TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

6.2 Deep Learning

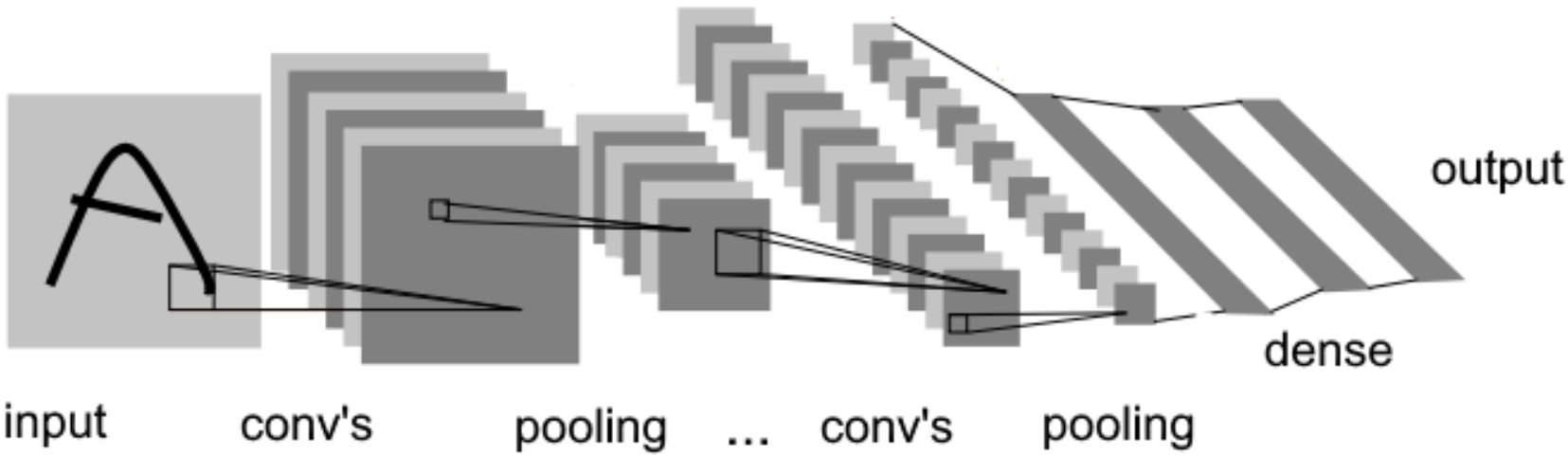
Machine Learning 1: Foundations

Marius Kloft *(TUK)*

Recap

CNNs

- learn a prediction model and an image representation at the same time



What is Deep Learning?

Many CNNs that are used for image processing are really deep: they consist of **8 layers** or more (sometimes >1000).

$\Rightarrow$ We speak of deep learning.

Definition

An ANN with many layers
(usually eight or more) is called **deep neural network**.

