



Datenzentrierte Intelligenz: Synthetische Bilddaten als Grundlage für KI-Erfolgsgeschichten



WAAM Anwendungsfall Vorstellung

Wire Arc Additive Manufacturing (Additive Fertigung mit Lichtbogen)

- Schmelzung des Metalldrahts mithilfe eines Lichtbogens, um das Material aufzutragen
- Individuelle, additive Fertigung



WAAM Anwendungsfall Vorstellung

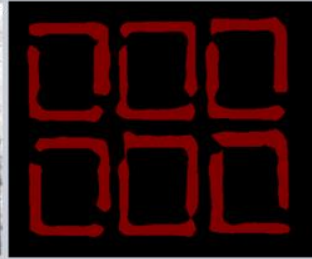
Aufgenommenes Bild



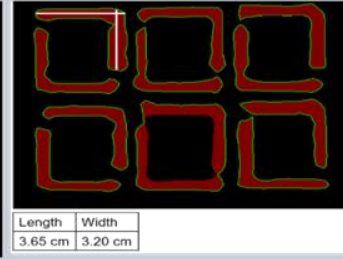
Materialträger Erkennung



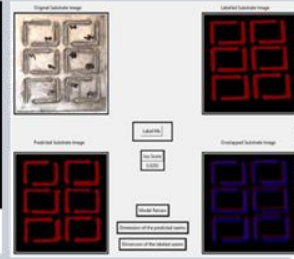
Schweißnähte Segmentierung



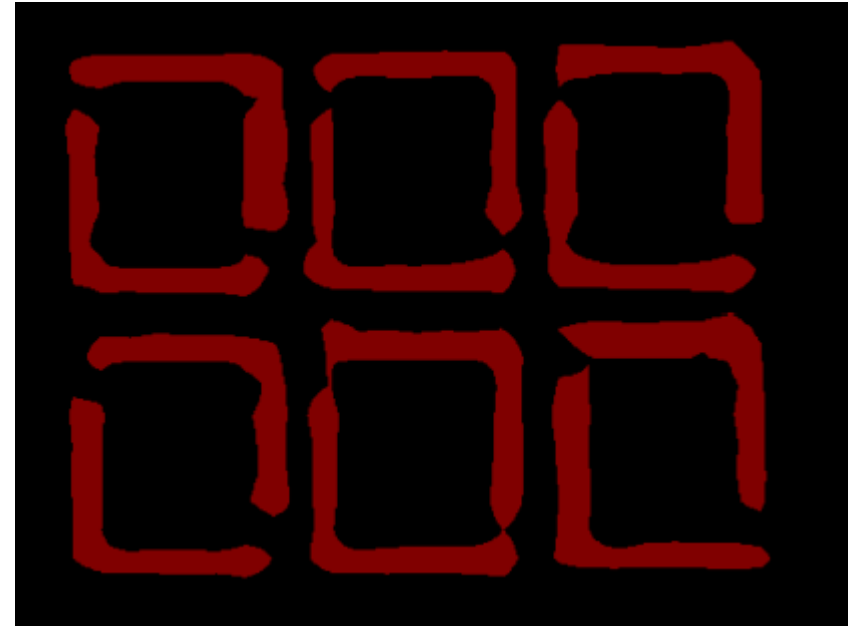
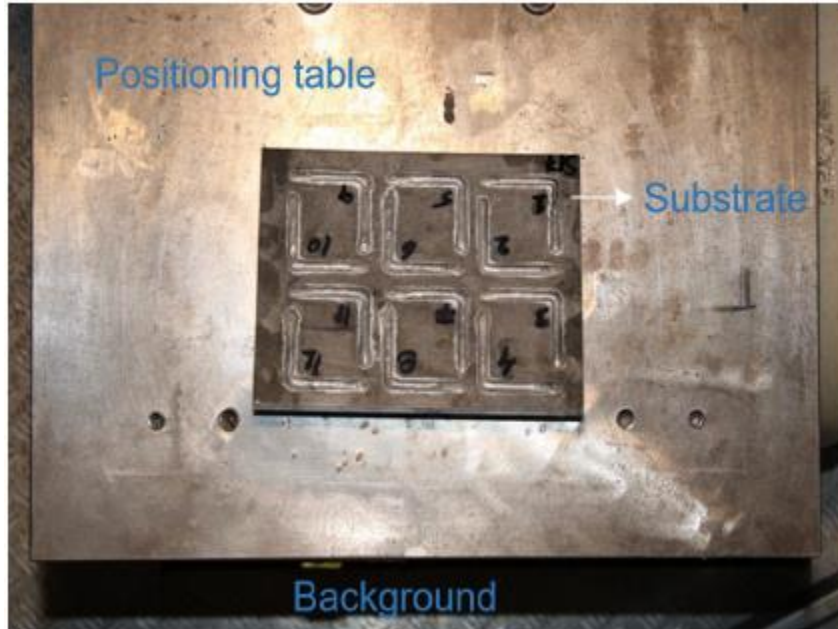
Vermessung der Geometrie



App



Segmentierung der Schweißnähte zur Messung der resultierenden Schweißnahtgeometrie



Was definiert einen guten Datensatz?

