

# 1 Projekt

## 1.1 Titel:

Automatische Erkennung und Vermessung von Eisenbahninfrastrukturelementen

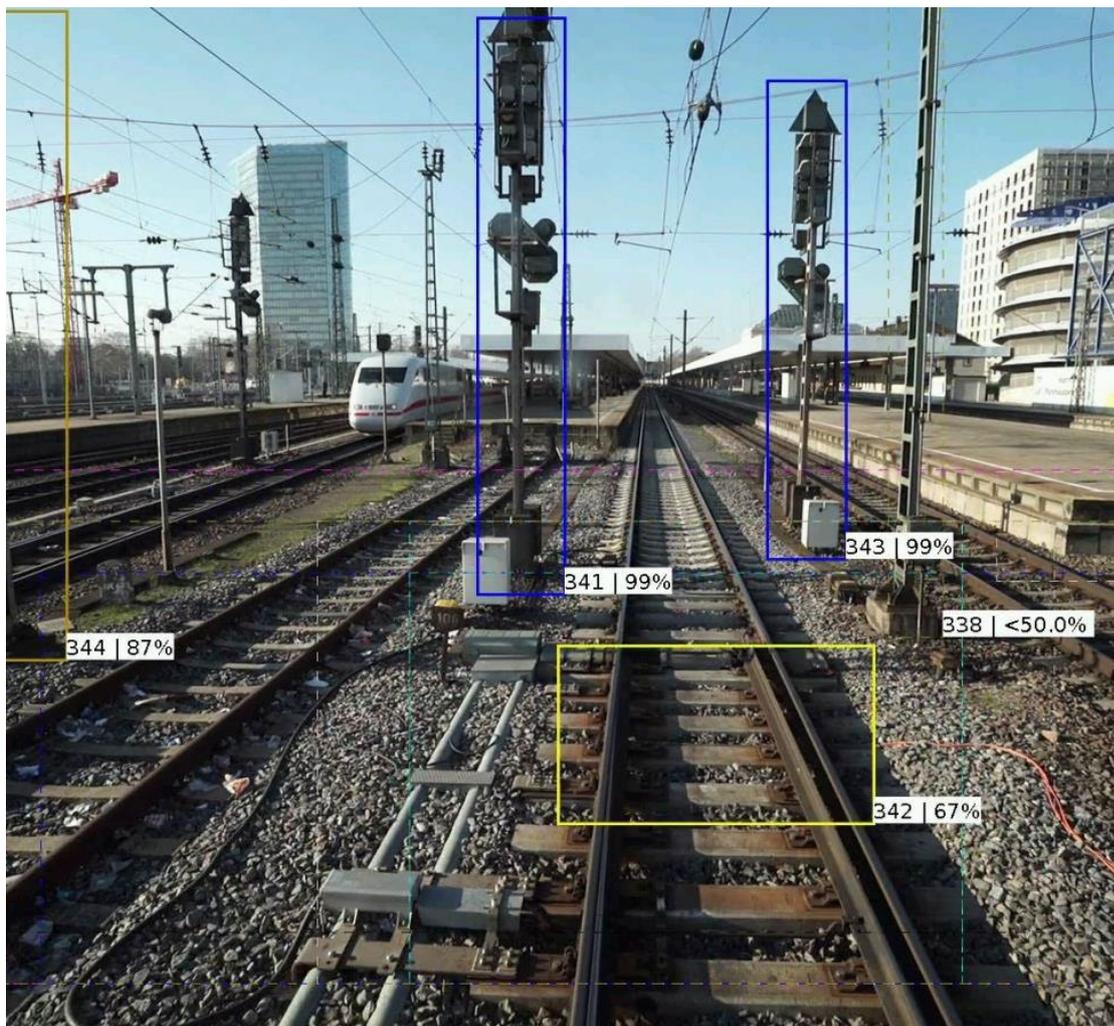
## 1.2 Laufzeit

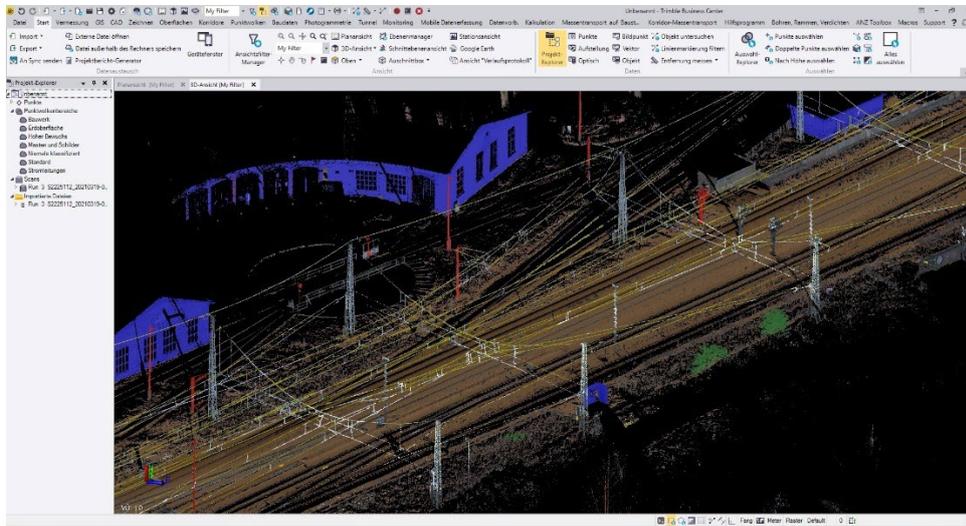
18 Monate Laufzeit, Einbindung in das Eisenbahn-Forschungsprojekt der FK07 der Hochschule München und einen Industriepartner

## 1.3 Motivation / Kurzbeschreibung der Ziele und Aufgaben

In der Vergangenheit wurde in Zusammenarbeit mit einem Forschungsinstitut das System SATengine entwickelt, um Bahninfrastrukturen zu vermessen und zu digitalisieren. Auf Basis von Streckenvideos und Positionsdaten werden