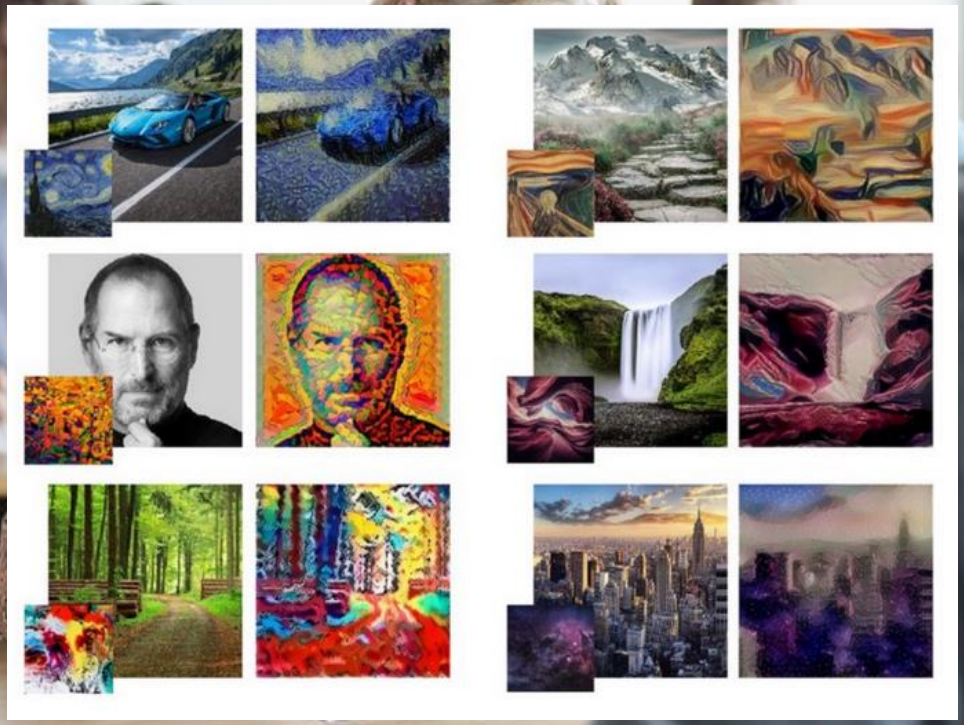
# Which machine learning technique is used in this case (and in Deepfakes)?

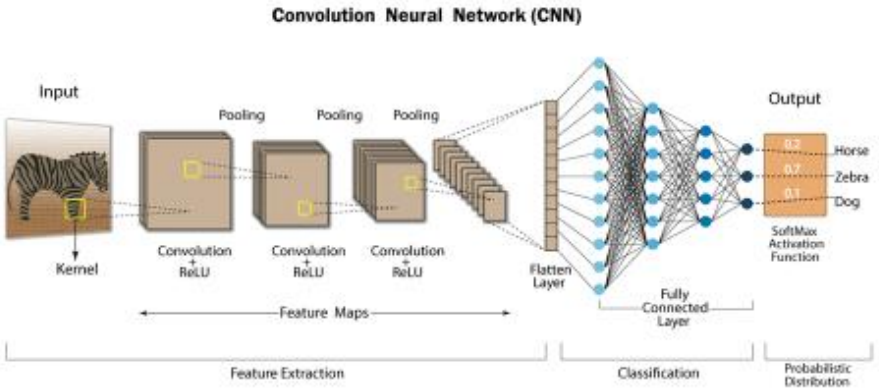
| Schwierigkeitsgrad \ Art des Wissens | Abfragewissen (Vorlesung) | Anwendungswissen (Literatur) |
|---|---|---|
| Einfach | | |
| Mittel | | |
| Schwierig | | |

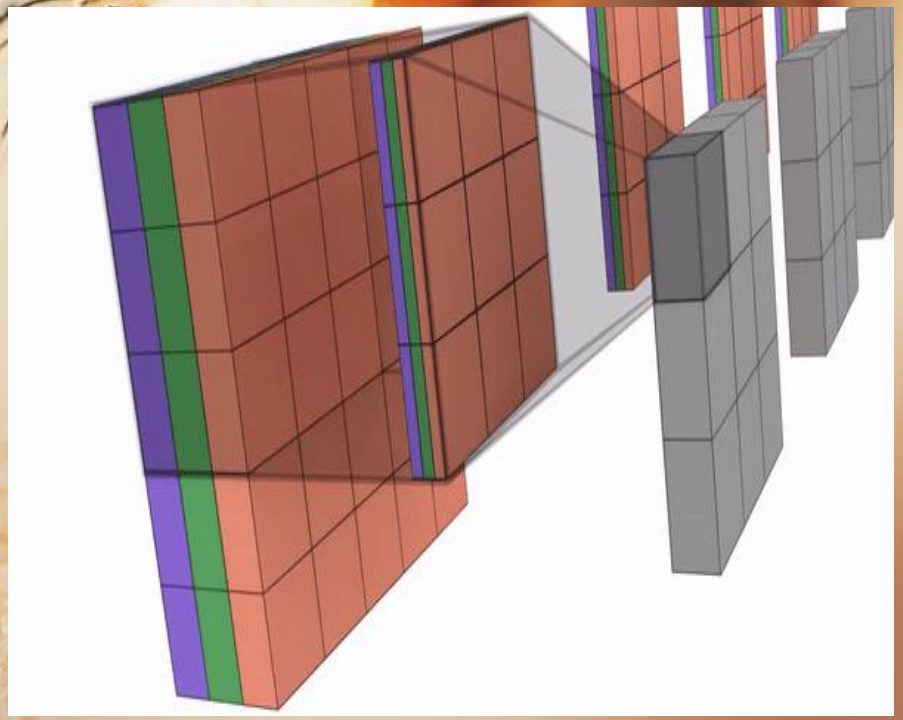


a) Convolutional Neural Networks
b) Transfer Learning
c) Generative Adversarial Networks
d) Reinforcement Learning
e) Transformer

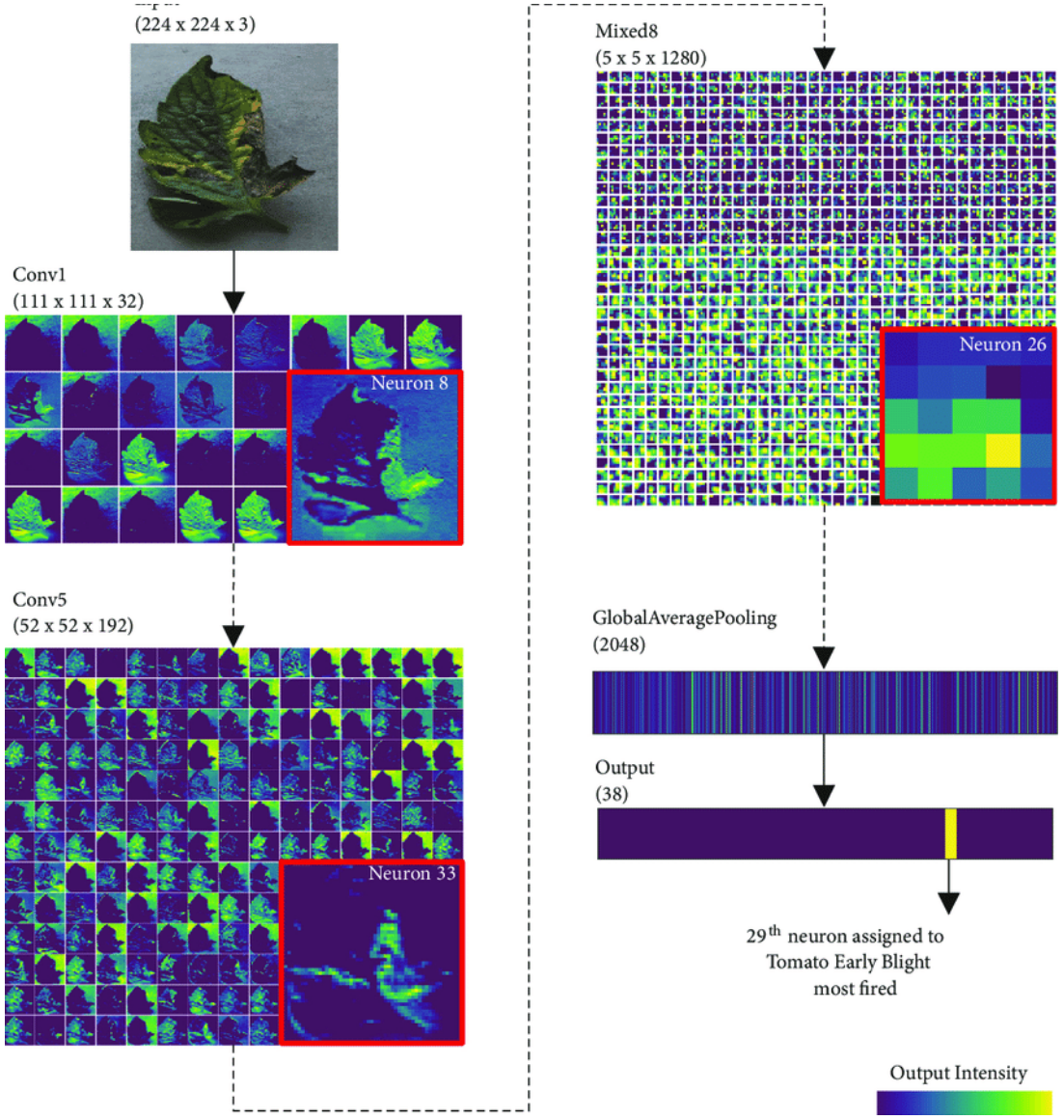
# 8 Machine Learning I
## - Neural Networks & Deep Learning
## (4) Basics of **Convolutional Neural Networks (CNN)**

