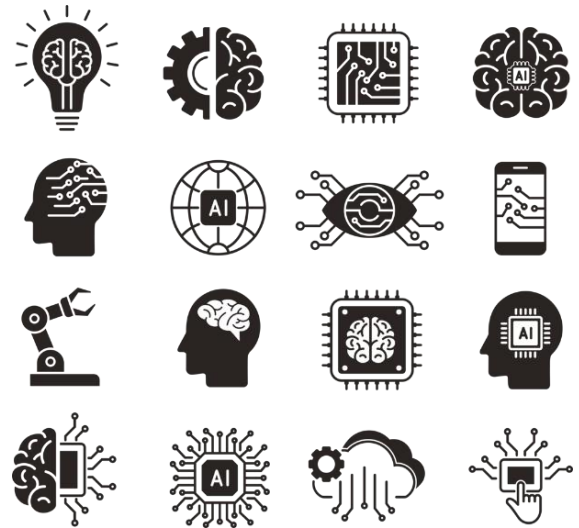


Training und Vergleich eines Vision-Language-Modells mit Chain-of-Thought-Daten zur Fehlererkennung in technischen Zeichnungen

Ausgangssituation:

Die automatisierte Erkennung von Fehlern in technischen Zeichnungen erfordert nicht nur eine zuverlässige Klassifikation, sondern zunehmend auch eine nachvollziehbare Begründung der getroffenen Entscheidung. Vision-Language-Modelle (VLMs) bieten hierfür durch ihre Fähigkeit zum Chain-of-Thought-Reasoning (CoT) einen vielversprechenden Ansatz, da sie Bildinhalte nicht nur klassifizieren, sondern den Erkennungsprozess auch schrittweise sprachlich begründen können. Zur Untersuchung dieses Ansatzes wurden bereits zwei unterschiedliche Arten von Chain-of-Thought-Daten zur Fehlererkennung in technischen Zeichnungen erhoben. Bislang ist jedoch unklar, wie gut sich diese Datensätze für das Training eines Vision-Language-Modells eignen, welchen Einfluss die jeweilige Datenart auf die erzielte Performance hat und wie hoch der Aufwand für ihre Erhebung im Verhältnis zum erzielten Nutzen ist.



Zielsetzung:

Ziel der Studienarbeit ist die Auswahl eines geeigneten Vision-Language-Modells sowie dessen Training mit den beiden erhobenen Chain-of-Thought-Datensätzen zur Fehlererkennung in technischen Zeichnungen. Dabei sollen insbesondere folgende Teilziele erreicht werden:

- Auswahl eines geeigneten Vision-Language-Modells sowie, falls erforderlich, Vorbereitung und Aufbereitung der beiden erhobenen Datensätze für das Training.
- Training des Modells mit beiden Datenarten sowie ein systematischer Vergleich der Ergebnisse.
- Durchführung einer Performance-Aufwand-Analyse, die aufzeigt, welche Datenart welchen Performancegewinn liefert und welcher Erhebungsaufwand jeweils damit verbunden ist.
- Ableitung einer Empfehlung, welche Art von Chain-of-Thought-Daten sich unter Abwägung von Nutzen und Aufwand für den weiteren Einsatz am besten eignet.

Voraussetzungen:

- Grundkenntnisse im Bereich CAD/technische Zeichnungserstellung
- Erfahrung mit KI-Methoden (Machine Learning / Computer Vision)
- Programmierkenntnisse in Python
- Selbstständige und strukturierte Arbeitsweise

Bei Interesse bitte einen Lebenslauf und aktuelle Leistungsnachweise an:

Sarah Steininger
Lehrstuhl für Fördertechnik Materialfluss Logistik (fml)
Tel.: +49 (89) 289 – 15919
sarah.steininger@tum.de