Infrastrukturelemente wie Weichen und Lichtsignale bestimmt und geographisch verortet. Dieser Prozess wurde in der Vergangenheit manuell durchgeführt, was zu hohen Kosten und langen Projektdauern geführt hat.

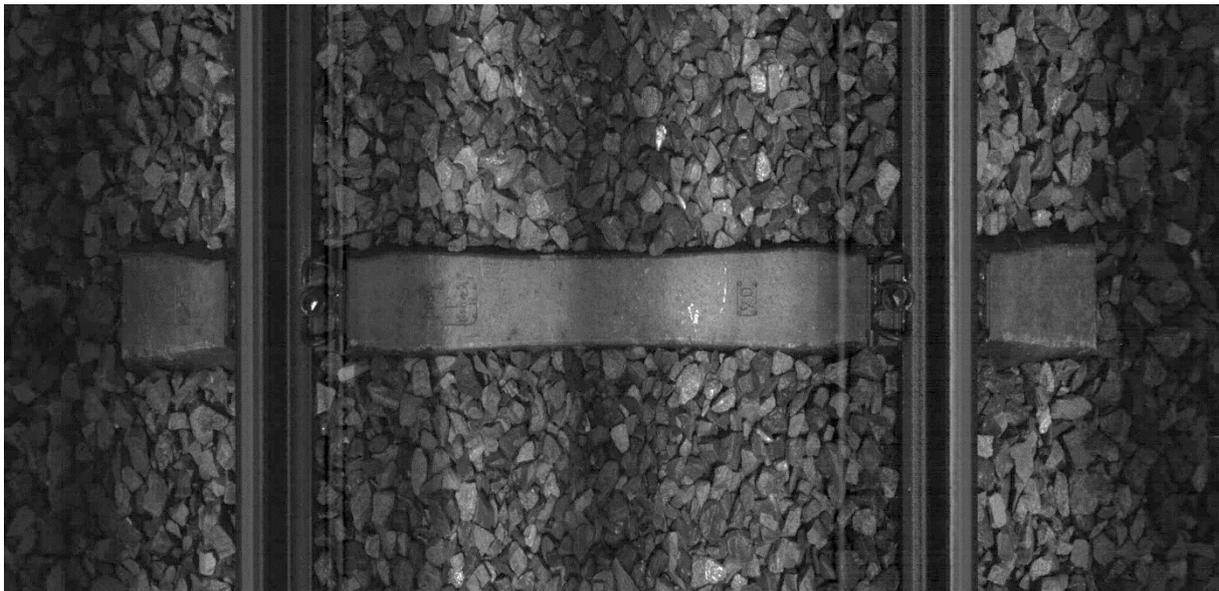
Daher wurde eine künstliche Intelligenz für die automatische Erkennung von Infrastrukturelementen in Befahrungsvideos entwickelt. Die Realisierung der Objekterkennung wurde mittels tiefer neuronaler Netze vorgenommen. Zusätzlich wurde ein umfangreiches Post-Processing entwickelt, das zur Fehlerreduktion verschiedene Plausibilitätsprüfungen durchführt.



