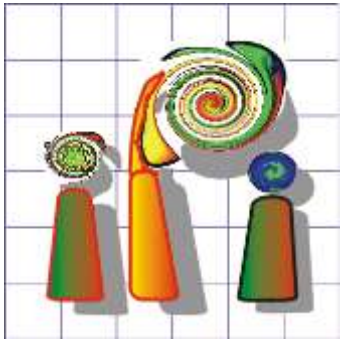
Deep learning – what it can and cannot do (with emphasize on image exploitation)



Christian Heipke

IPI - Institut for Photogrammetry and GeoInformation

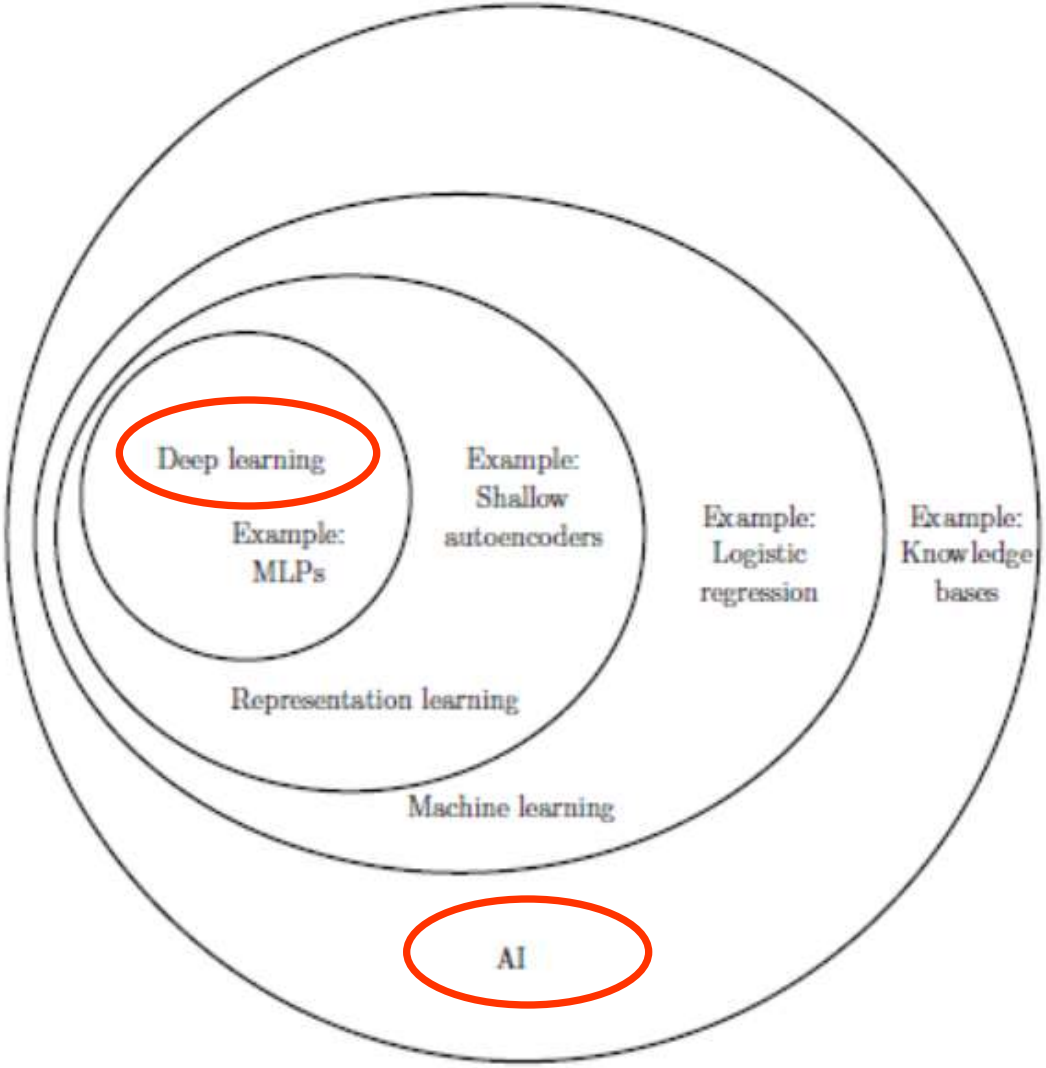
Leibniz Universität Hannover

President - ISPRS

YOLO: You Only Look Once (version 2) – a Convolutional Neural Network (CNN)