Quantität:

Mehr Daten sind meistens besser



Datenschutz und

Ethik:

möglichen
Auswirkungen auf
den Menschen

Qualität:

Relevant und
repräsentativ für das
Problem

Konsistenz:
in Format und
Struktur

Vielfältigkeit:

Diverse Szenarien
und Beispiele des
Problems

Ausgewogenheit:

Gleiche Anzahl für
jede Klassen

Herausforderungen mit Daten und Deep Learning

- Fehlende Standardisierung der Arbeitsabläufe bei der Datenkuration
- Ressourcenintensive, manuelle Datenaufbereitung
- Mangel an den richtigen Trainingsdaten



#Datacentric im WAAM -Projekt

1.

**Qualitätsprüfung der
Annotationen**

- Bericht zur Qualität
- Vorschläge zur Optimierung

2.

**Analyse der realen
Daten**

- Bericht zu möglichen Blindspots
- Vorschläge zur Optimierung

3.

Daten Anreicherung

- Ergänzung der realen Daten
- Statistik zu neuem Datensatz


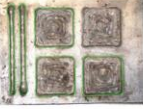

1. Qualitätsprüfung der Annotationen

Prüfung auf mögliche Fehler:

- Fehlklassifizierung
- Unvollständige Labels
- Ungenaue Lokalisierung
- Inkonsistenz der Labels
- ...

1. Annotation Quality:

⚠ The check was performed on the augmented, real dataset to find gross error.
90 images with issues were found, meaning 5.85 % of the datasets images have relevant annotation issues.

| Finding | Samples | File IDs |
|---|---|--|
| Shifted position (probably due to augmentation) |  | 721,722,723,724,725,726,728,729, 847,848,849,850,851,852,853,854, 855,856,857,858,859,860,861,862, 863,864,874,875,876,877,878,879, 880,881,882,1180,1181,1182,1183, ,1184,1185,1186,1187,1188,1234, 1235,1236,1237,1238,1239,1240, 1241,1242 |
| Missing label |  | 118,119,120,121,122,123,124,125, 126,208,209,210,211,212,213,214, 215,216,289,290,291,292,293,294, 295,296,297 |
| Incorrect polygon shapes? |  | 271,272,273,274,275,276,277,278, 279 |

