



Fußgängererkennung im Straßenverkehr mittels Deep Learning

PATRICK NAGORSKI

HAUPTSEMINAR SS 2018

HAW HAMBURG, INFORMATIK MASTER

BETREUER: PROF. DR. ANDREAS MEISEL

19.06.2018

Gliederung

- ▶ Motivation / Fragestellung
- ▶ Methodik
- ▶ Aktueller Stand
- ▶ Risiken
- ▶ Ausblick

Motivation



[01]

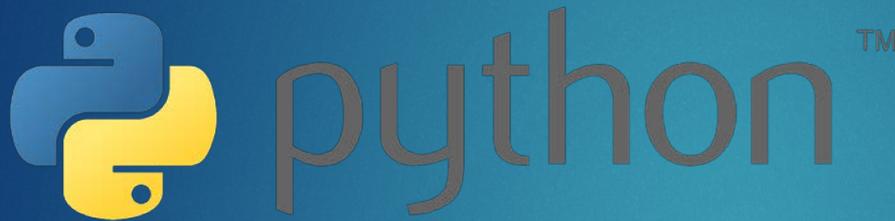
Fragestellung

5

- ▶ Wie wird eine zuverlässige Handlungsprognose aus der zeitlichen Veränderung von Fußgängerbewegungen entwickelt?

Technologien

6



[03]



[04]



[05]

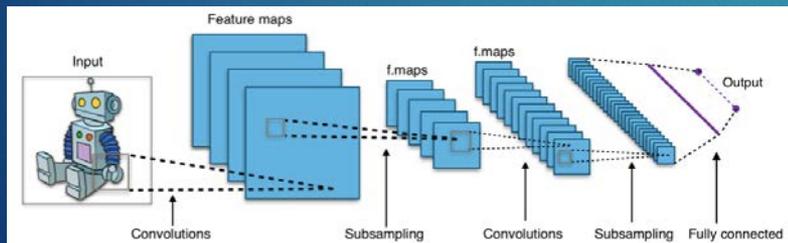


[06]

CNN vs. LSTM

CNN (Convolutional Neural Network)

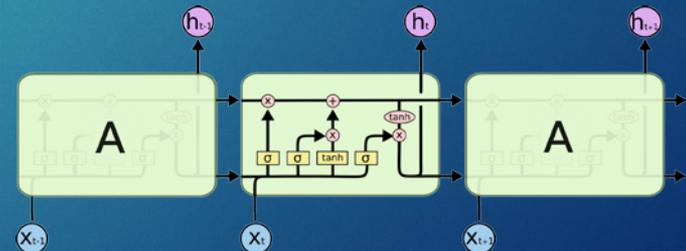
- ▶ Faltungsnetzwerk
- ▶ Hauptsächlich für Bild-Daten
 - ▶ Räumlicher Zusammenhang
- ▶ Robust gegen Rotation, Translation, Skalierung
 - ▶ Data Augmentation



[07]