One of the First Deep-learning Papers: AlexNet

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

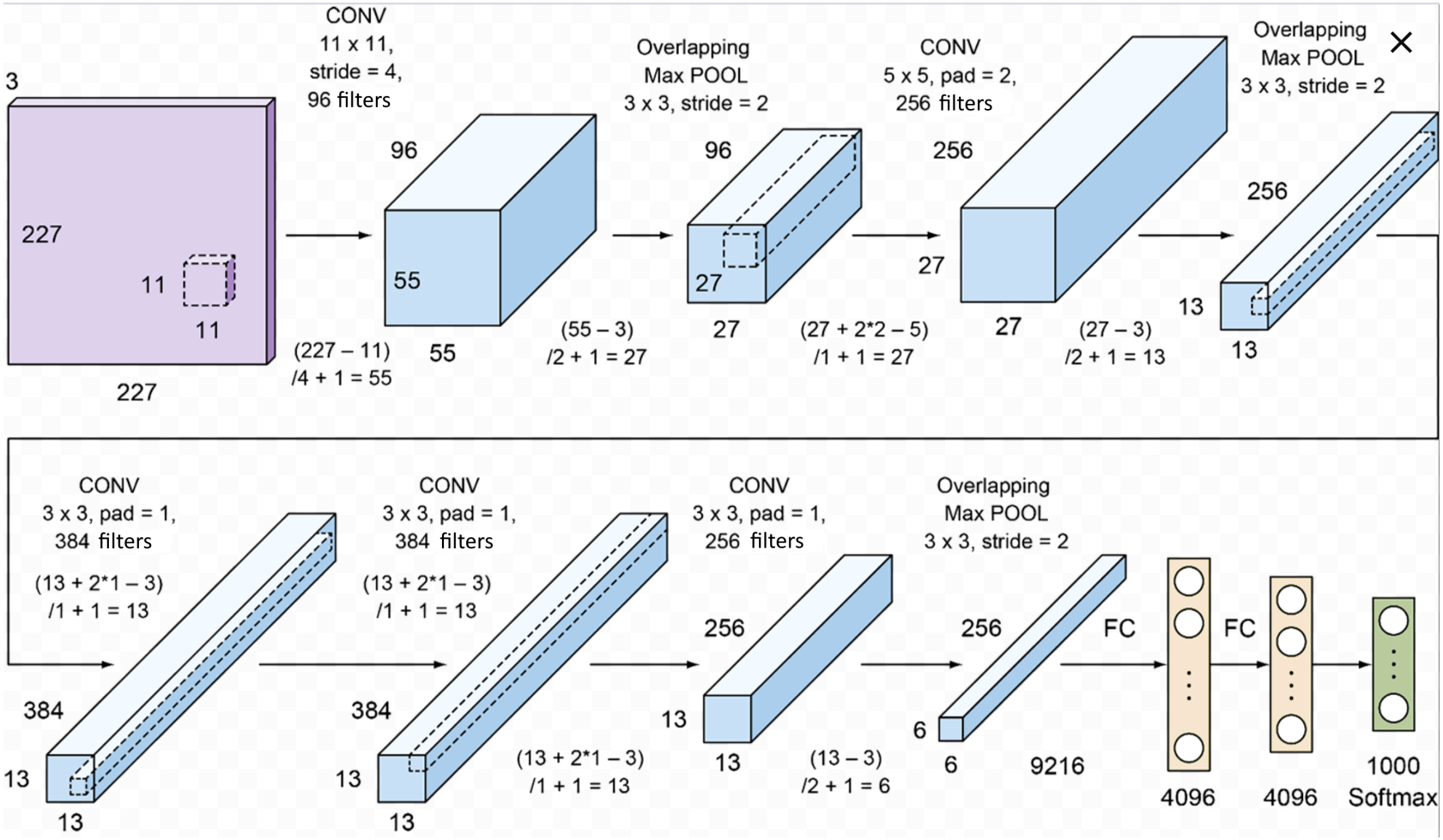
Geoffrey E. Hinton
University of Toronto
hinton@cs.utoronto.ca

