

Ostbayerische Technische Hochschule Amberg-Weiden
Fakultät Elektrotechnik, Medien und Informatik

Studiengang Künstliche Intelligenz

Studienarbeit Deep Vision

von

André Kestler

**Vergleich der YOLOX- und YOLOv8-Modelle für die
Objekterkennung im Kontext des Udacity Self Driving
Car-Datensatzes**

Bearbeitungszeitraum: von 21. Juni 2023
bis 19. Juli 2023

1. Prüfer: Prof. Dr. phil. Tatyana Ivanovska

Inhaltsverzeichnis

1	Einleitung	1
1.1	Aufgabenstellung	1
1.2	Übersicht	1
2	Datensatz	2
2.1	Beschreibung	2
2.2	Klassenaufteilung	2
2.3	Aufbau	2
2.3.1	Rohdaten	2
2.3.2	YOLOX	2
2.3.3	YOLOv8	2
3	Modell 1: YOLOX	5
3.1	Architektur	5
3.2	Verlustfunktion	5
3.3	Modellauswertung	5
4	Modell 2: YOLOv8	6
4.1	Architektur	6
4.2	Verlustfunktion	6
4.3	Modellauswertung	6
5	Zusammenfassung und Ausblick	7
	Literaturverzeichnis	8
	Abbildungsverzeichnis	9

6	Anhang	10
6.1	Ordnerstruktur	10
6.2	Dateien: Datensatz	10
6.3	Dateien: YOLOX	10
6.4	Dateien: YOLOv8	10

Abkürzungsverzeichnis

YOLO You only look once

Kapitel 1

Einleitung

1.1 Aufgabenstellung

Im Rahmen der Vorlesung Deep Vision ist ein Projekt im Themenbereich des Kurses zu bearbeiten. Die Bearbeitung erfolgt als Einzelarbeit. Als Projekt werden zwei YOLO (You only look once) Netzwerke mit dem Udacity Self Driving Car Datensatz trainiert und miteinander verglichen. In der folgenden Arbeit werden YOLOX und YOLOv8 verwendet.

1.2 Übersicht

In Kapitel 2 wird zunächst der Datensatz beschrieben. Dabei wird auf die Klasseneinteilung und die Datenstruktur eingegangen. In Kapitel 3 wird das YOLOX-Netzwerk vorgestellt. Dabei wird auf die verwendete Verlustfunktion, die Architektur und die Ergebnisse mit dem Datensatz eingegangen. Kapitel 4 beschreibt die verwendete Architektur von YOLOv8. Außerdem werden die Ergebnisse anhand des verwendeten Datensatzes präsentiert. Das letzte Kapitel Zusammenfassung und Ausblick befasst sich mit dem direkten Vergleich der Ergebnisse und gibt einen Ausblick über die weitere Bearbeitung. Im Anhang wird beschrieben, wie die angegebenen Skripte verwendet werden, um die Netzwerke selbst zu trainieren.

Kapitel 2

Datensatz

2.1 Beschreibung

Der Udacity Self Driving Car Dataset [2] ist eine umfangreiche Sammlung von Bildern, die von Kameras in Fahrzeugen aufgenommen wurden. Der Datensatz beinhaltet 15000 Samples mit einer Auflösung von 512x512 Pixeln. Er besteht aus Bildern und die zugehörigen Annotationen, die Informationen über die enthaltenen Objekte in der Umgebung enthalten.

Die Bilder in dem Datensatz umfassen verschiedene Szenarien im Straßenverkehr. Darunter befinden sich Stadt- und Landstraßen. Die Bilder wurden bei unterschiedlichen Lichtbedingungen aufgenommen, um eine breite Vielfalt an Situationen in den Daten abzudecken.

