

Visuelle Baudokumentation durch Kran-Kameras und Multimodale Sprachmodelle (VLMs)

Ausgangssituation

Auf modernen Baustellen ist die lückenlose Dokumentation von Baufortschritt, Anlieferungen und sicherheitsrelevanten Abläufen essenziell. Kamerasysteme bieten hier ein enormes Potenzial – insbesondere, wenn sie mit aktuellen KI-Methoden wie Visual Language Models (VLMs) ausgewertet werden. Diese multimodalen Modelle (z. B. LLaVA, GPT-4V) können visuelle Szenen interpretieren und in natürlichen Text umsetzen.

In dieser Arbeit soll untersucht werden, wie Kameras, die am Ausleger eines Krans montiert sind, zur automatisierten visuellen Baudokumentation verwendet werden können.

Ziel der Arbeit

Ziel ist es, ein prototypisches System zu entwickeln, das:

- Kamerabilder von einer simulierten oder echten Kranperspektive verarbeitet
- mittels eines VLM Bauprozesse (z. B. Betonage, Anlieferung) erkennt
- relevante Ereignisse als Textbeschreibung dokumentiert

Aufgabenstellung

Die Arbeit gliedert sich in mehrere Schritte: Zunächst sollen bestehende Ansätze zur Verwendung von VLMs in Bereichen wie autonomem Fahren oder Bauüberwachung recherchiert und bewertet werden. Darauf aufbauend soll ein technisches Konzept für die Anwendung auf Baustellen erarbeitet werden. Dies beinhaltet die Auswahl geeigneter Kamerapositionen (z. B. am Kran-Ausleger), die Definition typischer zu erkennender Szenen sowie die Auswahl eines geeigneten Modells



und Frameworks. Im Anschluss soll ein prototypisches System implementiert werden, das Kamerabilder verarbeitet, relevante Ereignisse erkennt und diese als Textberichte ausgibt. Die Erkennung kann entweder auf synthetischen Bilddaten (z.B. Unity Engine) oder auf vorhandenen realen Baustellendaten getestet werden. Abschließend erfolgt eine Evaluation der Erkennungsleistung und der praktischen Anwendbarkeit des entwickelten Systems.

Voraussetzungen

- Grundlagen in Python und Deep Learning
- Interesse an Computer Vision / Multimodal AI
- Optional: Erfahrung mit Bilddaten und -verarbeitung

Bei Interesse freue ich mich über Ihre Bewerbung, die Sie inklusive Lebenslaufs und Leistungsnachweis bitte an alexander.schock@tum.de senden.

Kontakt

Alexander Schock, M.Sc.

Raum: MW 0510
Tel.: +49 (89) 289 - 15913
E-Mail: alexander.schock@tum.de