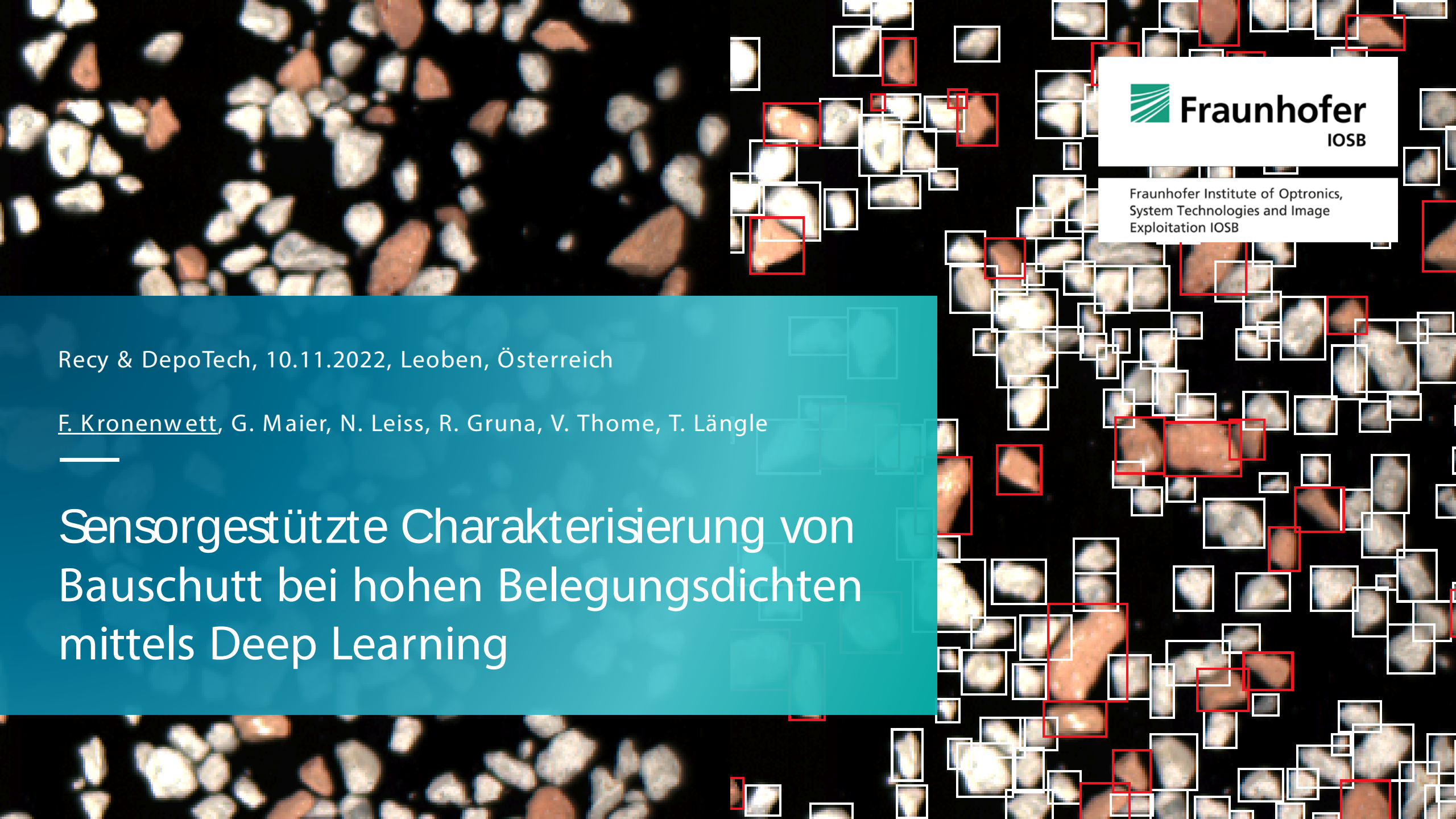


Fraunhofer IOSB

Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB

Recy & DepoTech, 10.11.2022, Leoben, Österreich

F. Kronenwett, G. Maier, N. Leiss, R. Gruna, V. Thome, T. Längle

Sensorgestützte Charakterisierung von Bauschutt bei hohen Belegungsdichten mittels Deep Learning

Motivation
Herausforderungen des Bauschuttrecyclings

Bausektor als ressourcenintensivster Wirtschaftssektor

- 2018 war die mengenmäßig größte Abfallgruppe in Deutschland die mineralischen Bauabfälle einschließlich des Bodenaushubs mit 218,8 Millionen Tonnen
  - 59,8 Mio. t davon Bauschutt
  - 46,6 Mio. t wurden davon recycelt, das entspricht 77,9%

Probleme beim Bauschuttrecycling:

- Qualität des RC-Materials entspricht nicht denen der Primärrohstoffe → häufig Downcycling
- Stark heterogene Zusammensetzung des Bauschutts
- Bauschutt enthält häufig Schadstoffe

→ Bauschuttcharakterisierung und sortenreine Trennung als wichtiger Schritt in der Bauschuttrecyclingkette aufgrund dargestellter Herausforderungen

 © Fraunhofer IOSB

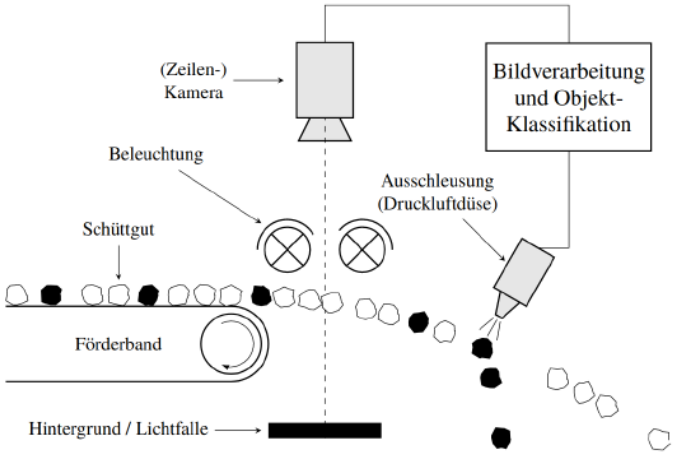
Fraunhofer IOSB

Sensorgestützte Schüttgutsortierung
Funktionsprinzip

Sortenreine sensorgestützte Bauschuttsortierung als Lösung zur Erzielung einer echten Kreislaufwirtschaft

