greatlearning
Learning for Life

Introduction to Neural Networks

Course Plan

- History
- Tensorflow and Keras - workflow
- Building Blocks of NN/Deep Learning
  - Feed forward
  - Back propagation
  - Fully connected layer
  - Activation functions
  - Softmax function
  - Cross-entropy loss
  - Optimization functions
  - Learning Rate
  - Batch normalization
- Hyper-parameters in deep learning
- Case Study

Sources

A lot of the material has been gratefully collected from:

- **http://cs231n.stanford.edu/**
- https://devblogs.nvidia.com/parallelforall/deep-learning-nutshell-history-training/
- https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html
- https://research.fb.com/learning-to-segment/
- https://research.fb.com/deep-learning-tutorial-at-cvpr-2014/
- https://www.cs.ox.ac.uk/people/nando.defreitas/machinelearning/practicals/practical4.pdf
- http://torch.ch/docs/developer-docs.html
- https://github.com/torch/nn/blob/31d7d2bc86a914e2a9e6b3874c497c60517dc853/doc/module.md
- https://web.stanford.edu/group/pdplab/pdphandbook/handbookch6.html
- http://neuralnetworksanddeeplearning.com/chap2.html

Brief History

greatlearning
Learning for Life

A bit of history:

Hubel & Wiesel,

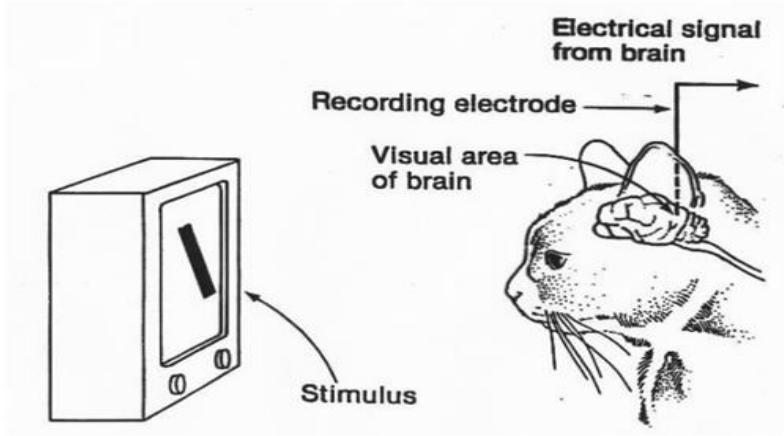
1959

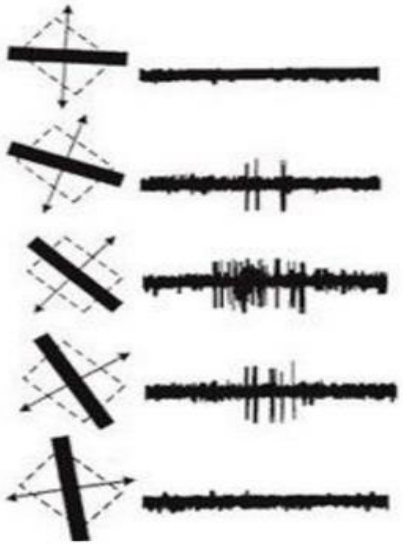
RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX

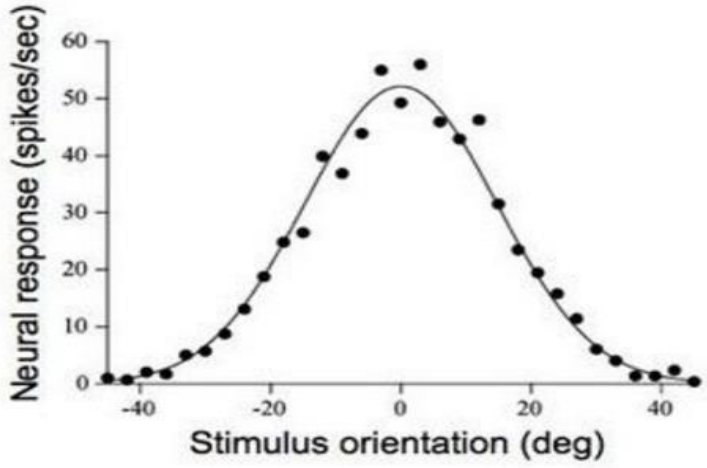
1962

RECEPTIVE FIELDS, BINOCULAR INTERACTION AND FUNCTIONAL ARCHITECTURE IN THE CAT'S VISUAL CORTEX

1968...





