

MIT KÜNSTLICHER INTELLIGENZ ZU MEHR SICHERHEIT AM BAHNGLEIS?!

Projektleitung

Prof. Dr.-Ing. Martin Schlüter
(Fachbereich Technik /i3mainz - Institut für
Raumbezogene Informations- und Messtechnik)

Beteiligte Personen

Bastian Pläß (M.Sc.)
Kira Zschiesche (M.Sc.)
Tamer Altinbas (M.Sc.)
Daniel Karla (M.Sc.)
Linda Rau (M.Sc.)

Laufzeit

01.04.2019 – 31.03.2022

Förderer

Carl-Zeiss-Stiftung, Bereitstellung der HALCON
Student Edition durch MVTec Software GmbH

Kontakt

i3mainz@hs-mainz.de

Motivation

Zum Bereich des spurgebundenen Baustellen-Monitorings gehört die Sicherung der Baustellen. Dazu zählt u. a. die lückenlose Erfassung und zuverlässige, ortsgerechte Weitergabe von Zugfahrten in den verschiedensten Gefahrensituationen, um dem Baustellenpersonal die rechtzeitige Räumung des Gleisbereichs zu ermöglichen. In diesem Anwendungsbeispiel soll experimentell Künstliche Intelligenz (KI) eingesetzt werden, um in Bilddaten automatisiert Züge zu detektieren und so für mehr Sicherheit im Gleisbereich zu sorgen. Als

„drittes Auge“ soll es das Portfolio an automatischen Warnanlagen und damit die sicherheitskritische Systemarchitektur erweitern. Durch den Einsatz einer Kamera in Verbindung mit einer trainierten KI kann die Bandbreite der bisher eingesetzten Sensorik praktikabel und kostengünstig ausgebaut werden. Mit Reduzierung der Aufwands- und Instandhaltungskosten bei gleichzeitiger Erweiterung des Sichtfeldes zwecks Erhöhung der Arbeitssicherheit sind die Vorteile des Computer-Vision-Systems allgemein beschrieben.

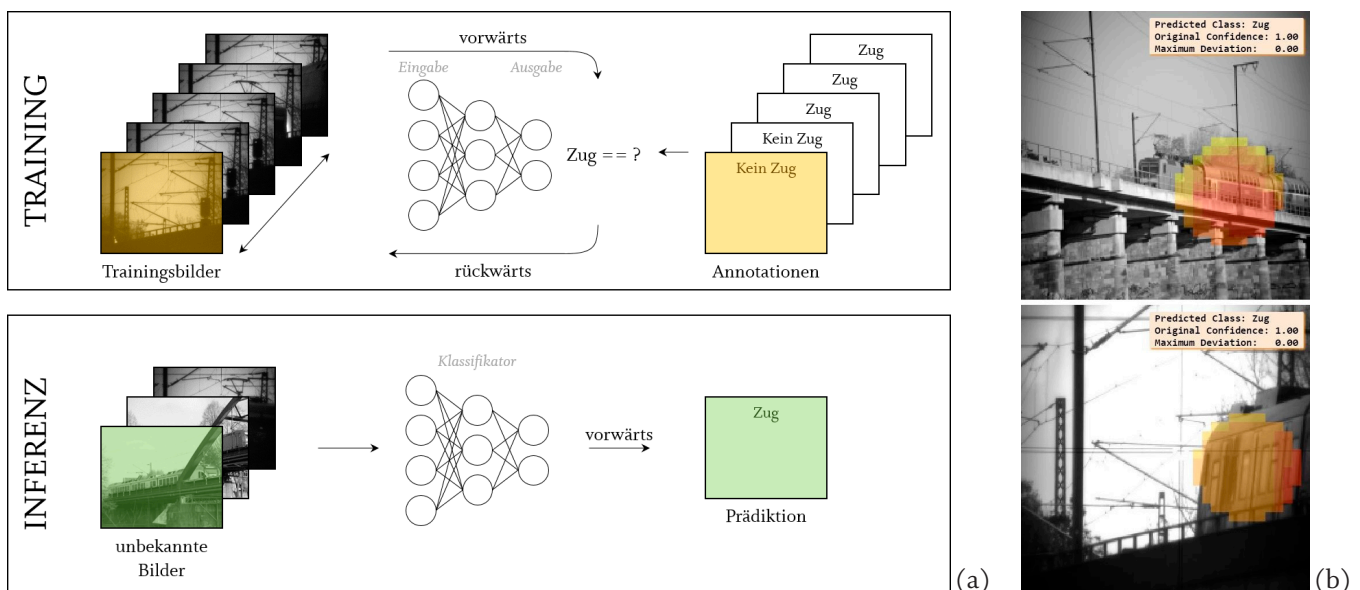


Abbildung (a): Während des Trainings werden die Verbindungen zwischen den in Schichten organisierten Neuronen (○) des ConvNets gewichtet, sodass diese während der Inferenzphase den Anwendungsfall bestmöglich beschreiben können.

Abbildung (b): Während des Trainings- und Inferenzprozesses werden diejenigen Bildregionen hervorgehoben, welche den maximalen Einfluss auf die Klassifizierungsentscheidung ausüben.

