Dr. Michael Vogt · 6. Dezember 2018 · Uni Siegen

Deep Learning –

Die Revolution der künstlichen Intelligenz

smiths detection

bringing technology to life

Künstliche Intelligenz in den Medien

Google AI algorithm masters ancient game of Go
Deep-learning software defeats human professional for first time.

"Deep learning is killing every problem in AI."

The AI Revolution: Why Deep Learning Is Suddenly Changing Your Life

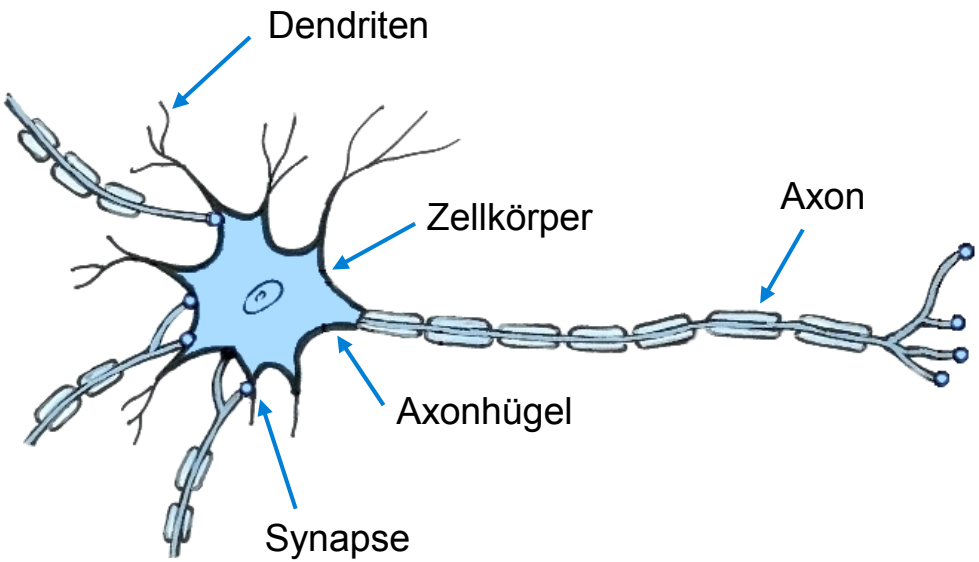
The First Person to Hack the iPhone Built a Self-Driving Car
George Hotz is taking on Google and Tesla by himself.

"Deep learning was like
ten breakthroughs at once –
and the best is still to come."

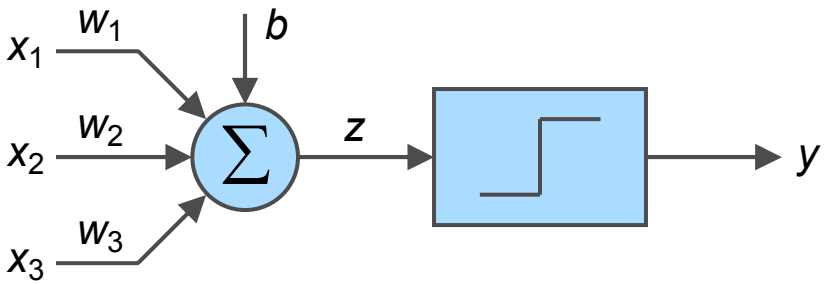
You're killing people!
Elon Musk attacks critics of self-driving cars

"The cost of a world-class deep learning expert was
about the same as a top NFL quarterback prospect"

Neuronen



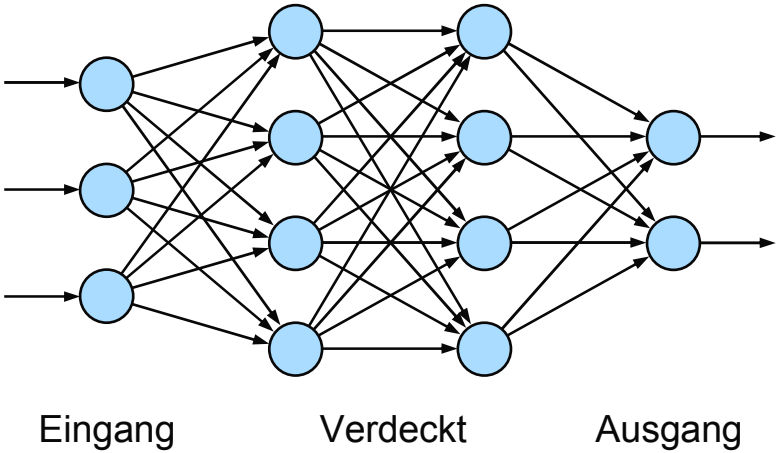
McCulloch-Pitts-Modell (1943)

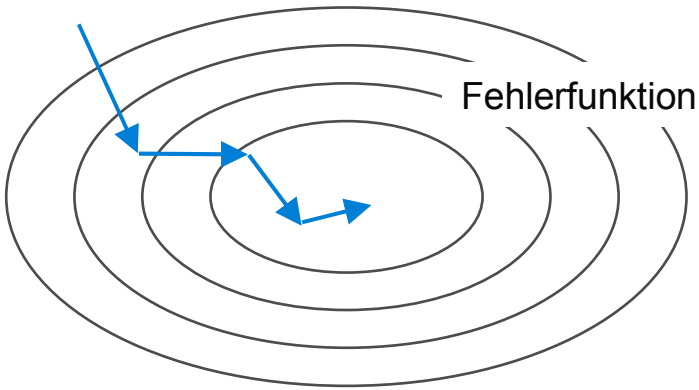


Künstliche neuronale Netze

Feed-Forward-Netze

- Nur Vorwärts-Verbindungen
- Perzeptron (Rosenblatt, 1958)
- „Eine Schicht ist genug!“ (1989)



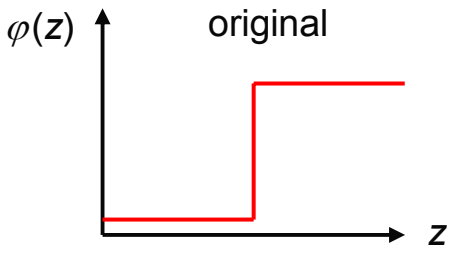


Trainieren von Netzen

- Lernen aus empirischen Daten
- Error Backpropagation & Co.
  (→ Gradientenabstieg)