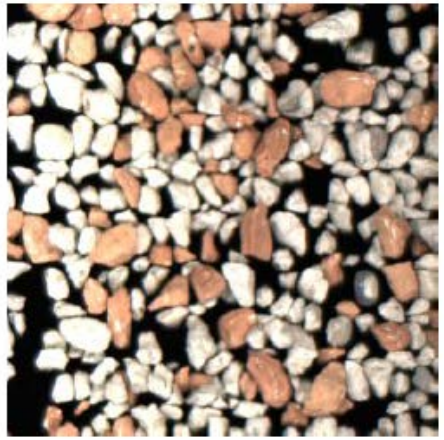
Detektion und Charakterisierung von Bauschutt als notwendiger Bestandteil einer Kreislaufwirtschaft

Beispiel einer sensorgestützten Schüttgutsortierung:



Beispiel: Aufnahme von Ziegel und Kalksandstein

- Farbzeilenkamera mit Aufnahme nach Abwurfkante
- Korngröße: ca. 3-5 mm



 © Fraunhofer IOSB

Fraunhofer IOSB

Problemstellung: Bauschuttcharakterisierung
Herausforderungen und Potential

Eigenschaften sensorgestützter Bauschuttcharakterisierung

- Viele kleine Objekte ähnlicher Größe
- Feste Kameraposition, gut separierbarer Hintergrund

→ Ziel: Objektdetektion, d.h. Lokalisierung und Klassifikation

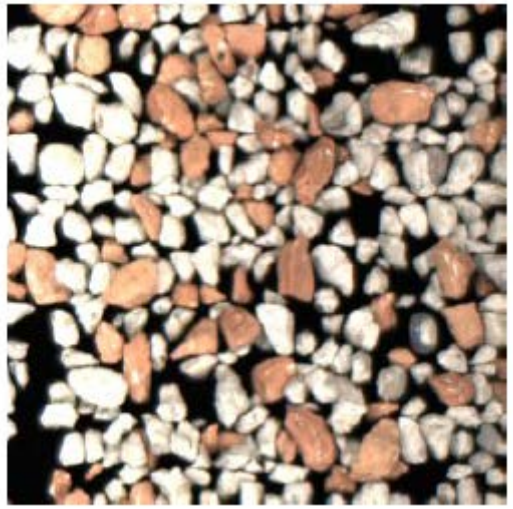
Probleme

- starke Überlappung bei hohem Materialdurchsatz

→ Klassische Bildverarbeitungs-Algorithmen mit Schwirigkeiten bei hohen Belegungsdichten

Lösungsansatz:
Deep Learning mit Potential Probleme zu lösen, aber

- Lange Berechnungszeiten / Inferenzzeiten
- Benötigt eine hohe Anzahl gelabelter Bilder für das Modelltraining



 © Fraunhofer IOSB

Fraunhofer IOSB

Problemstellung: Bauschuttcharakterisierung
Herausforderungen und Potential

Eigenschaften sensorgestützter Bauschuttcharakterisierung

- Viele kleine Objekte ähnlicher Größe
- Feste Kameraposition, gut separierbarer Hintergrund

→ Ziel: Objektdetektion, d.h. Lokalisierung und Klassifikation

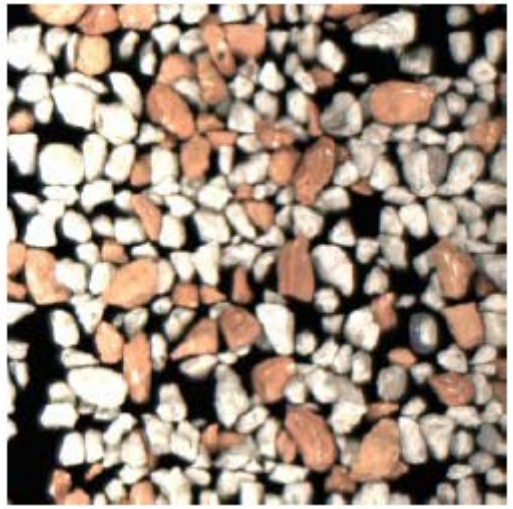
Probleme

- starke Überlappung bei hohem Materialdurchsatz

→ Klassische Bildverarbeitungs-Algorithmen mit Schwirigkeiten bei hohen Belegungsdichten

Lösungsansatz:
Deep Learning mit Potential Probleme zu lösen, aber

- Lange Berechnungszeiten / Inferenzzeiten
- Benötigt eine hohe Anzahl gelabelter Bilder für das Modelltraining



 © Fraunhofer IOSB

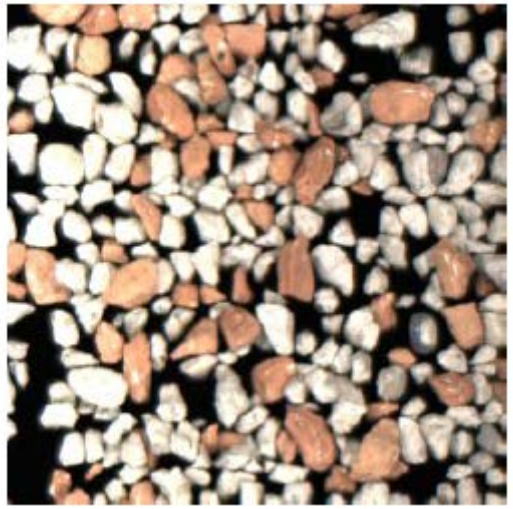
Fraunhofer IOSB

Problemstellung: Bauschuttcharakterisierung
Herausforderungen und Potential

Eigenschaften sensorgestützter Bauschuttcharakterisierung

- Viele kleine Objekte ähnlicher Größe
- Feste Kameraposition, gut separierbarer Hintergrund