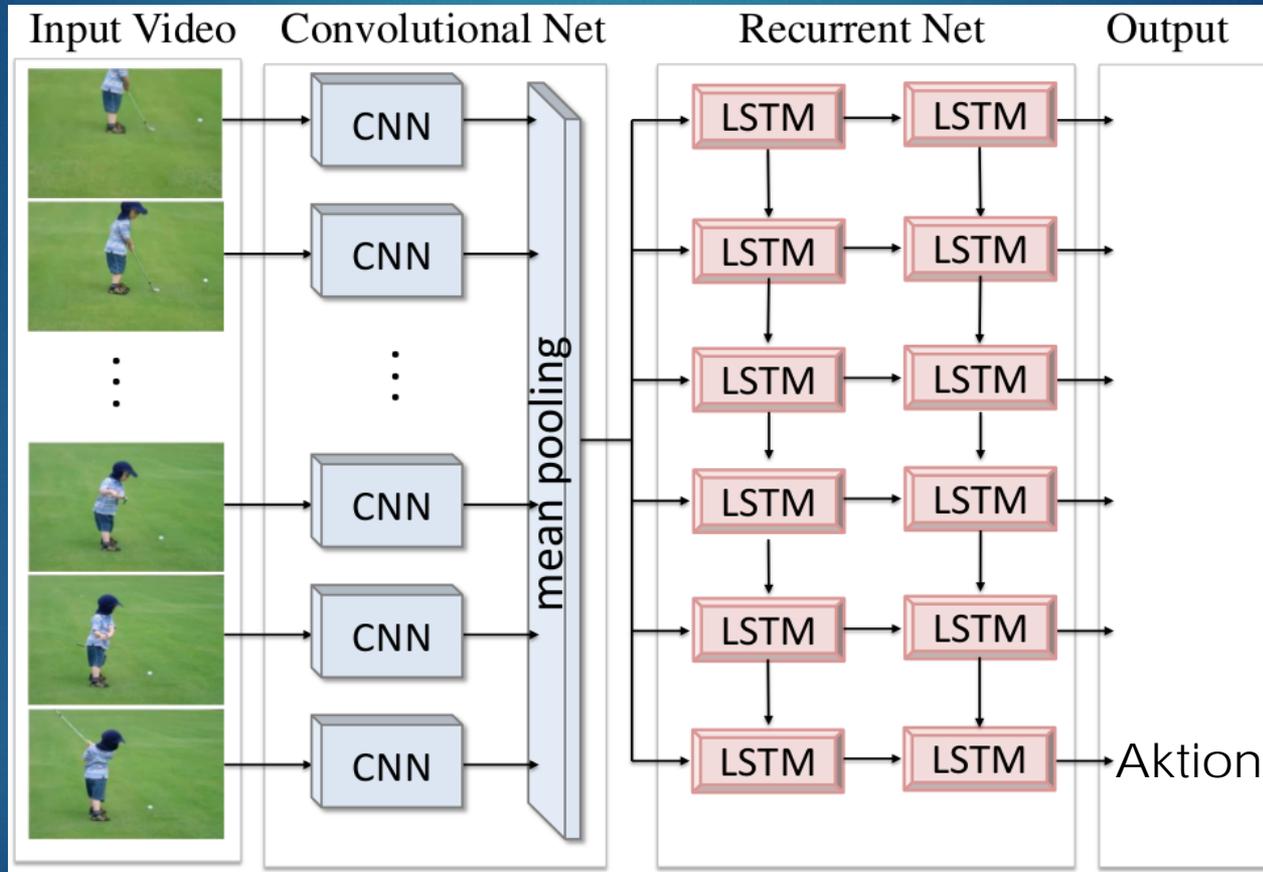
LSTM (Long short-term memory)

- ▶ Rekurrentes neuronales Netz
- ▶ Für zeitlich abhängige Daten
- ▶ Sequenzielles Lernen
- ▶ Langfristige Abhängigkeiten innerhalb der Sequenzen behandelbar



