

# Evaluation der Robustheit von Deep-Learning-Verfahren in der Qualitätssicherung am Beispiel des Rührreibschweißens

---

Forschungspraxis

Shaoxiang Tan  
Maschinenwesen

Garching, 20.09.2023

*Leitstand des iwb, TUM.*

# Gliederung

1. Einleitung
2. Generierung der Datenbasis
3. Bewertung der Robustheit
4. Zusammenfassung und Ausblick



# Einleitung

---



Die steigende Verwendung von FSW hat die Nachfrage nach zerstörungsfreien Prüfverfahren erhöht. Hierbei sind effektive und robuste Prüfverfahren unerlässlich.

## Einleitung (1/2) – Forschungsbedarf

### Rührreibschweißen (FSW)

- keine Zusatzstoffe, Schutzgas
- emissionsarm, umweltfreundlich
- gute Schweißnahtqualität



### Sicherheitskritische Anwendungen in

- Luftfahrtindustrie
- Raumfahrtindustrie
- Automobilindustrie

[1]

## Automatische Qualitätsüberwachung durch Deep-Learning-Verfahren

Klassifizierung																														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28		
A	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○		
B	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	○		
C	○	✗	✗	✗	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	○		
Legend	A	Schweißgrat					Keine Gratbildung							Signifikante Gratbildung																
	B	Innenbereich					Keine Defekte							Decklagenfehler							Oberflächenmangel									
	C	Temperatur					Richtige Temperatur							Hohe Temperatur							Niedrige Temperatur									
	✓	Richtige Klassifizierung								✗	Falsche Klassifizierung								○	Keine Evaluation möglich										

Von Hartl (2019) [2] wird eine Klassifizierungsgenauigkeit von 92,1% erreicht.