→ Ziel: Objektdetektion, d.h. Lokalisierung und Klassifikation



Probleme

- starke Überlappung bei hohem Materialdurchsatz

→ Klassische Bildverarbeitungs-Algorithmen mit Schwirigkeiten bei hohen Belegungsdichten

Lösungsansatz:
Deep Learning mit Potential Probleme zu lösen, aber

- Lange Berechnungszeiten / Inferenzzeiten
- Benötigt eine hohe Anzahl gelabelter Bilder für das Modelltraining

Untersuchung des Potentials, den aufwendigen Annotationsprozess mithilfe synthetischer Datenerzeugung zu verringern. Unter Verwendung synthetischer Daten werden ausgewählte Modelle trainiert und hinsichtlich ihrer Detektionsleistung bei verschiedenen Belegungsdichten verglichen.

 © Fraunhofer IOSB

Fraunhofer IOSB

Synthetische Datenerzeugung

 © Fraunhofer IOSB

Fraunhofer IOSB

Aufbau der Untersuchungen

Datenerzeugung, Modelltraining und Modelltest

Synthetische Datenerzeugung

Erzeugung von Trainingssetz verschiedener Belegungsdichten:

- 10 - 60 % Belegungsdichte
- Verhältnis Ziegel : Kalksandstein 30 : 70

Modelltraining und -test auf synthetischen Bildern

Training drei verschiedener Objektdetektions-Architekturen mit jeweils verschiedener Belegungsdichten der Trainingsbilder

- 4000 Trainings-, 500 Validierungs- und 500 Testbilder

Test der verschiedenen Modelle jeweils auf Testbildern verschiedener Belegungsdichte

- 10 - 60 % Belegungsdichte

Transfer und Modelltest auf realen Bildern

Auswahl des besten Modells und Test auf manuell gelabelten Schüttgutbildern verschiedener Belegungsdichten

- 10 - 60 % Belegungsdichte
- jeweils 10 Bilder pro Belegungsdichte

 © Fraunhofer IOSB

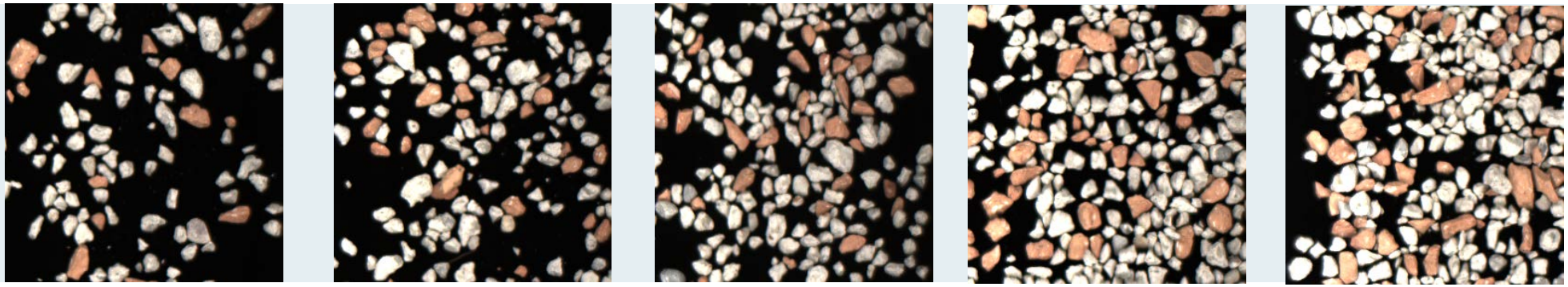
Fraunhofer IOSB

Synthetische Datenerzeugung

Reale Aufnahmen verschiedener Belegungsdichten

Reale Kameraaufnahmen von Ziegel und Kalksandstein bei verschiedenen Belegungsdichten

Reale Aufnahmen



20% Belegungsdichte | 30% Belegungsdichte | 40% Belegungsdichte | 50% Belegungsdichte | 60% Belegungsdichte

 © Fraunhofer IOSB

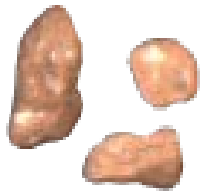
Fraunhofer IOSB

Synthetische Datenerzeugung
Vorgehen

Erzeugung synthetischer Datensätze verschiedener Belegungsdichten

Ausgehend einzelner Objektinstanzen werden durch zufällige Positionierung vollständige Schüttgutbilder erzeugt.

Ziegel



Kalksandstein

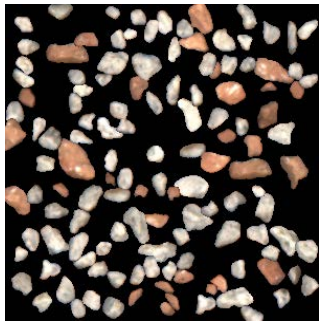
Einstellung über verschiedene Parameter:

- Belegungsdichte
- Maximale Überlappung
- Maximale Objektanzahl
- Mischverhältnis

Ablauf:

1. Position aus Gleichverteilung
2. Überprüfung der Koordinate anhand maximal zulässiger Überlappung bereits gesetzter Objekten
3. Terminierung über maximale Belegungsdichte oder maximale Objektanzahl