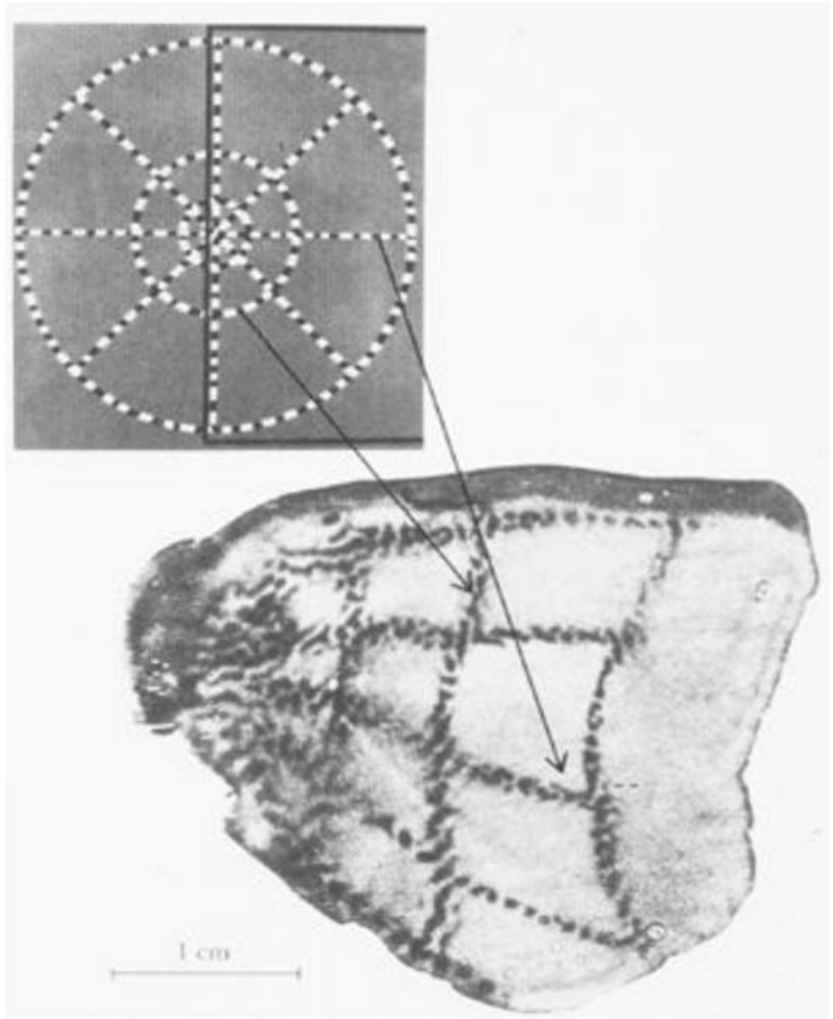


Sourced from: Receptive fields of single Neurons in the Cat's Striate Cortex by D.H.Hubel and T.N.Wiesel