**CNN: Grundstruktur am Beispiel**



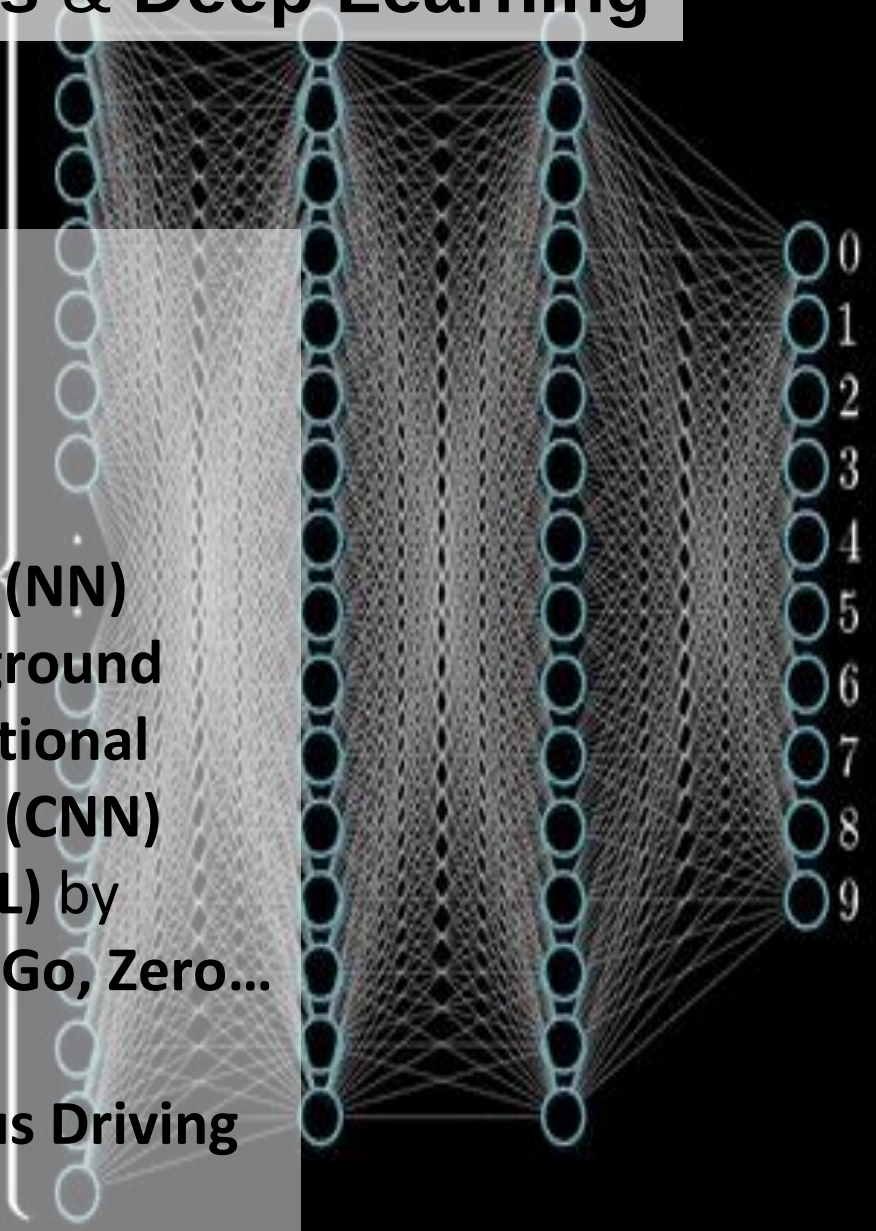



# CNN: Eine Beispielanwendung




Toda Y., Okura F.: **How Convolutional Neural Networks Diagnose Plant Disease**, 2019

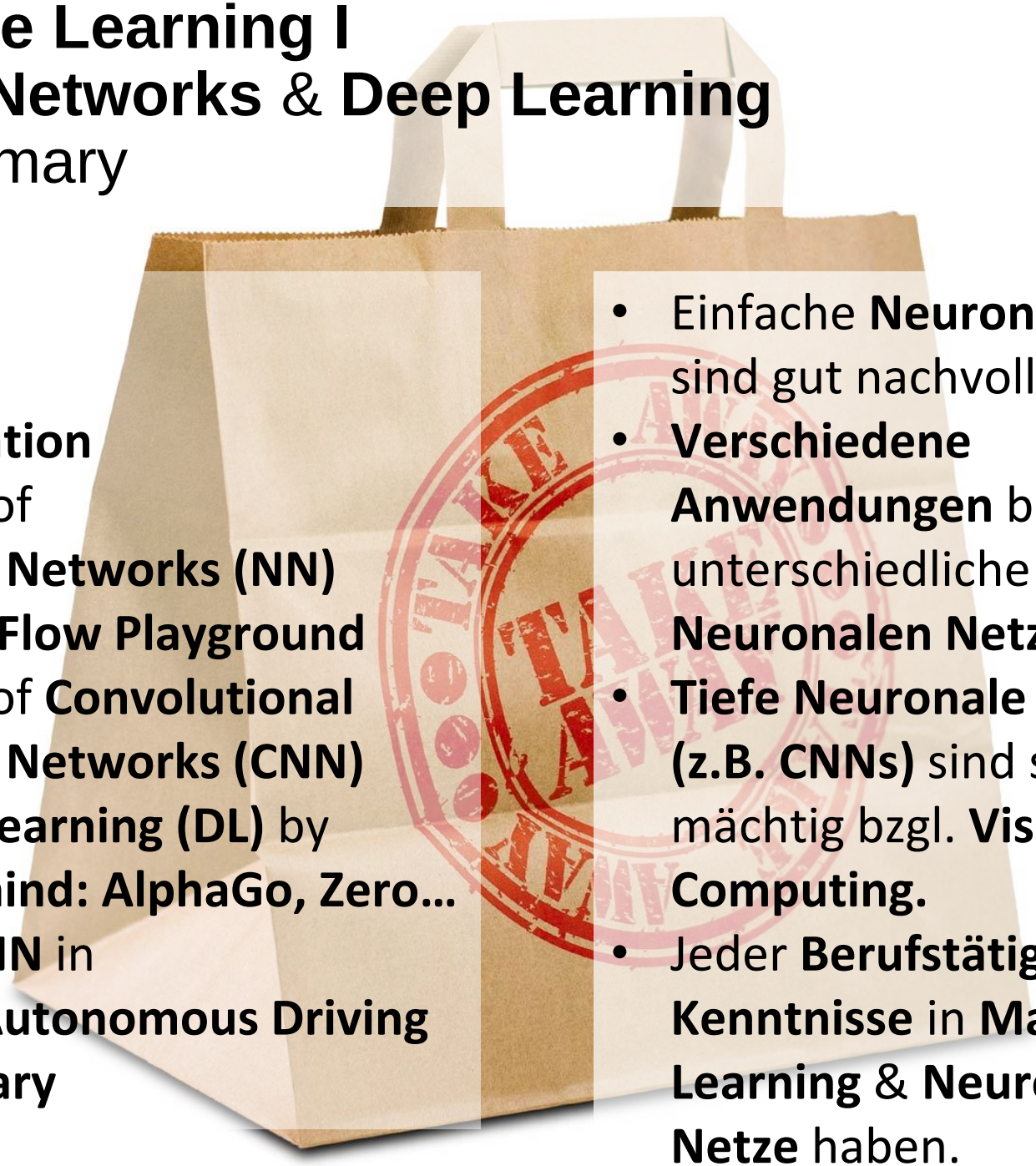
# 8 Machine Learning I
# - Neural Networks & Deep Learning




**Content**:

1. **Motivation**
2. Basics of **Neural Networks (NN)**
3. **TensorFlow Playground**
4. Basics of **Convolutional Neural Networks (CNN)**
5. **Deep Learning (DL)** by **Deepmind: AlphaGo, Zero…**
6. **Deep NN** in **Tesla Autonomous Driving**
7. **Summary**

# 8 Machine Learning I
# - Neural Networks & Deep Learning
## (7) Summary

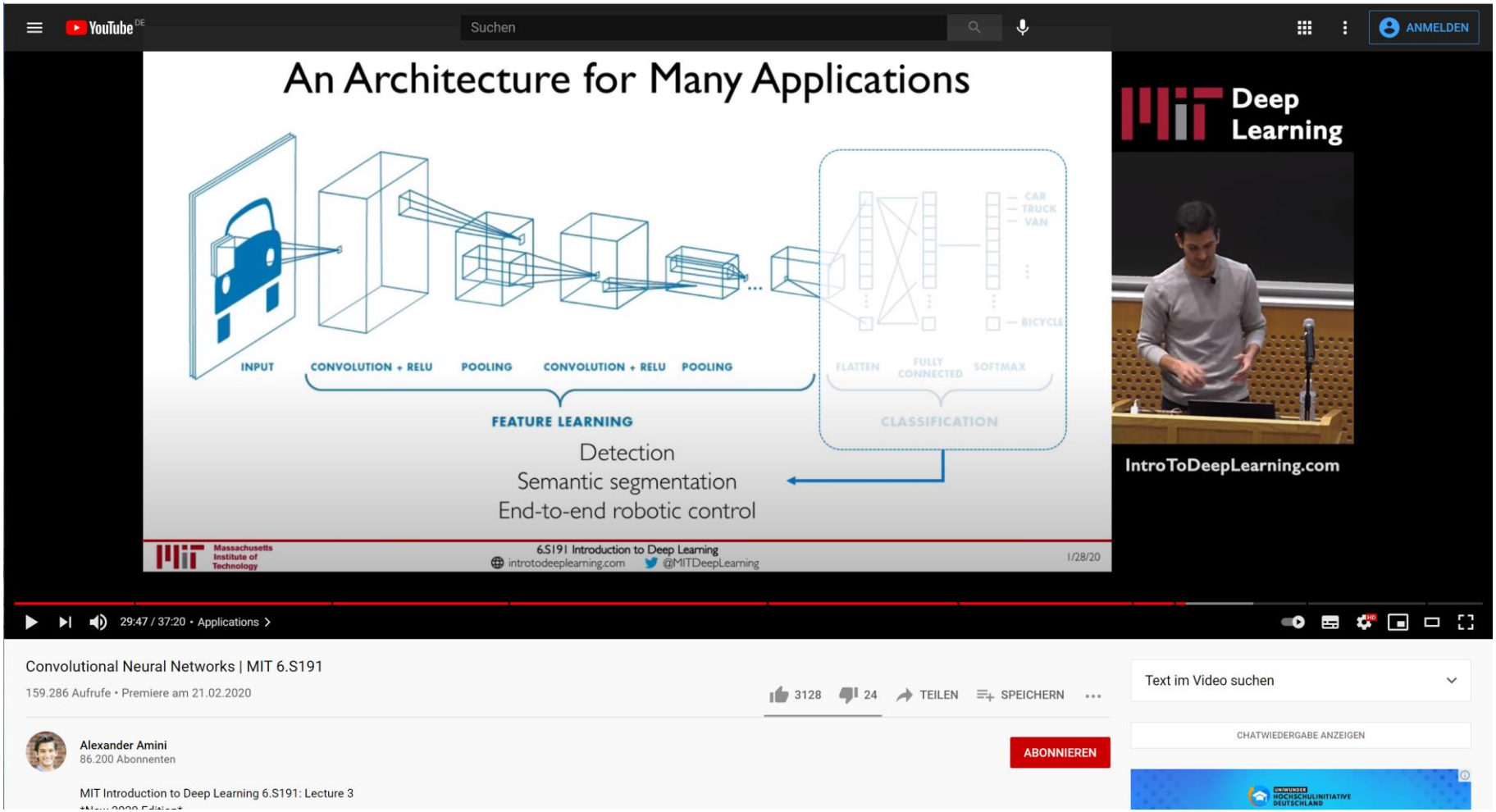


**Content**:

1. **Motivation**
2. Basics of **Neural Networks (NN)**
3. **TensorFlow Playground**
4. Basics of **Convolutional Neural Networks (CNN)**
5. **Deep Learning (DL)** by **Deepmind: AlphaGo, Zero…**
6. **Deep NN** in **Tesla Autonomous Driving**
7. **Summary**

- Einfache **Neuronale Netze** sind gut nachvollziehbar.
- **Verschiedene Anwendungen** benötigen unterschiedliche **Neuronalen Netze**.
- **Tiefe Neuronale Netze (z.B. CNNs)** sind sehr mächtig bzgl. **Visual Computing.**
- Jeder **Berufstätige** sollte **Kenntnisse** in **Machine Learning** & **Neuronale Netze** haben.

