



1/2 Bei der Überwachung von Infrastruktur müssen große Datenmengen ausgewertet werden. Automatisierte Objekterkennung kann diesen Prozess beschleunigen. Farbige Segmentierungsmasken markieren die erkannten Objekte im automatisch ausgewerteten Bild (rechts).



AUTOMATISIERTE 3D-DATEN-INTERPRETATION 3D-AI

Für die Vermessung von Infrastruktur werden heute leistungsfähige Kameras oder Laserscanner eingesetzt, die hochaufgelöste Bilder bzw. sehr präzise, georeferenzierte Messdaten liefern. Die Interpretation dieser Daten geschieht in der Regel manuell. Mit dem von Fraunhofer IPM entwickelten »Deep Learning Framework« 3D-AI kann dieser Prozess automatisiert werden.

Künstliche neuronale Netze

Die Auswertung der 3D-Daten übernehmen dabei komplexe Lernalgorithmen, die auf dem Konzept des »Deep Learning« mit künstlichen neuronalen Netzen (KNN) basieren und klassischen Methoden der Objekterkennung inzwischen überlegen sind. Noch vor wenigen Jahren dauerte das Training solcher Algorithmen Wochen oder gar Monate. Heute ist dieser Prozess dank massiver Parallelisierung in wenigen Stunden möglich. Die Auswertung neuer Datensätze auf Basis eines trainierten KNN erfolgt dann

sogar in Echtzeit. In KNN durchläuft die eingespeiste Information eine Vielzahl miteinander verknüpfter künstlicher Neuronen, wird verarbeitet und an weitere Neuronen weitergegeben. Mithilfe manuell annotierter Trainingsdaten werden für bestimmte Eingangsmuster zugehörige Ausgabemuster erlernt. Auf Basis dieser »Erfahrungswerte« können neuartige Eingangsdaten dann in Echtzeit analysiert werden. Dabei erweisen sich KNN als sehr robust gegenüber Variationen charakteristischer Farben, Kanten oder Formen.

Scannerdaten und Kamerabilder

Je detaillierter die Informationen im Datensatz, desto besser gelingen Objekterkennung und -klassifizierung. Datenbasis für die automatisierte Objekterkennung können sowohl Kamera- als auch Scannerdaten oder auch fusionierte Daten sein. Das von Fraunhofer IPM entwickelte Framework 3D-AI überführt dabei die georeferenzierten Punk-

Fraunhofer-Institut für Physikalische Messtechnik IPM

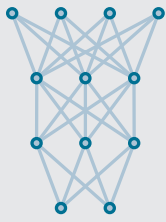
Heidenhofstraße 8
79110 Freiburg

Ansprechpartner

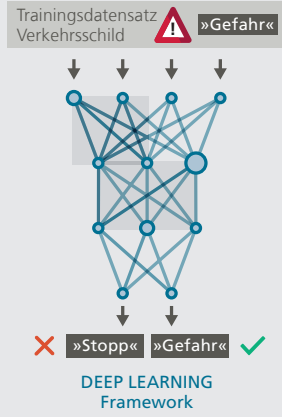
Prof. Dr. Alexander Reiterer
Abteilungsleiter
Objekt- und Formerfassung
Telefon +49 761 8857-183
alexander.reiterer@ipm.fraunhofer.de

Dr. Markus Leidinger
Geschäftsfeldentwickler
Telefon +49 761 8857-413
markus.leidinger@ipm.fraunhofer.de