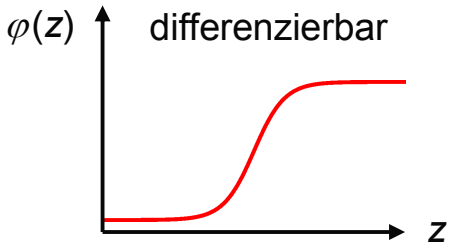
Überwachtes Lernen

Aufgabe: Approximiere $f_{\mathbf{w}}(\mathbf{x})$ auf $\{(\mathbf{x}_k, y_k)\}_{k=1}^{N}$



- Minimiere Fehler $E = \sum_k (y_k - f_{\mathbf{w}}(\mathbf{x}_k))^2$
- Gradient aus Error-Backpropagation

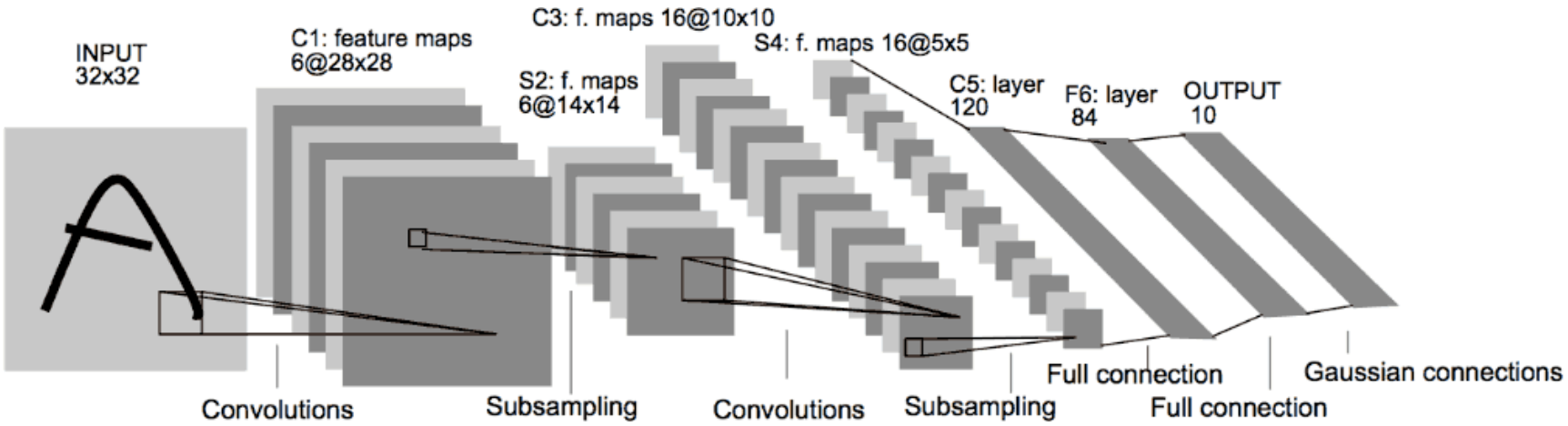
$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial f_{\mathbf{w}}} \cdot \underbrace{\frac{\partial f_{\mathbf{w}}}{\partial z}}_{=\varphi'(z)} \cdot \frac{\partial z}{\dots} \cdot \frac{\dots}{\partial w_{ij}}$$



- Parameter-Update: $w_{ij}^{\text{neu}} = w_{ij}^{\text{alt}} - \eta \cdot \frac{\partial E^{\text{alt}}}{\partial w_{ij}}$

Maßnahmen: Momentum, Regularisierung, Early-Stopping, …

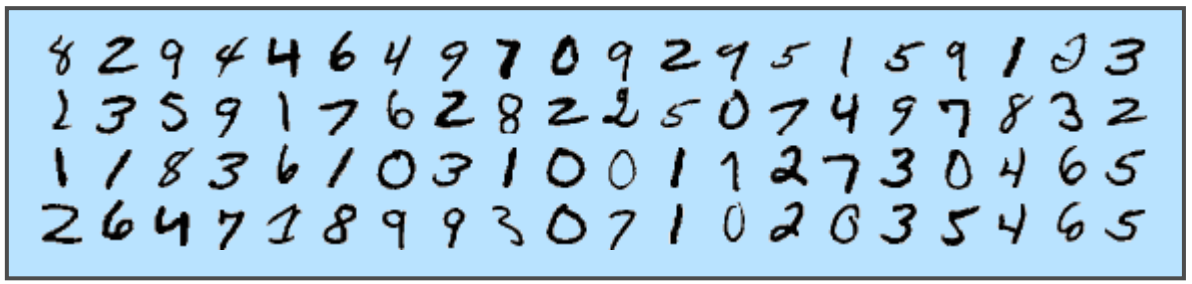
LeNet (1989)



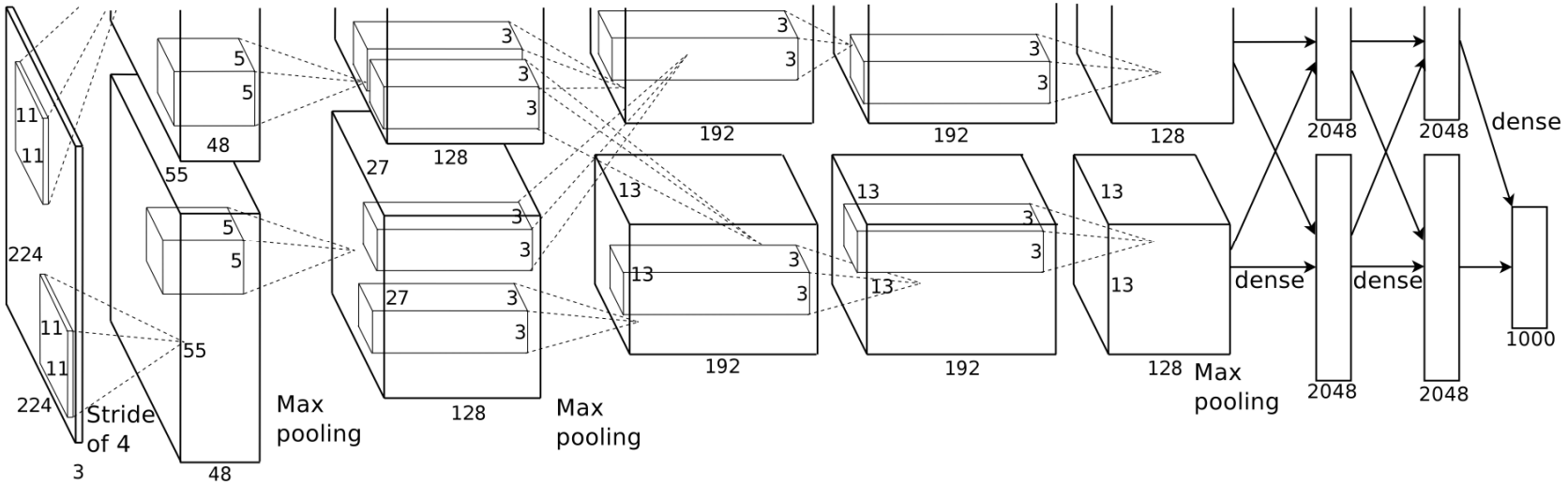
LeCun et al. (1998)