Qualitätsbericht
Annotationen

1. Qualitätsprüfung der Annotationen



Verarbeitungsfehler

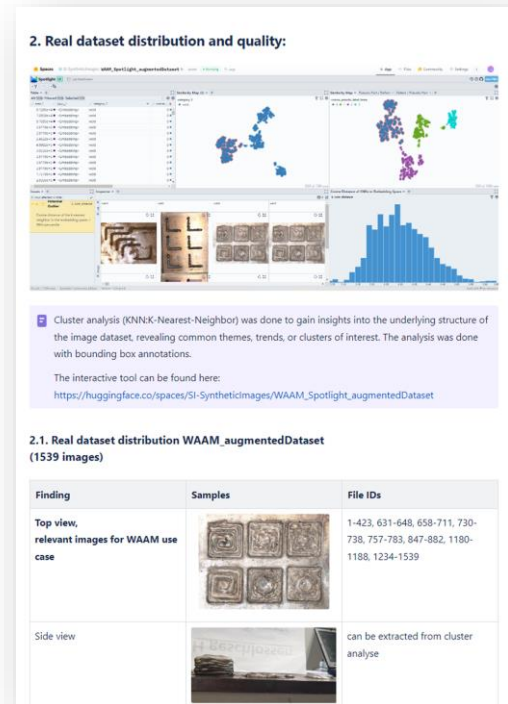


Fehlende
Labels

2. Analyse der realen Daten

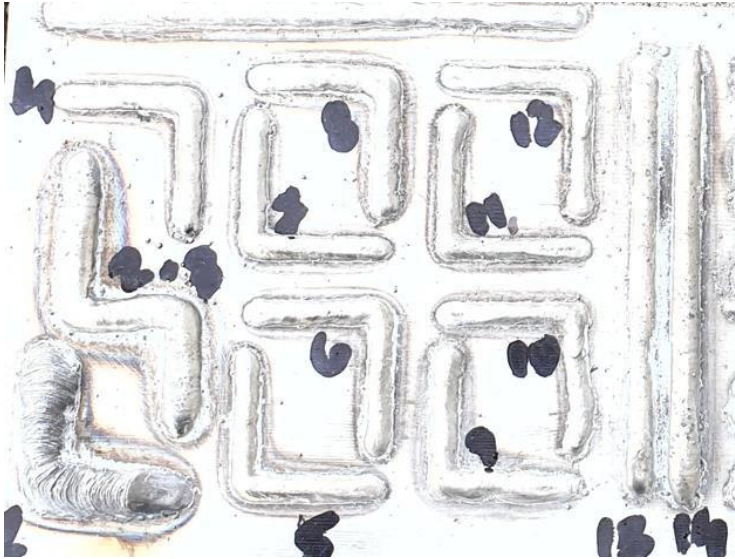
Ziele der Analyse:

- Was repräsentiert der Datensatz, bzw. was nicht?
- Was sieht “die Maschine” in den Bildern?
- Was beeinflusst das Model?
- Was sind außergewöhnliche Merkmale?
- Wie können wir dagegen steuern?