Brief History

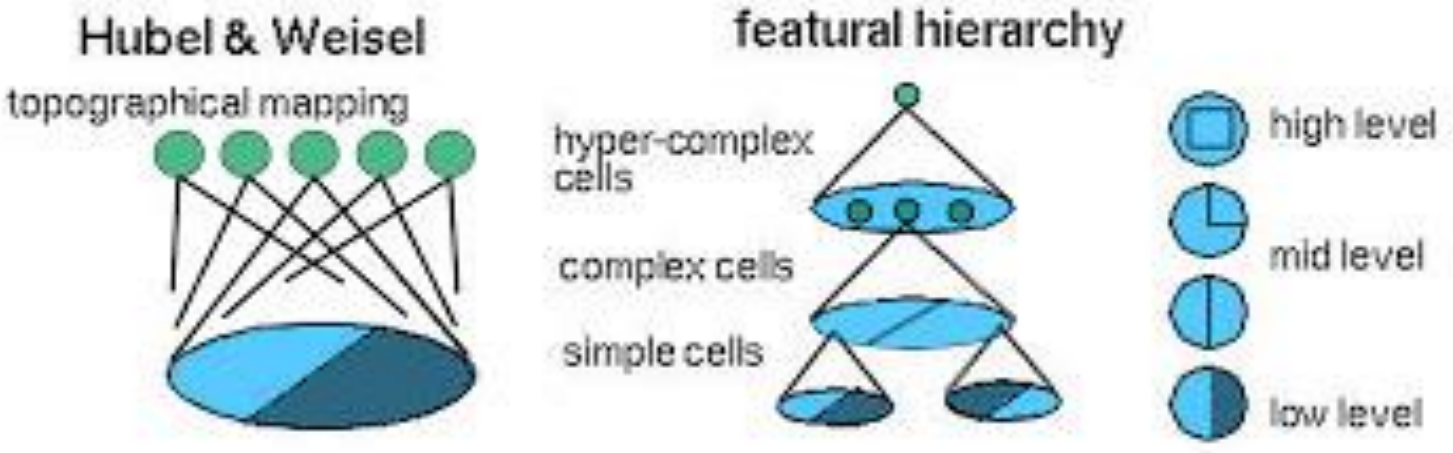
A bit of history

Topographical mapping in the cortex:
nearby cells in cortex represented
nearby regions in the visual field

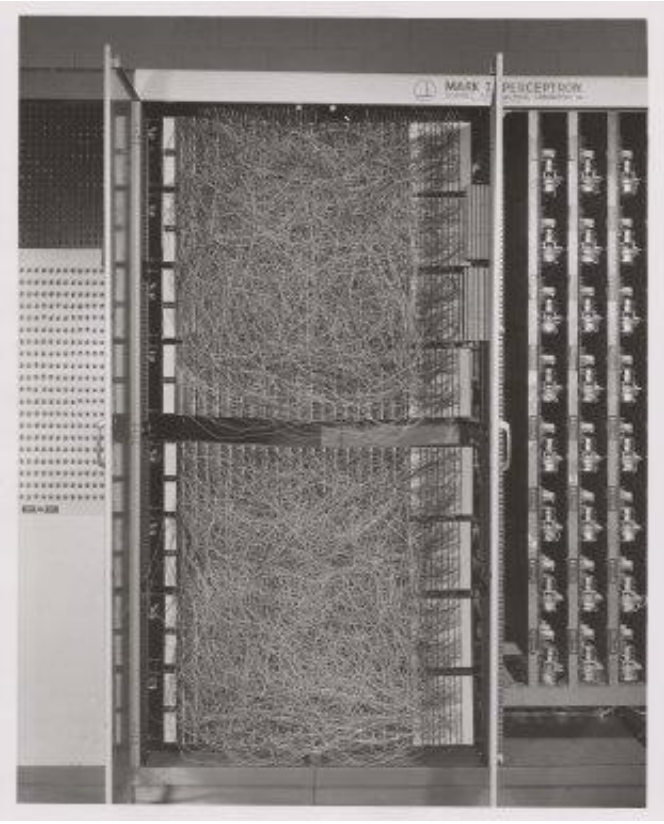


Sourced from: Receptive fields of single Neurons in the Cat's Striate Cortex by D.H.Hubel and T.N.Wiesel