MNIST-Datensatz:

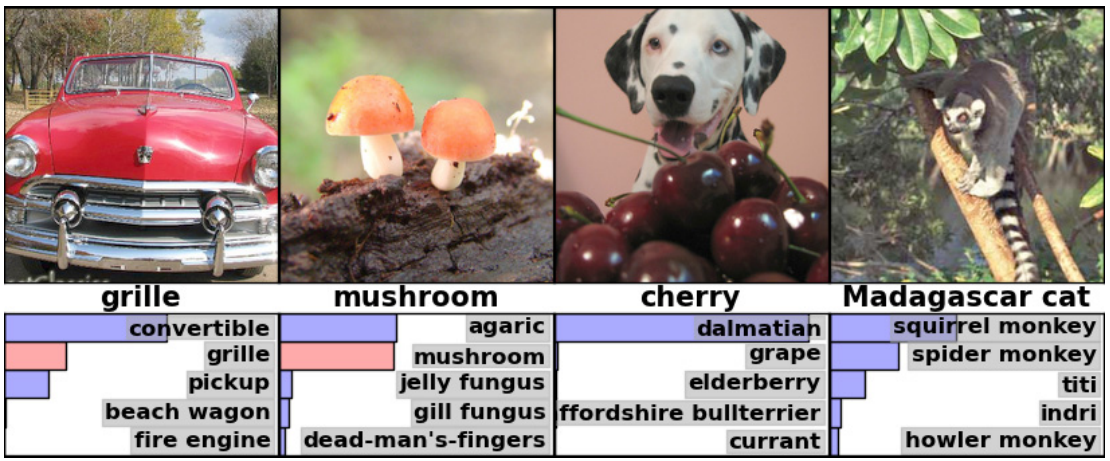
Handgeschriebene Ziffern



AlexNet (2012)

