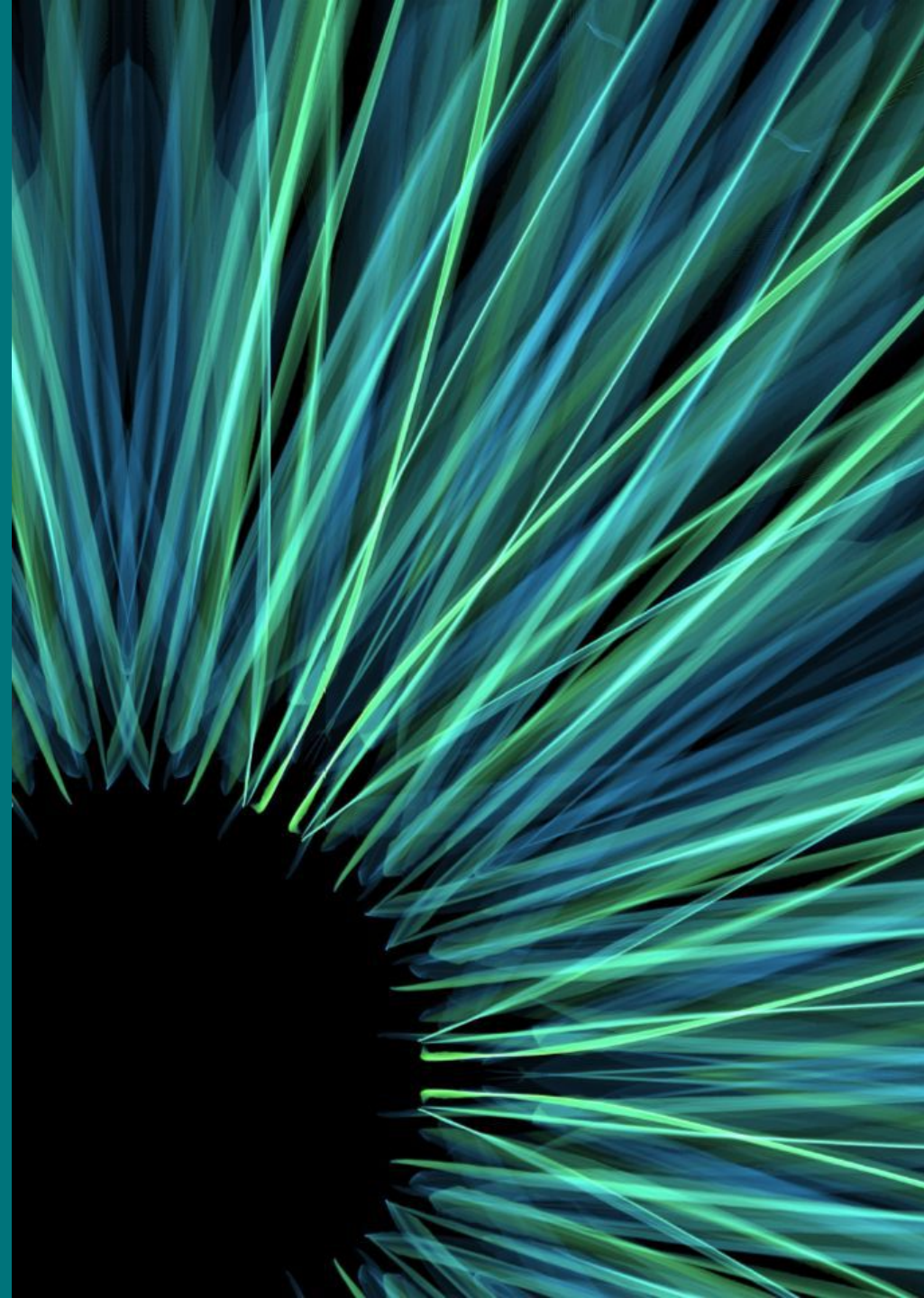


Einführung in Künstliche Intelligenz

Webinarreihe “KI für den Mittelstand”
28. Oktober 2020



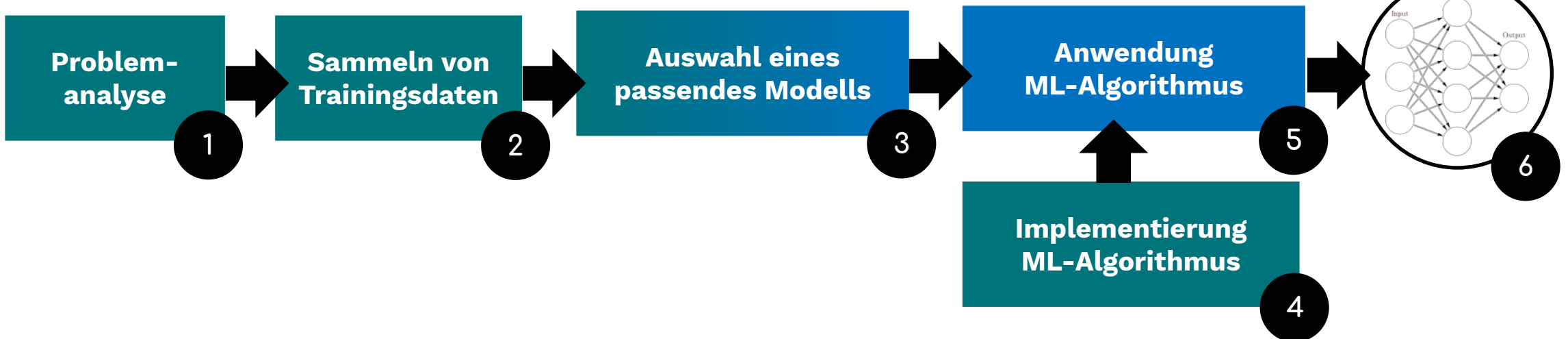
Maschinelles Lernen und Künstliche Neuronale Netze

Eine neue Herangehensweise an Softwareentwicklung

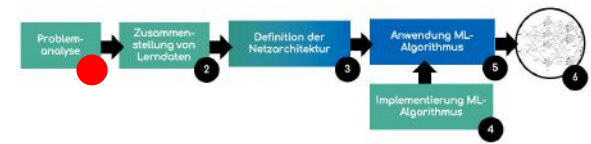
“Traditionelle“ Software



KI-Software basierend auf Künstlichen Neuronalen Netzen



Problemanalyse



” AI work requires carefully choosing A and B and providing the necessary data to help the AI figure out the A→B relationship. Choosing A and B creatively has already revolutionized many industries. It is poised to revolutionize many more. ”

Andrew Ng

- Verständnis der **Verarbeitungslogik von Input zu Ausgabe** verliert an Relevanz
- Fokus liegt auf Definition des gewünschten Outputs und **Identifikation der notwendigen Inputdaten**