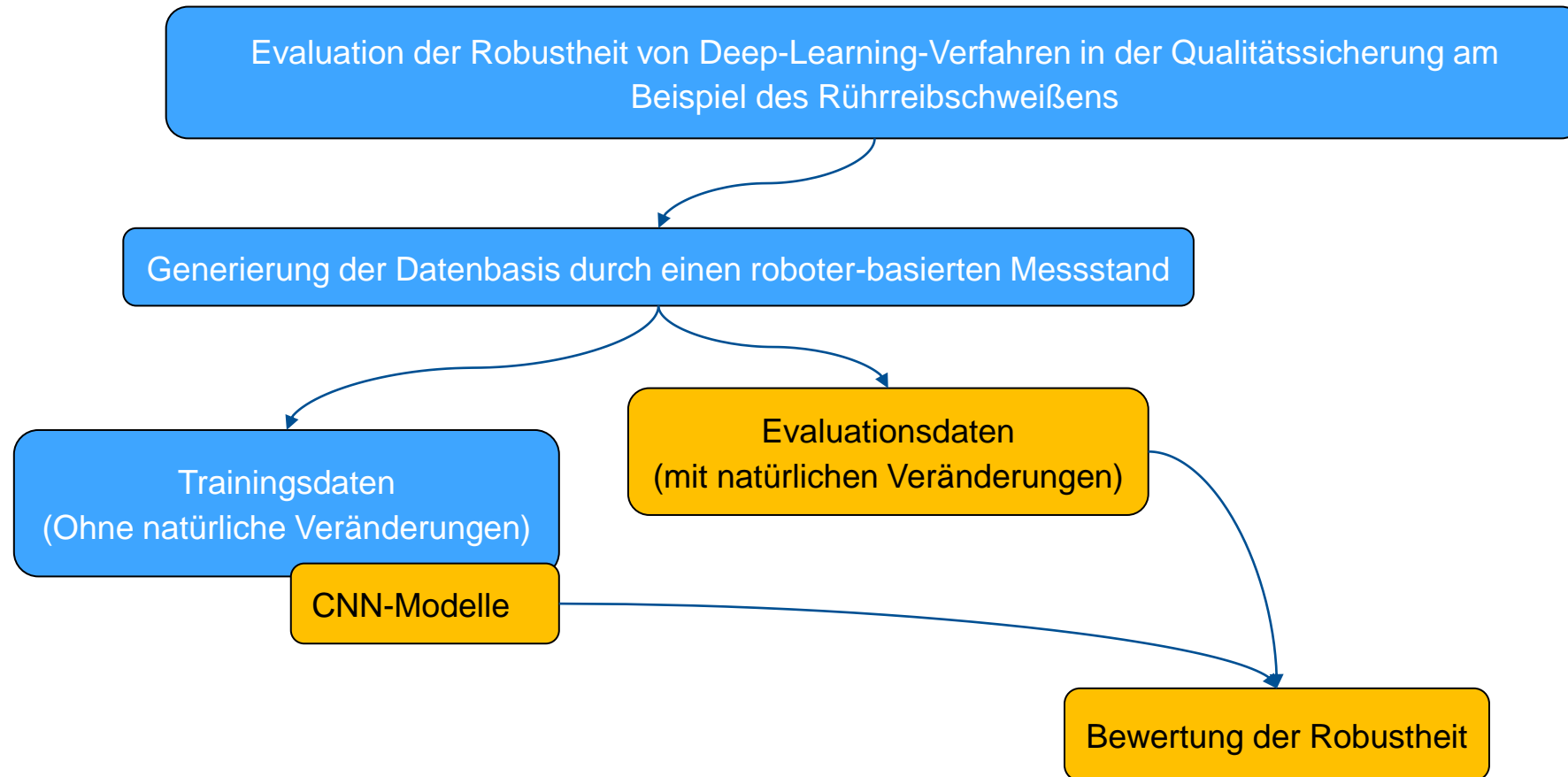


Was passiert, wenn es zu natürlichen Veränderungen oder Einflüssen kommt?

Studie von Drenkow (2022) [3] zeigt: es kann zu signifikanten Leistungsabfällen von Modellen um 30% bis 40% kommen, wenn sie mit natürlichen verschlechterten Daten konfrontiert werden.

# Das Ziel dieser Arbeit besteht in der Evaluation der Robustheit in der Qualitätssicherung am Beispiel des Rührreibschweißens