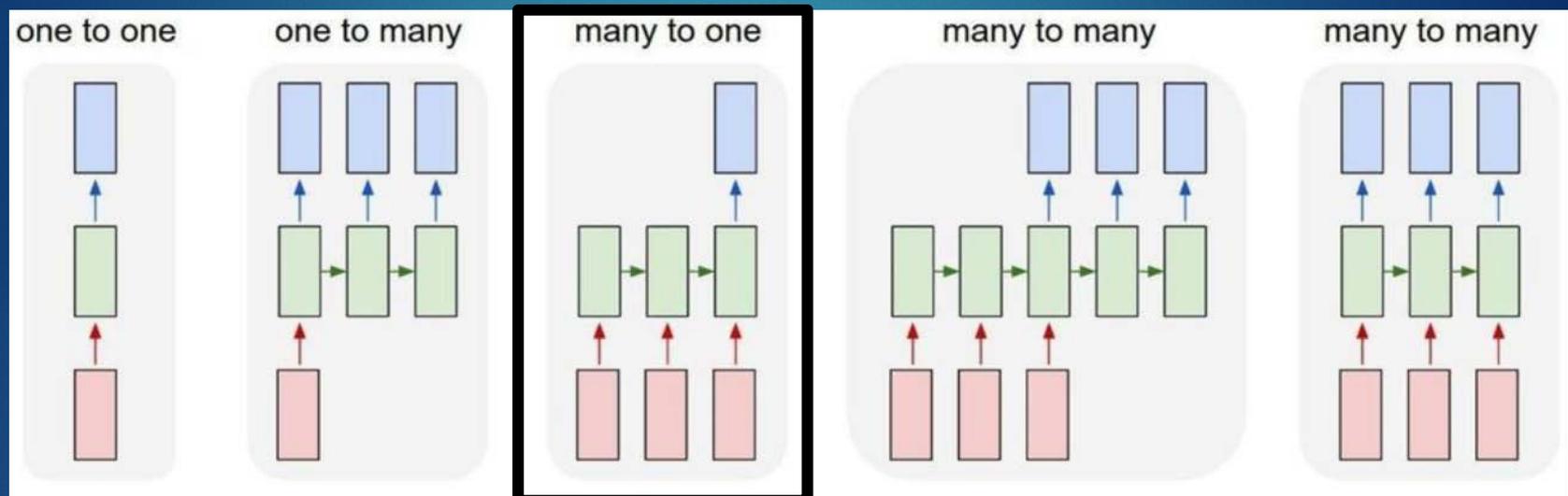
[08]

Convolutional LSTM (CNN+LSTM)



[09]

Arten von Rekurrenten neuronalen Netzen



[10]



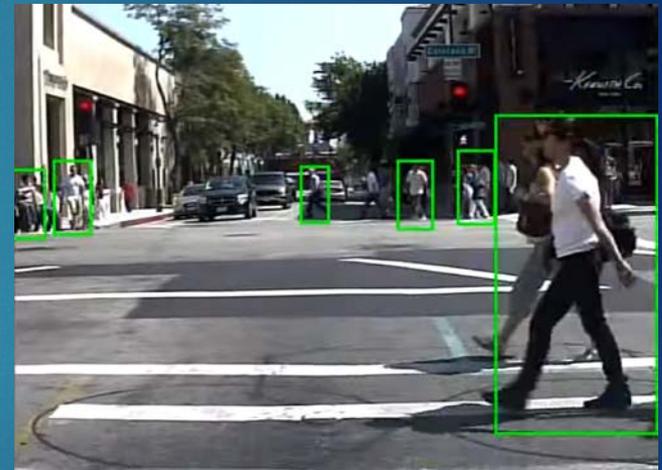
Datasets

10

- ▶ Caltech Pedestrian
- ▶ BDD100K

Caltech Pedestrian - Dataset

- ▶ Sehr bekanntes Dataset mit Fußgängern
- ▶ 10 Stunden 640x480 30Hz Videos aus fahrenden Auto
- ▶ 250.000 Frames
- ▶ 2300 einzigartige Fußgänger



[11]

