

Ausschreibung Bachelor- / Masterarbeit

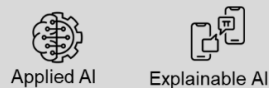
Transparenz für neuronale Netze in der Texterkennung an Güterwagen



Framework



Forschung



Domäne



Ausgangslage

Wir arbeiten gemeinsam mit der DB Cargo an der automatisierten Schaderkennung an Güterwagons. Der erste Schritt stellt dabei die Erkennung der Aufschriften auf den Wagons dar, um eine eindeutige Zuordnung der erkannten Schäden zu gewährleisten. Hierbei handelt es sich um ein Object Character Recognition (OCR) Problem, das in der Regel mit Deep Learning Modellen gelöst wird. Trotz der extremen Fortschritte im Bereich Deep Learning gibt es ein entscheidendes Problem mit diesen Modellen: Ihre mangelnde Transparenz, die vor allem bei unerwartet auftretenden Fehlern die Ursachenfindung erheblich erschwert

Problemstellung

Es soll ein Vergleich einer auf CNN's basierenden OCR Pipeline mit einer auf Vision Transformern basierenden Pipeline durchgeführt werden. Der Fokus der Arbeit soll auf der Erklärbarkeit der Ergebnisse unter Einbezug der Confidence der Modelle liegen.

Vorgehensweise und Erwartete Ergebnisse

- Training zweier verschiedener OCR Modelle/Pipelines (CNN und Transformer)
- Analyse der Modelle Bezug auf:
 - Wichtigkeit von Input Bereichen (mit Attention visualization bzw. Saliency methods)
 - Korrelation von Confidence und Korrektheit der Vorhersagen

Ansprechpartner


Nils Hütten | E-Mail: nhuetten@uni-wuppertal.de

Bachelor- / Master thesis

Transparency of Neural Networks in Text recognition on Freight Cars



Framework

 **python** **PYTORCH**

Forschung



Domäne



Initial Situation

We are working together with DB Cargo on automated damage detection on freight cars. The first step is the recognition of the inscriptions on the wagons in order to guarantee an unambiguous assignment of the recognized damages. This is an Object Character Recognition (OCR) problem, which is usually solved with Deep Learning models. Despite the extreme progress in Deep Learning, there is a crucial problem with these models: Their lack of transparency, which makes finding the cause of unexpected errors much more difficult.

Problem Definition

A comparison of an OCR pipeline based on CNN's with a pipeline based on vision transformers will be performed. The focus of the work will be on the explainability of the results including the confidence of the models.

Procedure and Expected Results

- Training of two different OCR models/pipelines (CNN & Transformer)
- Analysis of the models with regard to:
 - Importance of certain input regions
 - Correlation of prediction confidence and correctness of results

Contact Person

Nils Hütten| **E-Mail:** nhuetten@uni-wuppertal.de