# MIT Introduction to Deep Learning 6.S191




**MIT 6.S191 Convolutional Neural Networks, 2020**
www.youtube.com/watch?v=iaSUYvmCekI

# Michael Amberg

**Todays Content**:

1. **Motivation**
2. Basics of **Neural Networks (NN)**
3. **TensorFlow Playground**
4. Basics of **Convolutional Neural Networks (CNN)**
5. **Deep Learning (DL)** by **Deepmind: AlphaGo, Zero…**
6. **Deep NN** in **Tesla Autonomous Driving**
7. **Summary**

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Print/export
Download as PDF

# Deep learning

From Wikipedia, the free encyclopedia

**Deep learning** (also known as **deep structured learning**) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.[1][2][3]

Deep-learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, machine vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases surpassing human expert performance.[4][5][6]

Artificial neural networks (ANNs) were inspired by information processing and distributed communication nodes in biological systems. ANNs have various differences from biological brains. Specifically, neural networks tend to be static and symbolic, while the biological brain of most living organisms is dynamic (plastic) and analog.[7][8][9]

The adjective "deep" in deep learning comes from the use of multiple layers in the network. Early work showed that a linear perceptron cannot be a universal classifier, and then that a network with a nonpolynomial activation function with one hidden layer of unbounded width can on the other hand so be. Deep learning is a modern variation which is concerned with an unbounded number of layers of bounded size, which permits practical application and optimized implementation, while retaining theoretical universality under mild conditions. In deep learning the layers are also permitted to be heterogeneous and to deviate widely from biologically informed connectionist models, for the sake of efficiency, trainability and understandability, whence the "structured" part.

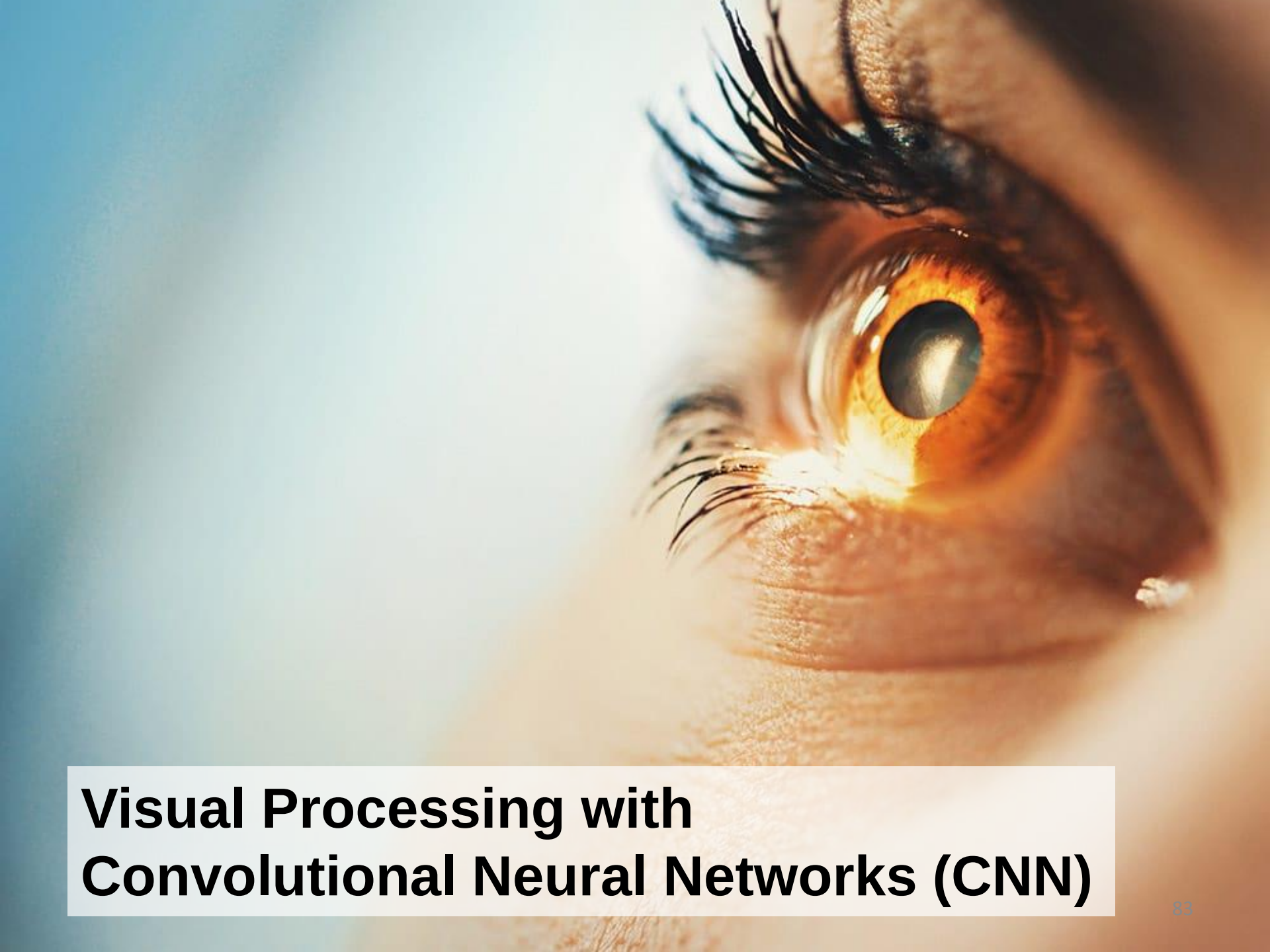
Part of a series on
**Machine learning and data mining**

| | |
|---|---|
| **Problems** | [show] |
| **Supervised learning** (classification • regression) | [show] |
| **Clustering** | [show] |
| **Dimensionality reduction** | [show] |
| **Structured prediction** | [show] |
| **Anomaly detection** | [show] |
| **Artificial neural network** | [show] |
| **Reinforcement learning** | [show] |
| **Theory** | [show] |
| **Machine-learning venues** | [show] |
| **Glossary of artificial intelligence** | [show] |
| **Related articles** | [show] |