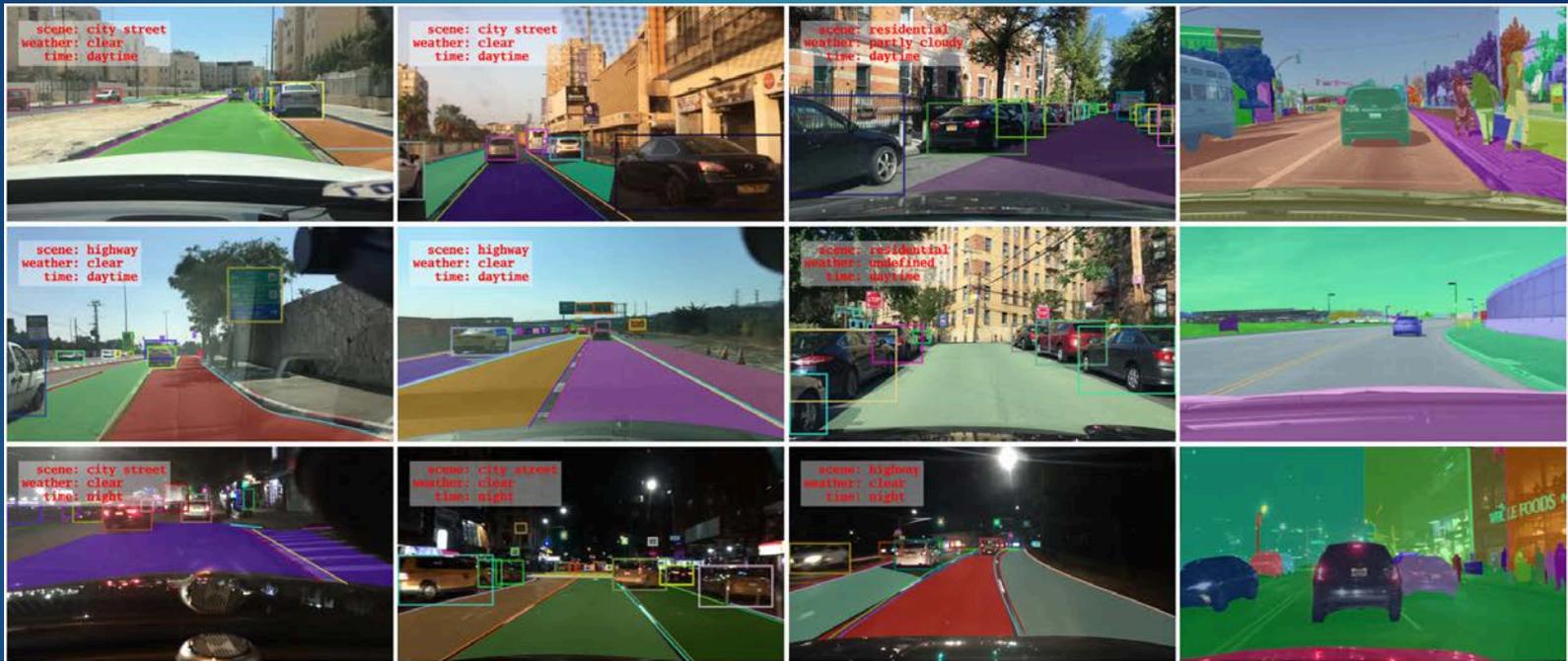
BDD100K - Dataset

12

- ▶ Größtes Dataset mit Aufzeichnungen aus Auto
- ▶ Seit 2018 zugänglich
- ▶ 100.000 HD Video Sequenzen
- ▶ 1.100 Stunden Videos
- ▶ Unterschiedliche Tageszeiten, Wetterbedingungen und Fahrscenarios
- ▶ Straßenobjekterkennung: (Label)
 - ▶ 100.000 Bilder mit 2D Rahmen versehen
 - ▶ Ampeln, Straßenschilder, Personen, Fahrräder, Autos, Züge, ...