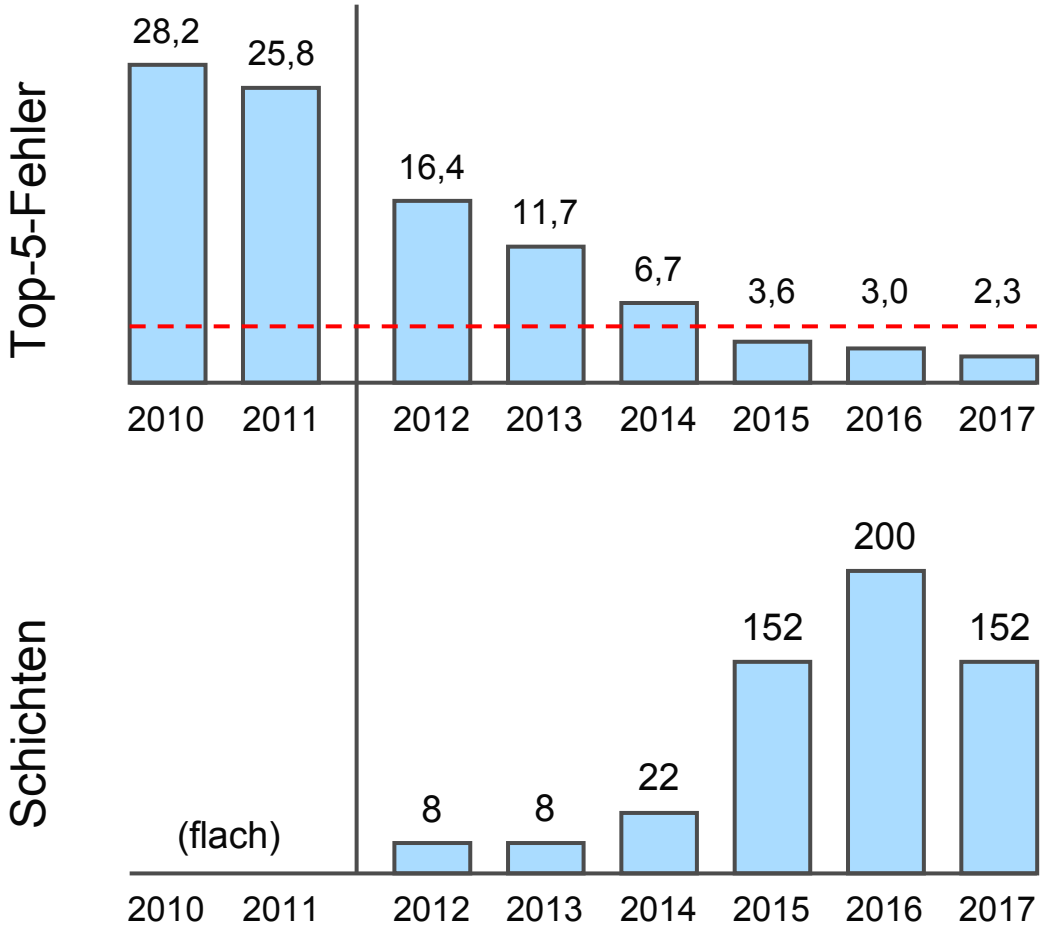


IMAGENET



http://www.image-net.org/

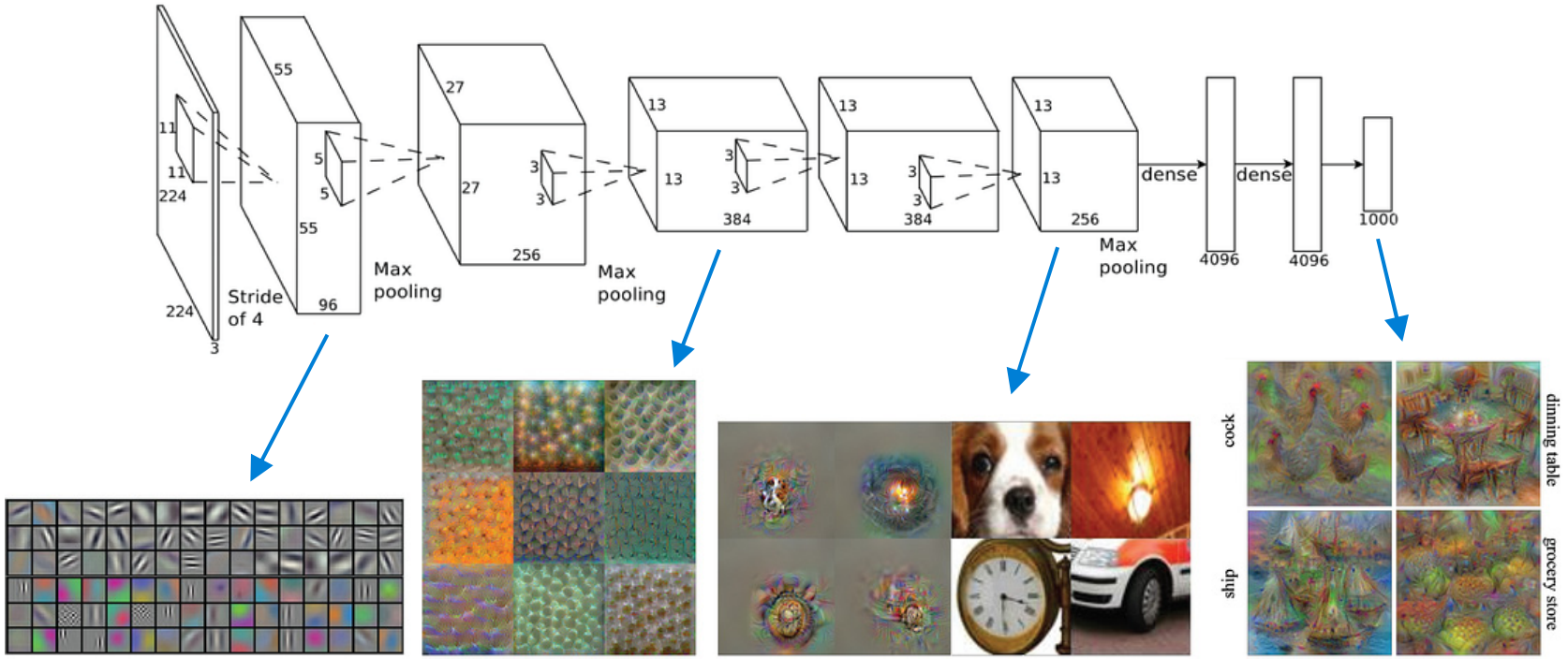
ImageNet Challenge (ILSVRC)



Aufgabe

- Klassifikation
- 1,4 Mio. Bilder
- 1000 Klassen
- Top-5-Fehler
- Proband: ~5%

Lernen von Merkmalen



Conv 1: Kanten, Flächen | Conv 3: Texturen | Conv 5: Teilobjekte | Fc 8: Objekt-Klassen

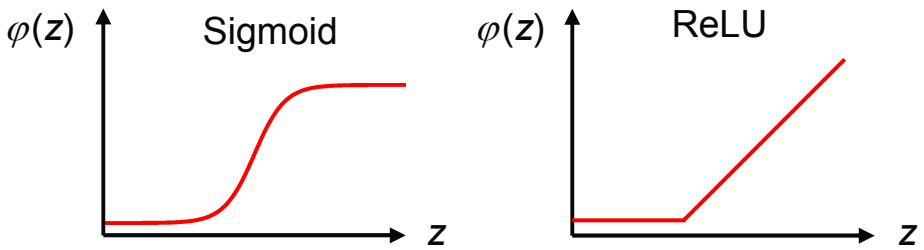
- Deep convolutional neural network (CNN)
- End-to-end learning