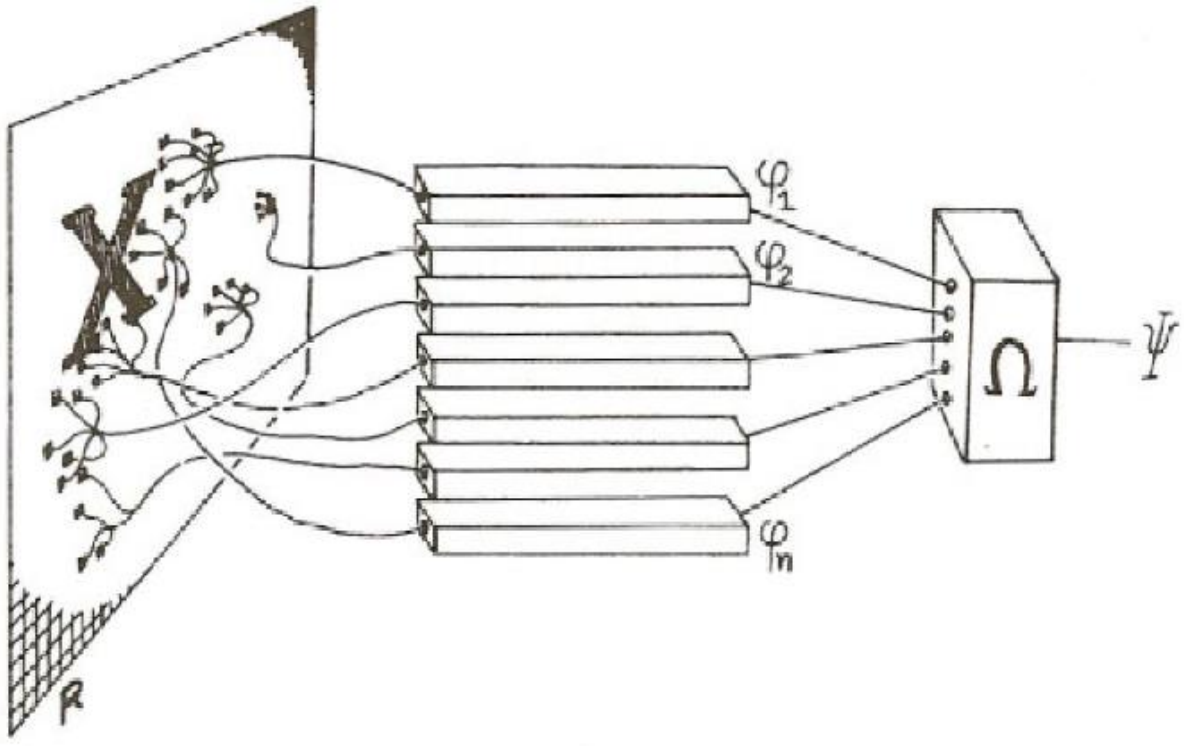
Hierarchical organization



Brief History – Mark I Perceptron – 1958



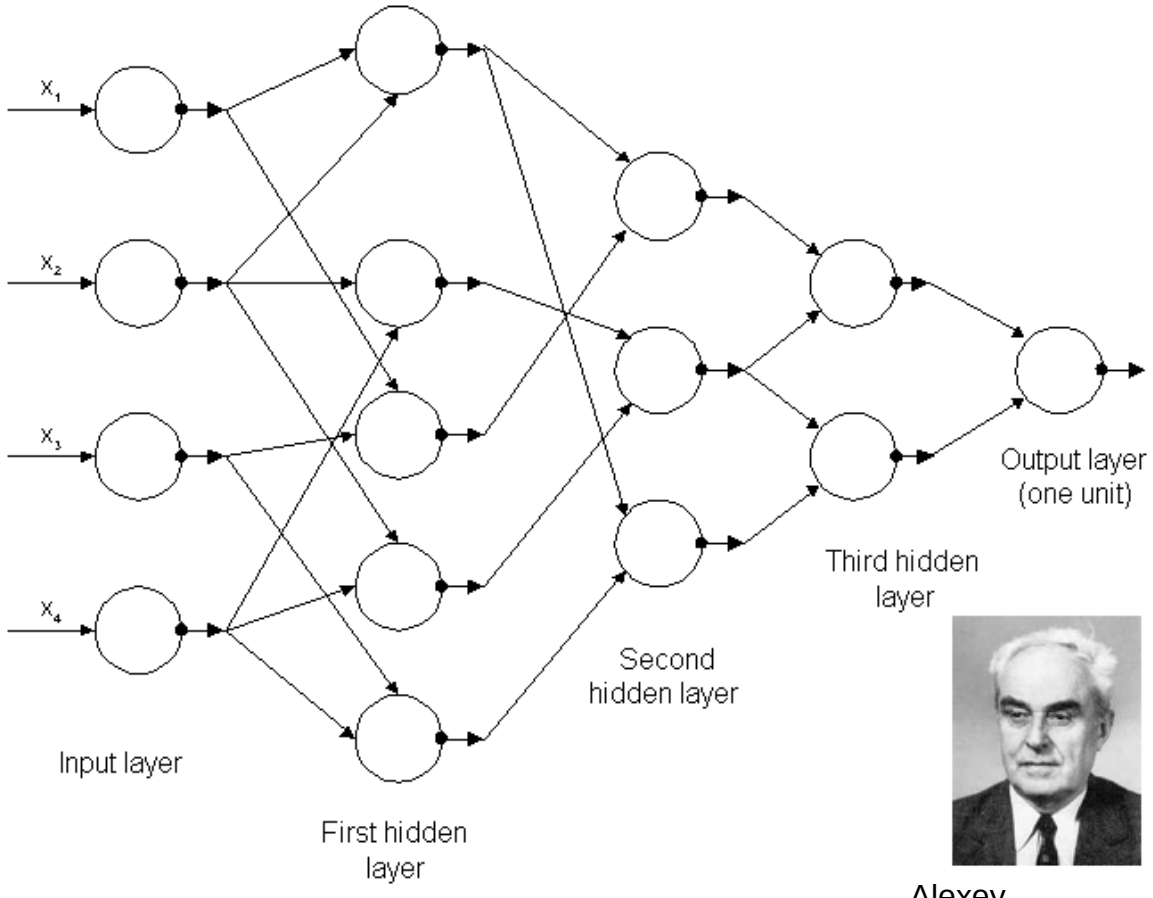
https://en.wikipedia.org/wiki/Perceptron