Ein Beispiel: Gestensteuerung

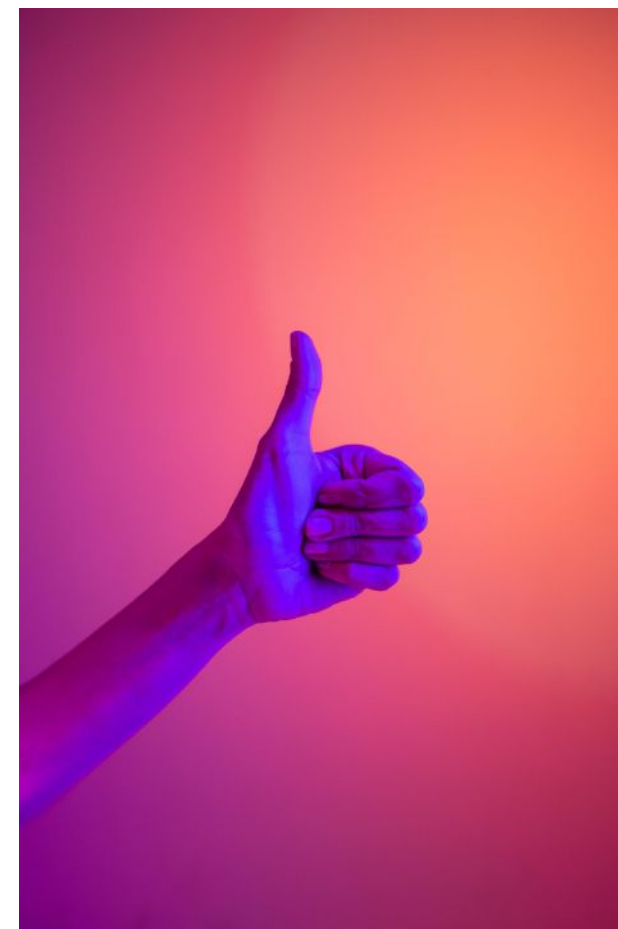
Welche Ausgabe oder “Aussage” soll unser System machen?

Gegeben eine Liste von möglichen Gesten, welche davon ist auf dem Bild zu sehen?

- Winken
- “OK”
- (keine davon)

Welchen Input braucht unser System, um diese Aussage machen zu können?

Bilder einer Kamera, die wir z.B. an der Front unserer Maschine anbringen könnten.



Warum wir nicht immer verstehen können, wie Daten verarbeitet werden müssen

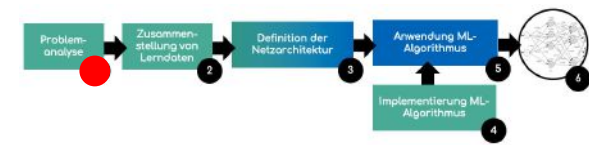


Image	Rot001	Grün001	Blau001	Rot002	Grün002	Blau002	...
001	157	135	122	142	118	102	...
002	63	106	20	95	143	51	...
003	101	151	208	100	152	207	...
004	64	52	48	63	55	54	...
005	110	100	90	113	104	94	...

So schaut ein Programmierer auf "Bilder"

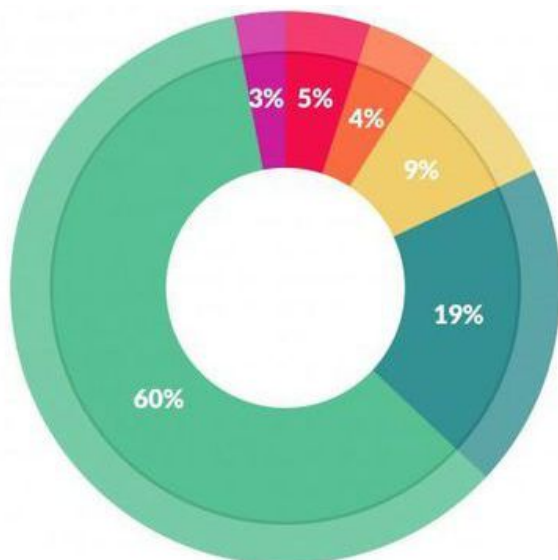




Sammeln von Trainingsdaten

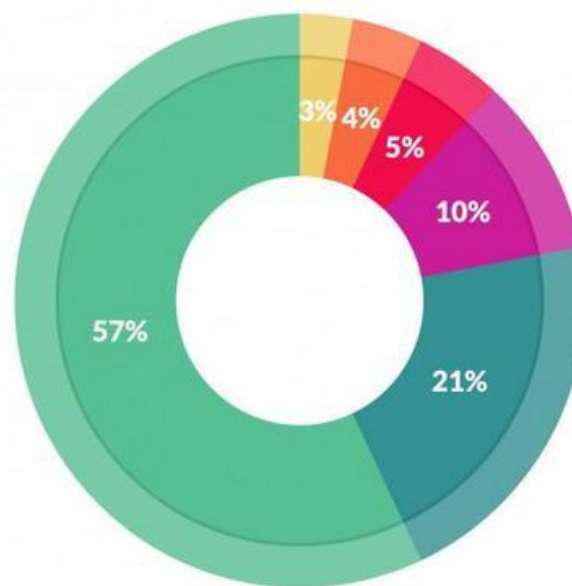
What data scientists spend the most time doing