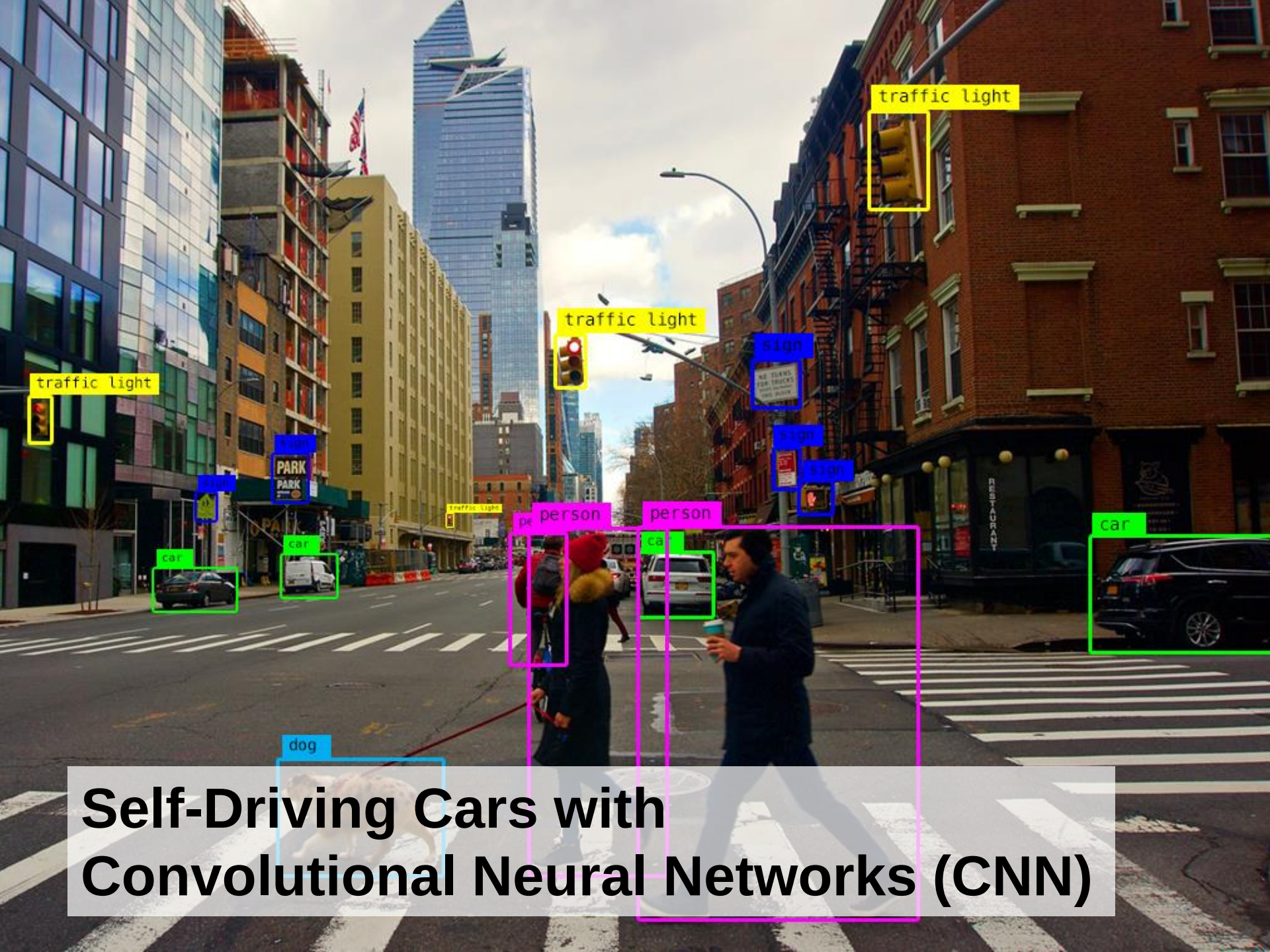
V • T • E

Part of a series on
**Artificial intelligence**

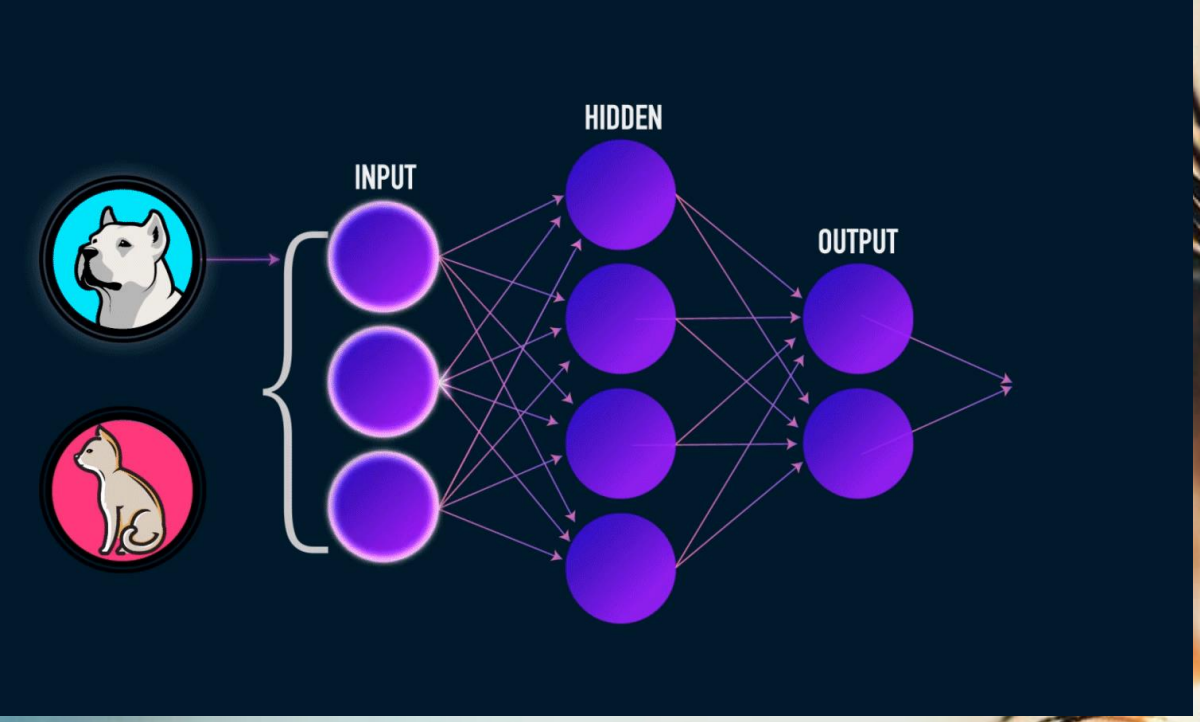
| | |
|---|---|
| **Major goals** | [show] |



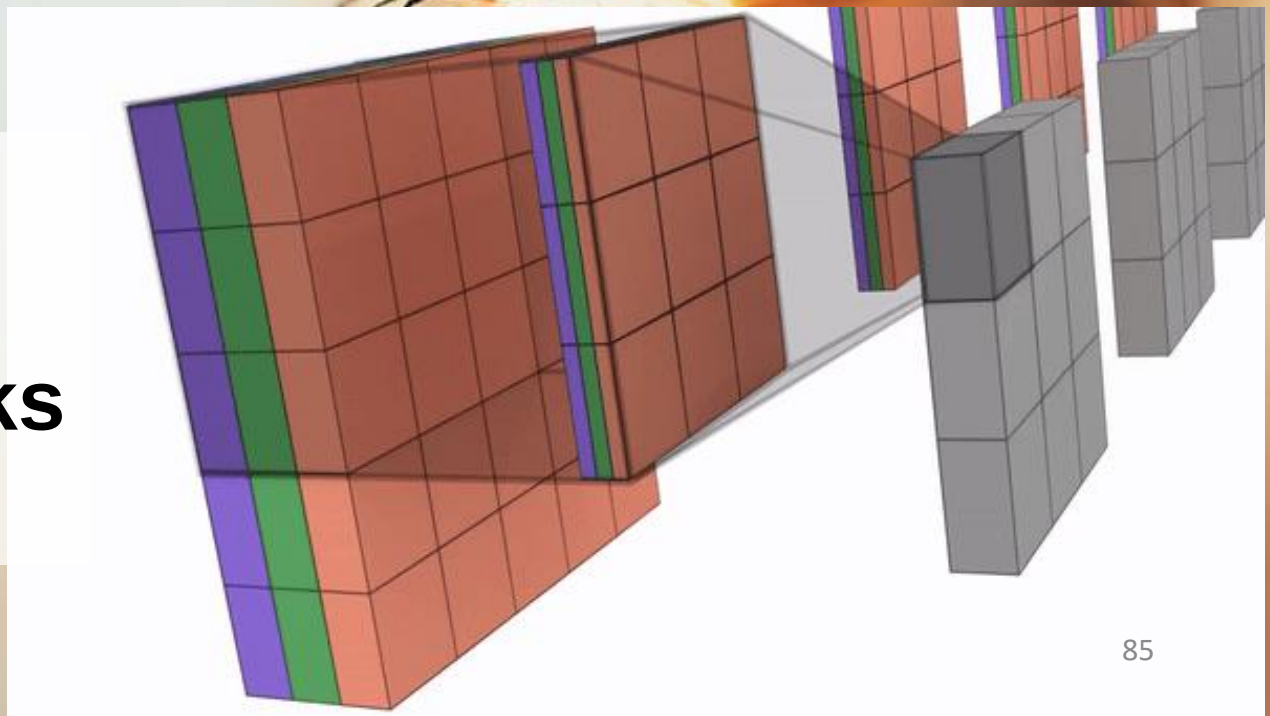
# Visual Processing with Convolutional Neural Networks (CNN)




# Self-Driving Cars with Convolutional Neural Networks (CNN)

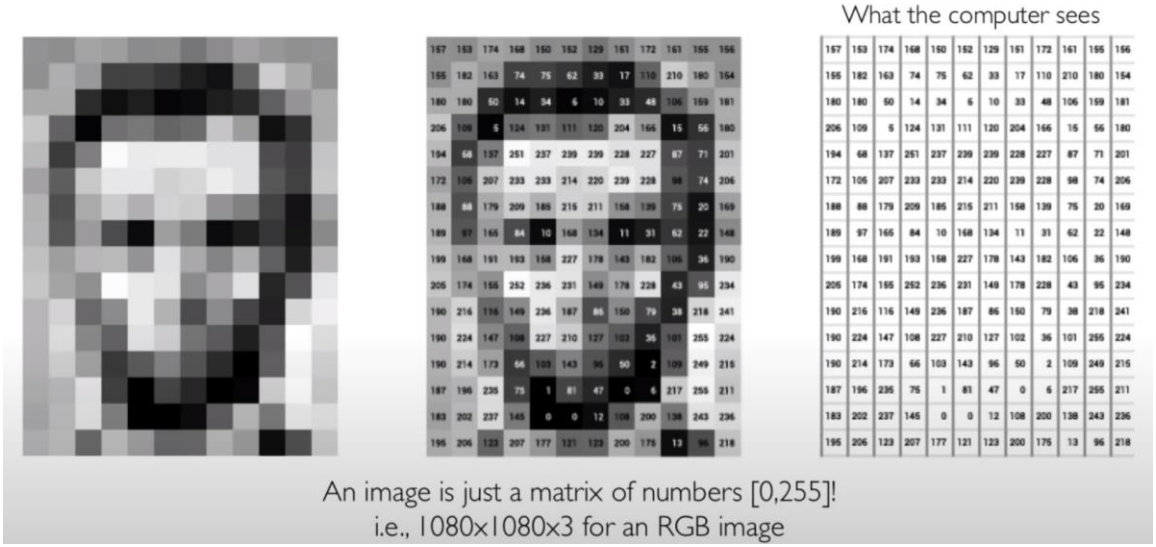


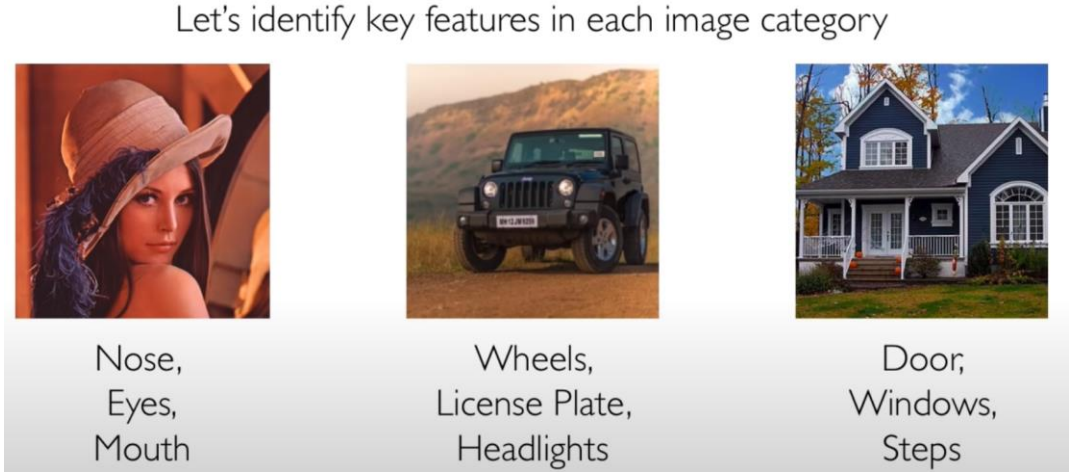

# Basics of Convolutional Neural Networks (CNN)



# CNN: Grundlegende Problemstellung (1/3)

Was **Computer sehen:** Ein **Bild** ist eine **Folge von Bildpunkten** (Pixel).




**MIT 6.S191 Convolutional Neural Networks, 2020**
www.youtube.com/watch?v=iaSUYvmCekI

# CNN: Grundlegende Problemstellung (2/3)

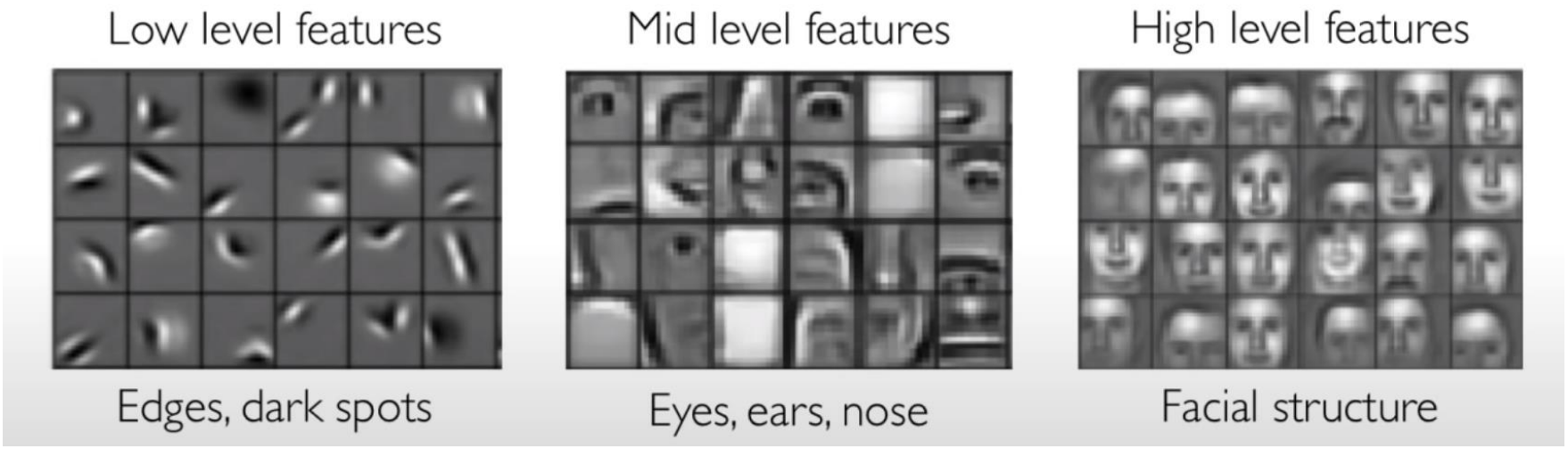
Eine **Objekterkennung** in **Bildern** ist **nicht einfach,**
da das **Umfeld** und **weitere Rahmenbedingungen** dies **erschweren.**


**MIT 6.S191 Convolutional Neural Networks, 2020**
www.youtube.com/watch?v=iaSUYvmCekI

# CNN: Grundlegende Problemstellung (3/3)

Kann ein **Computer** die **relevanten Aspekte (Features)** zur **Objekterkennung** aus **Daten selbständig erlernen,** ohne sich von den Rahmenbedingungen „**ablenken**“ zu lassen?