Abstract

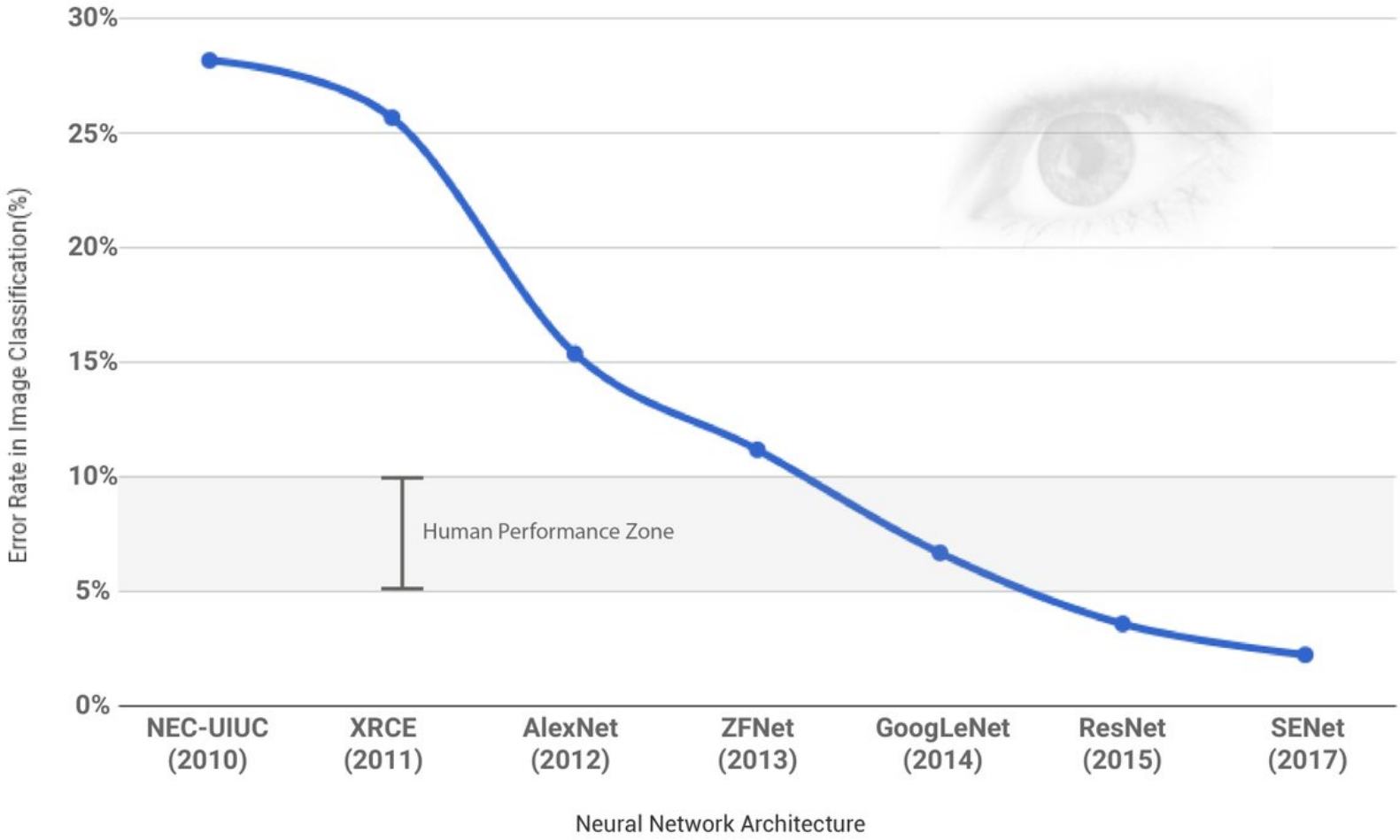
We trained a large, deep convolutional neural network to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and 17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation. To reduce overfitting in the fully-connected layers we employed a recently-developed regularization method called "dropout" that proved to be very effective. We also entered a variant of this model in the ILSVRC-2012 competition and achieved a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry.

Alex Krizhevsky et al.: Imagenet classification with deep convolutional neural networks, **NIPS** 2012