Analysebericht
Datensatz

2. Analyse der realen Daten

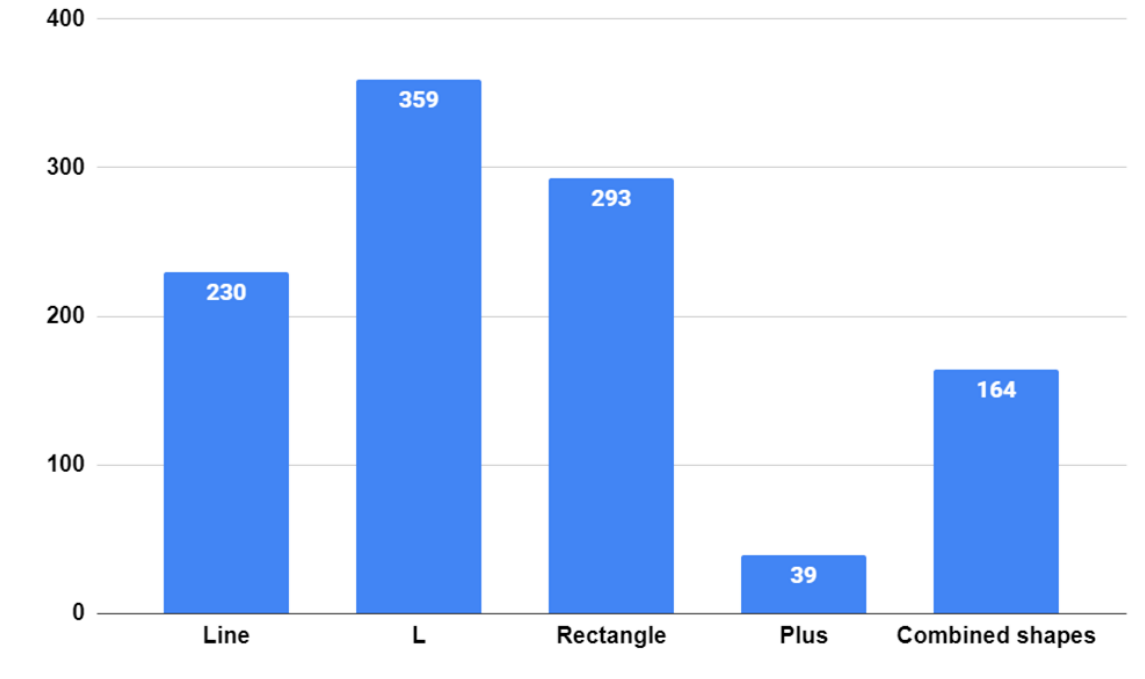


Überbelichtet



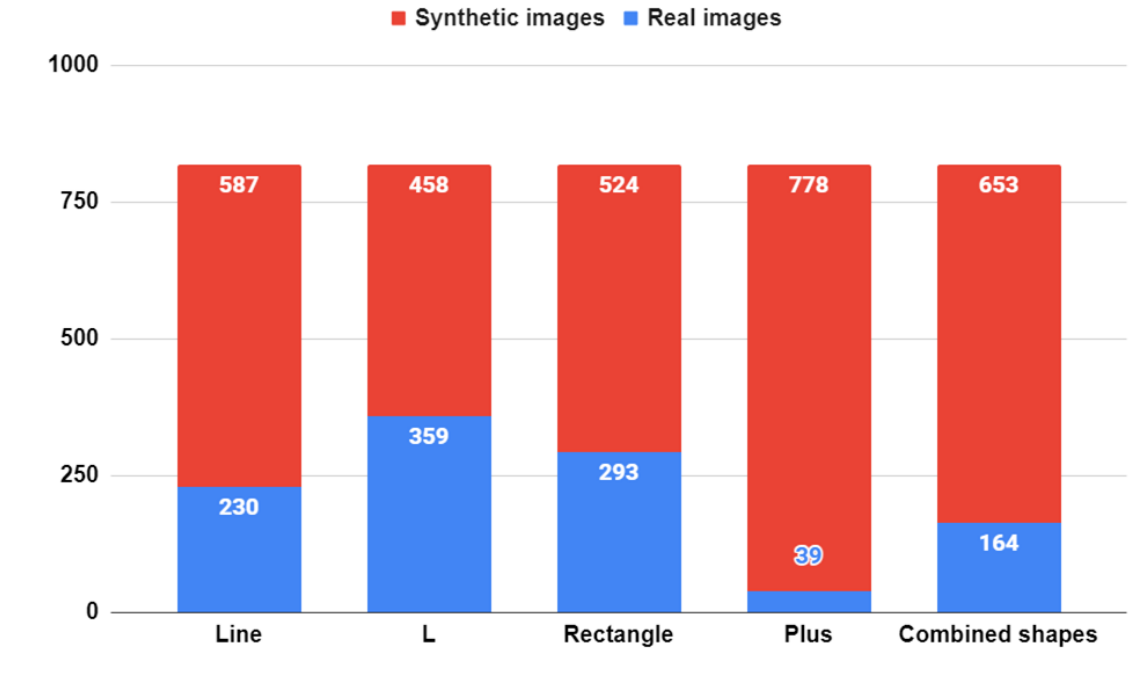