Realisierung

Computersysteme in die Lage intelligenten Wahrnehmens und Handelns zu versetzen, ist Gegenstand der KI-Forschung. Ein populärer Zweig der KI ist das Deep Learning (DL). Durch ein umfangreiches Training können Maschinen für eine eng definierte Aufgabe ausgebildet werden. Als ein effektives Werkzeug im Bereich der Bilderkennung haben sich KNN (künstliche neuronale Netze) herausgestellt, durch welche z. B. Züge in Bildern erkannt und diese als Gefahrenpotential klassifiziert werden können. Aus dem Trainingsprozess resultiert ein Modell, welches in der Lage ist, unbekannte Bilder zu klassifizieren. Dieser Vorgang des Schlussfolgerns aus trainierten Erfahrungswerten wird als Inferenz bezeichnet (vgl. Abbildung (a)). Hinter der maschinellen Bilderkennung verbirgt sich eine überwachte Lernmethode. Diese setzt annotierte Bilddaten, sog. Trainingsdaten, voraus, welche der Maschine das Lernziel vorgeben. Die binäre Bildklassifikation ist eine vergleichsweise einfache Aufgabe, wobei die Maschine lediglich lernen soll, ob das gesuchte Objekt dargestellt ist.

Aus der Aufnahme einer Gleissituation resultiert ein zeitbezogener Bildreihendatensatz. Für die Erzeugung eines adäquaten Trainingsdatensatzes wird ein klassisches Bildverarbeitungsverfahren entwickelt und auf den optischen Zeitreihendatensatz abgestimmt. Die Berechnung von Differenzbildern ist über einen Schwellwert und ein zu spezifizierendes Bildintervall steuerbar. Die Eigenevaluation des Verfahrens stellt im Vergleich zur manuellen Klassifizierung eine Beschleunigung um den Faktor 3 bei einer Fehlerquote im niedrigen, einstelligen Bereich dar. Zur Lösung der Bildklassifikation wird eine spezielle Form von KNN, sog. ConvNets, eingesetzt, welches mit Hilfe von mathematischen Filteroperationen (Convolutions) Strukturen innerhalb der Bilder erkennen kann. Die Entwicklung des Bildklassifikators basiert auf der DL-Bibliothek von MVTec HALCON. Die Software wurde im Rahmen des „MVTec on Campus“-Programms kostenlos zur Verfügung gestellt. Adaptiert wurde eine reduzierte Variante des AlexNet [Krizhevsky et al. 2012] mit 18 Schichten, einem Visual Geometry Group (VGG16)-nahen Encoder und einer linearen Aktivierung (ReLU). Das sog. SqueezeNet wird im Rahmen des Hyperparameter Tunings und mit augmentierten Trainingsdaten bestmöglich auf die Klassifizierungsaufgabe angepasst. Eine Batchsize von 32 Bildern, eine Anpassung der initialen Learning Rate von 0.01 um den Faktor 0.1

nach jeder 25. Epoche und ein Regularisierungsterm an Stelle des punktuellen Dropouts nach dem 17. Layer führen zu dem besten Modell. Der statistisch-basierte Trainingsprozess endet nach 100 Epochen. Der Top-k-Fehler dient bei der Evaluation als aussagekräftiger Score. Darüber hinaus visualisieren Heatmaps diejenigen Bildregionen, welche durch die trainierten Convolutions eine Aktivierung erfahren und folglich für die Bildklassifizierung verantwortlich sind (vgl. Abbildung (b)). Diese gekoppelte Evaluierungsstrategie erhöht die Transparenz und Zuverlässigkeit bei der Bewertung der Leistungsfähigkeit des trainierten Modells.

Ergebnisse

Der Workflow zur automatischen Erkennung von Zügen in Bilddatensätzen wurde erfolgreich getestet. Erste Erkenntnisse über die Einsetzbarkeit von KI in einer ingenieurgeodätischen Domäne konnten am vorliegenden Beispiel gesammelt werden. Die Problemstellung beschreibt eine binäre Bildklassifizierung. Dabei soll auf Grundlage eines trainierten DL-Modells automatisch erkannt werden, ob im Bild ein Zug vorhanden ist oder nicht. Durch die Verwendung einer KI kann vollständig auf die explizite Formulierung funktioneller Zusammenhänge verzichtet werden. Es zeigt sich, dass der eingesparte Aufwand für ein optimales Ergebnis in die Datenakquise und die Annotation fließen sollte, da die Erzeugung des Trainingsdatensatzes wichtige Erkenntnisse um das Paradigma DL erfordert und folglich die Güte des Bildklassifikators limitiert. Durch die softwareseitige Lösung der Trainingsdatenerzeugung kann der manuelle Arbeitsanteil deutlich gesenkt werden und somit erheblich zur Wirtschaftlichkeit des „dritten Auges“ beitragen. Der verwendete Differenzbildansatz ist jedoch nur bei vorhandenen Bildreihen sinnvoll einsetzbar. Im Rahmen dieses Forschungsprojektes wurde ein pre-trained ConvNet erfolgreich auf einen spezifischen Anwendungsfall adaptiert. Der trainierte Bildklassifikator erzielt einen Top-1-Score von 99,4% und erweist sich somit als praxistauglich. Bezugnehmend auf die Sicherung von Arbeitsstätten im Gleisbereich unterstreicht unser Ergebnis das Potenzial von DL-basierten Computer-Vision-Systemen.