Perceptrons by M. L Minsky and S. Papert, 1969

Brief History – The First Deep Networks

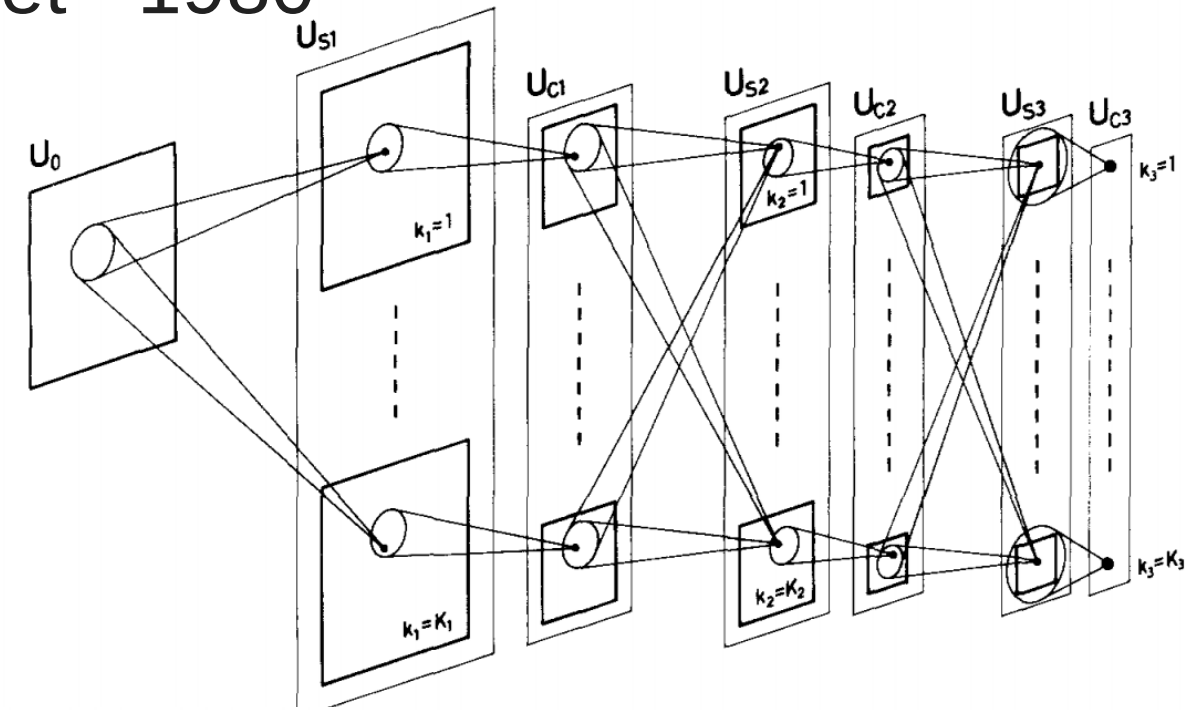
- Perceptron: single layer 1960s
- Multiple layers of non-linear features - Ivakhnenko and Lapa in 1965
- Thin but deep models with polynomial activation functions
- They did not use backpropagation