Bildfehler

2. Analyse der realen Daten



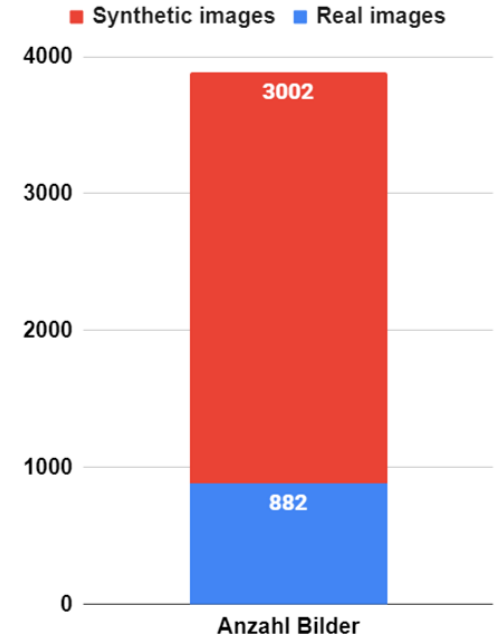
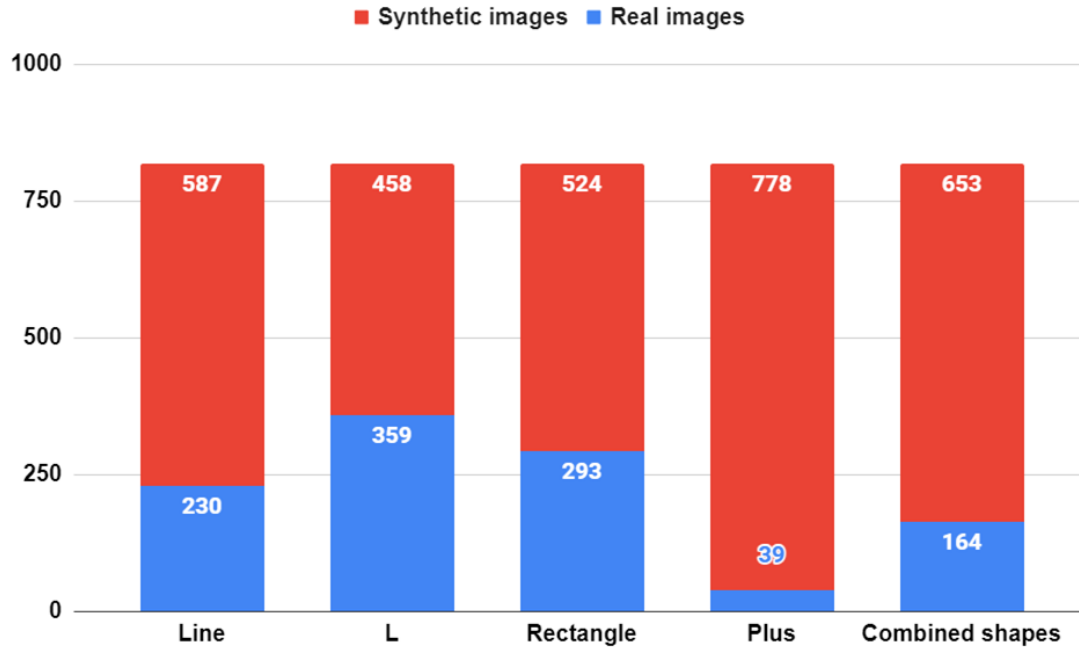
Form Verteilung über den Datensatz