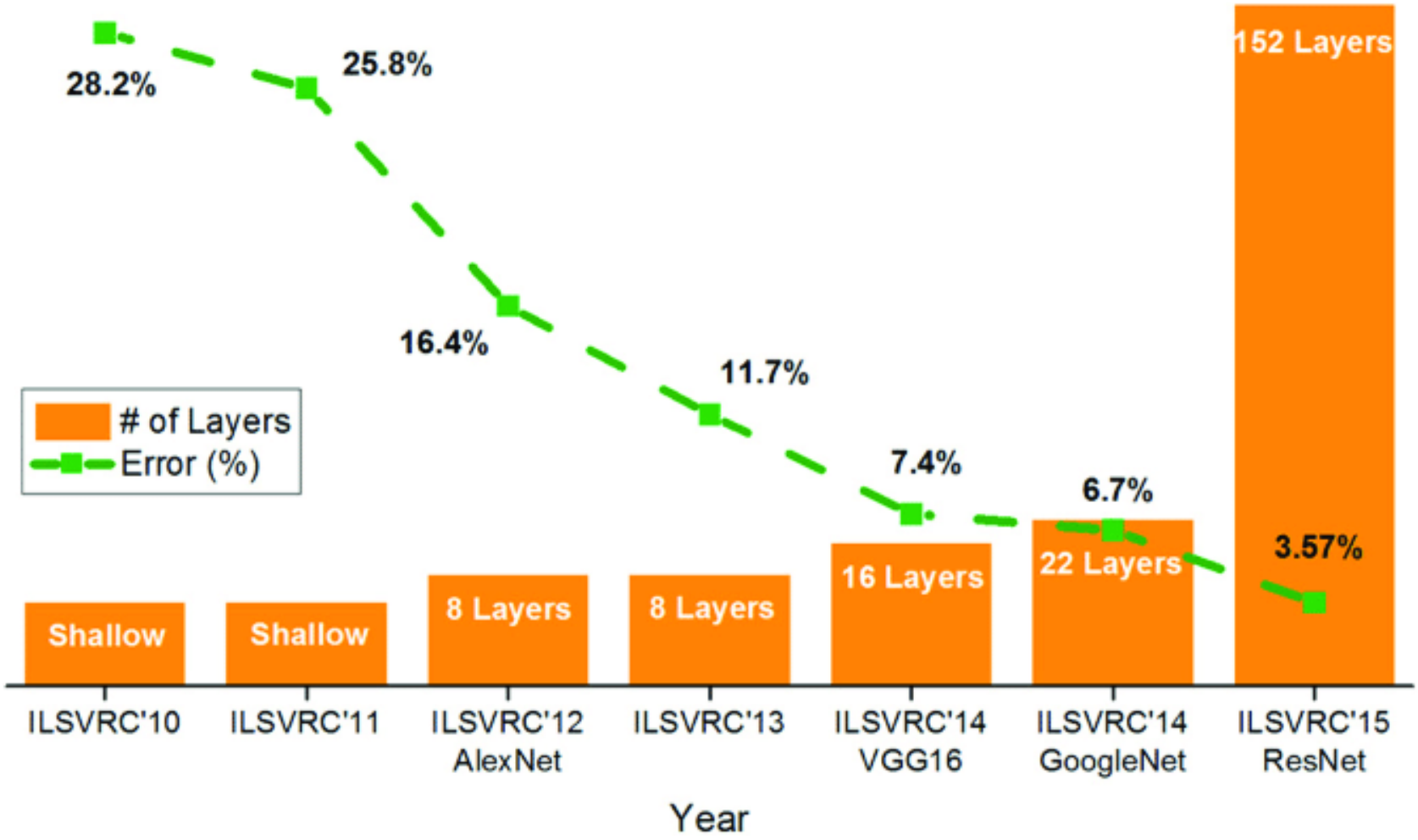
AlexNet Architecture



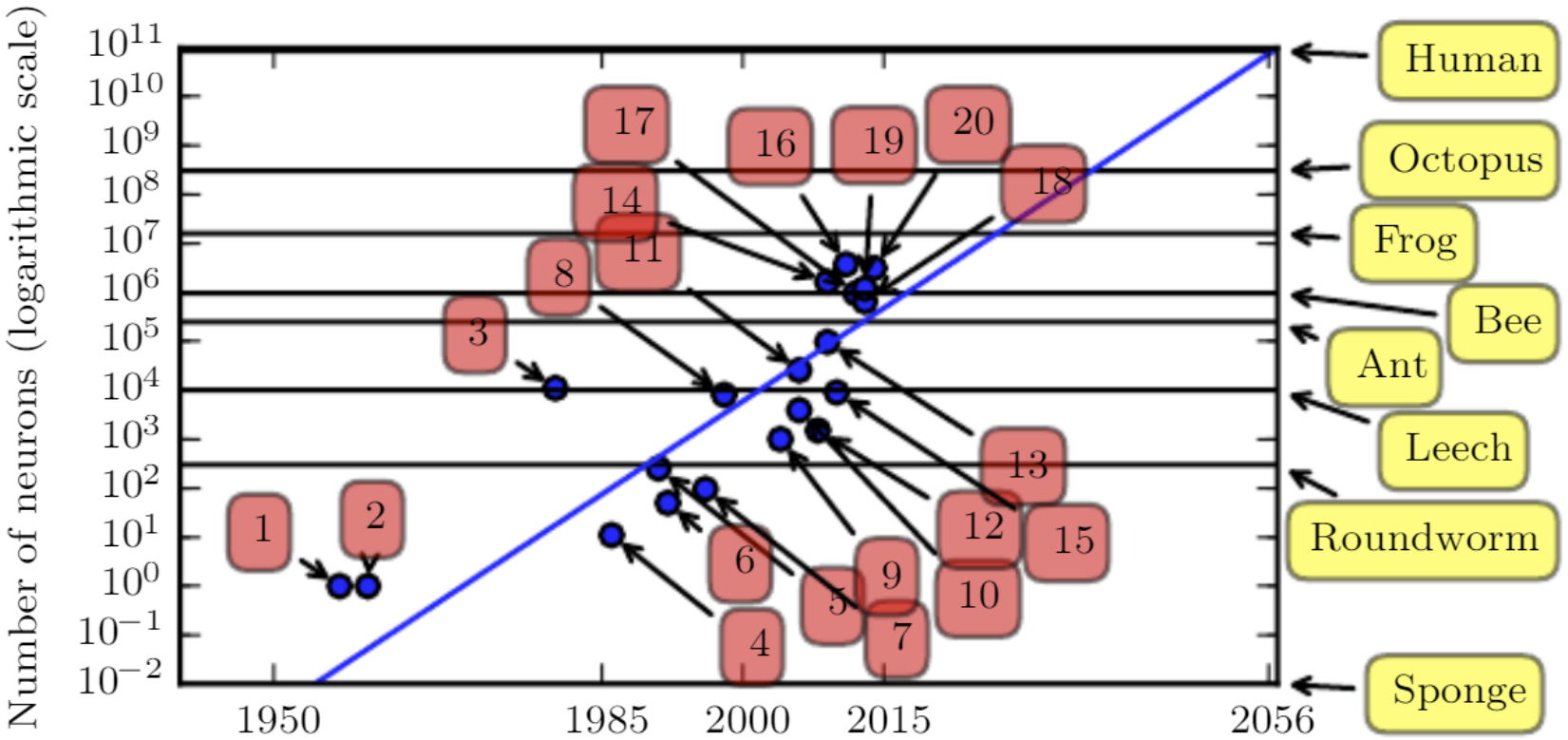
Super-human Performance of DL in Image Classification



Number of Layers Increasing



How Deep Can We Go?