http://vision03.csail.mit.edu/cnn_art/

Grund Nr. 1: Algorithmen

Verschwindende Gradienten

- Sättingungsfreie Aktivierung
- Initialisierung, Normierung

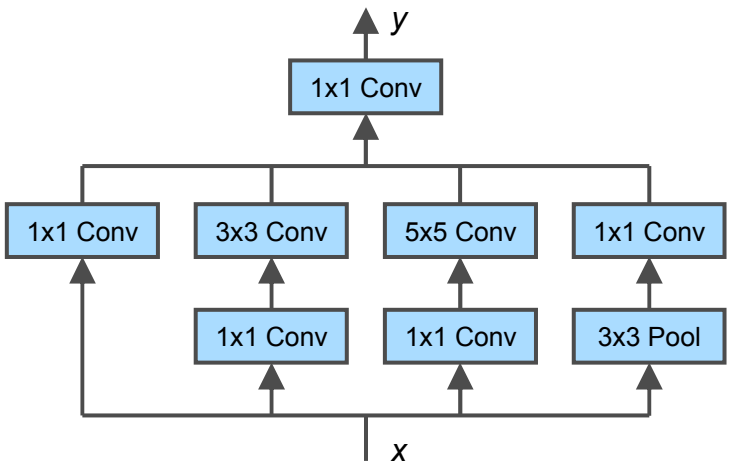


Parameter-Anzahl

- Stochastisches Training, Mini-Batches
- Regularisierung: Dropout

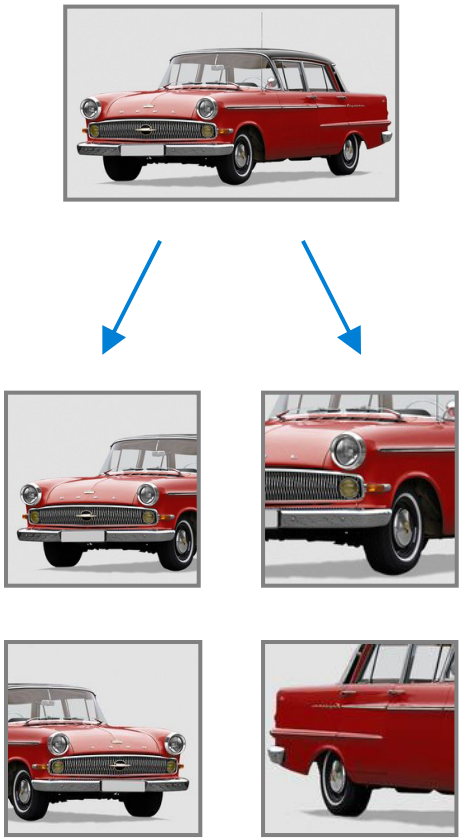
Netz-Architektur

- Faltung und Pooling
- Inception-Module, Residuen-Verbindungen



Grund Nr. 2: Daten

https://www.cityscapes-dataset.com/



https://pixabay.com/

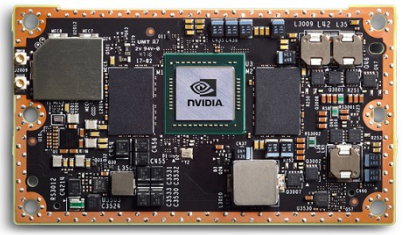
Bereitstellen von Trainingsdaten

- Automatisiertes Sammeln
- Labeling (Objekte / Klassen markieren)
- Künstliches Variieren und Überlagern
- Entwicklungsprozess: Daten ↔ Algorithmen