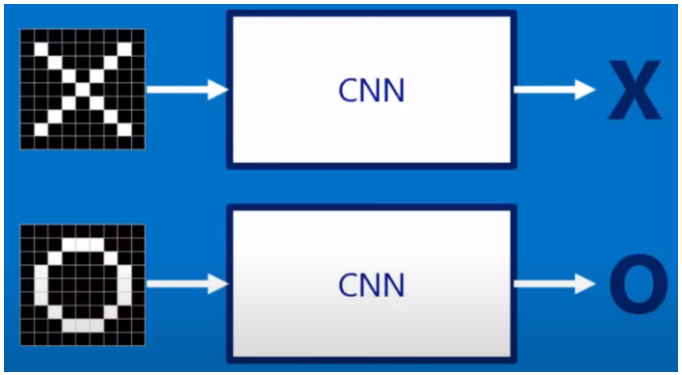
Can we learn a **hierarchy of features** directly from the data instead of hand engineering?



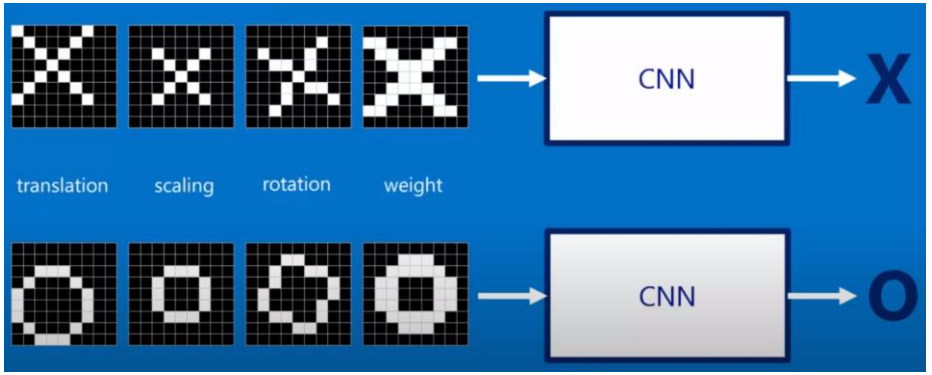

**MIT 6.S191 Convolutional Neural Networks, 2020**
www.youtube.com/watch?v=iaSUYvmCekI

# CNN: Kernidee am Beispiel Objekterkennung (1/5)

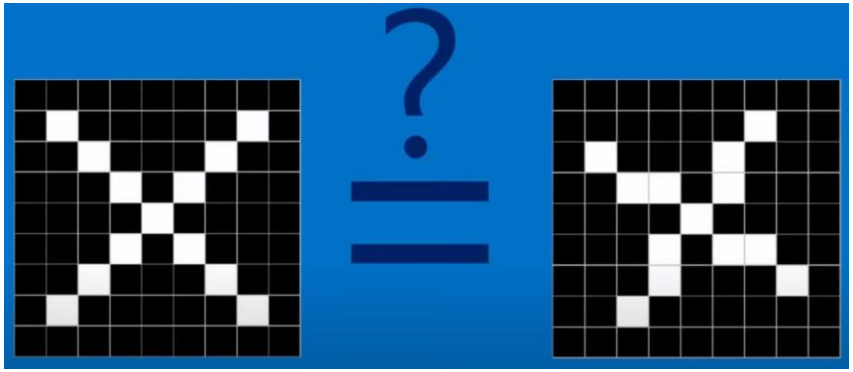
**Problemstellung**



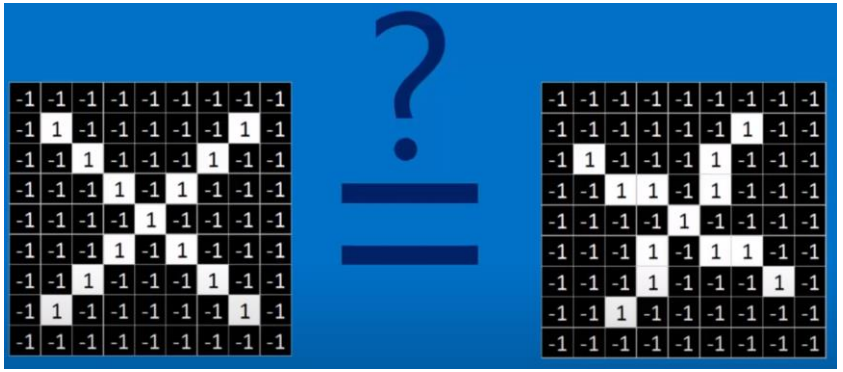

**Herausforderung**




**Problem**



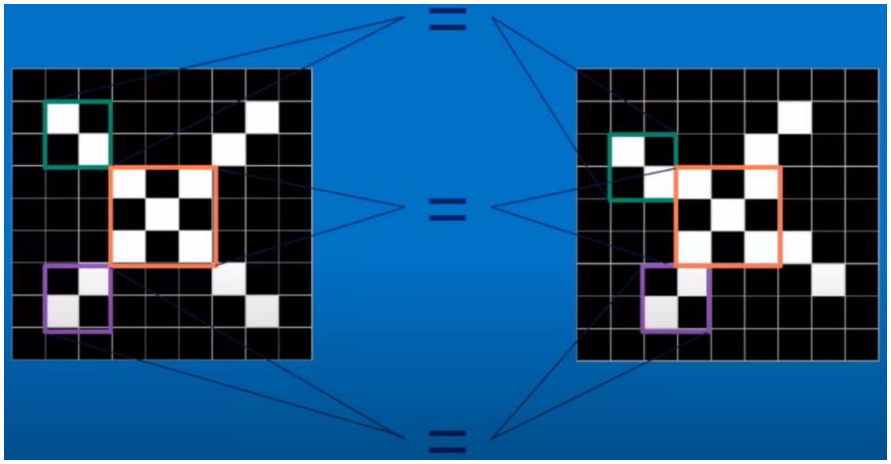

**Problem (Computerdarstellung)**



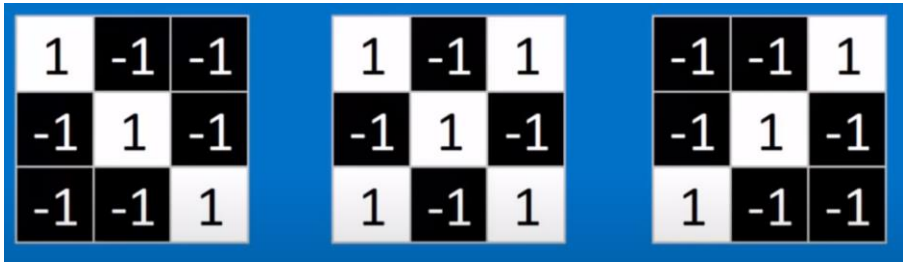

**How Convolutional Neural Networks work, 2016**
www.youtube.com/watch?v=FmpDIaiMIeA

# CNN: Kernidee am Beispiel Objekterkennung (2/5)

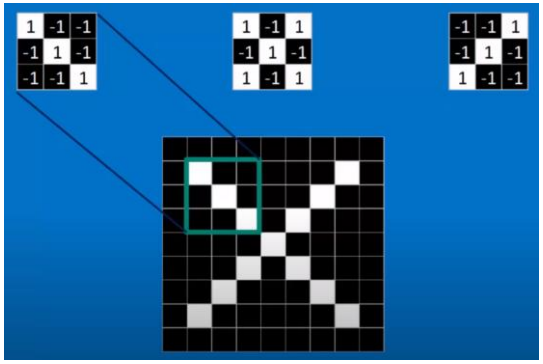
**Idee: Features (Filter) helfen**

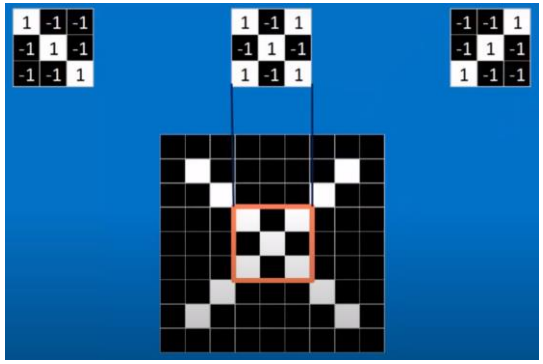


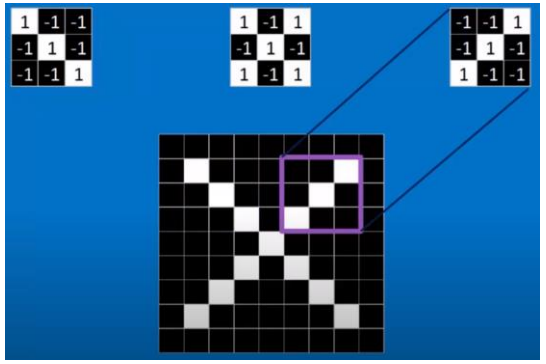
**Relevante Features (Filter) identifizieren**