J. Redmon, S. Divvalay, R. Girshick, A. Farhadiy, Uni. Washington, 2016

Deep learning – a part of AI



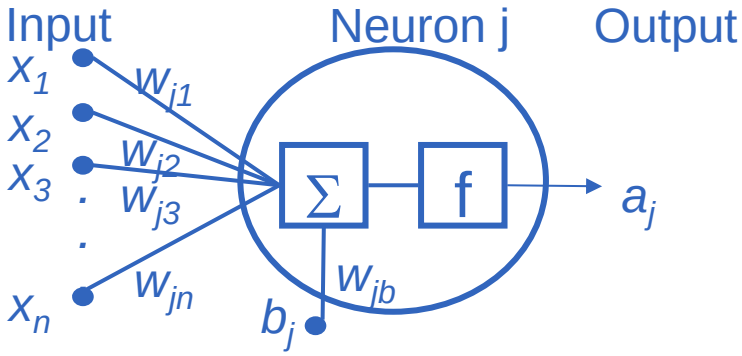
… for problems which are

- easy for people to perform, but
- difficult to describe formally

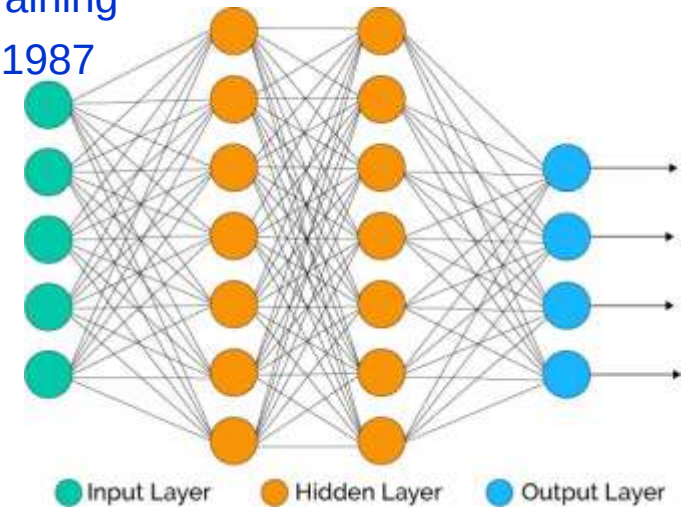
© Goodfellow et al., 2016

Development of deep learning

- Cybernetics („the perceptron“)
  - f: a **linear** function
  - McCulloch, Pitts, **1943,** Rosenblatt, 1958/62



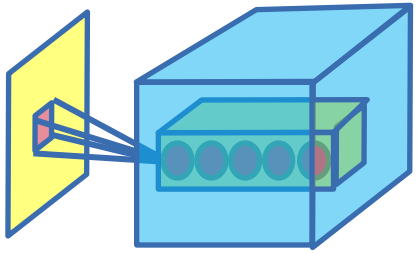
- Connectionism, ANN, MLPs
  - f: a **non-linear** function (sigmoid fct., ReLU
    distributed representation, back propagation for training
  - Rumelhart et al., **1986,** Hinton et al., 1986, LeCun, 1987

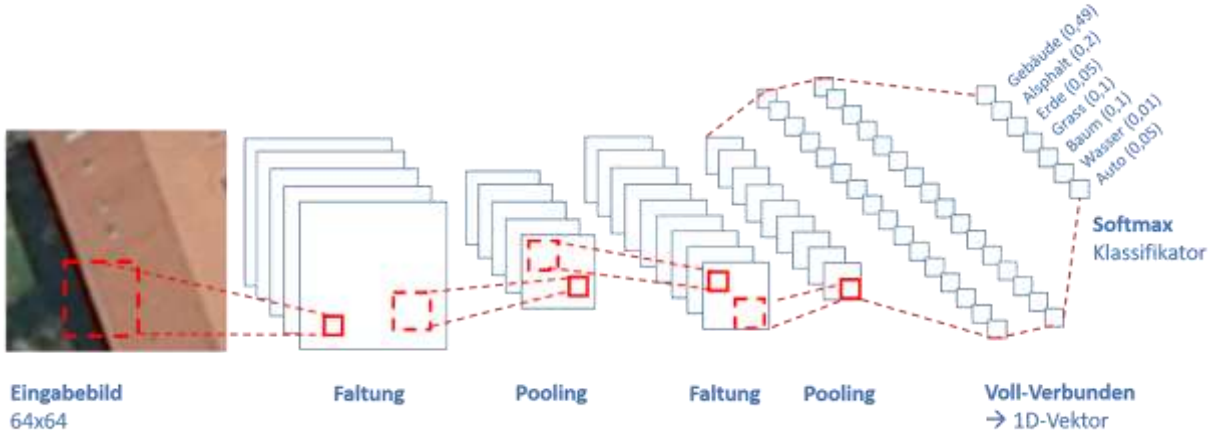


- Deep learning
  - **learn representation (features)**
  - **learn from actions** (e.g. to write SW)
  - faster computers, more training data:
    efficient training of networks with „many“ layers
  - Hinton et al., **2006**
  - Krizhevsky et al., 2012: CNN wins ILSVRC, decreasing the error by nearly 50%