2.2 Klassenaufteilung

2.3 Aufbau

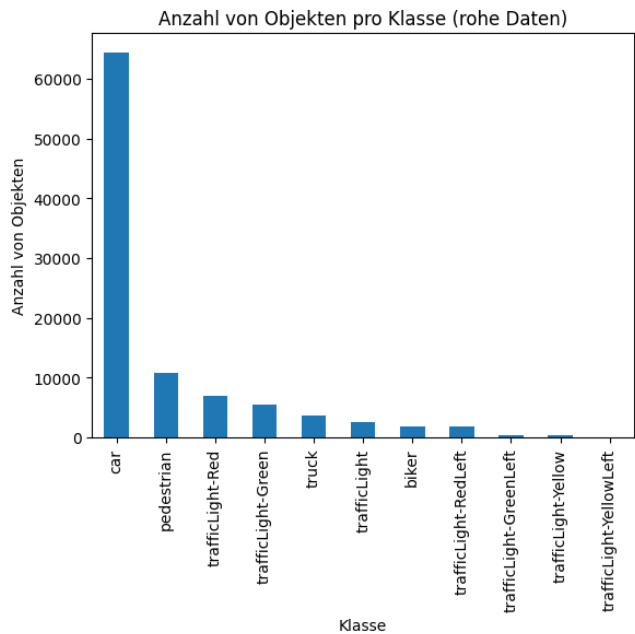
2.3.1 Rohdaten

2.3.2 YOLOX

2.3.3 YOLOv8

Klassen	Anzahl	Index
car	64399	0
pedestrian	10806	1
trafficLight-Red	6870	2
trafficLight-Green	5465	3
truck	3623	4
trafficLight	2568	5
biker	1846	6
trafficLight-RedLeft	1751	7
trafficLight-GreenLeft	310	8
trafficLight-Yellow	272	9
trafficLight-YellowLeft	14	10

(a) Klassenverteilung der rohen Daten als Tabelle

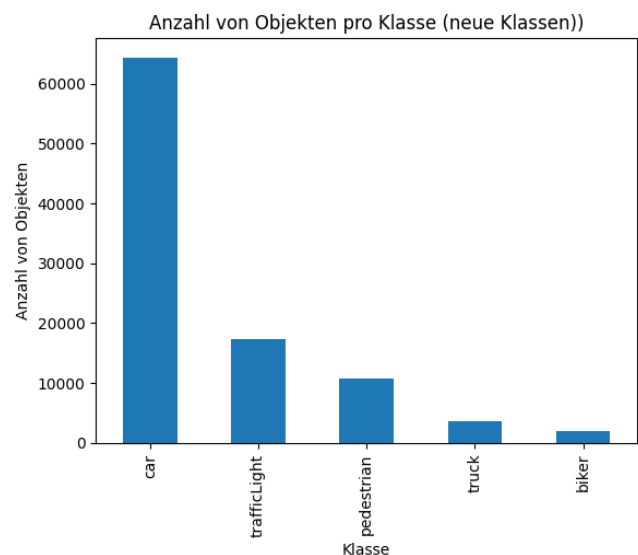


(b) Klassenverteilung der rohen Daten als Diagramm

Abbildung 2.1: Klassenverteilung der rohen Daten

Klassen	Anzahl	Index
car	64399	0
pedestrian	10806	1
trafficLight	17250	2
truck	3623	3
biker	1846	4