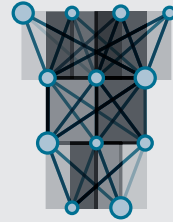
www.ipm.fraunhofer.de/of



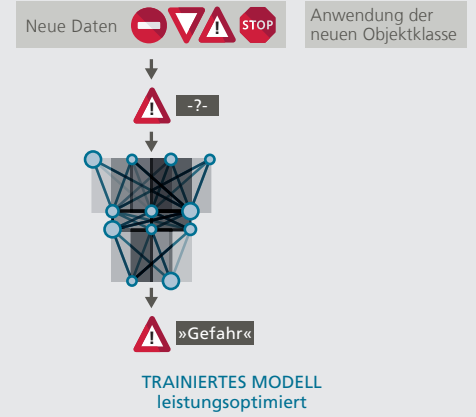
UNTRAINIERTES MODELL
Neuronales Netz



FRAUNHOFER IPM



TRAINIERTES MODELL
neue Objektklasse



TRAINIERTES MODELL
leistungsoptimiert

KUNDE

te der Scannerdaten zunächst in ein Rasterformat, das Tiefeninformationen enthält, und verknüpft diese dann mit den RGB-Kameradaten. Dieser pixelbasierte RGB-D(epth)-Datensatz enthält zu jedem RGB-Kamerabild ein entsprechendes Tiefenbild und bildet ein ideales Eingabeformat für KNN. Liegen hochauflösende Scannerdaten vor, so können diese im Framework direkt verwendet werden. Der Umweg über die Erzeugung von RGB-D-Daten entfällt. Letztlich entscheiden die Aufgabenstellung und die Art der verfügbaren Daten darüber, welcher Ansatz geeignet ist.

Klassifizierte Objekte in 3D

Die Architektur des Netzes, also die Anzahl der Netzschichten und die Art der hierarchi-

schen Verknüpfungen, ist an die jeweilige Aufgabe angepasst. In der 3D-Datenverarbeitung sind insbesondere die sog. faltenden neuronalen Netze (Convolutional Neural Networks) von Interesse. Diese machen sich die Tatsache zunutze, dass Informationen in Bildern und Scannerdaten meist lokal, d. h. auf einen kleineren Teilbereich des Datensatzes beschränkt sind. Deshalb können für die Analyse des Dateninhaltes Filter verwendet werden, die jeweils kleine Ausschnitte betrachten (beispielsweise 7×7 oder 11×11 Punkte/Pixel). Unabhängig vom Netzdesign wird für das Trainieren des Netzes ein repräsentativer Trainingsdatensatz erstellt. Dazu erfolgt zunächst eine manuelle semantische Segmentierung der Bilder, bei der jedes Pixel bzw. jeder 3D-Punkt einer bestimmten Objektklasse zugeordnet wird (Annotierung). Ein mit diesen Daten trainier-

tes Netz kann mithilfe eines neuen Trainingsdatensatzes jederzeit um weitere Objektklassen erweitert oder entsprechend verbessert werden. Anhand der Pixelkoordinaten der in den Bilddaten erkannten Objekte lässt sich die Segmentierung in die Punktwolke zurückprojizieren. Alternativ kann die Segmentierung direkt in den 3D-Daten vorgenommen werden. Nach der Segmentierung der Punktwolke lässt sich ein klassifizierter Datensatz exportieren.

Für das Erstellen der Trainingsdaten stellt Fraunhofer IPM ein eigenes Tool zur einfachen und schnellen Annotierung der 2D- und 3D-Daten zur Verfügung. Das trainierte Netz wird als ausführbares Programmpaket ausgegeben (Windows oder Linux). Alle Schnittstellen werden wunschgemäß angepasst.

3 Prozesskette im Fraunhofer IPM
Deep Learning Framework 3D-AI bestehend aus Network-Design, -Training und -Anwendung.

4 RGB-Bild (l.), Intensitätsbild (M.) und ausgewertetes Bild mit farbigen Masken für klassifizierte Objekte, übertragen in die Punktwolke (r.).

Deep Learning

Deep Learning ist als Methode des »Machine Learning« ein Teilbereich der künstlichen Intelligenz, die auf lernende Algorithmen setzt. Das Identifizieren und Klassifizieren von Objekten in einem Bild beispielsweise erfolgt anhand von Trainingsdatensätzen. »Deep Learning« basiert auf künstlichen neuronalen Netzen und erweist sich im Vergleich zu klassischen Methoden der Objekterkennung als robuster gegenüber variierenden Formen, Verdeckungen, Beschädigungen oder Ausbleichungen, wie sie in realen Szenen typisch sind.

