

Evaluation vortrainierter neuronaler Netzwerke zur Anwendung auf die Zählung von Personen mit Objektdetektion

Autor: Oskar Rudolf - Referentin: Prof. Dr. Elke Hergenröther - Koreferent: Prof. Dr. Sebastian Döhler
Fachbereich Mathematik und Fachbereich Informatik

Motivation

- Durch leistungsfähigere Prozessoren können immer mehr KI-Anwendungen im breiten Feld eingesetzt werden.
- Durch Bereitstellung vortrainierter KI-Modelle ist eine Expertise für Training und Datenakquise für die Anwendung nicht länger obligatorisch.
- Im Anwendungsfall Computer Vision können KI-Systeme in autonomen Smarthome-Umgebungen zu einer Reduzierung der Umweltbelastung und Einsparung von Ressourcen beitragen.
- Eine breite Untersuchung von vortrainierten Modelle ist erforderlich.
- Für den Proof-of-Concept soll eine Spezialisierung auf die Zählung von Personen durchgeführt werden.

Ziele

Evaluation vortrainierter neuronaler Netze für die Anwendung in Computer Vision

- ✓ Vergleich verschiedener State-of-the-Art Architekturen
- ✓ Vergleich der Laufzeit bei horizontaler Herunterskalierung der Hardware
- ✓ Leistung der Personenzählung in Realumgebung

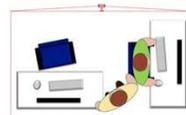
+ Subjektive Pilotevaluierung von Benutzerschnittstellen für die Anwendung vortrainierter Modelle.

Theorie

- Im Vergleich mit pixelbasierten Computer Vision Methoden sind spezielle neuronale Netze (CNN) in Bezug auf Objekterkennung deutlich leistungsfähiger.
- Zu den geeigneten Netzwerkarchitekturen gehören YOLO¹⁾, Resnet²⁾ EfficientDet³⁾, und CenterNet⁴⁾.

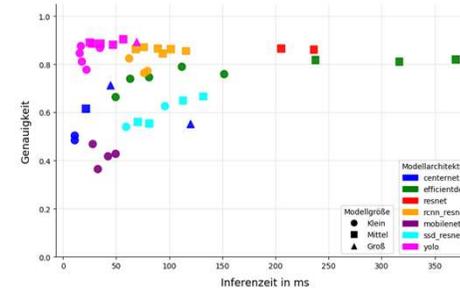
Methode

- Rasteranalyse mit 46 vortrainierten Modellen unter Nutzung von darauf zugeschnittenen Benutzerschnittstellen (*kein* Hyperparametertuning).
- Experiment 1: Evaluierung der Objekterkennungskapazitäten der Modelle anhand des COCO-Datensatzes⁵⁾ gefiltert auf Bilder mit Personen.
- Augmentierung der Daten zur Untersuchung der Generalisierungsfähigkeit der Modelle.
- Experiment 2 & 3 mit drei besonders performanten Modellen mit unterschiedlichen Modellgrößen:
 - Untersuchung der Laufzeit in einem Low-End-System
 - Untersuchung in einer simulierten Realumgebung.
- Zu den relevanten Metriken gehören Genauigkeit, Sensitivität, Präzision, Modellgröße und Inferenzzeit.



Ergebnisse

Vergleich der NN-Architekturen



- N = 45 vortrainierte Modelle (1 Ausschluss); Median Inferenzzeit 70ms; ØGenauigkeit 76%.
- YOLO-Modellarchitekturen waren am leistungsfähigsten.
- Die Analyse in Bezug auf Bildaugmentierungen lieferte keine Unterschiede.

Vergleich der Inferenzzeit* mit Low-End System

	High-End RTX A5000; 49-10900X; 64Gb RAM	Low-End Raspberry Pi5b	Anzahl Parameter in Millionen
Kleines Modell	16,79	18,89	25,89
Mittleres Modell	27,63	112,01	68,20
Großes Modell	69,40	103,19	155,48

*Angaben in Sekunden

Leistung in Realumgebung (Echtzeit)

	Genauigkeit	Sensitivität	Präzision
Kleines Modell	95,50 %	0,95,9 %	99,5 %
Mittleres Modell	95,40 %	0,98,0 %	97,3 %
Großes Modell	97,10 %	0,97,1 %	100,00 %

Diskussion

- Die hohe allgemeine Leistung der vortrainierten Modelle ist ein Indikator für deren Einsetzbarkeit in einem breiten Anwendungsspektrum ohne ressourcen-aufwendiges Training eigener Modelle.
- Im Anwendungskontext der Personenzählung sind die Objekterkennungsarchitekturen der YOLO-Modellfamilie besonders gut geeignet.
- Solange genügend RAM zur Verfügung steht, erhöht sich die Laufzeit in niedrigskalierten Systemen unwesentlich.
- In der Realumgebung konnten die Modelle auch bei Abweichungen der Kameraperspektive und den Lichtverhältnissen eine bemerkenswert hohe Prognosegenauigkeit liefern.
- Benutzerschnittstellen sind teilweise noch fehlerbehaftet oder haben obsoletere Dokumentation.

Fazit

- Der leichte Zugang zu performanten Modellen ohne Expertenkenntnisse kann helfen, KI-Technologie gesellschaftstauglicher zu machen.
- Neuronale Netze können insbesondere gut eingesetzt werden, um video-basierte Verfahren zu optimieren oder autonome Prozesse zu steuern.
- Zusätzliche Forschung zur Optimierung des Speicherbedarfs der Modelle kann zu einer besseren Einsetzbarkeit in eingebetteten Systemen führen.