- Building training sets: 3%
- Cleaning and organizing data: 60%
- Collecting data sets; 19%
- Mining data for patterns: 9%
- Refining algorithms: 4%
- Other: 5%



What's the least enjoyable part of data science?

- Building training sets: 10%
- Cleaning and organizing data: 57%
- Collecting data sets: 21%
- Mining data for patterns: 3%
- Refining algorithms: 4%
- Other: 5%

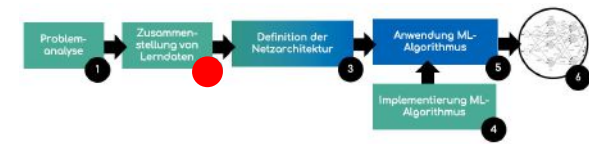




Demo: Gestensteuerung im Handumdrehen

<https://teachablemachine.withgoogle.com/train>

Sammeln von Trainingsdaten



” Among leading AI teams, many can likely replicate others’ software in, at most, 1–2 years. But it is exceedingly difficult to get access to someone else’s data. Thus data, **rather than software, is the defensible barrier for many businesses.** ”

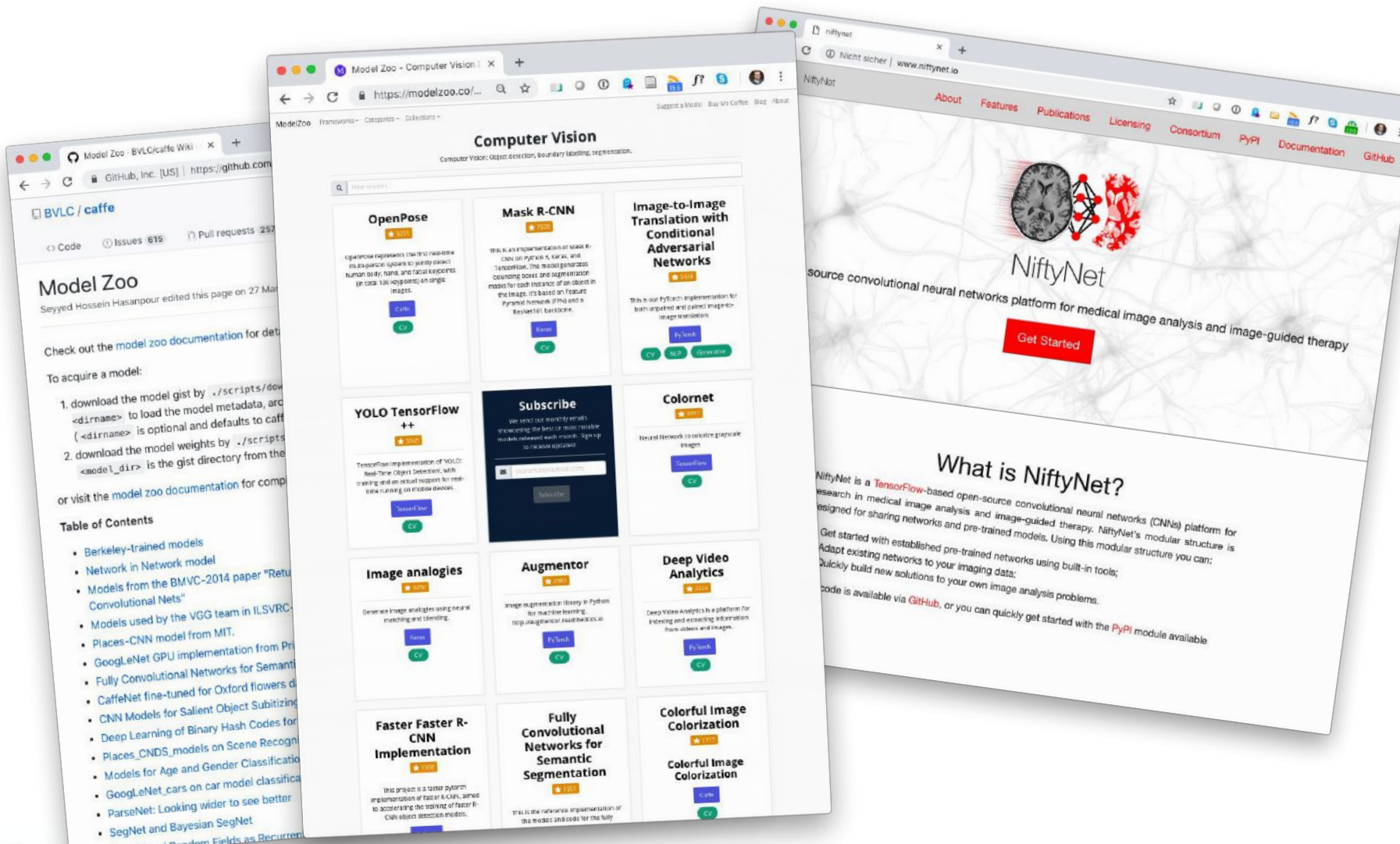
Andrew Ng

- Wachsender **Wert von Daten**
- **Digitalisierung Voraussetzung** für skalierte KI-Lösungen
- **Datenstrategie auf Basis von Anwendungspotenzialen** sollte erarbeitet werden!

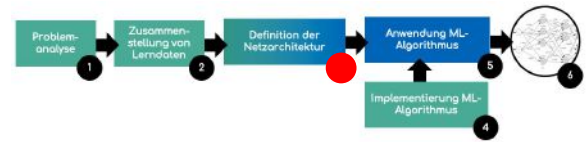


Auswahl eines passendes Modells

- Zahlreiche **Ressourcen sind kostenfrei** und sollten als Grundlage genutzt werden (z.T. sogar bereits fertig trainiert!)



Auswahl eines passendes Modells



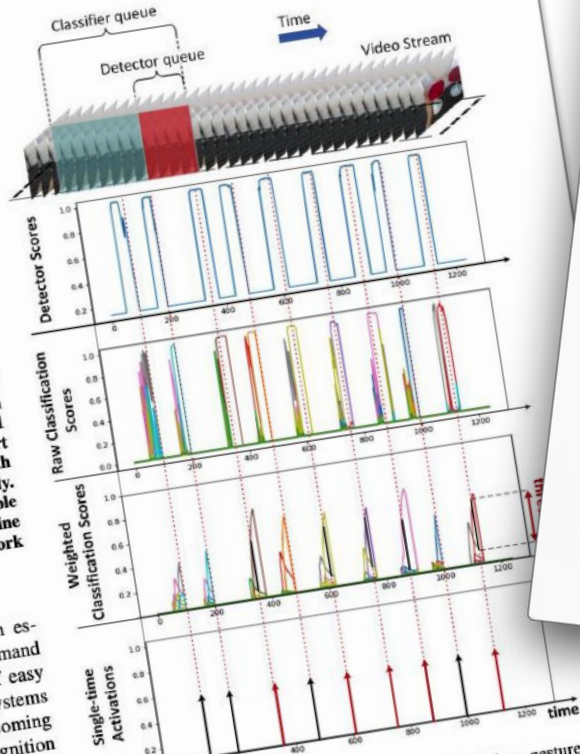
Real-time Hand Gesture Detection and Classification Using Convolutional Neural Networks

Okan Köpüklü¹, Ahmet Gunduz¹, Neslihan Kose², Gerhard Rigoll¹
¹ Institute for Human-Machine Communication, TU Munich, Germany
² Dependability Research Lab, Intel Labs Europe, Intel Deutschland GmbH, Germany