Annotationsdatei



Synthetischen Bild

 © Fraunhofer IOSB

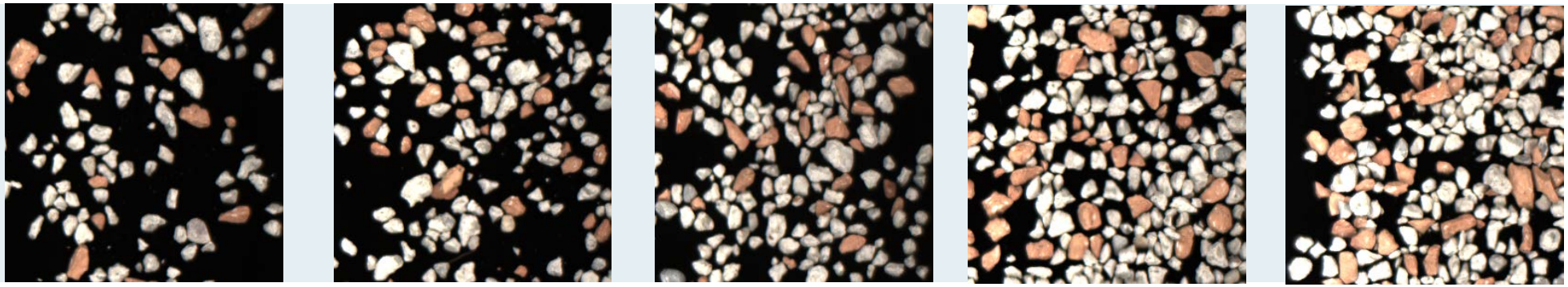
Fraunhofer IOSB

Synthetische Datenerzeugung

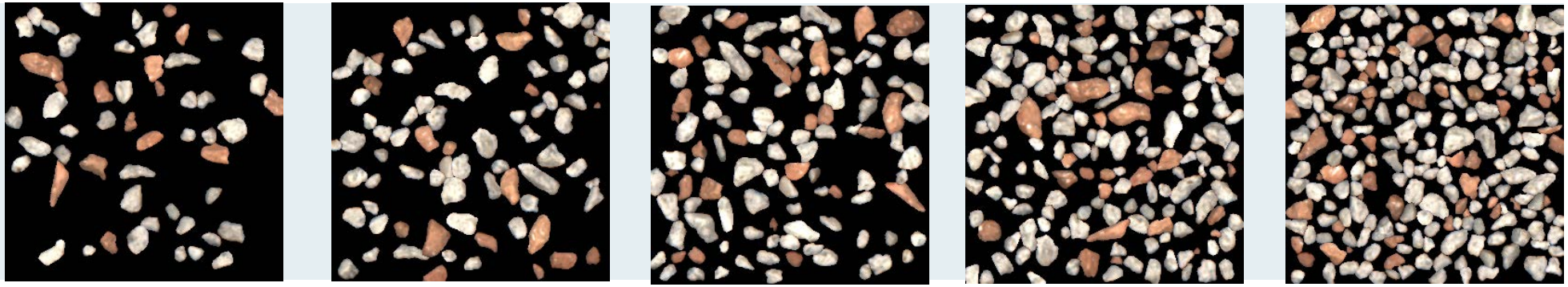
Erzeugte synthetische Schüttgutbilder von echtem Schüttgut

Vergleich zw. realen und synthetische Aufnahmen von Ziegel und Kalksandstein bei verschiedenen Belegungsdichten

Reale Aufnahmen



Synthetische Aufnahmen



20% Belegungsdichte | 30% Belegungsdichte | 40% Belegungsdichte | 50% Belegungsdichte | 60% Belegungsdichte

 © Fraunhofer IOSB

Fraunhofer IOSB

Modelltraining und -test auf synthetischen Bildern

 © Fraunhofer IOSB

Fraunhofer IOSB

Aufbau der Untersuchungen

Datenerzeugung, Modelltraining und Modelltest

Synthetische Datenerzeugung

Erzeugung von Trainingssetz verschiedener Belegungsdichten:

- 10 - 60 % Belegungsdichte
- Verhältnis Ziegel : Kalksandstein 30 : 70

Modelltraining und -test auf synthetischen Bildern

Training drei verschiedener Objektdetektions-Architekturen mit jeweils verschiedener Belegungsdichten der Trainingsbilder

- 4000 Trainings-, 500 Validierungs- und 500 Testbilder

Test der verschiedenen Modelle jeweils auf Testbildern verschiedener Belegungsdichte

- 10 - 60 % Belegungsdichte

Transfer und Modelltest auf realen Bildern

Auswahl des besten Modells und Test auf manuell gelabelten Schüttgutbildern verschiedener Belegungsdichten

- 10 - 60 % Belegungsdichte
- jeweils 10 Bilder pro Belegungsdichte

 © Fraunhofer IOSB

Fraunhofer IOSB

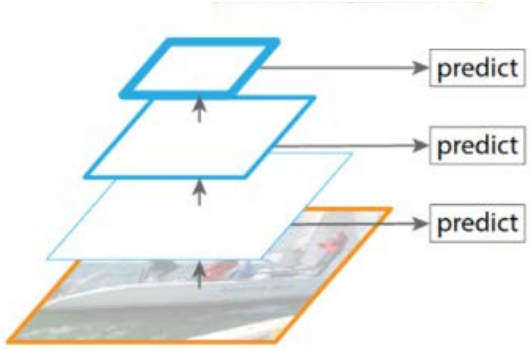
Vergleich und Topologie verwendeter Architekturen

Unterschiede und Eignung

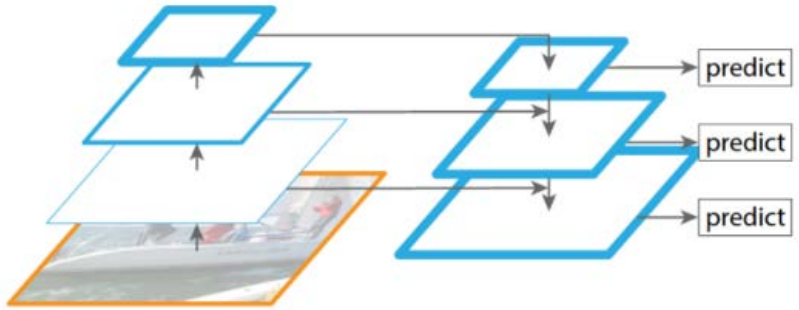
Verwendete Objektdetektions-Architekturen:

	Faster R-CNN	SSD	YOLOv3
Architektur	Two-Stage	One-Stage	One-Stage
Backbone	RestNet-50	VGG-16	Darknet-53
Topologie	Feature Pyramid Network	Pyramid Feature Hierachy	Feature Pyramid Network