3. Daten Anreicherung



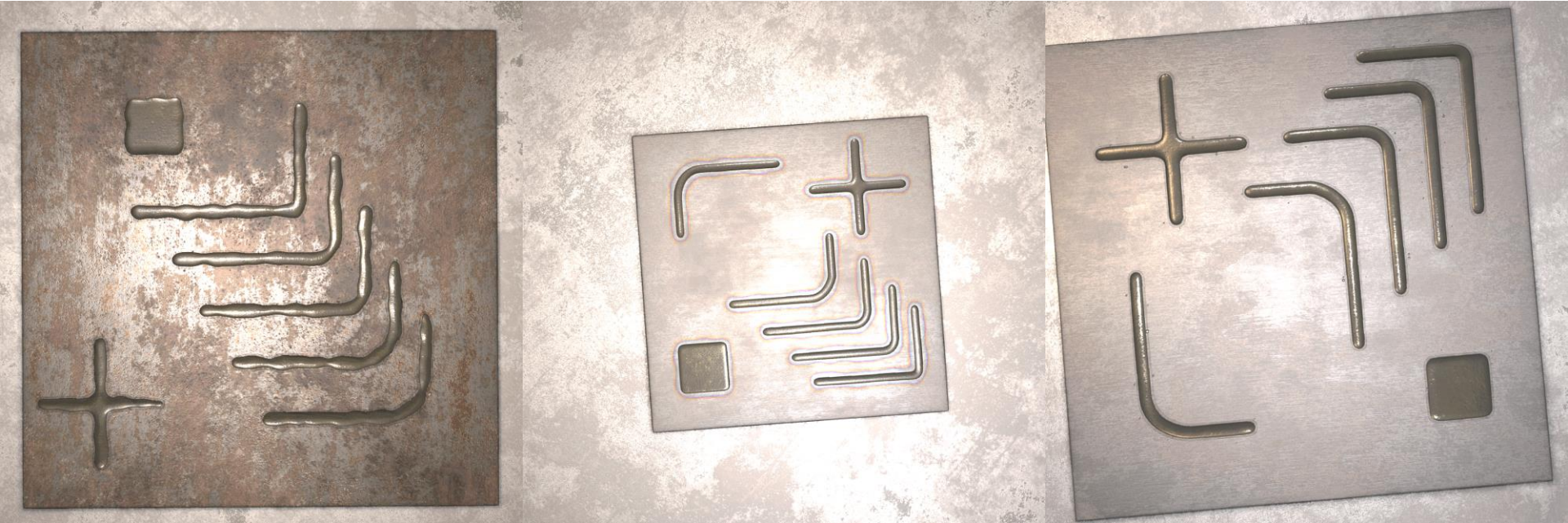
Ideale Form Verteilung über den Datensatz

3. Daten Anreicherung



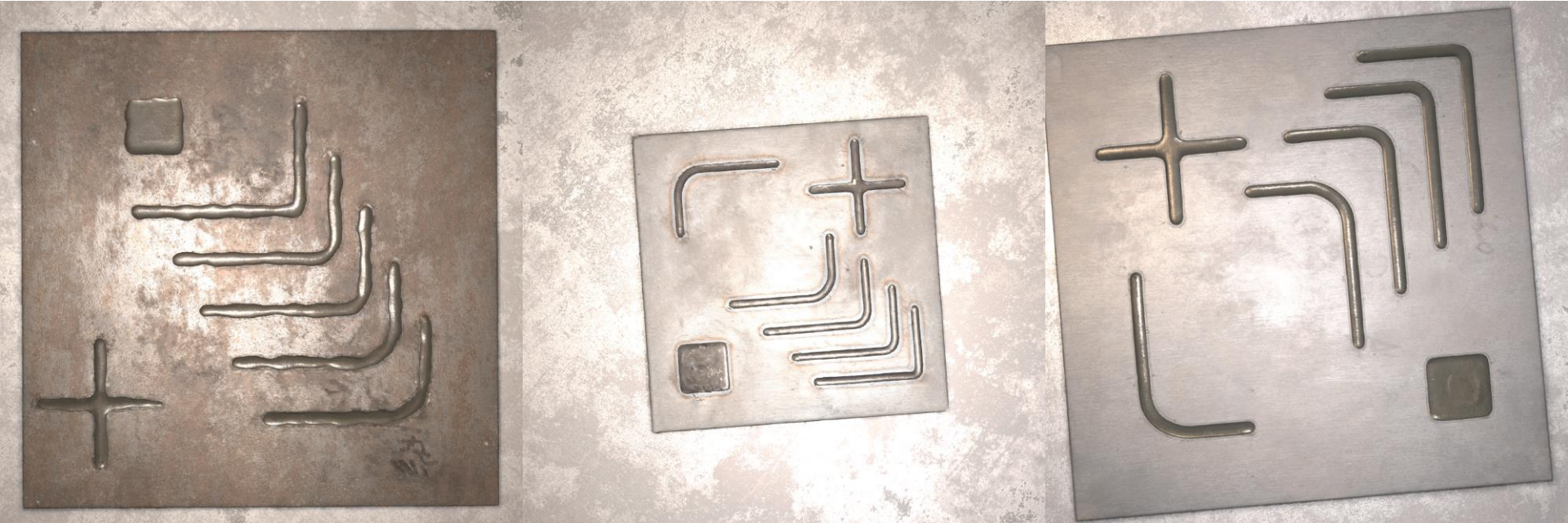
Ideale Form Verteilung über den Datensatz