Abstract—Real-time recognition of dynamic hand gestures from video streams is a challenging task since (i) there is no indication when a gesture starts and ends in the video, (ii) performed gestures should only be recognized once, and (iii) the entire architecture should be designed considering the memory and power budget. In this work, we address these challenges by proposing a hierarchical structure enabling offline-working convolutional neural network (CNN) architectures to operate online efficiently by using sliding window approach. The proposed architecture consists of two models: (1) A detector which is a lightweight CNN architecture to detect gestures and (2) a classifier which is a deep CNN to classify the detected gestures. In order to evaluate the single-time activations of the detected gestures, we propose to use Levenshtein distance as an evaluation metric since it can measure misclassifications, multiple detections, and missing detections at the same time. We evaluate our architecture on two publicly available datasets - EgoGesture and NVIDIA Dynamic Hand Gesture Datasets - which require temporal detection and classification of the performed hand gestures. ResNeXt-101 model, which is used as a classifier, achieves the state-of-the-art offline classification accuracy of 94.04% and 83.82% for depth modality on EgoGesture and NVIDIA benchmarks, respectively. In real-time detection and classification, we obtain considerable early detections while achieving performances close to offline operation. The codes and pretrained models used in this work are publicly available¹.

I. INTRODUCTION

Computers and computing devices are becoming an essential part of our lives day by day. The increasing demand for such computing devices increased the necessity of easy to use human-computer interfaces. For this reason, systems and control are becoming more and more important in the field of human-computer interaction.



modelzoo.co
Model Zoo - OpenPose Caffe Model

OpenPose

Linux | build error

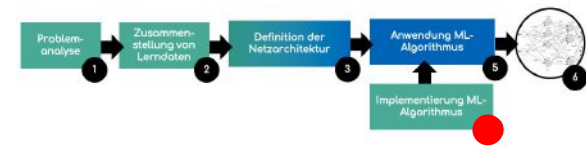
OpenPose represents the first real-time multi-person system to jointly detect human body, hand, and facial keypoints (in total 130 keypoints) on single images.