**Features (Filter) anwenden: Suche nach Features im Bild**







**How Convolutional Neural Networks work, 2016**
www.youtube.com/watch?v=FmpDIaiMIeA

# CNN: Kernidee am Beispiel Objekterkennung (3/5)

**Stapel von mehreren Features (Filtern)**

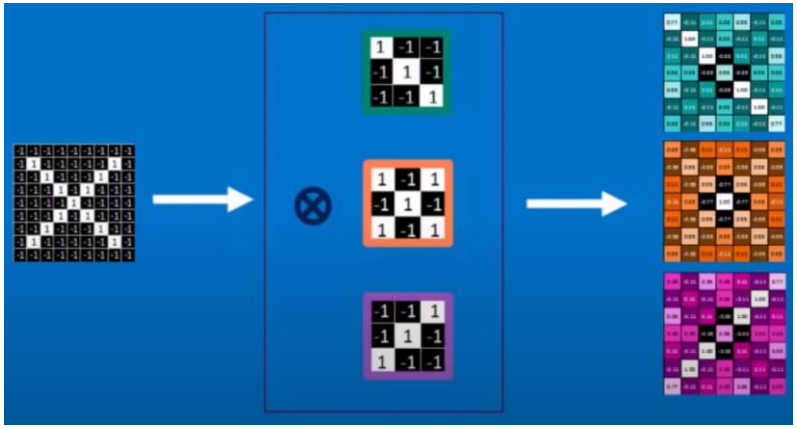
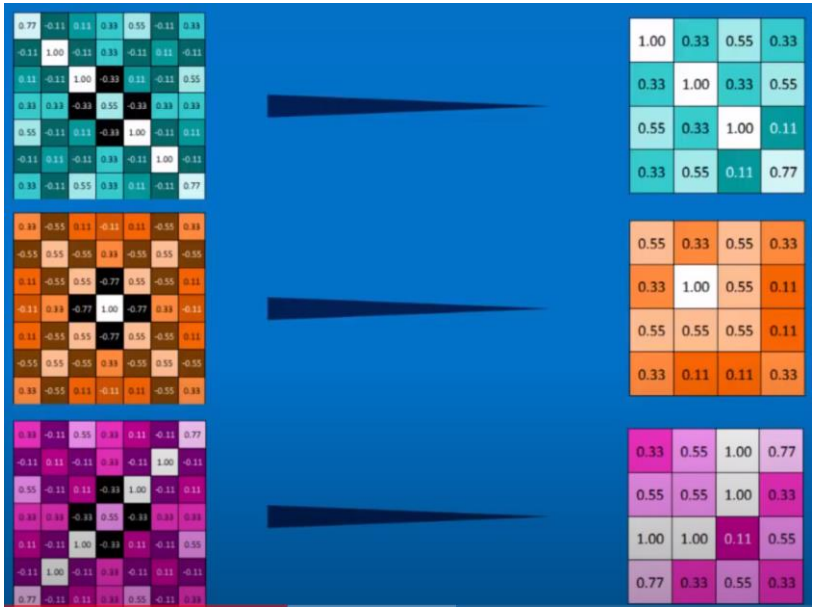
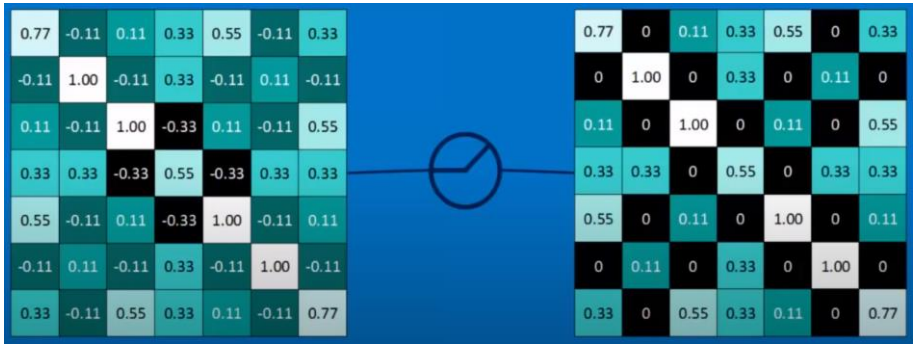


Bild Filter Feature

**Pooling (Komplexität reduzieren & ungenauer werden)**



**ReLU (Unwichtiges entfernen)**

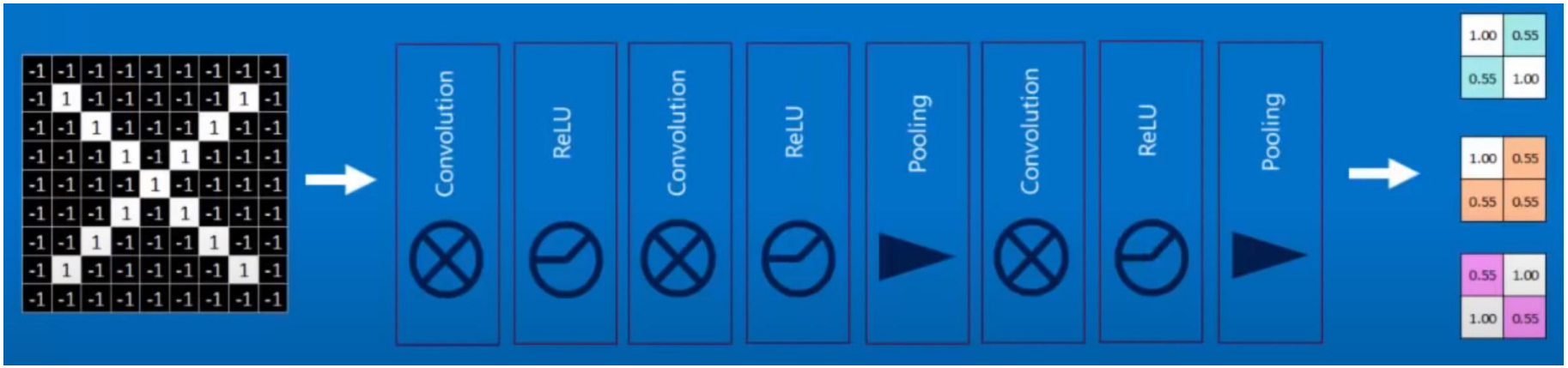


**How Convolutional Neural Networks work, 2016**

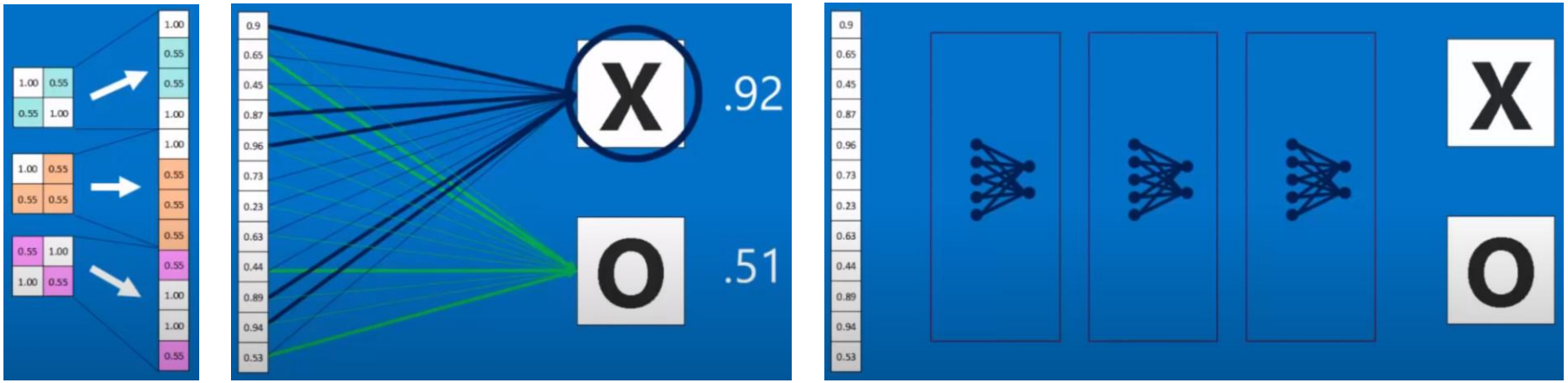
www.youtube.com/watch?v=FmpDIaiMIeA

# CNN: Kernidee am Beispiel Objekterkennung (4/5)

**Diese Operationen können mehrfach wiederholt werden:**
**Features suchen (Conv./Falten), Ergebnis bereinigen (ReLU) und komprimieren (Pooling).**



