Tensorflow Object Detection API

# Allgemeine Infos – Wissenswertes (gut in Vortrag einbringbar)

Unsere bisher verwendeten Neuronalen Netze haben immer gesamte Bilder klassifiziert. Die Tensorflow Object Detection API hingegen versucht Objekte im Bild zu finden und zu lokalisieren.

**Idee**: Bilder vom Seil nicht mehr in Kategorien Links, Mitte, Rechts, etc. einteilen, sondern das Netz generell auf Seile trainieren, sodass die eigentliche Position des Seils ermittelt wird.

## Zur Detektion geeignete Neuronale Netze

**R-CNN (Region-based Convolutional Neural Networks)**

Einfache Ansätze (R-CNN) legen dabei lediglich eine Art Suchfenster über das Bild und sind daher sehr langsam, da teilweise über 1000 einzelne Bildausschnitte klassifiziert werden müssen.

**Fast R-CNN**

In neueren Ansätzen (Fast R-CNN) wird dem eigentlichen Klassifikations-Netz noch ein weiteres, kleines Neuronales Netz vorangestellt, welches nur dafür da ist zu ermitteln, wo in einem Bild überhaupt generell interessante Bildausschnitte sein könnten. Diese vorgeschlagenen Ausschnitte werden dann wieder einzeln über das eigentliche, dahinterliegende Netz klassifiziert.

**Faster R-CNN**

Im aktuell effizientesten Ansatz (Faster R-CNN) müssen die vom vorgeschalteten Netz für interessant befundenen Bilder nicht mehr einzeln klassifiziert werden, sondern es genügt das gesamte Bild ein einziges Mal zu klassifizieren und dann Teilregionen einzeln zu betrachten. Dadurch ist Faster RCNN auch für Videostreams, etc. geeignet.

**SSD**

Google bietet in der Object Detection API neben Faster R-CNN auch noch ein ähnliches vorgeschaltetes Netz namens SSD (Single Shot Multibox Detector) an. Dieses nutzt zur Klassifizierung ein MobileNet und soll nochmal schneller als Faster RCNN sein. Beide Architekturen könnten für uns interessant sein.

Wichtige Infos, Links und Anleitungen zur API

* Respository: <https://github.com/tensorflow/models/tree/master/research/object_detection>

# How-To-Anleitung

Google selbst bietet unter dem oben angegebenen Link Skripte zum Training solcher Netze an. Dabei müssen die Trainings- und Testdaten im Tensorflow-eigenen Format **TFRecords** vorliegen. In diesem Dateiformat werden alle Bilder und Labels zusammen in eine Datei gebracht, welche dann für das Training genutzt werden können.

Eine Schritt-für-Schritt-Anleitung habe ich in „Object Detection API – Anleitung.docx“ verfasst.