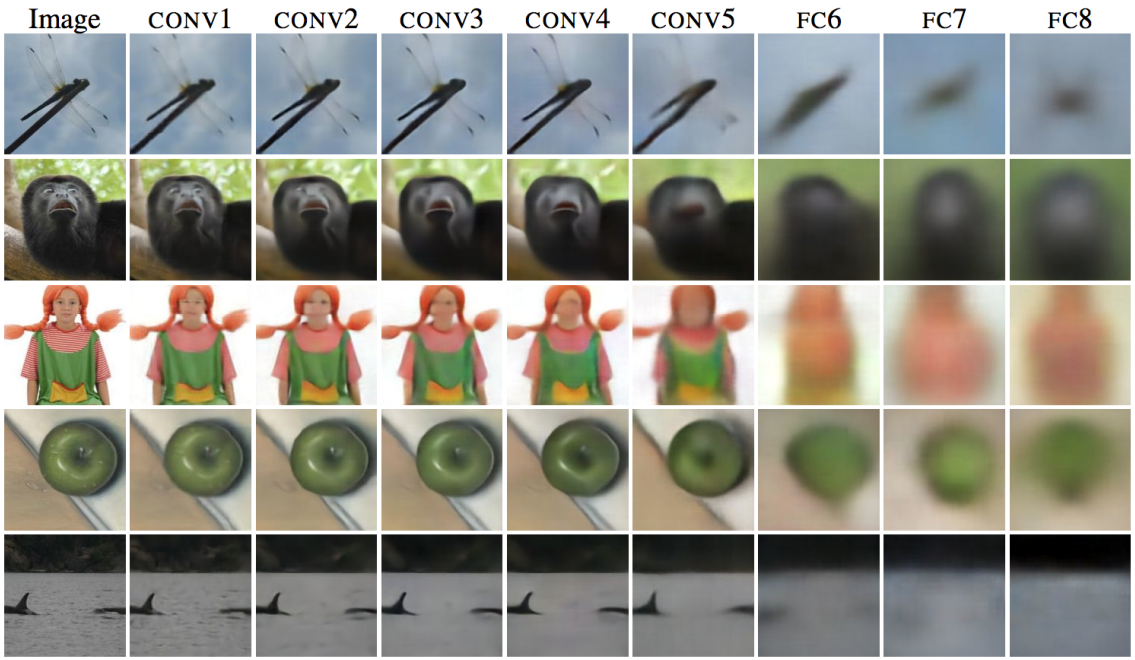
Goodfellow et al., Fig. 1.11

Legend

1. Perceptron (Rosenblatt, 1958, 1962)
2. Adaptive linear element (Widrow and Hoff, 1960)
3. Neocognitron (Fukushima, 1980)
4. Early back-propagation network (Rumelhart *et al.*, 1986b)
5. Recurrent neural network for speech recognition (Robinson and Fallside, 1991)
6. Multilayer perceptron for speech recognition (Bengio *et al.*, 1991)
7. Mean field sigmoid belief network (Saul *et al.*, 1996)
8. LeNet-5 (LeCun *et al.*, 1998b)
9. Echo state network (Jaeger and Haas, 2004)
10. Deep belief network (Hinton *et al.*, 2006)
11. GPU-accelerated convolutional network (Chellapilla *et al.*, 2006)
12. Deep Boltzmann machine (Salakhutdinov and Hinton, 2009a)
13. GPU-accelerated deep belief network (Raina *et al.*, 2009)
14. Unsupervised convolutional network (Jarrett *et al.*, 2009)
15. GPU-accelerated multilayer perceptron (Ciresan *et al.*, 2010)
16. OMP-1 network (Coates and Ng, 2011)
17. Distributed autoencoder (Le *et al.*, 2012)
18. Multi-GPU convolutional network (Krizhevsky *et al.*, 2012)
19. COTS HPC unsupervised convolutional network (Coates *et al.*, 2013)
20. GoogLeNet (Szegedy *et al.*, 2014a)

How Deep Learning Works

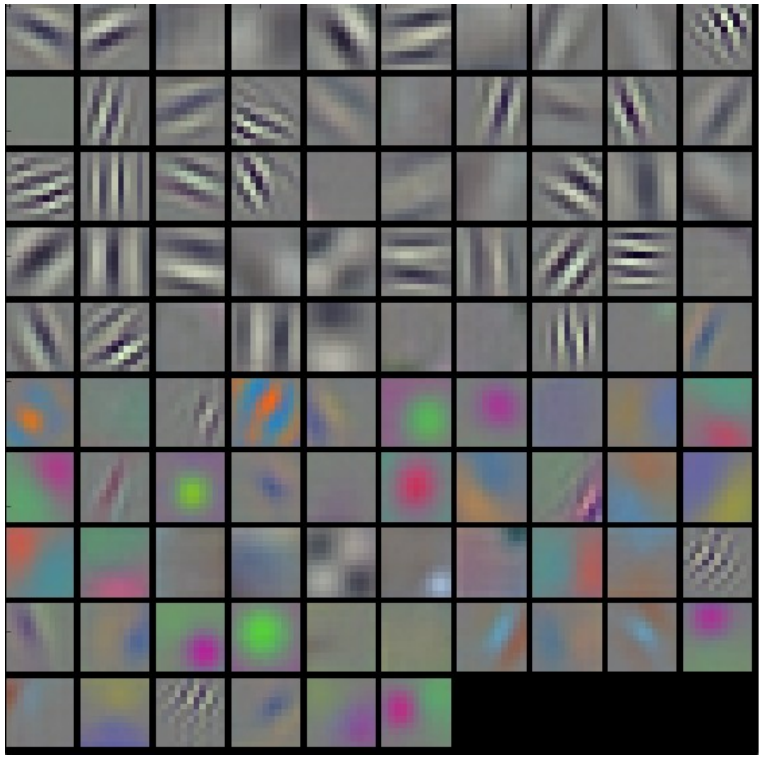
- A central concept of deep learning is that lower layers extract basic features (e.g., edge detectors), while higher layers compose them to complex features (complex cells, invariant object detectors).
  - This is in rough correspondence with our understanding of how the visual cortex processes images.