Aufgrund ähnlicher Objektgrößen, Skalierungen in Feature-Pyramiden für Detektionsleistung nicht zwangsweise notwendig: Verschlankung der Architekturen möglich



Pyramid Feature Hierachy



Feature Pyramid Network

Grafik aus: Lin, Tsung-Yi, et al. " Feature pyramid networks for object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

 © Fraunhofer IOSB

Fraunhofer IOSB

Vergleich und Topologie verwendeter Architekturen

Unterschiede und Eignung

Verwendete Objektdetektions-Architekturen:

	Faster R-CNN	SSD	YOLOv3
Architektur	Two-Stage	One-Stage	One-Stage
Backbone	RestNet-50	VGG-16	Darknet-53
Topologie	Feature Pyramid Network	Pyramid Feature Hierachy	Feature Pyramid Network

Anpassungen an spezifischer Detektionsaufgabe:

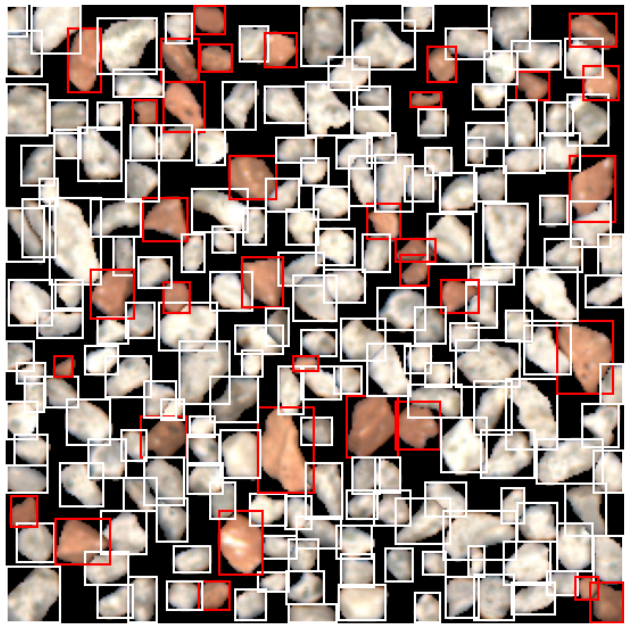
- Klassenanzahl (zwei)
- Anchorgrößen an zu erwartende Objektgrößen
- Maximale Anzahl an Objekten pro Bild
- Lernraten

 © Fraunhofer IOSB

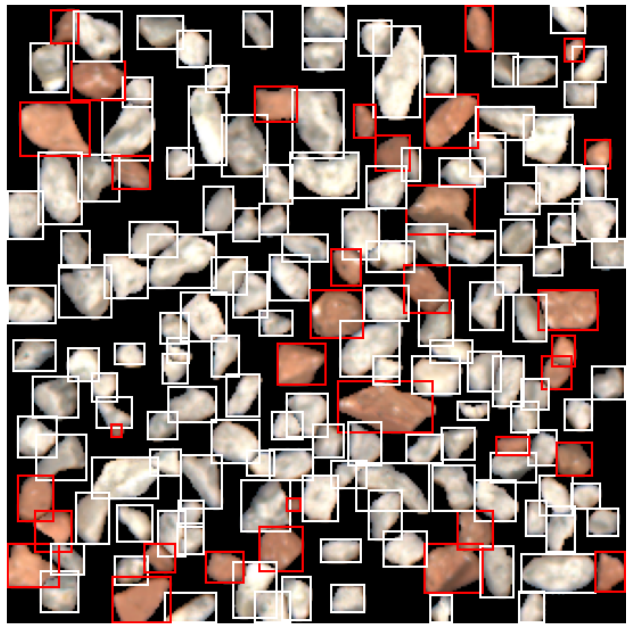
Ergebnisse auf synthetischen Daten

Detektionsleistung der verschiedenen Architekturen

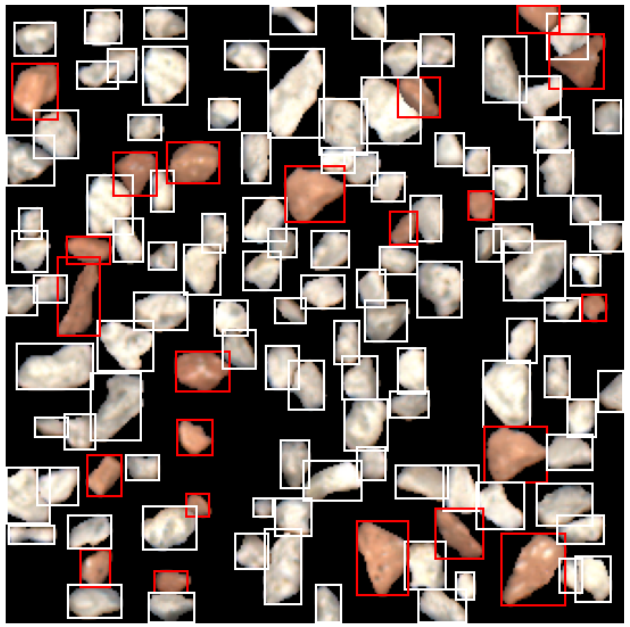
Beispielbilder:



Faster R-CNN



SSD

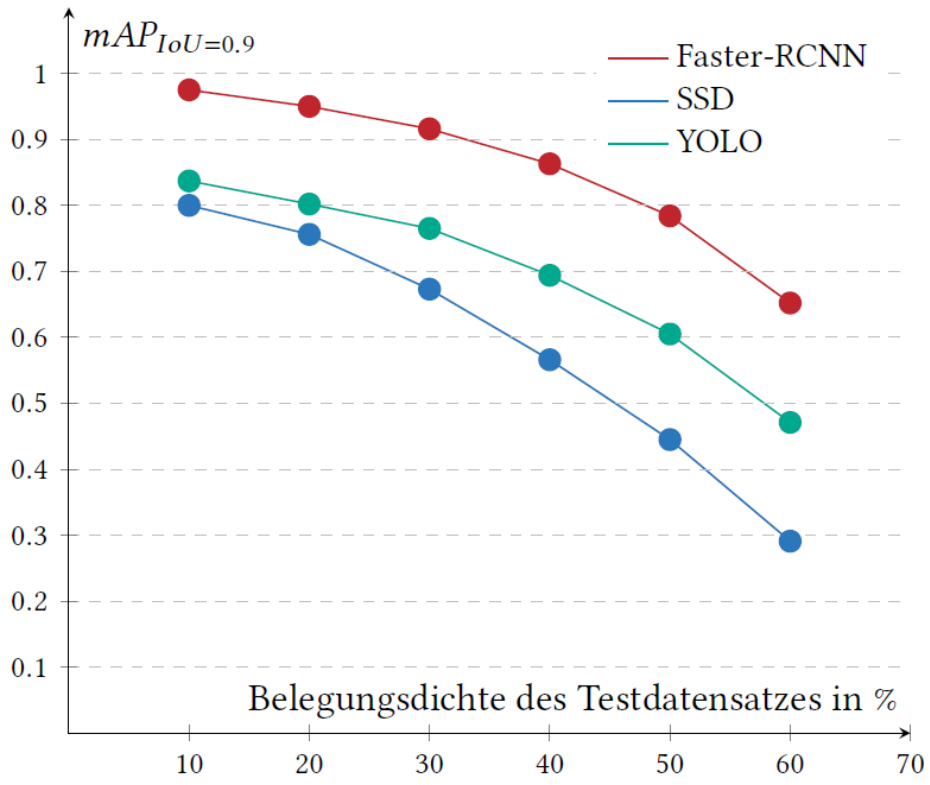


YOLOv3

 © Fraunhofer IOSB

Ergebnisse auf synthetischen Daten

Detektionsleistung auf verschiedenen Belegungsdichten



Beispieltabelle: Faster R-CNN

- Verwendung der restriktiven Metrik $mAP_{IoU=0,9}$
  → Detektion gilt als TP bei 90% Überlappung mit GT

$mAP_{IoU=0.9}$		Belegungsdichte des Testdatensatzes					
		10%	20%	30%	40%	50%	60%
Belegungsdichte des Trainingsdatensatzes	10%	0.968	0.932	0.881	0.795	0.692	0.526
	20%	0.968	0.944	0.903	0.836	0.750	0.607
	30%	0.968	0.949	0.909	0.852	0.768	0.638
	40%	**0.975**	**0.950**	**0.916**	**0.863**	**0.784**	0.652
	50%	0.966	0.948	0.912	0.859	**0.784**	0.658
	60%	0.964	0.939	0.902	0.850	0.781	**0.660**

 © Fraunhofer IOSB

Fraunhofer IOSB

Ergebnisse auf synthetischen Daten

Detektionsleistung auf verschiedenen Belegungsdichten

Erkenntnisse aus synthetischer Testdaten:

Ergebnisse:

- Alle drei Architekturen mit guter Detektionsleistung
  - Faster R-CNN am genausten, gefolgt von YOLOv3 und SSD
- Modelle, die mit höheren Belegungsdichten trainiert wurden, mit besserer Detektionsleitung auf niedrigeren Belegungsdichten

Erkenntnisse:

- Anforderungen an den Trainingsdatensatz:
  - Belegungsdichte etwas höher wählen (ca. 10%) als die zu erwartende Auslastung
- Overfitting auf überrepräsentierter Klasse beachten (Datensatz, Hyper-Trainings-Parameter,…)

 © Fraunhofer IOSB

Transfer und Modelltest auf realen Bildern

 © Fraunhofer IOSB

Fraunhofer IOSB

Aufbau der Untersuchungen

Datenerzeugung, Modelltraining und Modelltest

Synthetische Datenerzeugung

Erzeugung von Trainingssetz verschiedener Belegungsdichten:

- 10 - 60 % Belegungsdichte
- Verhältnis Ziegel : Kalksandstein 30 : 70

Modelltraining und -test auf synthetischen Bildern

Training drei verschiedener Objektdetektions-Architekturen mit jeweils verschiedener Belegungsdichten der Trainingsbilder

- 4000 Trainings-, 500 Validierungs- und 500 Testbilder

Test der verschiedenen Modelle jeweils auf Testbildern verschiedener Belegungsdichte

- 10 - 60 % Belegungsdichte

Transfer und Modelltest auf realen Bildern

Auswahl des besten Modells und Test auf manuell gelabelten Schüttgutbildern verschiedener Belegungsdichten

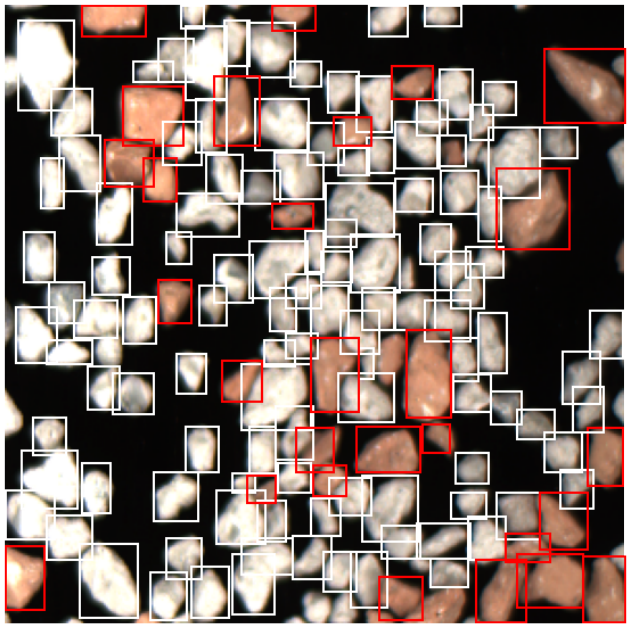
- 10 - 60 % Belegungsdichte
- jeweils 10 Bilder pro Belegungsdichte

