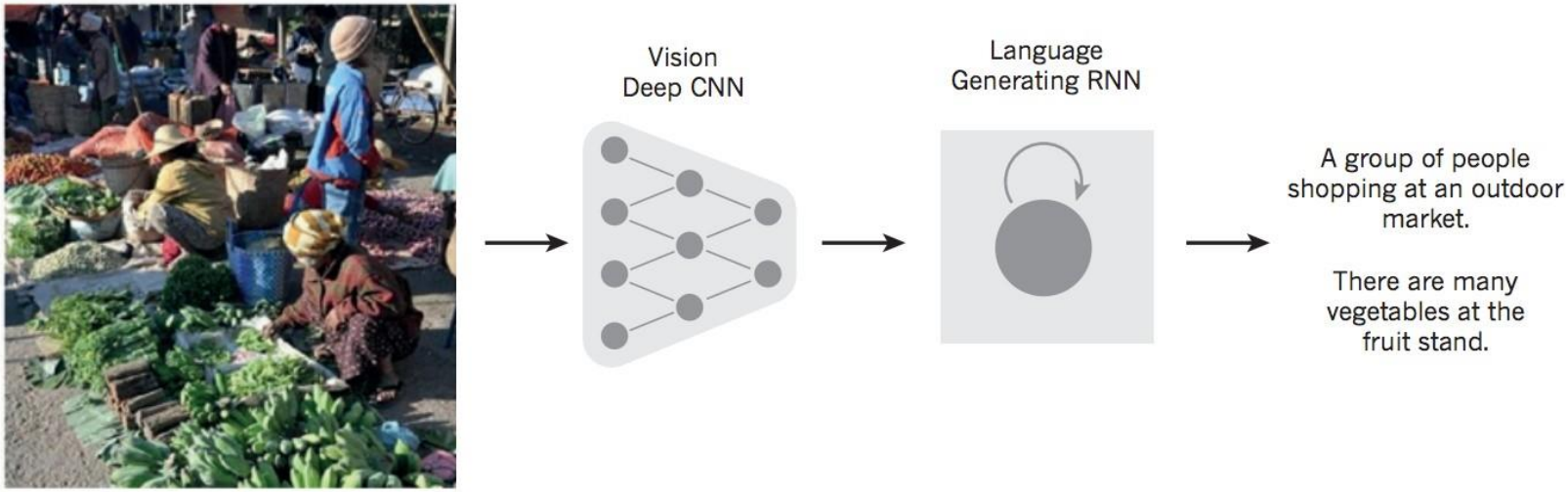


https://deepmind.com/blog/deep-reinforcement-learning/

Breakthrough achievements ++: Image Captioning

A man is riding a horse next to a building.

Breakthrough achievements ++: Image Captioning



https://pdollar.wordpress.com/2015/01/21/image-captioning/

Breakthrough achievements ++: Image Captioning

A woman is throwing a **frisbee** in a park.

A **dog** is standing on a hardwood floor.

A **stop** sign is on a road with a mountain in the background

A little **girl** sitting on a bed with a teddy bear.

A group of **people** sitting on a boat in the water.

A giraffe standing in a forest with **trees** in the background.

https://pdollar.wordpress.com/2015/01/21/image-captioning/

References

- Book:
  - Ian Goodfellow, Yoshua Bengio, Aaron Courville. **Deep Learning.** MIT Press, 2016
- Overviews:
  - Yann LeCun, Yoshua Bengio & Geoffrey Hinton. **Deep learning**. Nature 521, 436–444 (28 May 2015)
  - Andrew L. Beam. Deep Learning 101 - Part 1: History and Background, Blog post, Feb 23, 2017

- Breakthrough:
  - Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. **ImageNet Classification with Deep Convolutional Neural Networks.** Advances in Neural Information Processing Systems 25 (NIPS 2012)
  - Yaniv Tagiman, Ming Yang, Marc Aurelio Ranzato, Lior Wolf. **DeepFace: Closing the Gap to Human-Level Performance in Face V erification**, CVPR 2014
  - Volodymyr Mnih, et al. **Human-level Control through Deep Reinforcement Learning** In Nature, 518: 529–533, 2015.
  - Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya K hosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. **ImageNet Large Scale Visual Recognition Challenge.** IJCV, 2015.
  - Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent V anhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. **Deep Neural Networks for Acoustic Modeling in Speech Re cognition: The Shared Views of Four Research Groups.** IEEE Signal Processing Magazine, Vol 29(6):82 - 97, 2012
  - https://pdollar.wordpress.com/2015/01/21/image-captioning/