

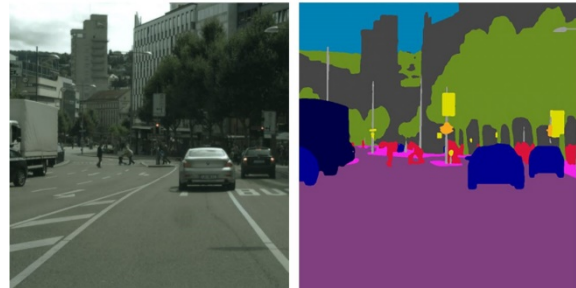
Pressemitteilung der Technischen Universität Braunschweig
10. August 2021

Konsistenz für Bilderkennung von Fahrzeugen Best Paper Award geht zum dritten Mal in Folge an das Institut für Nachrichtentechnik

Die CVPR (Conference on Computer Vision and Pattern Recognition) ist die wichtigste Konferenz im Bereich des maschinellen Sehens. Mit Professor Tim Fingscheidt, Serin Varghese, Andreas Bär und Marvin Klingner gewannen zum dritten Mal in Folge Wissenschaftler der TU Braunschweig einen Best Paper Award bei einem der Konferenz-Workshops. Die Forschenden des Instituts für Nachrichtentechnik entwickelten zusammen mit der Volkswagen Group Automation ein Training, mit dem Fahrzeuge sicherer ihre Umgebung einschätzen können.

In dem Workshop ging es um den sicheren Einsatz künstlicher Intelligenz für das autonome Fahren.

Künstliche neuronale Netzwerke verarbeiten vor allem die Bilder, die autonome Fahrzeuge von ihrer Umgebung aufnehmen. Sie lernen anhand von Daten automatisch, was sie umgibt und wie weit es entfernt ist. Zentrale Methode für die Umgebungswahrnehmung ist die semantische Segmentierung. Dabei klassifizieren die neuronalen Netzwerke in Sekundenbruchteilen jeden Pixel in Straße, Auto oder Passant*in.



Links: Eine Szene im Straßenverkehr, wie sie das menschliche Auge sieht. Rechts: Die verschiedenen Farbflächen visualisieren, wie eine künstliche Intelligenz mit semantischer Segmentierung die selbe Szene klassifiziert. Bildnachweis: The Cityscapes Dataset, IfN/TU Braunschweig.

Grundsätzlich ist die Leistungsfähigkeit von künstlicher Intelligenz bei der semantischen Segmentierung sehr schnell und zuverlässig. Bei einem Video, also einer schnellen Abfolge vieler Einzelbilder, kann es jedoch zu Widersprüchen kommen. „Ein Fahrradfahrer kann beispielsweise auf einem Bild erkannt und auf dem nächsten völlig übersehen werden“, erklärt Marvin Klingner, einer der Autoren. „Wenn der Computer dann die Informationen zusammenführt, muss er aus den widersprüchlichen Informationen eine eindeutige Entscheidung treffen. Besonders schwierig ist für den Computer dabei die Frage, welche Information die richtige ist.“

Mit dem passenden Training zur Konsistenz

Die Wissenschaftler lösten das Problem, indem sie das neuronale Netz während des Lernprozesses zur zeitlichen Konsistenz zwangen. In der Regel trainiert man hier Bild für Bild und optimiert das neuronale Netz so, dass es auf jedem einzelnen Bild möglichst gute Ergebnisse liefert. Zusätzlich nutzten die Wissenschaftler Bildsequenzen. Dabei sanktionierten sie, wenn das neuronale Netz bei der Beschreibung der Sequenzen inkonsistent war. Dadurch lernt das Netzwerk zeitliche konsistente Beschreibungen zu liefern und verbessert sich kontinuierlich im Verlauf des Trainings.

Serin Varghese, ebenfalls einer der Autoren: „Diese Art des Lernprozesses ist an sich nicht völlig neu. Aber normalerweise bräuchte die künstliche Intelligenz dann auch Videos in der Eingabe, was sehr rechenaufwendig ist. Der besondere Fall in unserem Paper: Die Videos braucht man nur fürs Training und in der Anwendung funktioniert es dann trotzdem auch mit Einzelbildern.“

Kontakt:

Prof. Dr.-Ing. Tim Fingscheidt
Technische Universität Braunschweig
Institut für Nachrichtentechnik
Abteilung Signalverarbeitung
Schleinitzstraße 22
38106 Braunschweig
Tel.: 0531 391-2485
E-Mail: t.fingscheidt@tu-braunschweig.de
www.ifn.ing.tu-bs.de