 © Fraunhofer IOSB

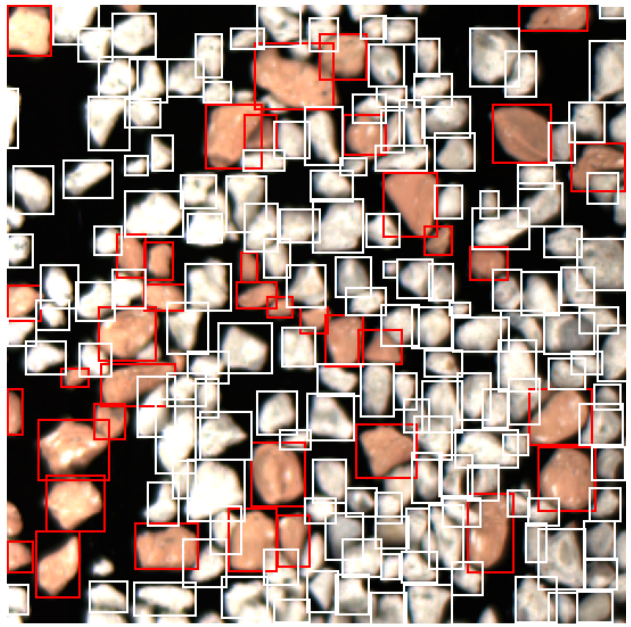
Ergebnisse auf realen Daten

Detektionsleistung der verschiedenen Architekturen

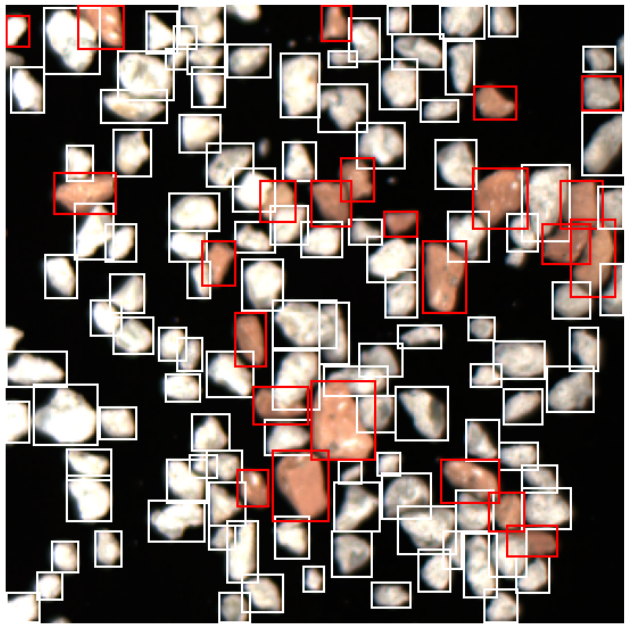
Beispielbilder:



Faster R-CNN



SSD



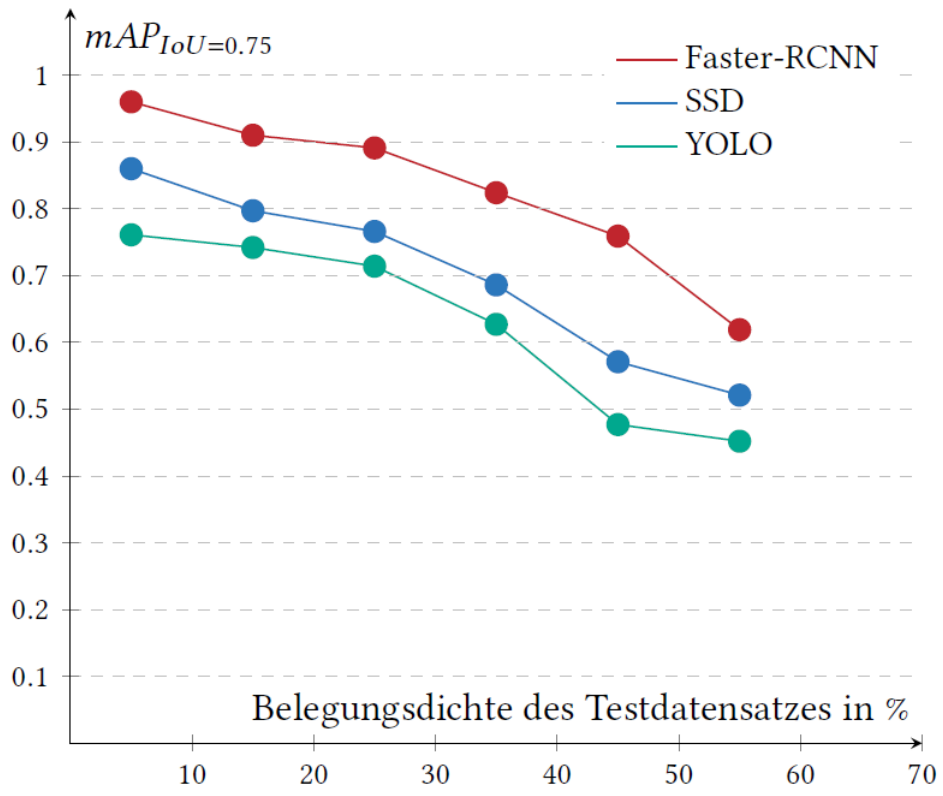
YOLOv3

 © Fraunhofer IOSB

Fraunhofer IOSB

Transfer und Ergebnisse auf echten Schüttgutaufnahmen

Detektionsleistung auf verschiedenen Belegungsdichten



Beispieltabelle: Faster R-CNN

- Verwendung der restriktiven Metrik $mAP_{IoU=0.75}$
  → Detektion gilt als TP bei 75% Überlappung mit GT