OpenPose

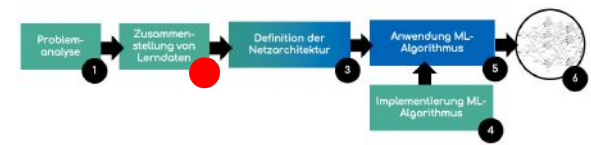
Features

10323v3 [cs.CV] 18 Oct 2019

Implementierung ML-Algorithmus



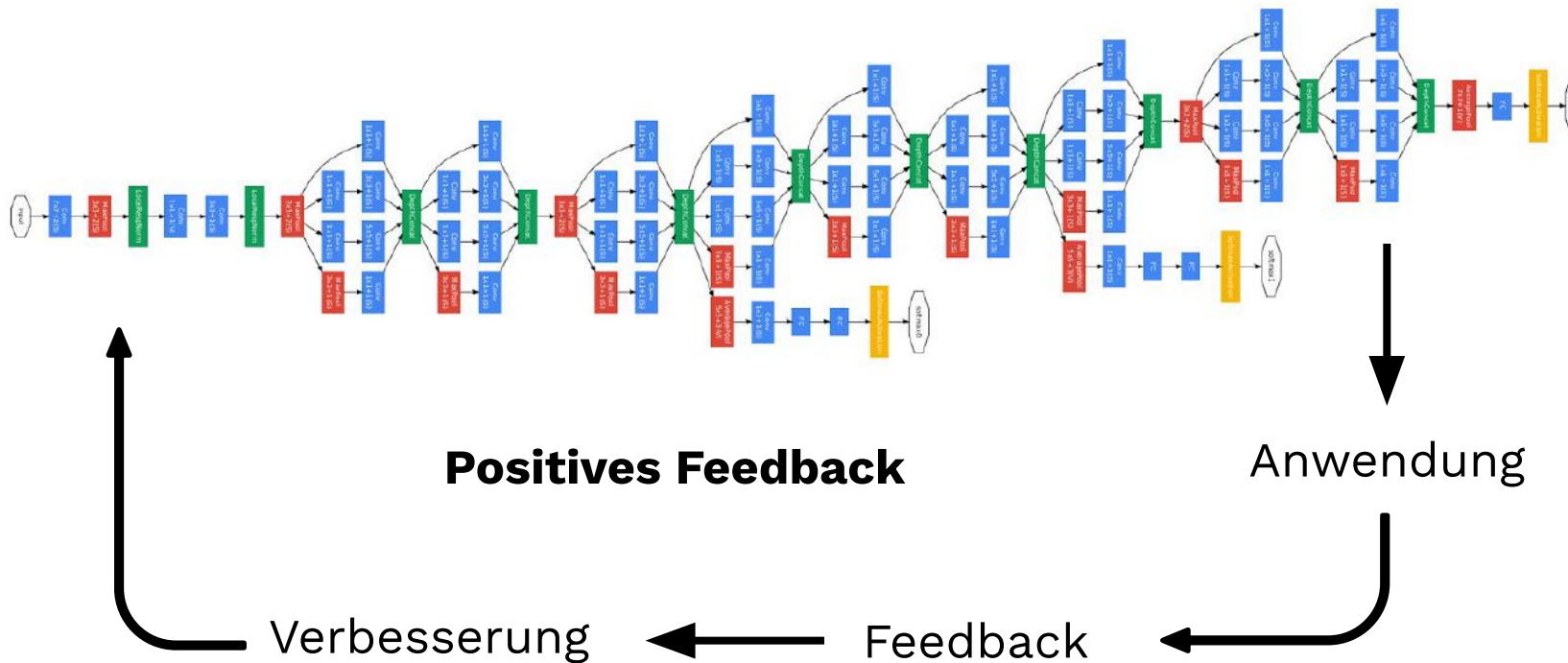
- Hier gilt: **Das Rad nicht neu erfinden!**
- Verschiedene sog. “Frameworks” als **Open Source Software frei verfügbar**
- Zahlreiche **Online-Communities** bieten Raum zum Austausch und Hilfestellungen



Demo: Gestensteuerung im Handumdrehen

<https://teachablemachine.withgoogle.com/train>

Das Ergebnis: Eine trainierte KI!



- Lernende Systeme sind **in der Regel nicht fehlerfrei**
- In der Anwendung sollten weiter Daten gesammelt werden, um das **System kontinuierlich zu verbessern**

KI Mythen

- KI lernt **selbständig**
- KI's entwickeln ist **schwierig**
- Der **Wert** liegt im **Algorithmus**
- Maschinen **denken wie Menschen**
- KI ist "nur noch ein **digitales Werkzeug**"
- Wir kennen die **Grenzen** des Möglichen

vs. Realität

- KI bedeutet **induktives Lernen** und erfordert nur **wenig Programmierung**
- KI arbeitet für **bestimmte Probleme schneller und besser** als wir Menschen
- Prototypen sind häufig einfach, **Skalierung in die Anwendung sehr komplex**
- KI hat **keinen "gesunden Menschenverstand" oder ein "Bewusstsein"**, das über die Trainingsdaten hinausgeht
- Prozesse müssen für die Bearbeitung durch **Mensch und Maschine** optimiert werden
- **Sicherheit** ist große Herausforderung bei automatisierten Entscheidungen
- KI erfordert **umfassende Aus- und Weiterbildung**

Vielen Dank für die Aufmerksamkeit

Dr. Hendrik Brakemeier

Head of AI Journey

brakemeier@unternehmertum.de

<https://www.linkedin.com/in/hbrakemeier>

appliedAI Initiative

UnternehmerTUM GmbH

Lichtenbergstraße 6

685748 Garching

info@appliedAI.de

<https://appliedAI.de>