## Einleitung (2/2) – Zielsetzung

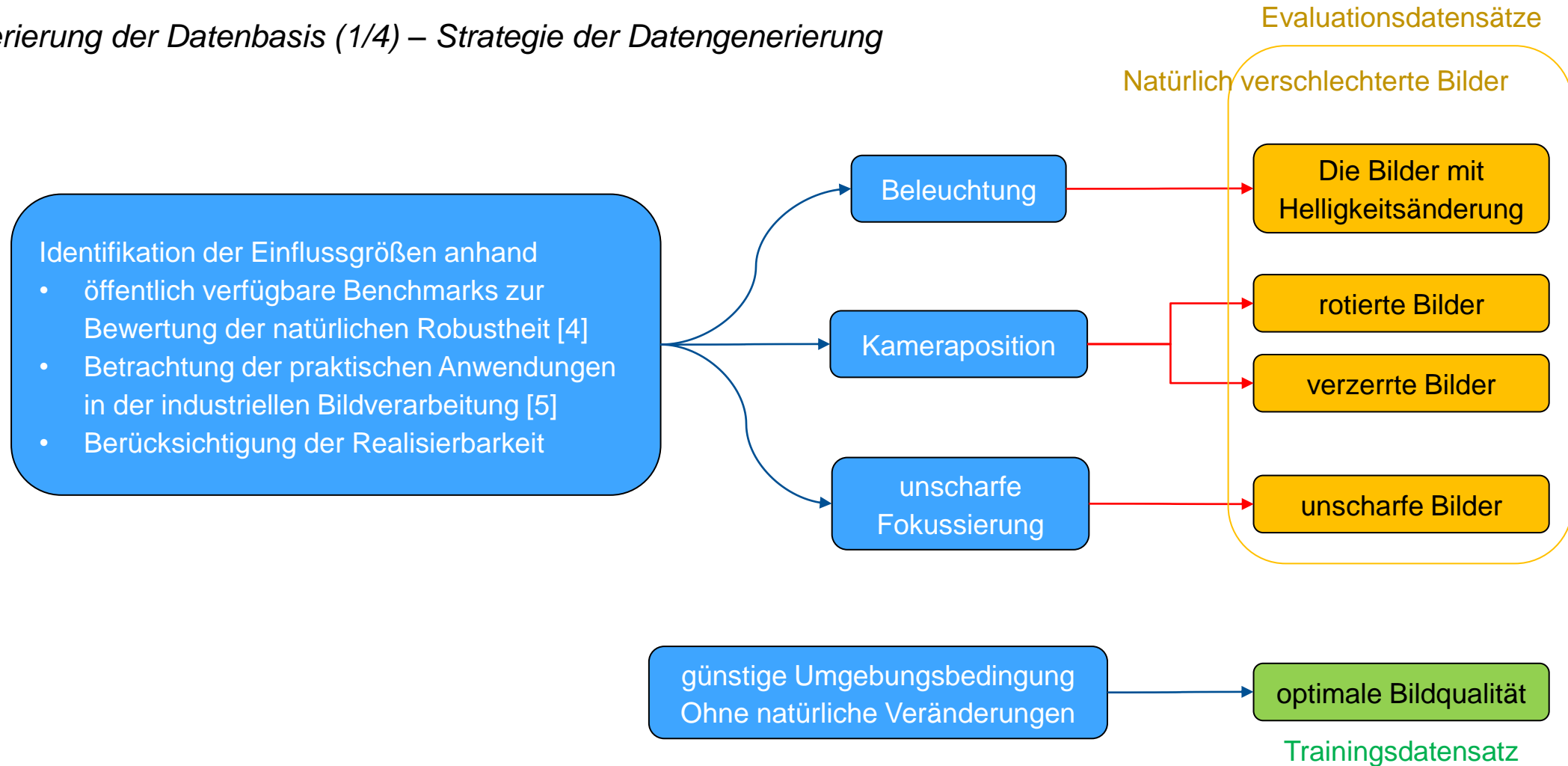


## Generierung der Datenbasis

---

Die untersuchten Einflussgrößen umfassen Beleuchtungsänderung, variierte Kameraposition und unscharfe Fokussierung.