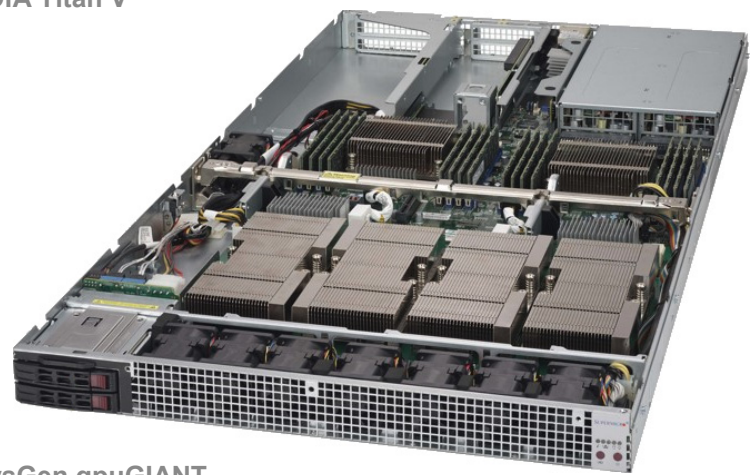
Grund Nr. 3: Hardware

NVIDIA Titan V

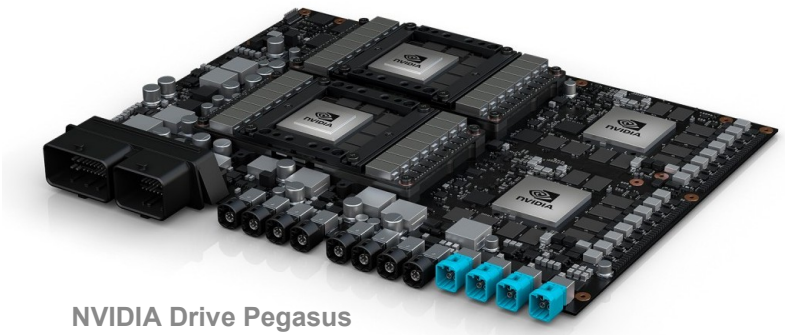
sysGen devCUBE



NVIDIA Jetson TX2

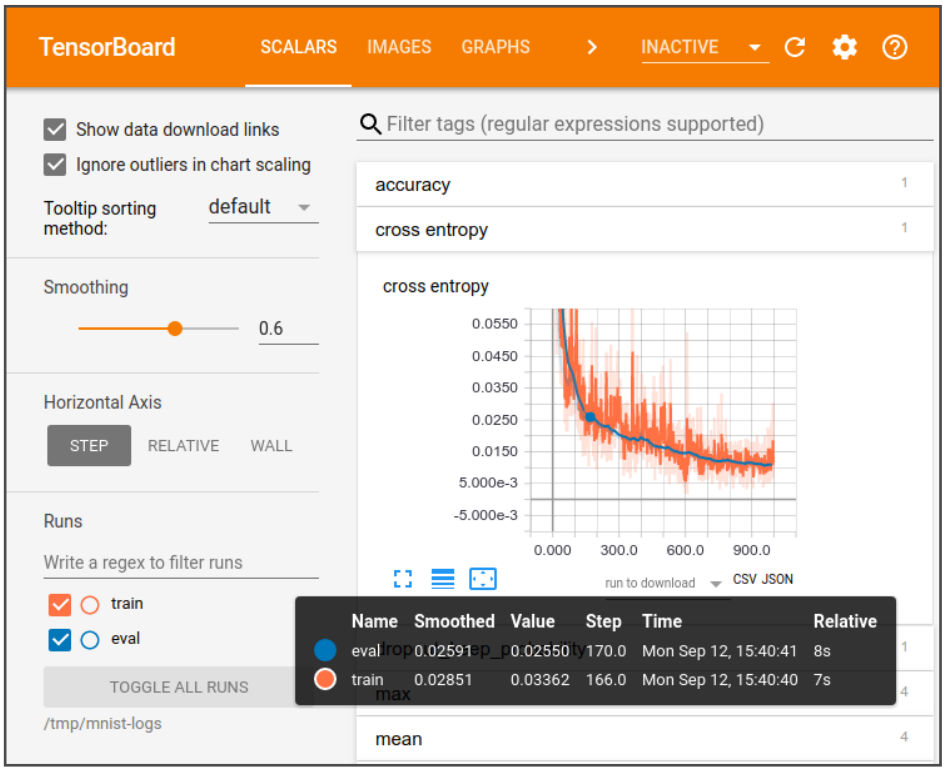


sysGen gpuGIANT



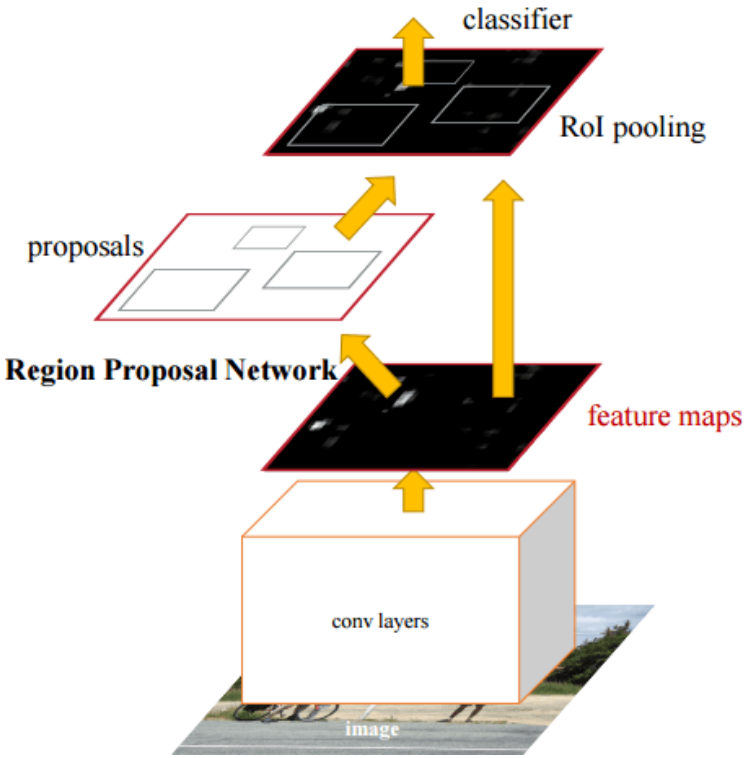
NVIDIA Drive Pegasus

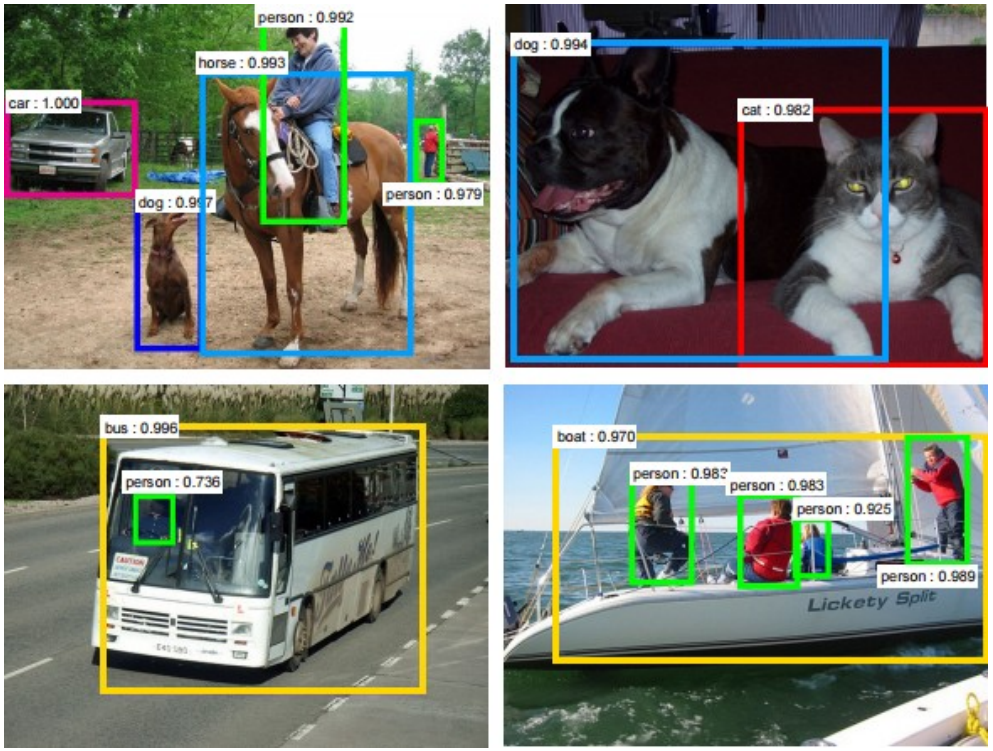
Grund Nr. 4: Software



http://www.tensorflow.org/

Objekt-Detektion

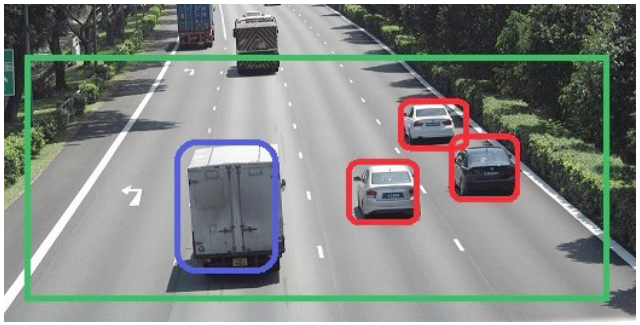




Ren et al. (2016): Faster R-CNN

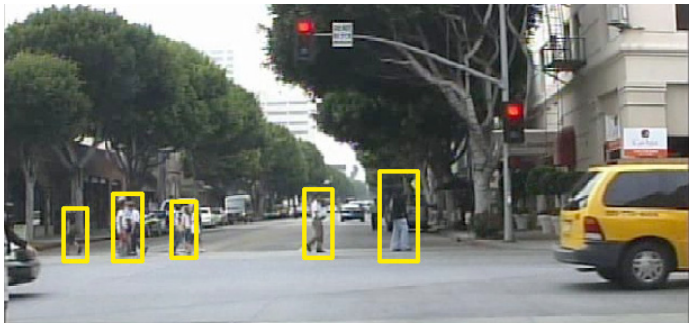
Erfassen von Verkehrsszenen

Fahrzeuge



Zhou et al. (2016)