(a) Klassenverteilung der neuen Klassen als Tabelle



(b) Klassenverteilung der neuen Klassen als Diagramm

Abbildung 2.2: Klassenverteilung der neuen Klassen

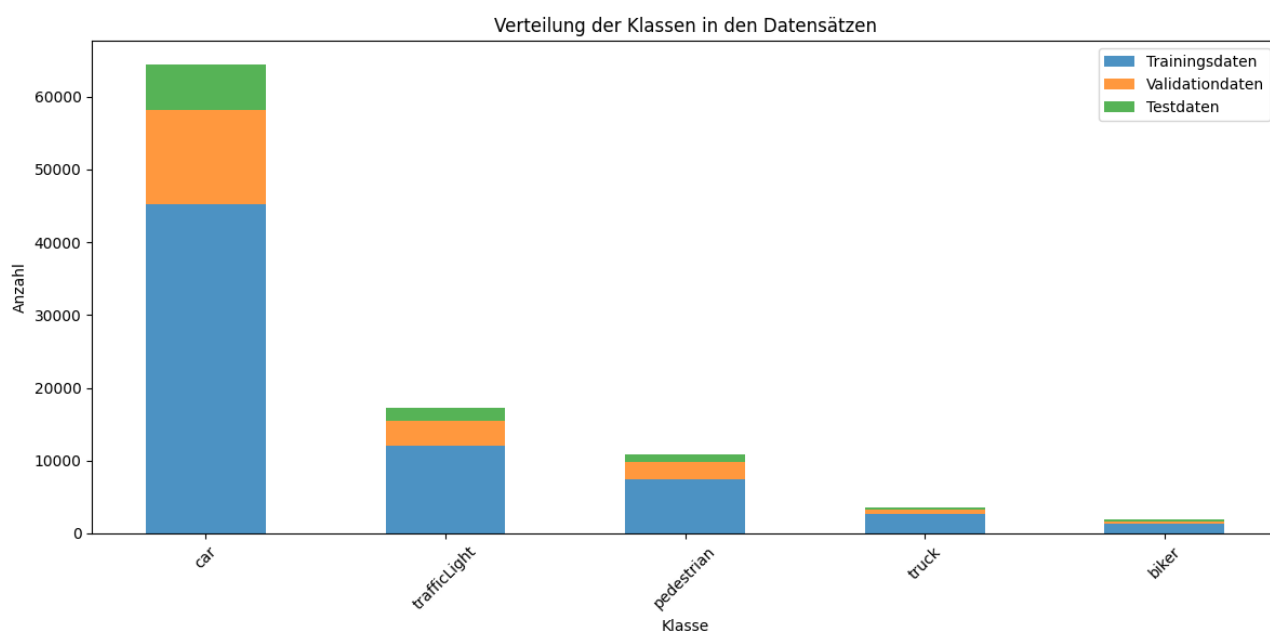


Abbildung 2.3: Aufteilung in Trainings-, Validation- und Testdaten

Kapitel 3

Modell 1: YOLOX

3.1 Architektur

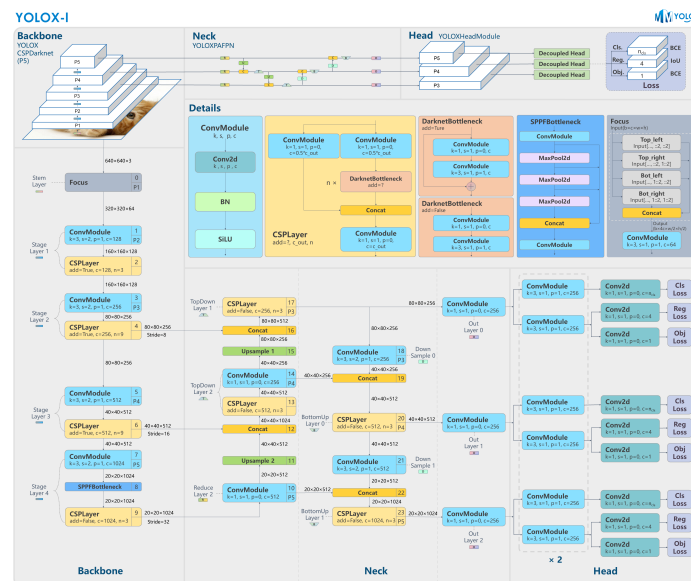


Abbildung 3.1: Übersicht über die Architektur von YOLOX. Quelle: [1]

3.2 Verlustfunktion

3.3 Modellauswertung

Kapitel 4

Modell 2: YOLOv8

4.1 Architektur

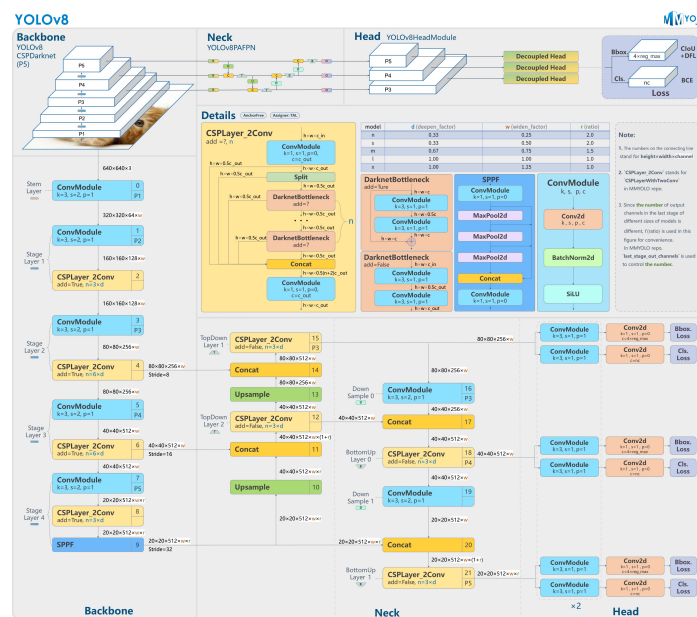


Abbildung 4.1: Übersicht über die Architektur von YOLOv8. Quelle: [1]

4.2 Verlustfunktion

4.3 Modellauswertung

Kapitel 5

Zusammenfassung und Ausblick

Literaturverzeichnis

- [1] CONTRIBUTORS, MMYOLO: *Udacity Self Driving Car Dataset*. <https://github.com/open-mmlab/mmyolo>, Abruf: 10. Juli 2023
- [2] ZHANG, Edward: *Udacity Self Driving Car Dataset*. <https://www.kaggle.com/datasets/sshikamaru/udacity-self-driving-car-dataset>, Abruf: 23. Juni 2023

Abbildungsverzeichnis

2.1	Klassenverteilung der rohen Daten	3
2.2	Klassenverteilung der neuen Klassen	3
2.3	Aufteilung in Trainings-, Validation- und Testdaten	4
3.1	Übersicht über die Architektur von YOLOX	5
4.1	Übersicht über die Architektur von YOLOv8	6

Kapitel 6

Anhang

6.1 Ordnerstruktur

6.2 Dateien: Datensatz

6.3 Dateien: YOLOX

6.4 Dateien: YOLOv8