3. Daten Anreicherung



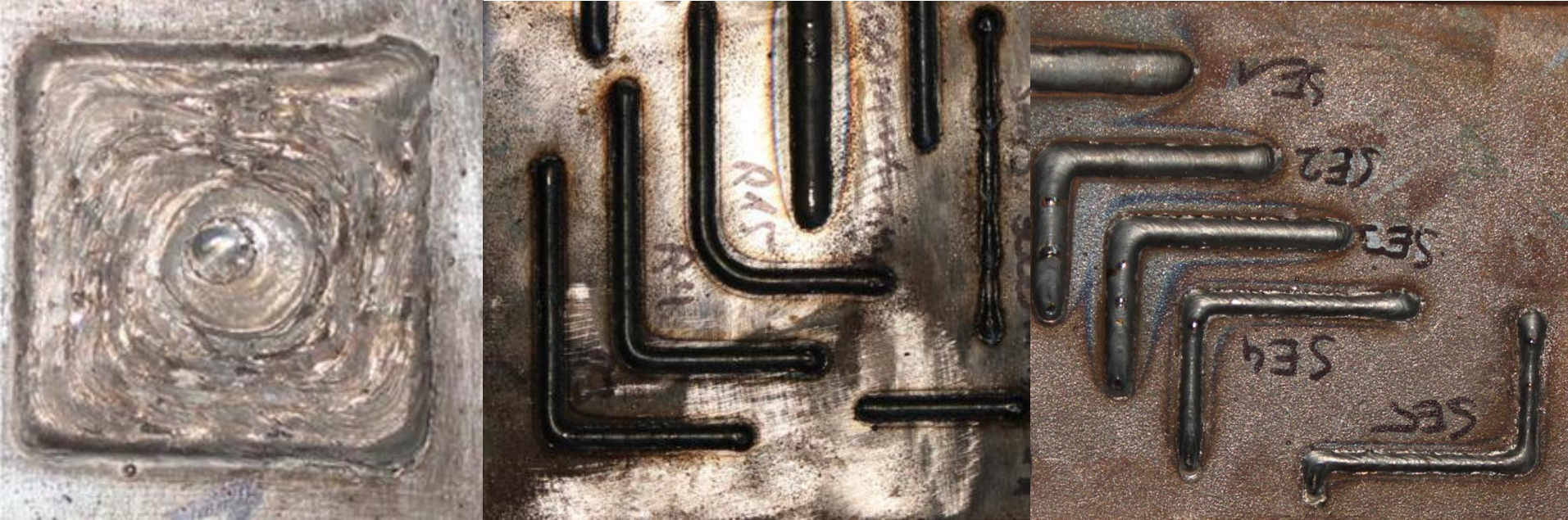
3D simuliert und gerendert

3. Daten Anreicherung



Nach Styletransfer

3. Daten Anreicherung



Oberflächen Struktur

Oberflächen Struktur

Schrift ähnliche Features

3. Daten Anreicherung



Oberflächen Struktur

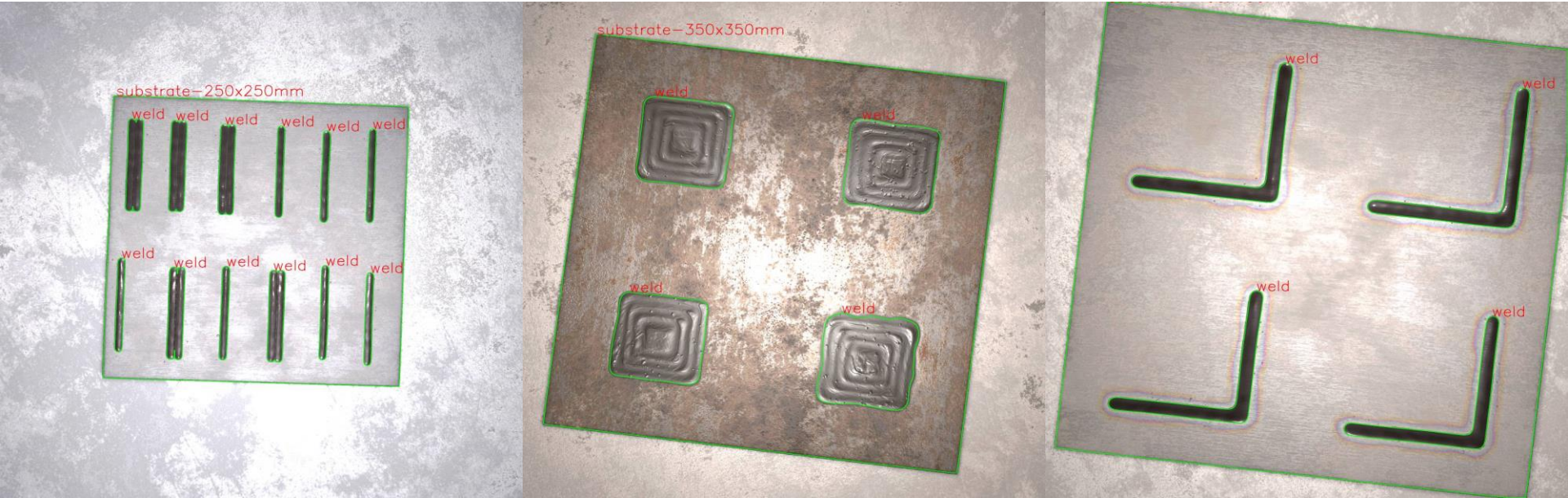


Oberflächen Struktur



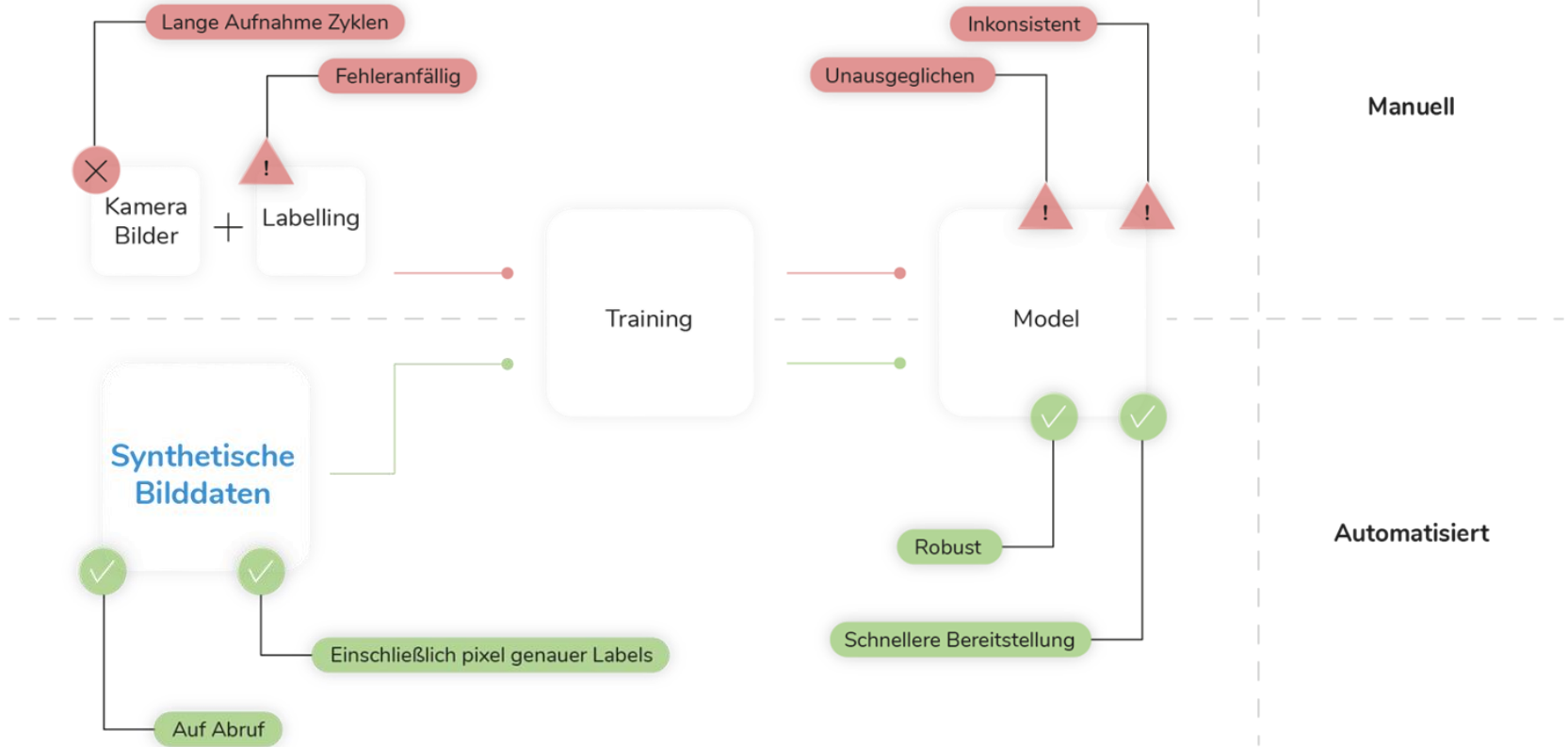
Schrift ähnliche Features

3. Daten Anreicherung

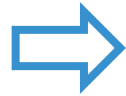


Annotation

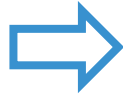
KI-Implementierung mit **synthetischen Daten**



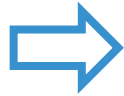
Wirtschaftliche Vorteile von synthetischer Daten



Geringere Zeit bis zur
Bereitstellung



Verkürzte Zeit bis zum
Erreichen des ROI

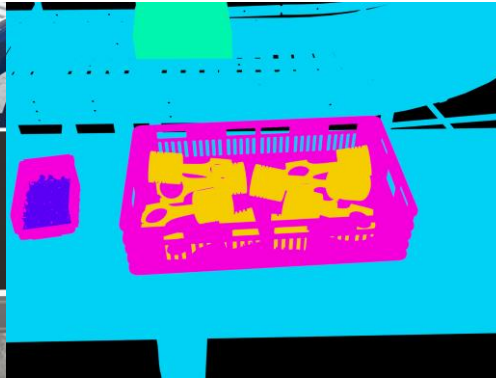


Reduzierter Bedarf an