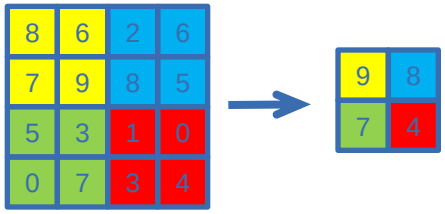
CNN: Convolutional neural networks

- convolution



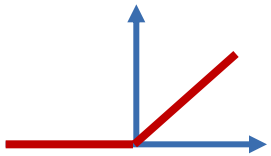


- (max-)pooling



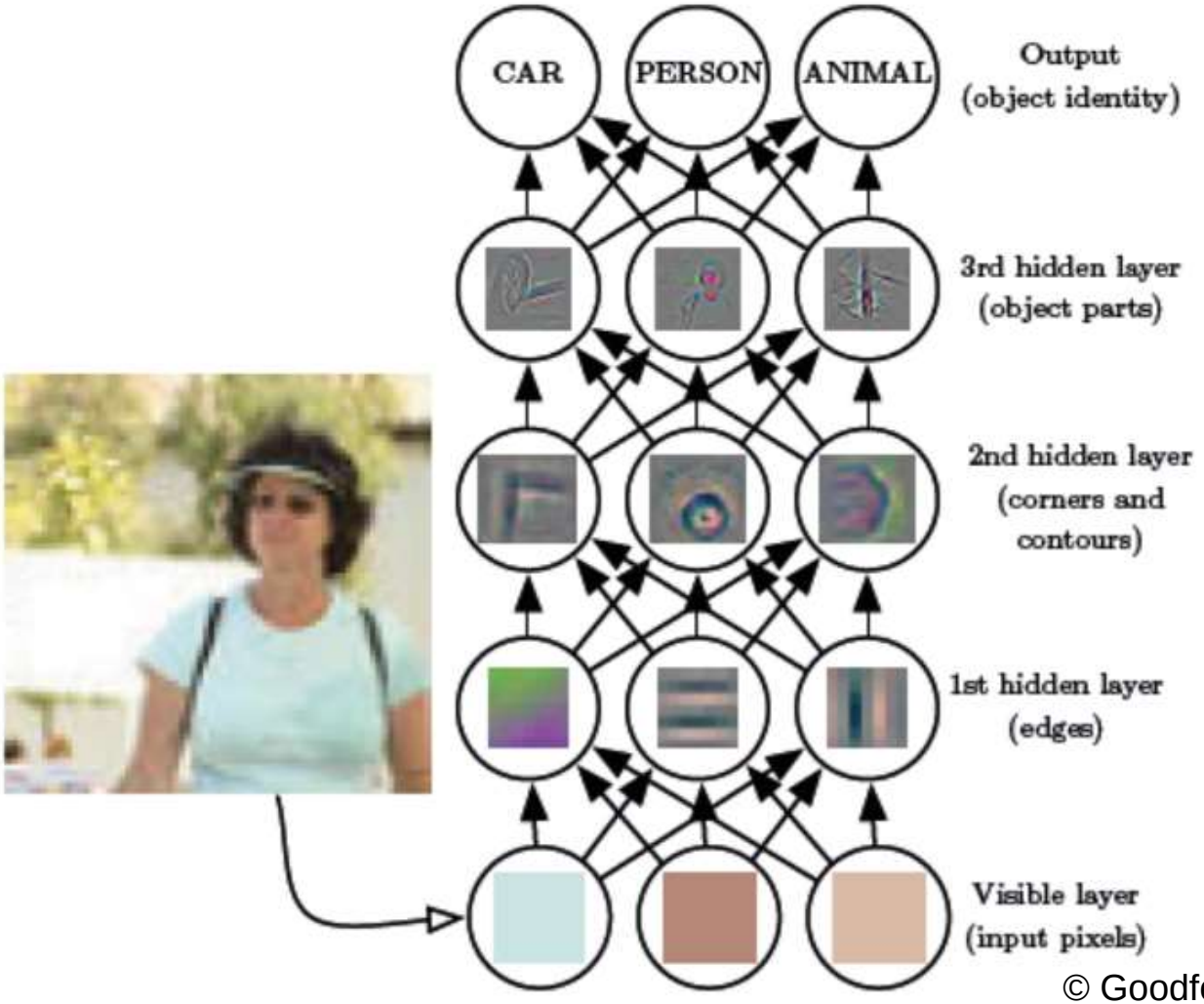
- activation function

e.g. ReLU(x) = max(0,x)