BDD100K - Dataset

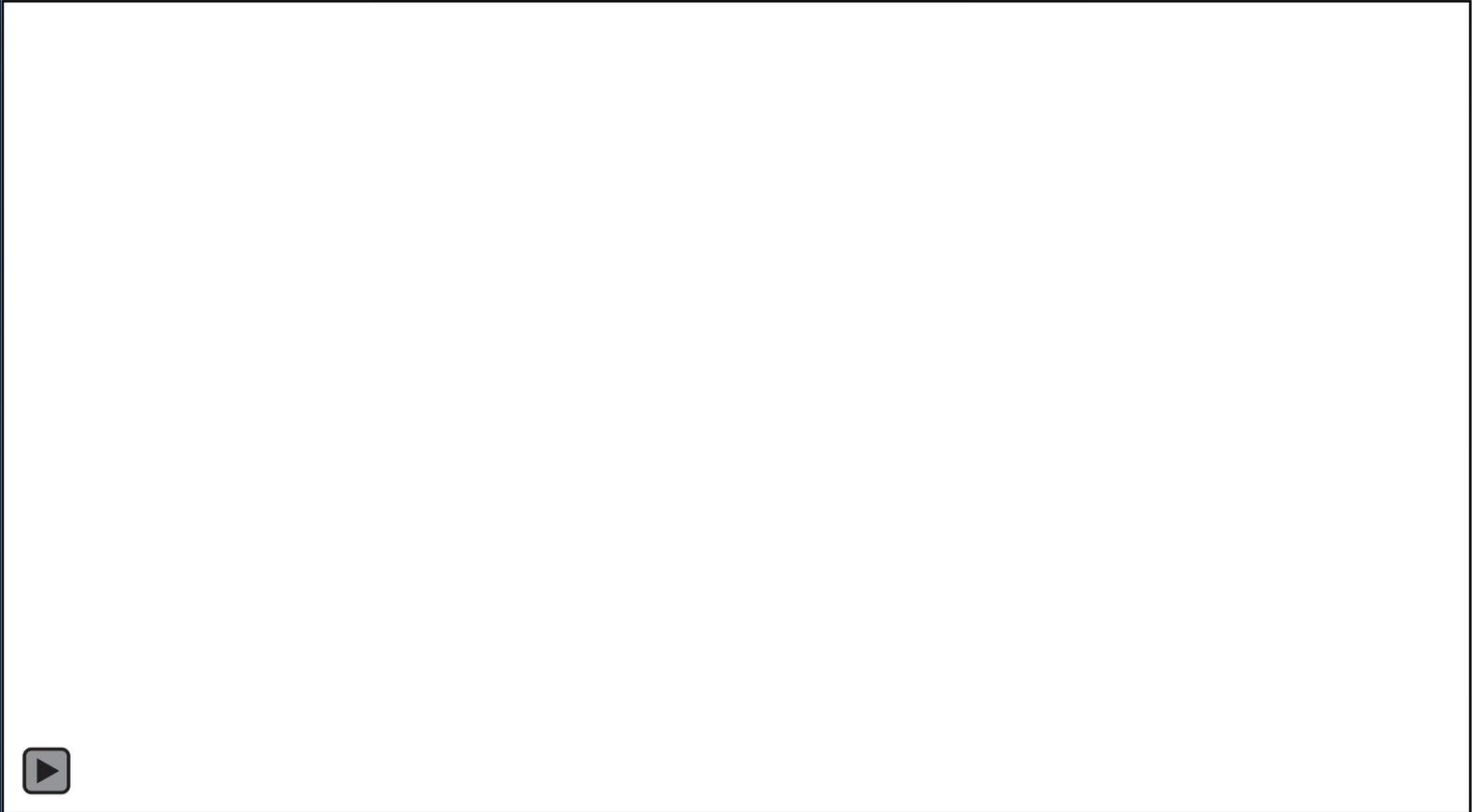
13



[12]