Projektressourcen

3. Daten Anreicherung - Annotationen



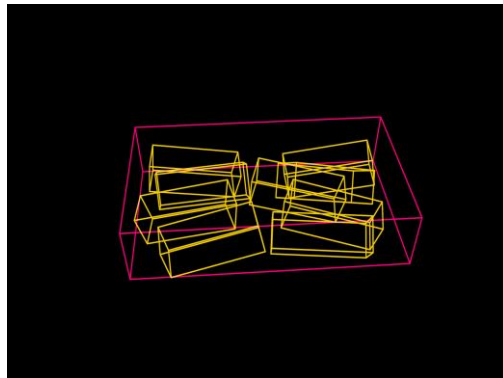
Ground truth



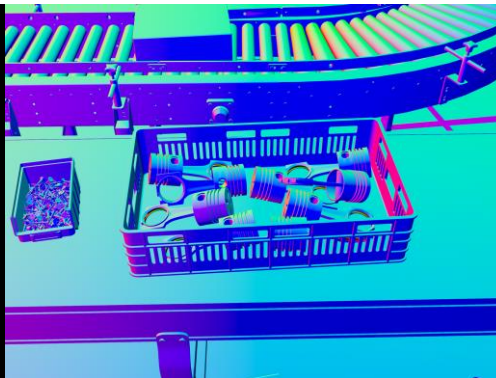
Class Segmentation



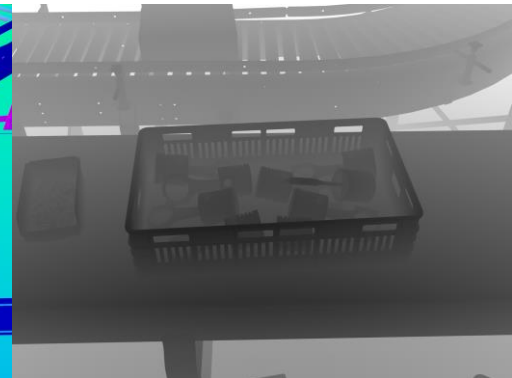
Instance Segmentation



3D Bounding Box



Surface Normal



Depth

Automatisierte, visuelle Inspektion (AVI)



Augenommenes Bild



Synthetischer Bild-Datensatz



Inclusive Annotation

About Si

Tools & Services:

- Data consulting
- Dataset curation (analytics)
- Datasets
- Data generation
- Feasibility study
- Model Maintenance

Tech capabilities:

- Data analytics
- Image processing
- Machine learning
- Deep Learning
- Generative AI



Tim Schäfer,
Co-Founder



[linkedin.com/in/tim-schaefer/](https://www.linkedin.com/in/tim-schaefer/)



tim.schaefer@synthetic-images.com