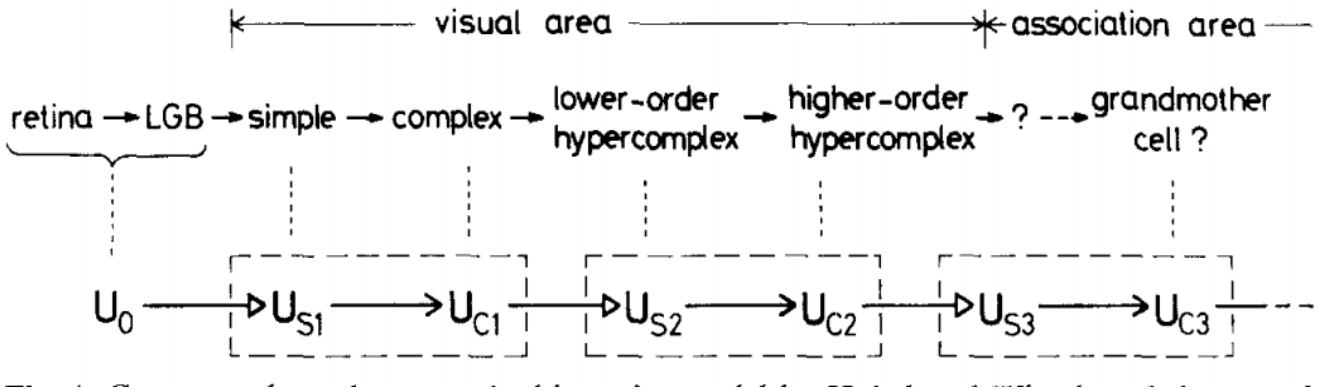


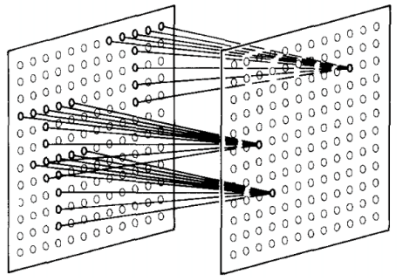
Alexey Ivakhnenko

Brief History – The First ConvNet - 1980

- Neocognitron: multiple convolutional and pooling layers similar to modern networks, but the network was trained by using a reinforcement scheme
- Did not still use backpropagation
- Translational invariant







Kunihiko Fukushima

Image Credits : Google

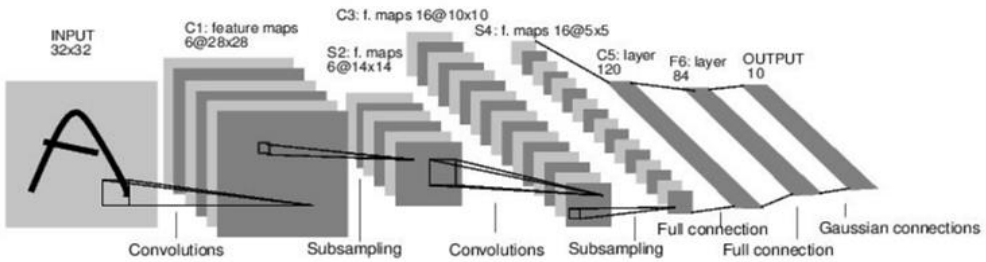
Brief History

greatlearning
Learning for Life

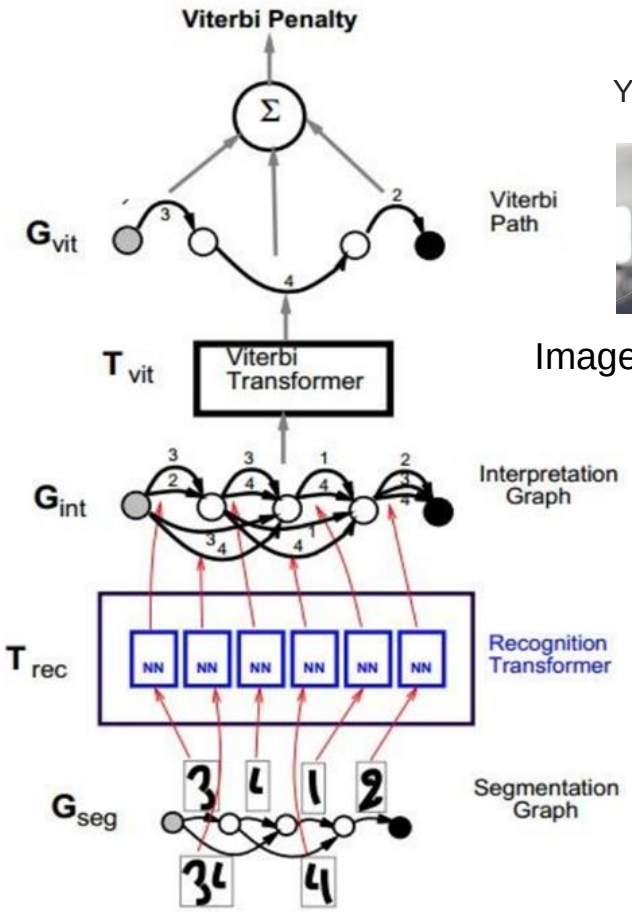
A bit of history:

