



Masterarbeit: Repurpose-GAN und Long-Tails

Problemstellung

Long-Tail Verteilungen sind insbesondere bei komplexeren Aufgabenstellungen schwierig in Modellhypothesen zu erfassen, da die unterrepräsentierten Observationen meist aufwändiger zu finden sind oder der Aufwand die Vielfalt an selteneren Mustern zu annotieren teuer ist. In dieser Masterarbeit soll die gesteigerte Label-Effizienz eines neuen Transfer-Learning Verfahrens vor dem Kontext einer Long-Tail-Verteilung geprüft werden.

Daten

Als Basis existiert ein Bilddatensatz von mehreren Millionen ungelabelten Fahrzeugaufnahmen sowie ein Datensatz von etwa 10.000 dieser Aufnahmen mit gelabelten Bauteilen der Fahrzeuge.

Die Fahrzeuge auf den Aufnahmen sind präzise anhand einer Taxonomie beschrieben und es bestehen Verteilungen zur Auftrittshäufigkeit der Fahrzeuge und deren äußere Erscheinung.

Methodik

In der Masterarbeit soll zunächst auf dem ungelabelten Bilddatensatz ein GAN trainiert werden und anschließend der Generator für die Bauteilsegmentierung auf den gelabelten Bildern adjustiert werden. Dieses Verfahren soll im Wesentlichen auf dieser Publikation aufbauen: <https://repurposegans.github.io>

Die Forschungsfrage soll anschließend mit Analysen der Fehlerverteilungen bewertet werden. Dazu können die Verteilungen in der Taxonomie oder weitere Attribute der Fahrzeuge in den Aufnahmen verwendet werden. Es können aber auch andere Verteilungsaspekte direkt aus den Bilddaten wie Kameraposition, Lichtverhältnisse oder Eigenschaften der Segmentierungsmasken selbst analysiert werden.

Ansprechpartner:

Julian Marstaller
+49 157 7806 5550
julian.marstaller@dat.de
Hellmuth-Hirth-Straße 1
73760 Ostfildern
Germany