Reconstructions from different layers of AlexNet.

Further Reading

More very interesting aspects of visualization (multiple methods and examples) can be found at `http://cs231n.github.io/understanding-cnn/`, where we find also a nice visualization of a typical, well-learned filter by a CNN (1st layer left, 2nd layer right):



We Stop Here

There is a lot more to know about deep learning (some of it in ML2):

- Autoencoders (later in this course)
- Residual neural networks
- Deep Boltzmann machines
- Deep belief networks
- Recurrent neural networks (see TUK course *Very Deep Learning*)
- Deep generative models (ML2)
- Applications in computer vision
  - Didier Stricker's Courses at TUK
- Various kinds of other applications:
  - e.g., AI art, speech recognition, and natural language processing (see Very Deep Learning)

Conclusion

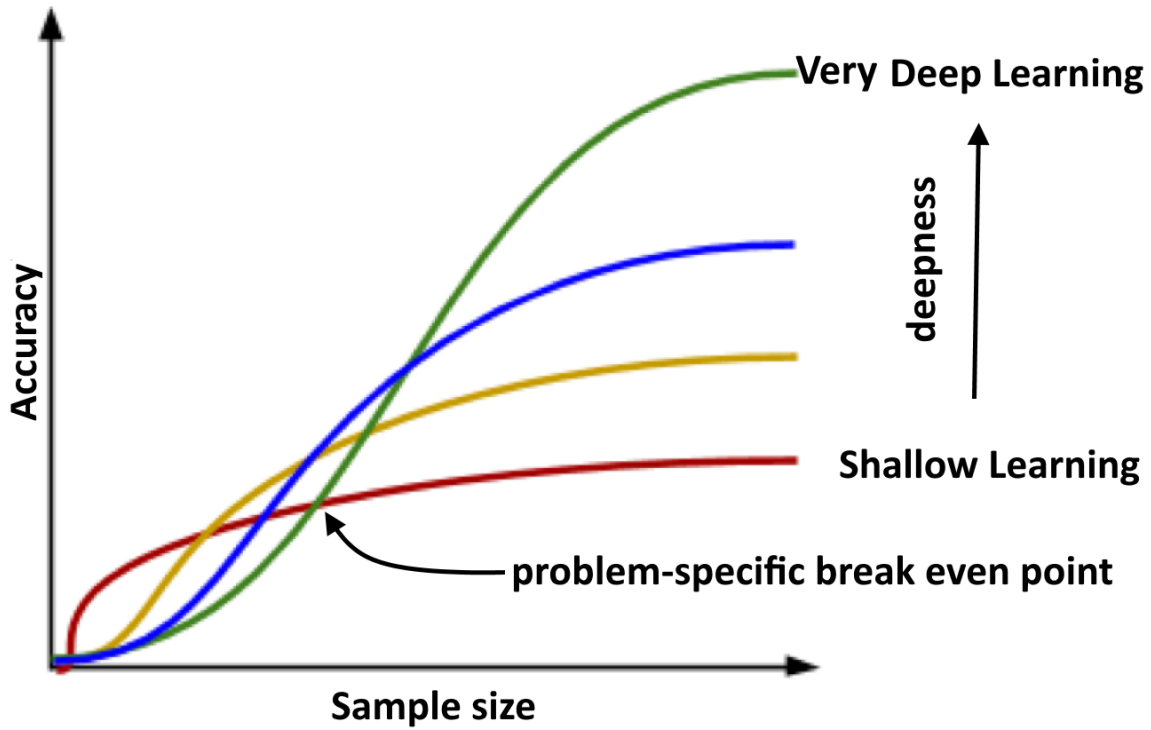
Deep learning

- ANNs using eight or more layers

Has led to drastic improvements in many applications

- e.g., computer vision

Deep vs. shallow learning:



References I

A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. The MIT Press, 2016, ISBN: 0262035618.

ML Group