Gradient-based learning applied to document recognition

[LeCun, Bottou, Bengio, Haffner 1998]



LeNet-5

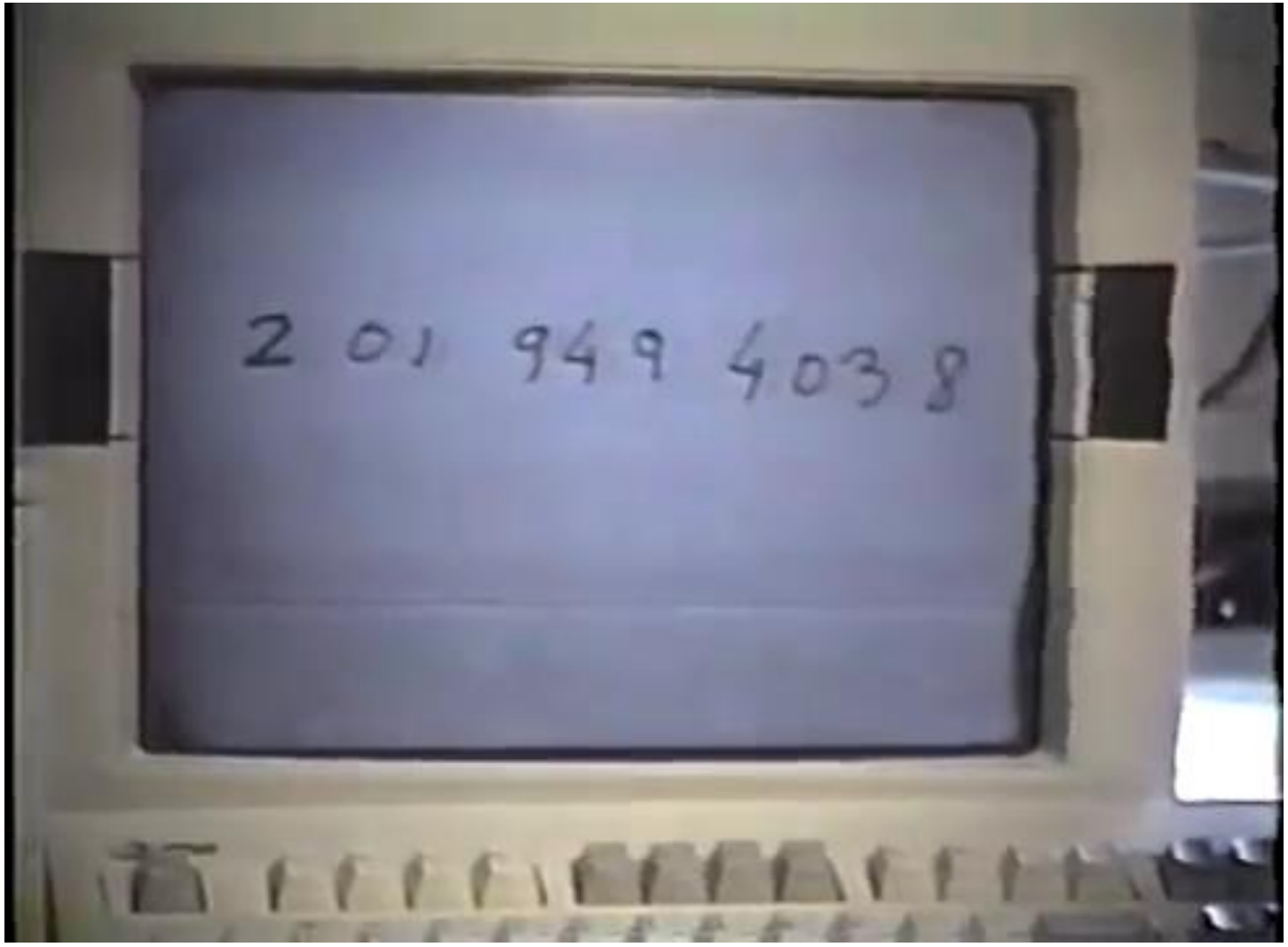


Yann LeCun

ImageSource: Google

Sourced from: GradientBased Learning Applied to Document Recognition, LeCun et al. (1998)

