

Aktivitätserkennung im Schienenbau mit Convolutional Neural Networks

Studiengang: MAS Data Science

Auf Gleisbaustellen ist Tempo angesagt, damit Fahrgäste ihr Ziel wie geplant erreichen. Organisatorisch ist dies herausfordernd, da Verzögerungen häufig auftreten. Die Rhomberg Sersa Rail Group will Mitarbeitenden vor Ort datenbasierte Entscheidungsgrundlagen über den Baufortschritt bieten. Dafür sind unterschiedliche Sensoren im Einsatz. Die Arbeit untersucht die Frage, ob eine Videokamera mit einem trainierten Modell typische Aktivitäten des Gleisbaukrans erkennen kann.

Ausgangslage

Das Erkennen von Aktivitäten auf Baustellen ist wenig erforscht. Jede Baustelle unterscheidet sich, und es fehlt an vorhandenen Video-Datensätzen. Die Rhomberg Sersa Rail Group hat im Jahr 2023 zusammen mit der Fachhochschule Nordwestschweiz das Innosuisse-Projekt „Real Time Construction Monitoring“ (RTCM) gestartet. Ziel dieses Projekts ist es, Bauaktivitäten vollständig automatisiert zu erfassen. Plandaten, kombiniert mit dem Baufortschritt in Echtzeit, sollen es Mitarbeitenden vor Ort ermöglichen, datenbasierte Entscheidungen zu treffen, um Verzögerungen zu vermeiden. Der Gleisbaukran Kirow ist für den Baufortschritt eine wichtige Maschine. Im Projekt erfassen verschiedene Sensoren und eine OAK-D Pro W Videokamera die Bewegungen des Krans.

Ziel

In der Master-Thesis liegt der Fokus auf dem Videomaterial des Gleisbaukrans. Die Autorin untersucht, wie genau Convolutional Neural Networks (CNN) Aktivitäten in Videosegmenten klassifizieren können. Ziel der Arbeit ist es, die Leistung der Modelle aufzuzeigen sowie Erkenntnisse und Empfehlungen zu erarbeiten. Damit leistet die Arbeit einen Beitrag zum Projekt und erweitert das Wissen in diesem Gebiet.

Umsetzung

Aus dem 15-stündigen Videomaterial und dem definierten Prozess wurden sechs Aktivitäten identifiziert und in ihren Bewegungen beschrieben. Daraus entstand ein Datensatz mit 870 Videosegmenten. Beim Training wurde der Datensatz durch Methoden der Bild- und Video-Augmentierung, zum Beispiel durch unterschiedliche Wetterverhältnisse wie Regen, Schnee oder Nebel, erweitert. Das Training und die Evaluierung der Modelle S3D, X3D-S und X3D-M erfolgten mithilfe von PyTorch und Transfer Learning. Die Bewertung basierte auf der Accuracy.

Fazit

Mit einer Accuracy von 82,58 % liefert das Modell X3D-S die beste Performance. Das Modell erkennt vier der sechs Aktivitäten mit einem Recall-Wert über 0,8. Unerwartete Bewegungen und Personen im Bild beeinträchtigen die korrekte Klassifizierung. Daher sollte der Datensatz erweitert werden, um die Robustheit des Modells zu erhöhen. Es ist empfehlenswert, längere Segmente zu verwenden, um die charakteristischen Bewegungsabläufe einer Aktivität zu enthalten. Diese Erkenntnisse fliessen nun zurück in das Vorhaben, wobei das trainierte Modell X3D-S die automatische Erkennung typischer Kranaktivitäten verbessert und somit die datenbasierte Analyse des Baufortschritts im RTCM-Projekt unterstützt.



Christina Thierwächter
christina@thierwaechter.ch



Bildausschnitt der Videokamera auf dem Dach des Gleisbaukrans