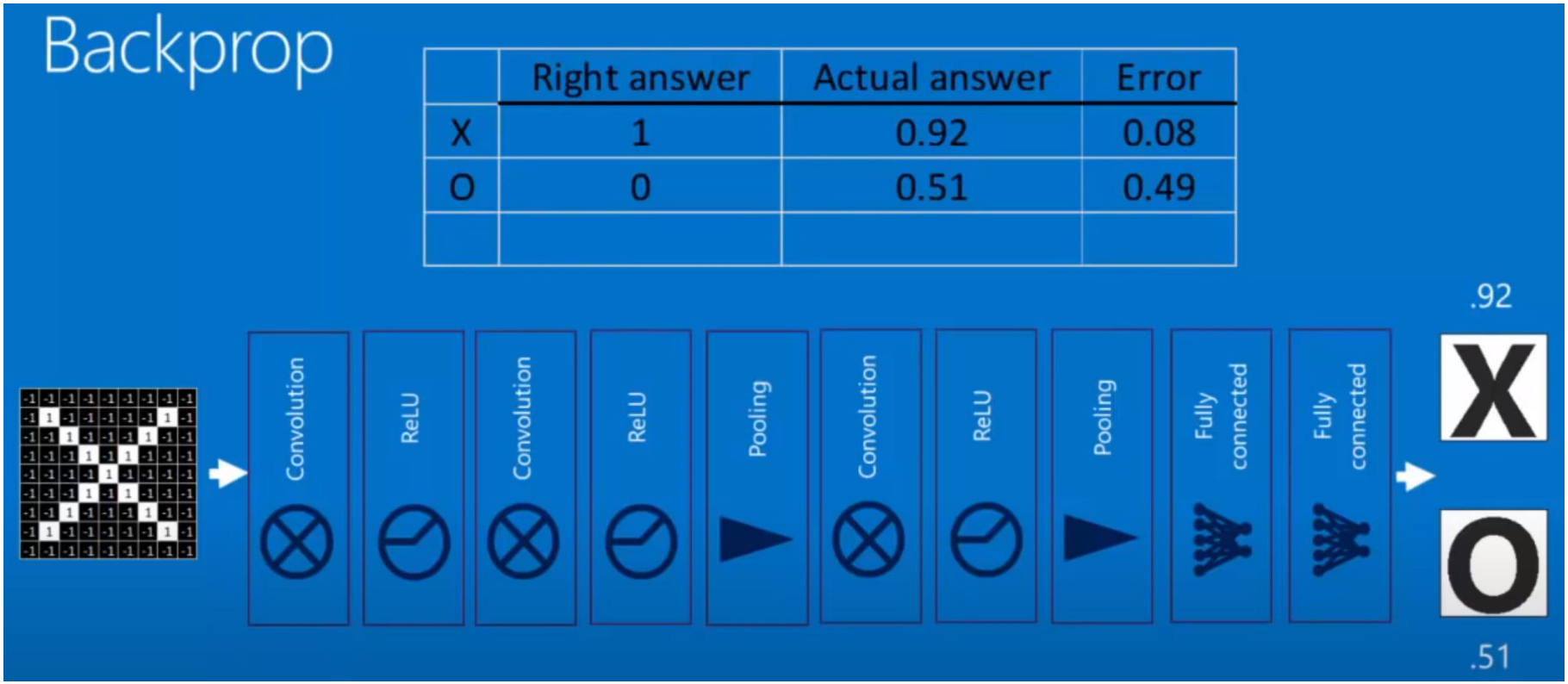

**Ergebnis (die finalen Features) ebnen (engl. flatten) und bewerten (evtl. mehrstufig)**




**How Convolutional Neural Networks work, 2016**
www.youtube.com/watch?v=FmpDIaiMIeA

# CNN: Kernidee am Beispiel Objekterkennung (5/5)

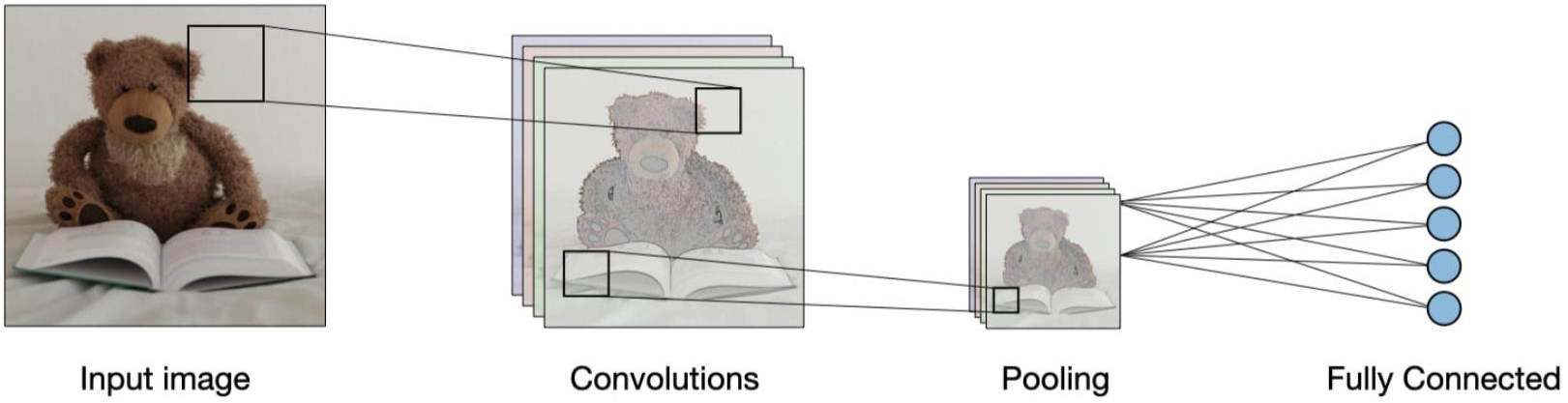
**Backpropagation (Trainieren)**




**How Convolutional Neural Networks work, 2016**
www.youtube.com/watch?v=FmpDIaiMIeA

# Zentrale Bausteine klassischer CNN am Beispiel

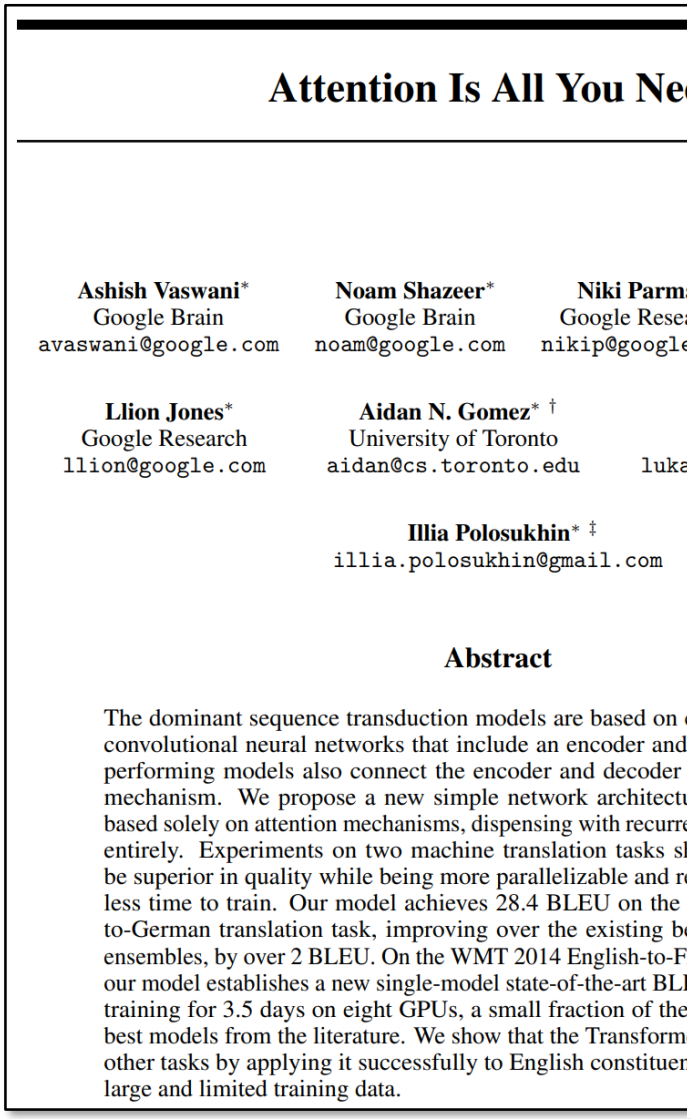
❒ **Architecture of a traditional CNN** — Convolutional neural networks, also known as CNNs, are a specific type of neural networks that are generally composed of the following layers:




The convolution layer and the pooling layer can be fine-tuned with respect to hyperparameters that are described in the next sections.

stanford.edu/~shervine/teaching/cs-230/cheatsheet-convolutional-neural-networks

# Google, 2017: Attention is All You Need



Attention Is All You Ne


**Ashish Vaswani*** Google Brain avaswani@google.com

**Noam Shazeer*** Google Brain noam@google.com

**Niki Parm** Google Rese nikip@google

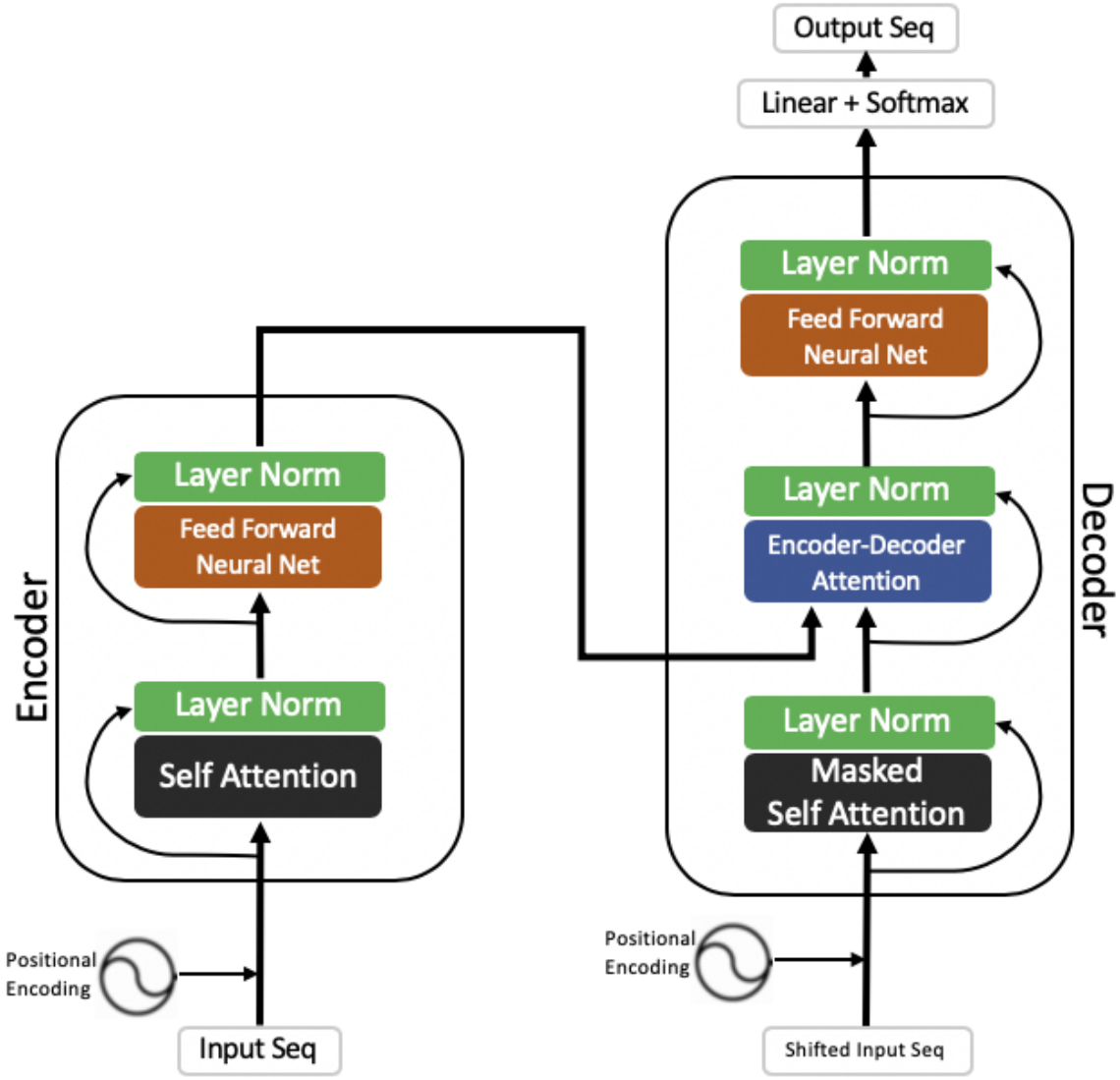
**Llion Jones*** Google Research llion@google.com

