

Videoanalyse und Objekttracking

1st Bartolovic Eduard
Hochschule München
München, Deutschland
eduard.bartolovic0@hm.edu

2nd Thomas Willeit
Hochschule München
München, Deutschland
XXXXX@hm.edu

3rd Schäfer Julia
Hochschule München
München, Deutschland
j.schaefer0@hm.edu

Zusammenfassung—

I. KONZEPT

Ziel des Projektes ist es klassische Verfahren mit neuere DeepLearning Verfahren zu vergleichen. Dafür ist für beide arten eine Pipeline geschaffen worden.

Für die Klassische Verfahren wurde für die Objekterkennung Gausmixture verwendet und für das Tracking Sort. Für die DeepLearnig Verfahren wurde für die Objekterkennung ein Neuronales Netz und für das Tracking wurde Deepsort verwendet.

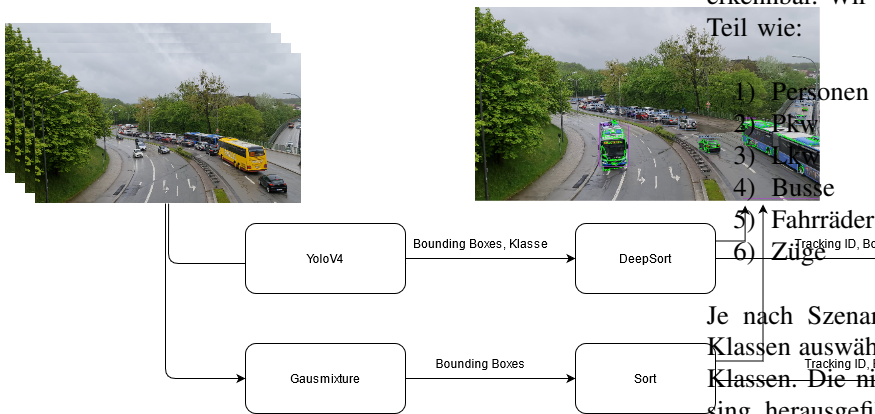


Abbildung 1. Konzept für das Projekt

II. OBJEKTERKENNUNG: GAUSS MIXTURE

III. OBJEKTERKENNUNG: YOLOV4

YOLO ist ein Neuronal Netz für die Echtzeit-Objekterkennung. YOLO ist die Abkürzung für 'You Only Look Once' was übersetzt 'Man sieht nur einmal hin' heißt. Es zerlegt die Aufgabe der Objekterkennung in zwei Teile. Einmal in die Objektposition über Bounding Boxes und dann in Klassifizierung zur Bestimmung der Objektklasse. Dieses System basiert auf Residualblöcken, Bounding-Box-Regression und IOU-Techniken. Das Bild in ein $S \times S$ -Gitter mit den Residualblöcken aufgeteilt. Wenn der Mittelpunkt eines Objekts in eine Gitterzelle fällt, ist diese Gitterzelle für die Erkennung dieses Objekts zuständig. Jede Gitterzelle sagt B Bounding Boxes und Konfidenzwerte für diese Boxen voraus [1].....

Die Klassifikation Pipeline besteht aus einem Feature Extrator

und einem Klassifikator....

Da Objekte meist immer ähnliche Formen haben kann man eine gewisse Menge an sogenannten Ankerboxen definieren welche als Basis Boundingboxen fungieren. Diese Ankerboxen kann man mittels Clusteralgorithmen wie K-Means und den Boundingboxen aus dem Datensatz berechnen....

In diesem Projekt verwendeten wir die 4te und damit aktuell neueste Version von Yolo [2]. So bietet jede Version inkrementelle Verbesserung zum jeweiligen Vorgänger.

Unser Modell wurde bereits mit dem Microsoft COCO Datensatz trainiert. Es sind 80 verschiedene Klassen erkennbar. Wir interessieren uns aber nur für einen kleineren Teil wie:

- 1) Personen
- 2) Pkw
- 3)
- 4) Busse
- 5) Fahrräder
- 6) Züge

Je nach Szenario lassen sich per Parameter die relevanten Klassen auswählen. Das Modell prädiziert trotzdem noch alle Klassen. Die nicht relevanten werden einfach im Postprocessingschritt herausgefiltert. Das Ergebnis einer Prädiktion ist eine Liste von Objekten welches aus den x,y Koordinate der oberen linken Punktes der Boundingbox, der Breite w, der Höhe h, der Klasse und des Konfidenzwerts. Ein weitere Postprocessingschritt ist die Non-Maxima-Suppression. BEISPIELBILD? Der Output kann überschneidende Boundingboxen besitzen die eigentlich zur selben Klasse gehören. Diese Duplikate können mittels der Non-Maxima-Suppression entfernt werden. Heirfür wird die IOU der Boxen berechnet. Sollte die IOU einen Threshold überschreiten dann wird das Duplikat entfernt.

A. Fehler in der Erkennung

...falsch Klassifizierung wegen schlechter Datensatz.... Zu großer Datensatz...

Problem zu viele Elemente nebeneinander.... wegen zu großen matschigen Grid zu wenig Ankerboxen???

IV. TRACKING: SORT

V. TRACKING: DEEPSORT

VI. ZÄHLEN VON OBJEKTEN

VII. ZUSTANDSERKENNUNG ZÜGE

LITERATUR

[1] <https://arxiv.org/pdf/1506.02640.pdf>

[2] <https://arxiv.org/pdf/2004.10934.pdf>

VIII. ANHANG