Aktueller Stand - Pong

14



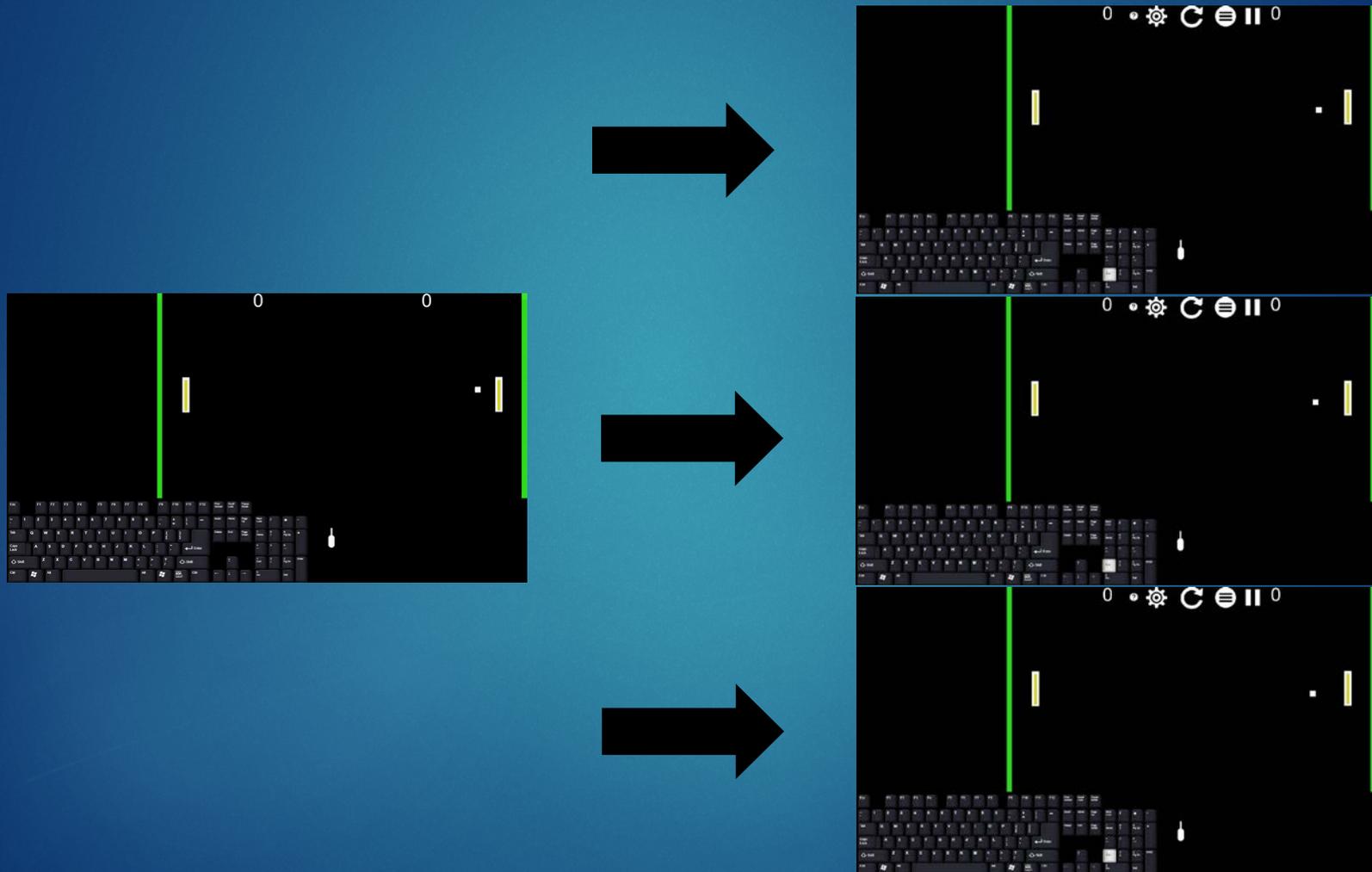
Aktueller Stand - Pong

15

- ▶ Datenvorverarbeitung
 - ▶ Video -> Frames
 - ▶ Labeln
 - ▶ Zuschneiden
 - ▶ Verkleinern
 - ▶ Data Reshape für LSTM