Fußgänger

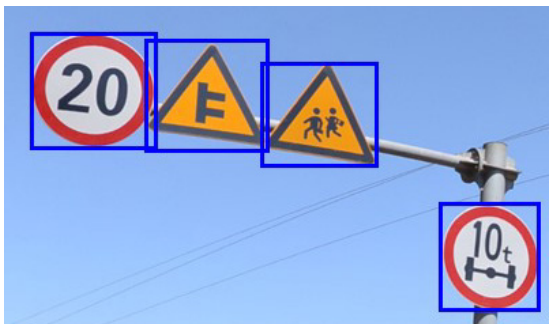


Angelova et al. (2015)

Straßen-Segmentierung

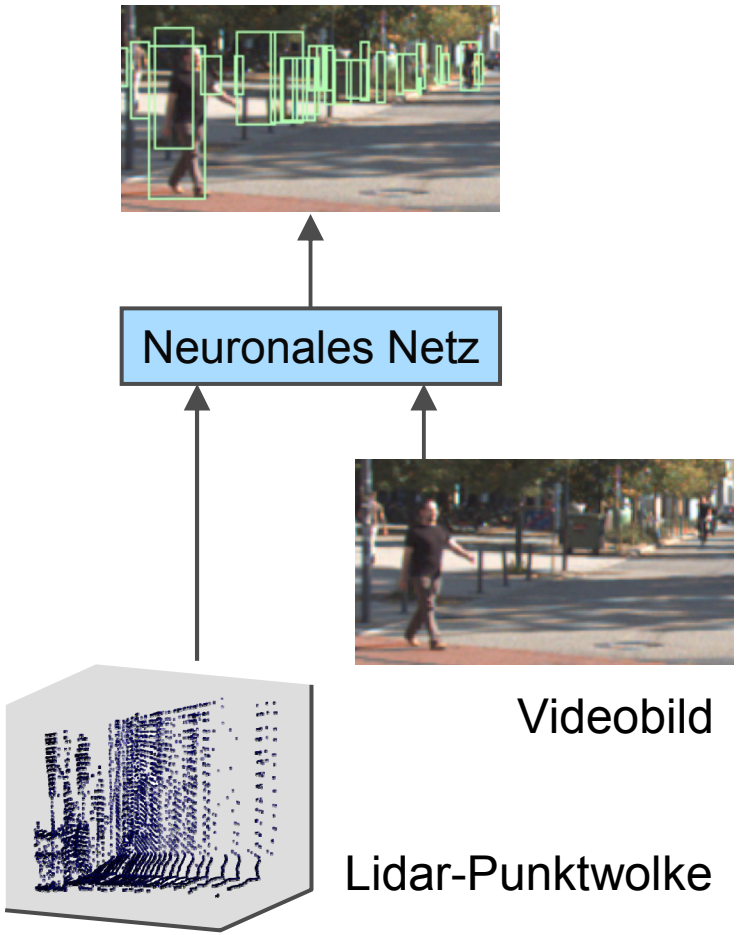
Wang et al. (2018)

Verkehrszeichen

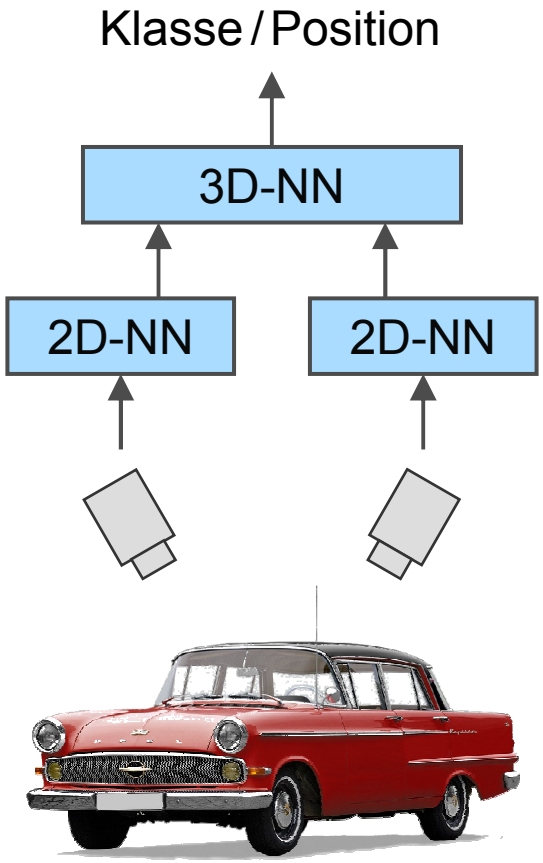


Zhu et al. (2016)

Fusion mehrerer Datenquellen

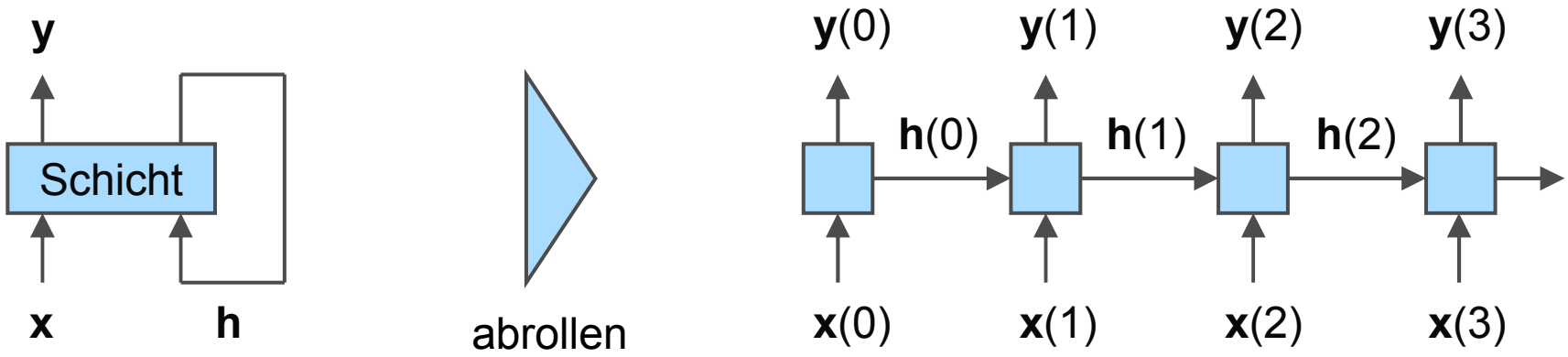


Matti et al. (2017)