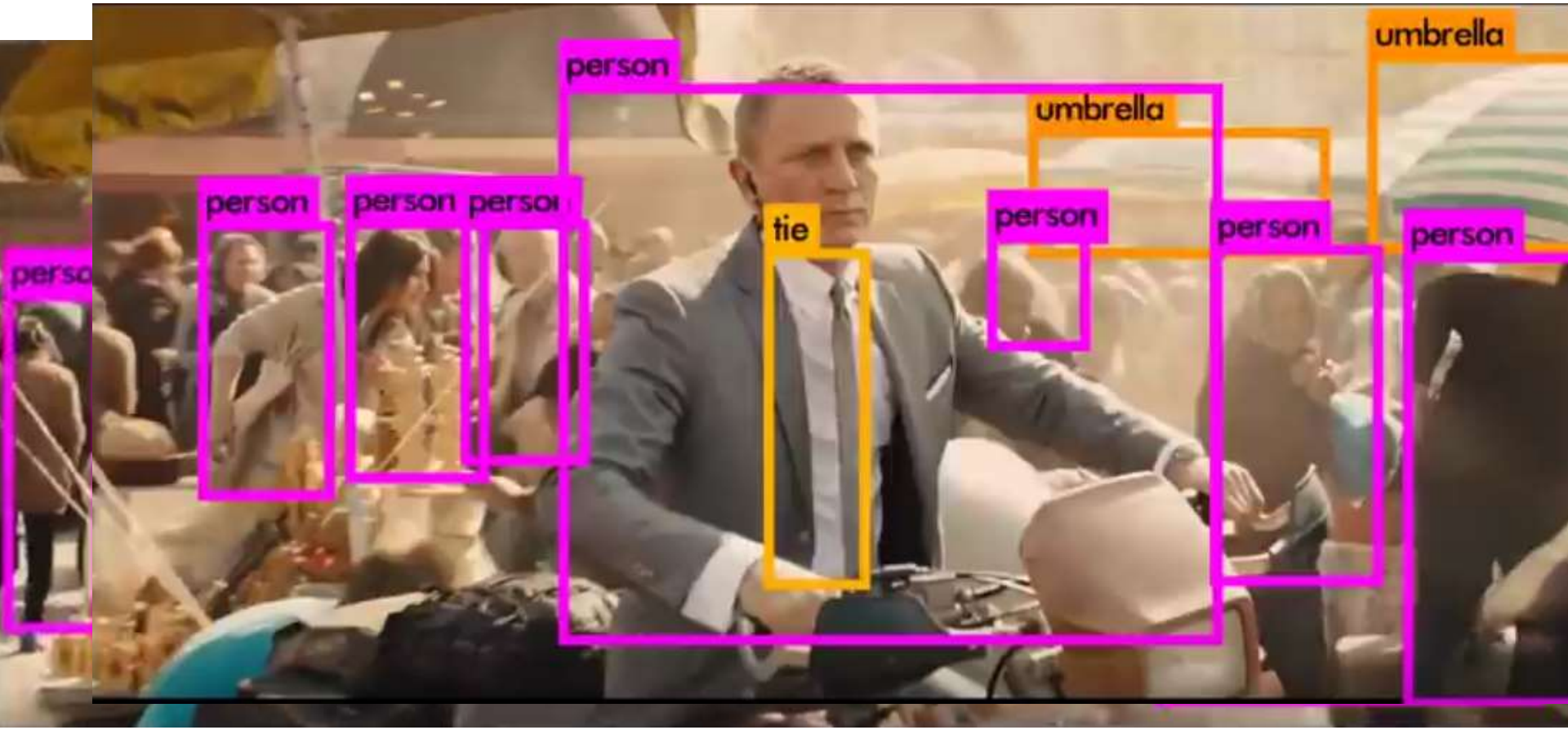


- for regularly structured data
- repeated execution of those steps based on “some “ architecture
- training for parameter learning (millions of filter coefficients)
- classification

CNN: multi-layered, „deep“ networks



© Goodfellow et al., 2016



YOLO v2: **Object detection** – state-of-the-art CNN system

% of correctly retrieved items: **70 – 80%** (real-time)

geometric accuracy not assessed

PETS[1] data set: **Object tracking** – ID switches do occur

Precision, robustness and accuracy **lower**

[1] Performance evaluation of tracking and surveillance, http://www.cvg.reading.ac.uk/PETS2009/a.html.

Potential – and limitations

- Very successful applications shown
  - Image and video analysis
  - Text and Language processing, automatic coding, …
- Limitations
  - unclear how/why deep learning (in part. CNN) performs well
    - many ad hoc parameters (network architecture, …)
    - little possibility for prediction, no self diagnosis
  - Precision and robustness need improvement for many app‘s.
  - Extremely large amount of (labelled !) training data needed (pre-trained nets can be used, but reduce adaptability)
- A current hype or a technology to stay?
  - Temporal behaviour, processing of streaming data?
  - Self organisation?

Slide credit: T. Darrell

Every picture tells a story …