**Aidan N. Gomez*** † University of Toronto aidan@cs.toronto.edu luka

**Illia Polosukhin*** ‡ illia.polosukhin@gmail.com

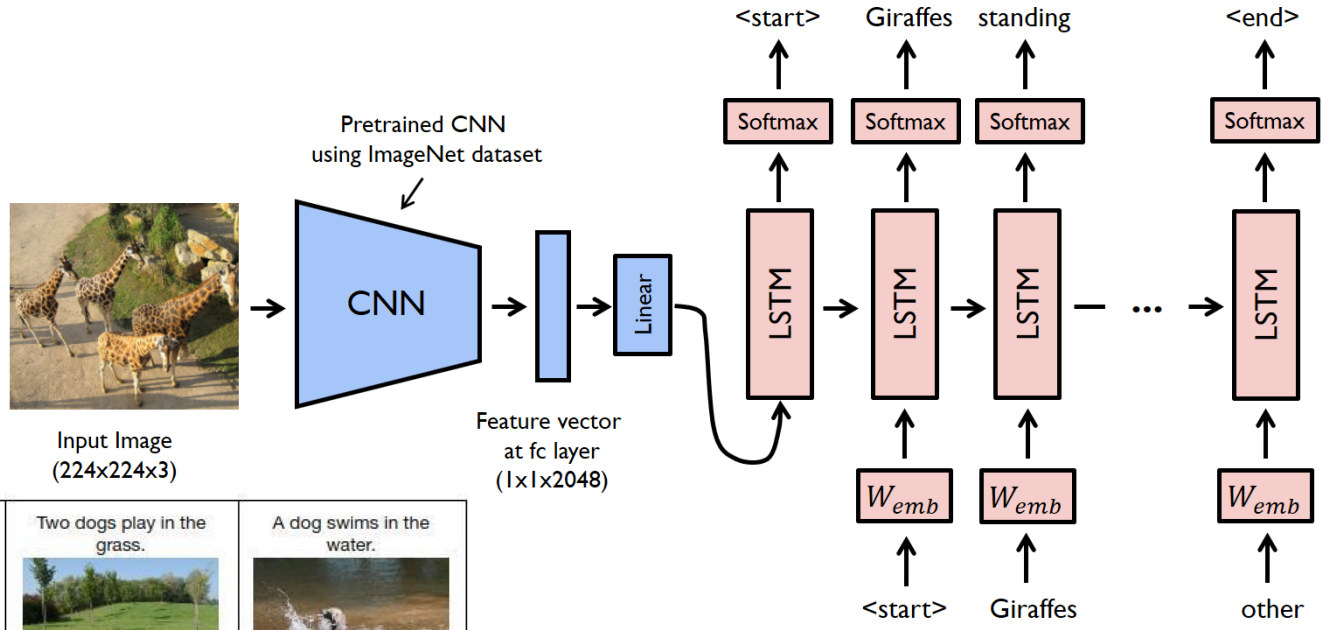

**Abstract**


The dominant sequence transduction models are based on
convolutional neural networks that include an encoder and
performing models also connect the encoder and decoder
mechanism. We propose a new simple network architectu
based solely on attention mechanisms, dispensing with recurr
entirely. Experiments on two machine translation tasks sl
be superior in quality while being more parallelizable and r
less time to train. Our model achieves 28.4 BLEU on the
to-German translation task, improving over the existing b
ensembles, by over 2 BLEU. On the WMT 2014 English-to-F
our model establishes a new single-model state-of-the-art BL
training for 3.5 days on eight GPUs, a small fraction of the
best models from the literature. We show that the Transform
other tasks by applying it successfully to English constituer
large and limited training data.


**Google et.al., 2017: Attention Is All You Need**
arxiv.org/pdf/1706.03762.pdf

# Kombinierter Einsatz von RNN und LSTM

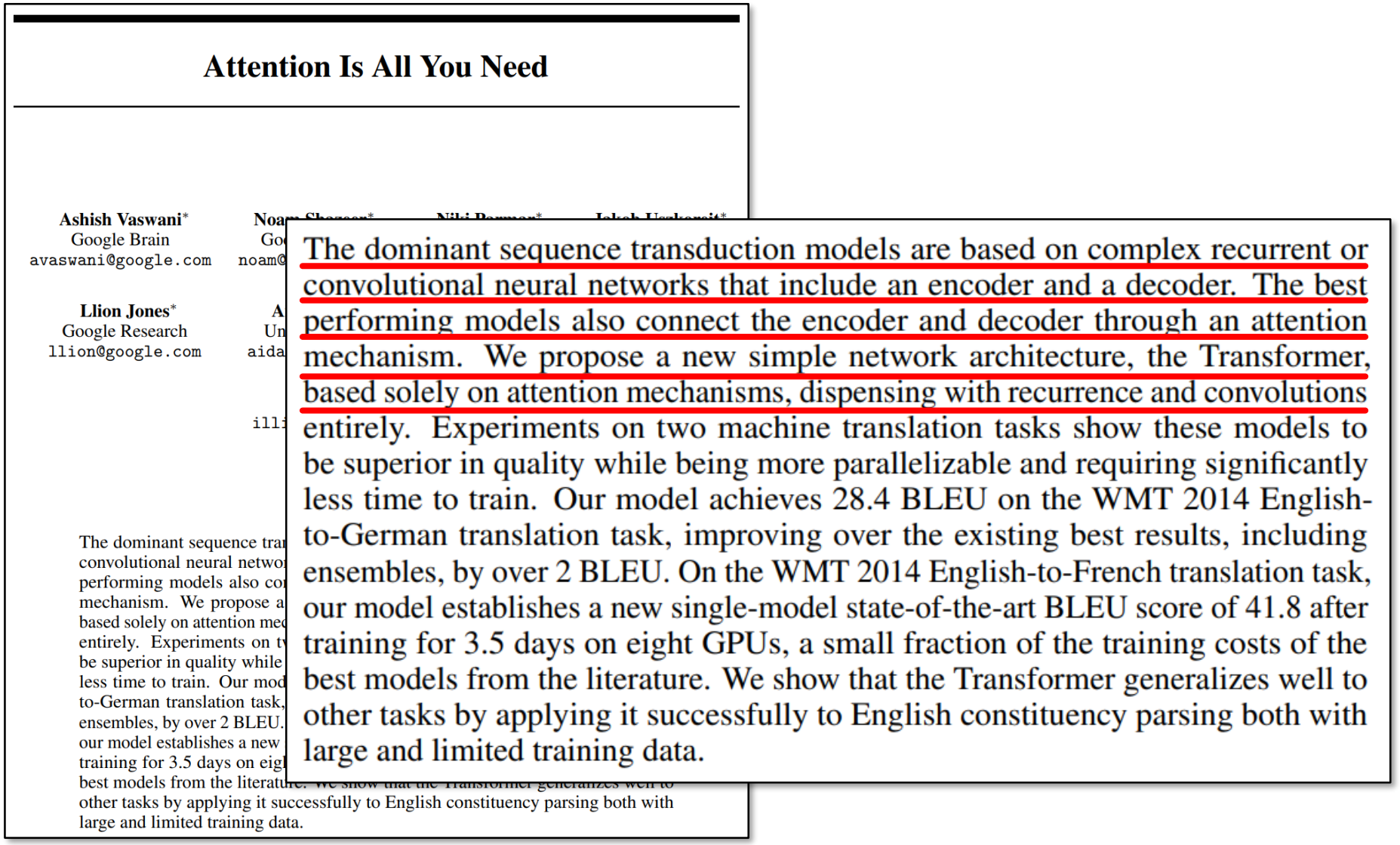




**RNN** und **LSTM** können auch mit **Feed-Forward-Netze** (z.B. **Convolutional Neural Networks)** kombiniert werden. Sie bringen damit „**Memory**“ hinein.

**Anwendungen** sind z.B.:
Textuelle Beschreibung von Bildern (**Image Captioning**) oder Automatische Generierung von Untertiteln (**Video Captioning**).

# Google, 2017: Attention is All You Need



**Attention Is All You Need**

**Ashish Vaswani***
Google Brain
avaswani@google.com

**Llion Jones***
Google Research
llion@google.com

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

**Google et.al., 2017: Attention Is All You Need**
arxiv.org/pdf/1706.03762.pdf

# Google, 2017: Attention is All You Need

## Attention Is All You Need


**Ashish Vaswani*** Google Brain avaswani@google.com

**Noam Shazeer*** Google Brain noam@google.com

**Niki Parmar*** Google Research nikip@google.com

**Jakob Uszkorei** Google Researcl usz@google.co

**Llion Jones*** Google Research llion@google.com

**Aidan N. Gomez*** † University of Toronto aidan@cs.toronto.edu

**Łukasz Kaiser*** Google Brain lukaszkaiser@google.com

**Illia Polosukhin*** ‡ illia.polosukhin@gmail.com



### Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

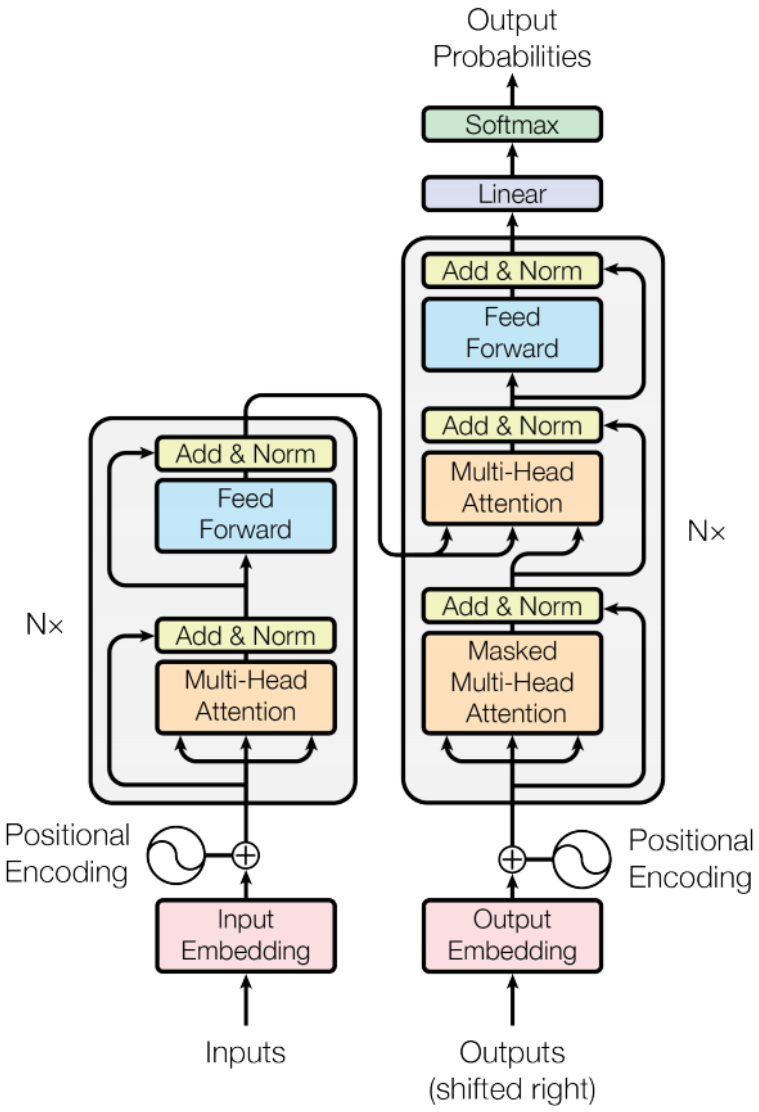




Figure 1: The Transformer - model architecture.

**Google et.al., 2017: Attention Is All You Need**
arxiv.org/pdf/1706.03762.pdf