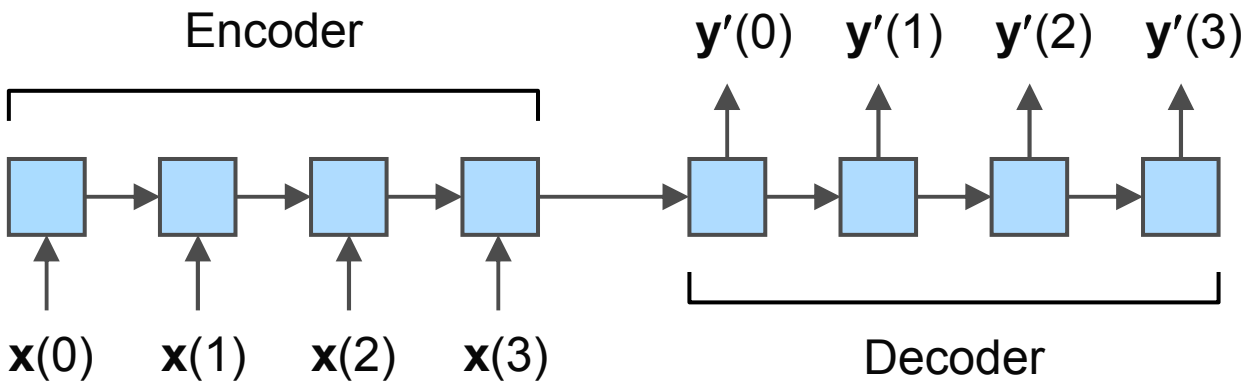
Multiview- / Stereokamera

Rekurrente neuronale Netze (RNNs)



Training: „Backpropagation Through Time“ (BPTT)

Sequenz-Verarbeitung:



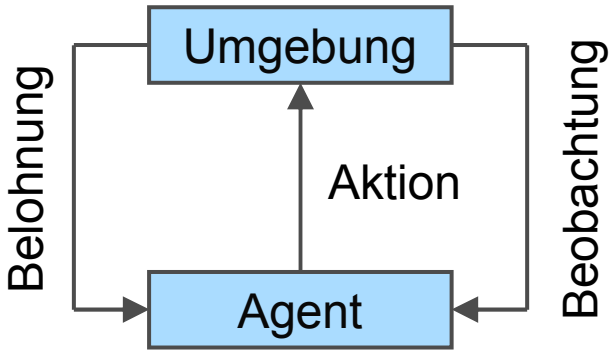
Bildbeschreibung (CNN-RNN)

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

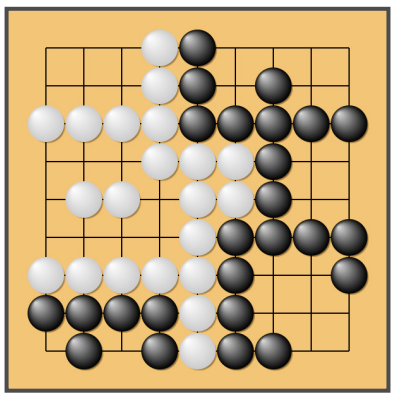
http://research.googleblog.com/2014/11/

Reinforcement Learning

Lernen durch Belohnung



Spielstrategien



https://de.wikipedia.org/

Roboter-Bewegung

Levine et al. (2017)

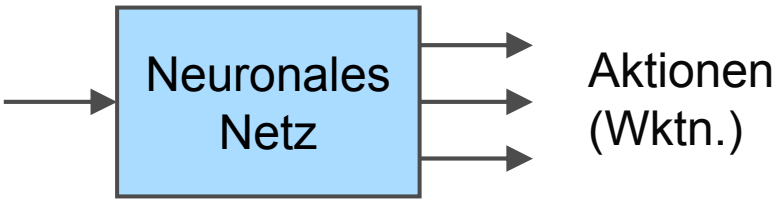
Autonomes Fahren

https://waymo.com/

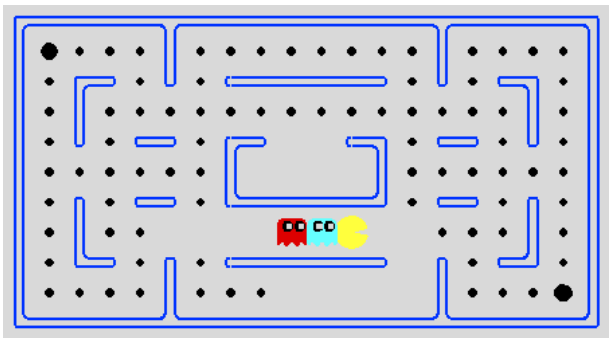
Deep Reinforcement Learning

Neuronale Netze als Policy

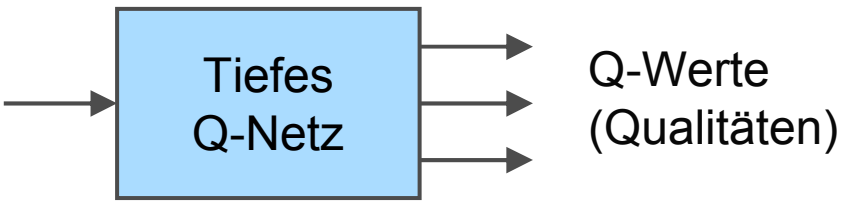
https://devblogs.nvidia.com/



Deep Q-Learning



https://www.cs.washington.edu/



Zusammenfassung

Deep Learning

- Neuronale Netze mit vielen Schichten
- Lernen aller Merkmale „end-to-end“
- Gründe: Algorithmen, Daten, Hardware, Software

Anwendungen

- Bilder / Videos: Objekt- und Gesichtserkennung, Segmentierung, Retusche, Superresolution, …
- Sequenzen: Spracherkennung, Übersetzung, Dialog
- Reinforcement Learning: Spiele, Robotik, Regelung

Konsequenzen

- Mehr Automatisierung bei besserer Qualität
- Verschiebung von Algorithmen zu Daten