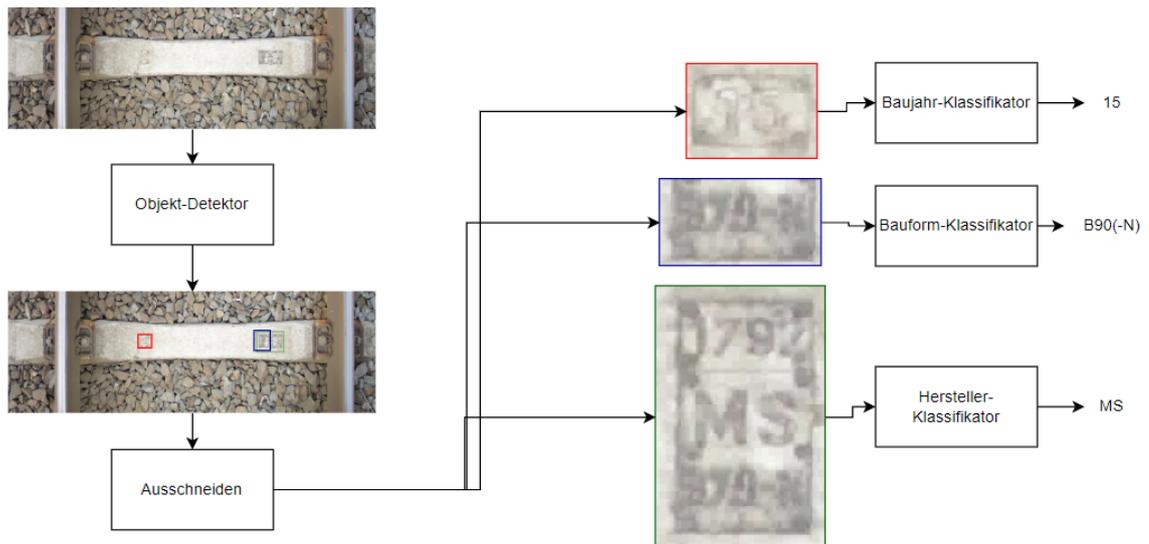


## Schwellenerkennung

Seit 2023 wurde SATengine erweitert, um zusätzlich Videos des Gleis-Oberbaus aufzunehmen.



Analog zum oben beschriebenen RIMcognize Projekt wurde eine Software zum Kartografieren und Klassifizieren von Bahnschwellen entwickelt. Es werden die Bahnschwellen detektiert, kategorisiert und aufgrund der auf Betonschwellen vorhandenen Prägungen genauer klassifiziert.



Da in den bisher befahrenen Strecken nicht alle möglichen Ausprägungen der Schwellen und derer Merkmale vorkamen, soll die Erkennungen für mögliche neue Erscheinungsformen vorbereitet und verbessert werden.

## Mögliche Forschungsprojekte 2024ff.

Die beschriebenen KI-Software Lösungen sollen in den nächsten Jahren fortlaufend weiterentwickelt werden. Hierzu sollen verschiedene mögliche Features vorab in Forschungsprojekten untersucht werden:

### Themen im Rahmen der Schwellenerkennung:

#### 1. Unsupervised Exploration für neue Labels

*Ziel: Die momentane manuelle Sichtung von neuen Videos auf das eventuelle Vorkommen neuer Typen bzw. Prägungen erfolgt heute stichprobenartig.*

*Durch einen möglichen neuen Ansatz soll eine höhere Sicherheit erreicht werden, sodass unbekannte Klassen nicht falsch klassifiziert werden, sondern als solche erkannt werden und für die Verbesserung der Schwellenerkennungssoftware zur Verfügung stehen.*

Wir lokalisieren bereits Prägungen, wie von Hersteller oder Baujahr, zuverlässig auf Schwellen. Zukünftig sollen weiterhin Videoaufnahmen von Strecken durchgeführt werden, die bisher mit dem System noch nicht befahren wurden, sodass potenziell der KI unbekannte Typen vorkommen können.