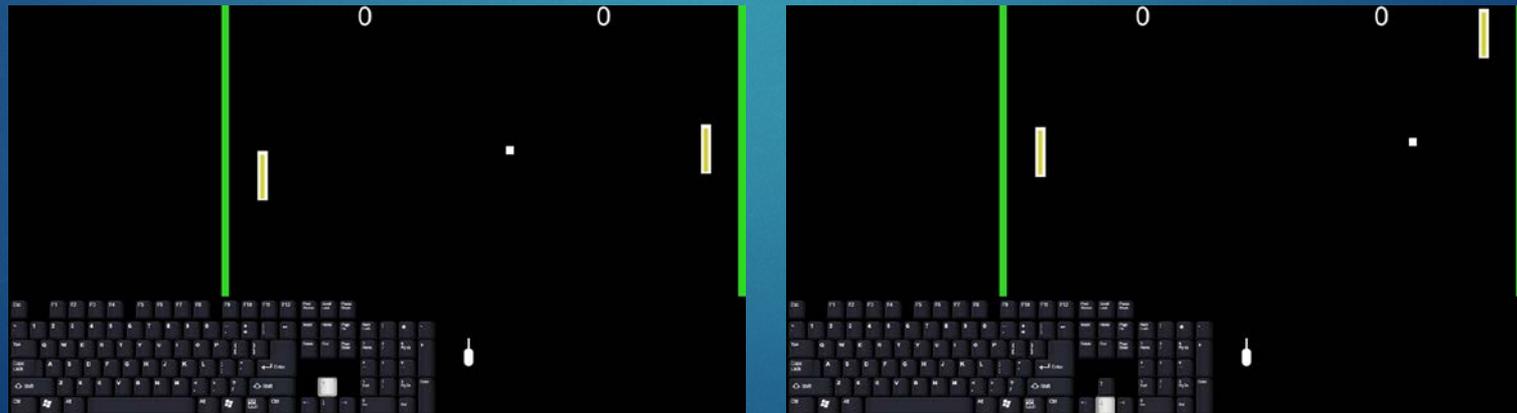
Video -> Frames

16



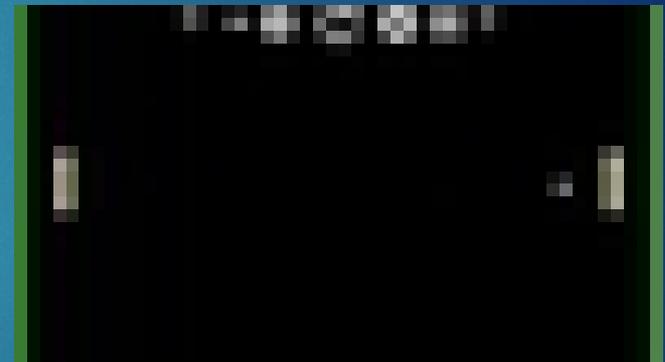
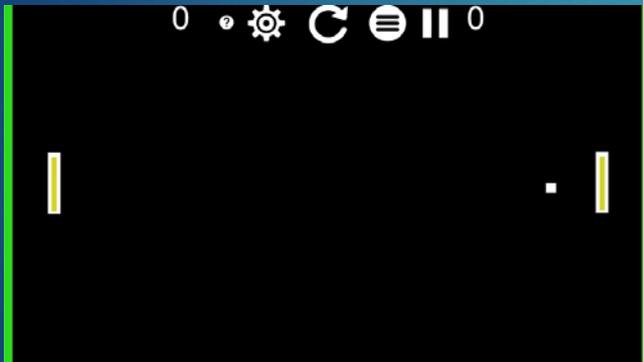
Motivation / Fragestellung - Methodik - **Aktueller Stand** - Risiken - Ausblick

- ▶ Einlesen der Bilder (als Graubild)
- ▶ Überprüfung des Grauwertes auf den Tasten OBEN und UNTEN
- ▶ Grauwert gedrückter Taste > 200
- ▶ Grauwert nicht gedrückter Taste < 200



Zuschneiden und Verkleinern

- ▶ Tastatur aus den Bildern entfernen
- ▶ Auflösung der Bilder drastisch reduzieren
910x510 -> 50x28



Data Reshape für LSTM - Allgemein

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

[13]