Metrik	Belegungsdichten des Testdatensatzes in %					
	0-10	10-20	20-30	30-40	40-50	50-60
$mAP_{IoU=0.75}$	0.960	0.910	0.891	0.824	0.759	0.619
AP_{Ziegel}	0.790	0.748	0.747	0.686	0.625	0.593
$AP_{\text{Kalksandstein}}$	0.707	0.713	0.689	0.655	0.622	0.507
$mAR_{IoU=0.75}$	0.974	0.928	0.931	0.890	0.802	0.746

 © Fraunhofer IOSB

Fraunhofer IOSB

Transfer und Ergebnisse auf echten Schüttgutaufnahmen

Erkenntnisse beim Test auf realen Daten

Erkenntnisse aus realen Testdaten:

Transfer der auf synthetischen Daten trainierten Modell auf reale Schüttgutaufnahmen

Ergebnisse:

- Abfall der Metrik beim Transfer auf realen Daten → Abnahme der Genauigkeit geschätzter Boudning Boxen
  - Anpassung auf Objektinsanzen während des Trainings führt zu Problemen der Detetkion nicht repräsentierter Objektformen
- Faster R-CNN mit bester Detektionsleistung, gefolgt von SSD und YOLOv3
  - Faster R-CNN: falsche zusätzliche Detektionen der überrepräsentierten und stark überlappenden Kalksandsteinklasse
  - YOLOv3 und SSD: starke Zunahme falscher Detektionen, speziell bei dichten Objektregionen

Erkenntnisse:

- Synthetische Datenerzeugung muss das Erscheinungsbild realer Schüttgutaufnahmen möglichst realistisch nachbilden
- Varianz in den Trainingsdaten erhöhen:
  - Abdeckung verschiedener Objektgrößen, -formen, etc.

 © Fraunhofer IOSB

Abschätzung der Echtzeitfähigkeit

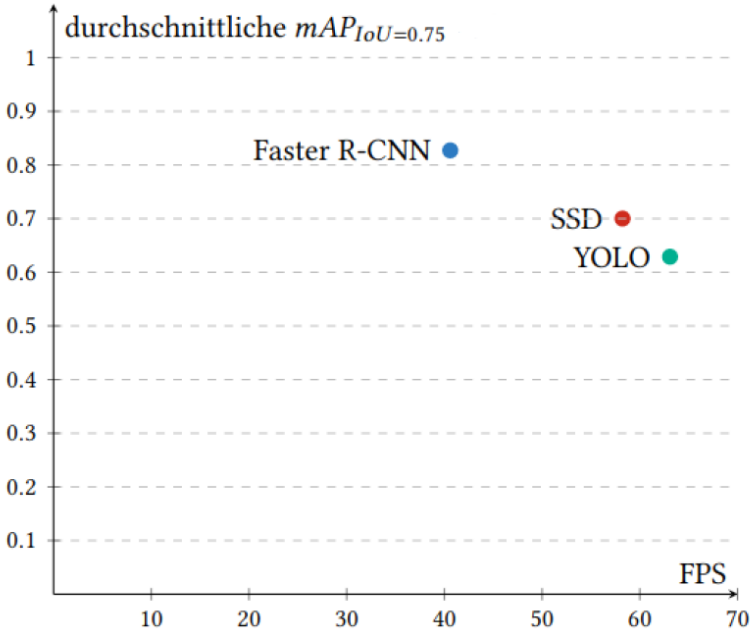
Trade-off zwischen Detektionsleistung und Geschwindigkeit

Aufbau der Inferenzzeitmessungen:

- Export und Optimierung mittels TensorRT
  - Nvidia GeForce GTX-970
  - Bildgröße: 300 x 300
- Verwendung von Open-Source Deployment Toolsets

Inferenzzeitmessungen:

Faster R-CNN		SSD		YOLO	
ms	FPS	ms	FPS	ms	FPS
24.43	40.94	17.07	58.59	15.57	63.24



Durchschnittliche $mAP_{IoU=0,75}$ des jeweils besten Modells auf allen Belegungsdichten realer Daten

 © Fraunhofer IOSB

Fraunhofer IOSB

Zusammenfassung

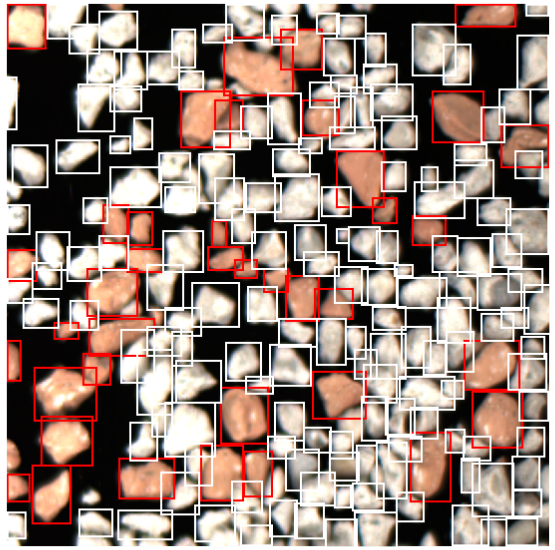
Sensorgestützte Charakterisierung von Bauschutt bei hohen Belegungsdichten mittels Deep Learning

Erkenntnisse:

Grundsätzliche Eignung gängiger Objektdetektions-Architekturen zur Lokalisierung und Klassifikation von Bauschutt

- Training auf synthetisch generierten Daten minimiert Label-Aufwand und erzielt ausreichend gute Ergebnisse
  - Zusammensetzung des Trainingsdatensatzes relevant:
    - → Variabilität der Objektinstanzen erhöhen, d.h. Form, Größe, Struktur etc.
    - → Optimierung der Objektpositionierung bei der Datenerzeugung
  - Verschiedene Belegungsdichten im Trainingsset beachten

- Trade-off zwischen Geschwindigkeit und Genauigkeit



Ausblick:

- Zusätzliche Optimierung und Anpassung der Architekturen
- Erhöhung der Variabilität der Trainingsdaten

 © Fraunhofer IOSB

Kontakt

Felix Kronenwett, M. Sc.
Sichtprüfsysteme
Telefon +49 721 6091-411

felix.kronenwett@iosb.fraunhofer.de