Da es regionale Unterschiede geben kann, welche Schwellen welchen Herstellers verbaut wurden, rechnen wir damit, dass zwar die Erkennung der Prägung zuverlässig funktionieren wird, die genaue Klassifizierung aber nicht möglich ist, da unsere Modelle noch nicht für den neuen Typ trainiert werden konnten.

Eine manuelle Sichtung der neuen Videos ist aufgrund der immensen Anzahl der vorkommenden Schwellen nicht zielführend. Deshalb sollen Methoden untersucht werden, die die Bildähnlichkeit

zu den Trainingsdaten identifizieren, um automatisch neue (untrainierte) Typen von Schwellen bzw. Prägungs-Klassen zu finden.

### Mögliche Vorgehensweisen

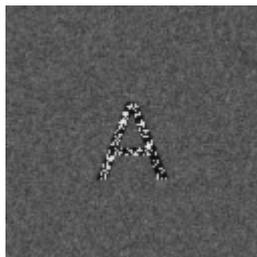
- Auswertung der Classification Scores. Annahme: Niedrige Scores können auf unbekannte Bilder hindeuten
- Erweiterung eines vorhandenen Clustering- bzw. Bildähnlichkeit-Ansatzes: Momentan wird mittels Autoencoder ein CNN-Encoder trainiert, um latente Repräsentationen von Schwellen bzw. Prägungen zu lernen. Die Ähnlichkeit zwischen zwei Bildern wird über die Ähnlichkeit der latenten Merkmals-Vektoren dieser beiden Bilder gebildet (MSE oder CosineSimilarity). Um nun neue Daten zu finden, können diejenigen Bilder, mit größtmöglicher Distanz (Unähnlichkeit) zu vorhandenen Trainingsdaten ausgewertet werden. Weiterhin besteht bis zur Umsetzung vermutlich ein Ansatz, der den Encoder mittels Contrastive Learning trainiert. Die bestehenden Ansätze den Encoder zu trainieren können verbessert werden, oder ein Vergleich mit einem weiteren Ansatz durchgeführt werden.

Zur Auswertung können Labels ausgewählter Klassen aus dem Training der bestehenden Modelle für die Schwellenerkennung herausgefiltert werden. Die Ansätze zum Finden neuer Labels bewerten sich dann danach, wie viele eben dieser ‚neuen‘ Labels sie detektieren.

## **2. Synthetische Datenerzeugung Schwellen-Merkmale**

**Ziel:** *Durch die Verwendung synthetischer Trainingsdaten sollen die in der Schwellenerkennung eingesetzten Deep Learning-Modelle verbessert werden. Dies sollte zu einer Verbesserung der korrekten Klassifizierungen der Schwellen führen. Zumindest sollte dieser Effekt für Klassen mit einer geringen Anzahl an Trainingsdaten nachweisbar sein.*

1. Da viele Merkmale wie z.B. Jahreszahlen oder die Bauform aus simplen Textzeichen bestehen, bietet sich zunächst eine simple Generierung mittels Bildverarbeitungssoftware oder klassischer Computer Vision Methoden an. Somit könnten wir z.B. die Jahreszahl „24“ für 2024 bereits trainieren, bevor wir die ersten Bilder solcher Schwellen gesehen haben. Die so generierten Bilder können per Image2Image in die Domäne der real aufgenommenen Prägungen und Schwellenbilder transferiert werden (StyleTransfer, StableDiffusion, andere?) Z.B. Primitive Synthetisierung einer Prägung des Herstellers „A“:



2. Um die Erkennung für Klassen zu verbessern, für die nicht genügend Trainingsdaten existieren, bzw. um die Erkennung robuster gegenüber Variationen in der Erscheinung einzelner Merkmale zu machen, kann auch ausgewertet werden, welchen Effekt eine Datengenerierung ohne ‚Photoshop‘ direkt über ein generatives Verfahren hat (ggfs. in Kombination mit „Unsupervised Representation Learning“ umzusetzen)

Auswertung: Für die Forschungsarbeit kann eine Beschränkung auf einen Prägungstyp sinnvoll sein. Z.B. Jahreszahl-Prägung. Als Auswertungsmetrik dient zum einen die Korrektheit der Bild-Klassifizierungen auf einem statistisch gewählten Testset, zum anderen der Ende-Zu-Ende Test der Schwellenerkennungssoftware, in dem ein separate erzeugtes Testvideo ausgewertet wird.