Data Reshape für LSTM - Konkret

Einzelbilder:

```
(996, 4, 1400)
[[[ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]]

[[ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]]

[[ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]]

...

[[ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]]

[[ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]]

[[ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  0.  1. 109.]
 [ 95.  0.  1. ...  1.  0. 109.]]]
```

Dazugehörige Label:

```
(996, 3)
[[[1. 0. 0.]
 [1. 0. 0.]
 [1. 0. 0.]
 ...
 [1. 0. 0.]
 [0. 0. 1.]
 [1. 0. 0.]]]]
```

Aktueller Stand - Pong

21

- ▶ Neuronales Netz
 - ▶ Reines LSTM Netz
 - ▶ Input: Videosequenz ; Output: UP, DOWN, NONE
 - ▶ Aus drei Frames => Klassifikation der Aktion
 - ▶ Genauigkeit $\approx 82\%$

- ▶ Trainingsdaten
 - ▶ Nicht genug Trainingsdaten
 - ▶ Qualität der Trainingsdaten nicht ausreichend
- ▶ Alternativen
 - ▶ TORCS (The Open Racing Car Simulator)
 - ▶ Daten selbst erzeugen
- ▶ Probleme der Alternativen:
 - ▶ Simulation in die Realität übertragen
 - ▶ Zu aufwändig

- ▶ Grundprojekt:
 - ▶ „PONG-LSTM-Projekt“ abschließen => Genauigkeit > 95%
- ▶ Hauptprojekt:
 - ▶ „PONG-CNN/LSTM-Projekt“ starten
 - ▶ Komplexere Videosequenzen
- ▶ Master Arbeit:
 - ▶ Das eigentliche Problem betrachten => Fußgängererkennung
 - ▶ Auswahl geeigneter Trainings- und Testdaten -> BDD100K?
 - ▶ Konkretisierung des Themas -> Fußgängererkennung zu allgemein

- ▶ F. Chollet, Deep Learning with Python, Manning, 2017
- ▶ A. Meisel, Vorlesungsfolien - Modellierung dynamischer Systeme, 2018
- ▶ J. Brownlee, CNN Long Short-Term Memory Networks, 2017: <https://machinelearningmastery.com/cnn-long-short-term-memory-networks/>
- ▶ The Unreasonable Effectiveness of Recurrent Neural Networks, 2015: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- ▶ S. Yin et al., Multi-CNN and Decision Tree Based Driving Behavior Evaluation, 2017
- ▶ S. Ren et al., Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, 2017
- ▶ X. Zhao et al., A Faster RCNN-based Pedestrian Detection System, 2016

Bildquellen

25

- ▶ [01] <https://www.automobil-produktion.de/technik-produktion/forschung-entwicklung/fahrerloses-auto-laechelt-fussgaenger-an-115.html>
- ▶ [02] <https://www.vorsicht-toter-winkel.de/index.php?id=1406>
- ▶ [03] <https://www.h-its.org/pso-lehre/vorlesung-python-programming-for-scientists/>
- ▶ [04] <https://anaconda.org/ijstokes/data-science-apps-with-anaconda/notebook>
- ▶ [05] <https://keras.io/>
- ▶ [06] <https://www.analyticsvidhya.com/blog/2018/01/tensorflow-1-5-google-released/>
- ▶ [07] <https://www.pcwelt.de/a/kuenstliche-intelligenz-eine-standortbestimmung,3450114>
- ▶ [08] https://leonardoaraujosantos.gitbooks.io/artificial-intelligence/content/recurrent_neural_networks.html
- ▶ [09] <http://www.cs.utexas.edu/users/ml/papers/venugopalan.naacl15.pdf>
- ▶ [10] <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- ▶ [11] <https://www.youtube.com/watch?v=P4p7SlziGEs>
- ▶ [12] <http://bdd-data.berkeley.edu/wad-2018.html>
- ▶ [13] http://tutorial.math.lamar.edu/Classes/DE/LA_Matrix.aspx