Brief History – LeNet-5 In Action



Brief History – AI Winter

- Rapid advances led to a hype of artificial intelligence (similar to the buzz around deep learning today)
- Researchers made promises to solve AI and received lots of funding
- In the 1970s it became clear that those promises could not be kept, funding was cut dramatically
- The field of artificial intelligence dropped to near pseudo-science status
- Research became very difficult (little funding; publications almost never made it through peer review)
- Further advances such as SVMs with nice properties in terms of training, provable error bounds were preferred and took the front seat
- However, a handful of researchers continued further down this path

Brief History – AI Winter

Geoffrey Hinton: University of Toronto & Google

Yann LeCun: New York University & Facebook

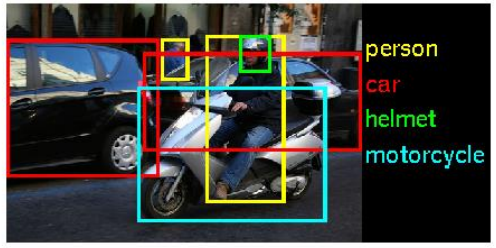
Yoshua Bengio: University of Montreal

Jurgen Schmidhuber: Swiss AI Lab & NNAISENSE

Andrew Ng: Stanford &Baidu

Brief History – The Tipping Point

- 2012 ILSVRC: ImageNet Large-Scale Visual Recognition Challenge – Annual World Cup of Computer Vision
- More than a million training images and 1000 categories



ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky
University of Toronto
kriz@cs.utoronto.ca

Ilya Sutskever
University of Toronto
ilya@cs.utoronto.ca

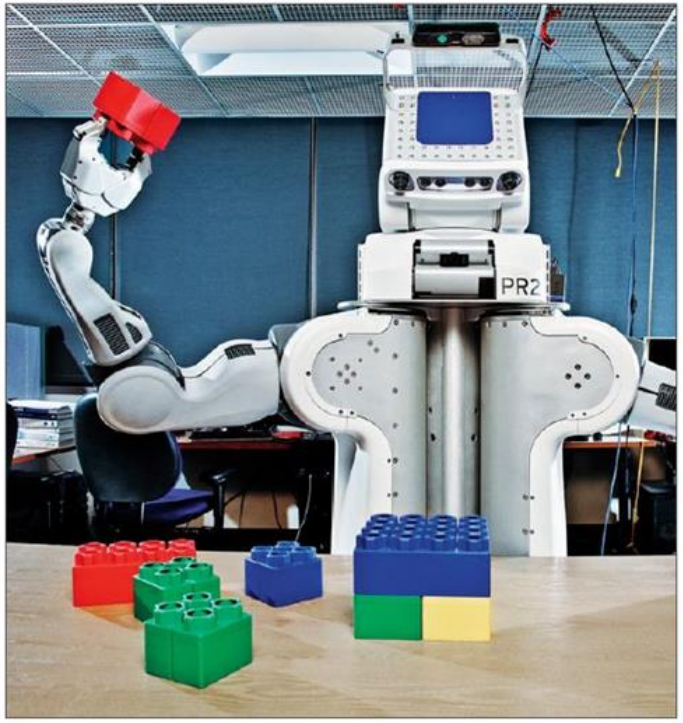
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Brief History – The Tipping Point

- Reported 15.4% Top 5 error rate. The next best entry achieved an error of 26.2%
- > 8000 citations (last year), by today >19000!
- The coming out party for CNNs in the computer vision community
- Shocked the computer vision community. Trained end-to-end on raw pixels, without using any feature engineering methods
- From here it was apparent that deep learning would take over computer vision and that other methods would not be able to catch up

Deep Learning – Today – One Net To Rule Them All

- Deep Learning == AI
- Solves problems previously unsolvable



greatlearning
Learning for Life

Thank you!