*Generierung der Datenbasis (1/4) – Strategie der Datengenerierung*



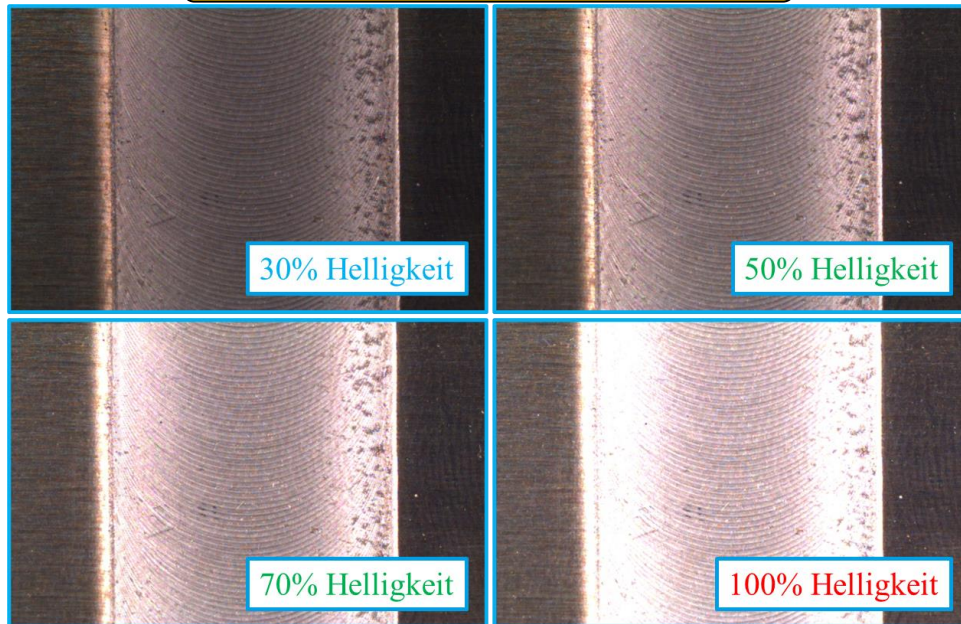


Die Bilddaten mit natürlichen Veränderungen werden dargestellt.

## Generierung der Datenbasis (1/4) – Strategie der Datengenerierung

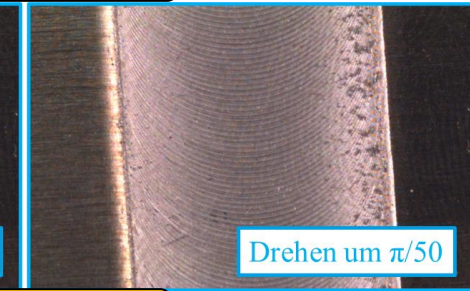
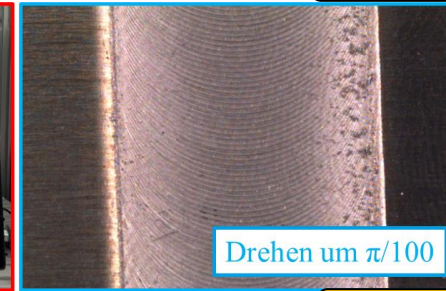
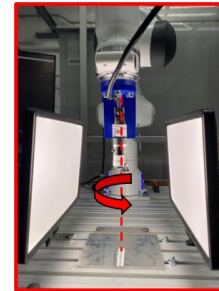
Beispielbilder, die durch verschiedene einzelne Einflüsse verändert werden

Die Bilder mit Helligkeitsänderung

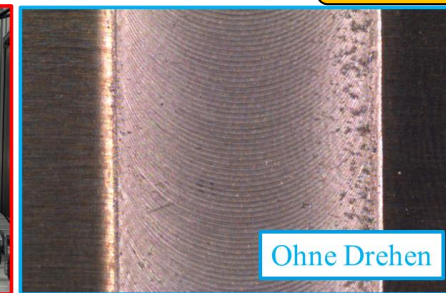
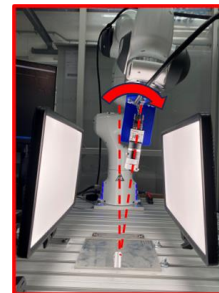


17 Szenarien:  
Helligkeitsstufe + Rotation, Verzerrung oder Unschärfe

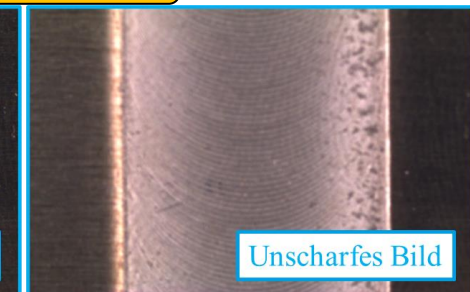
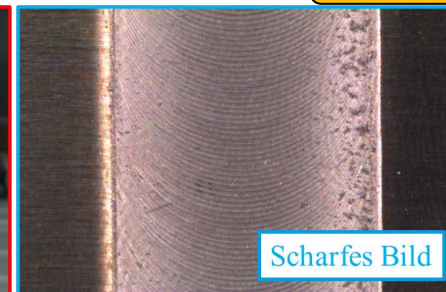
rotierte Bilder



verzerrte Bilder



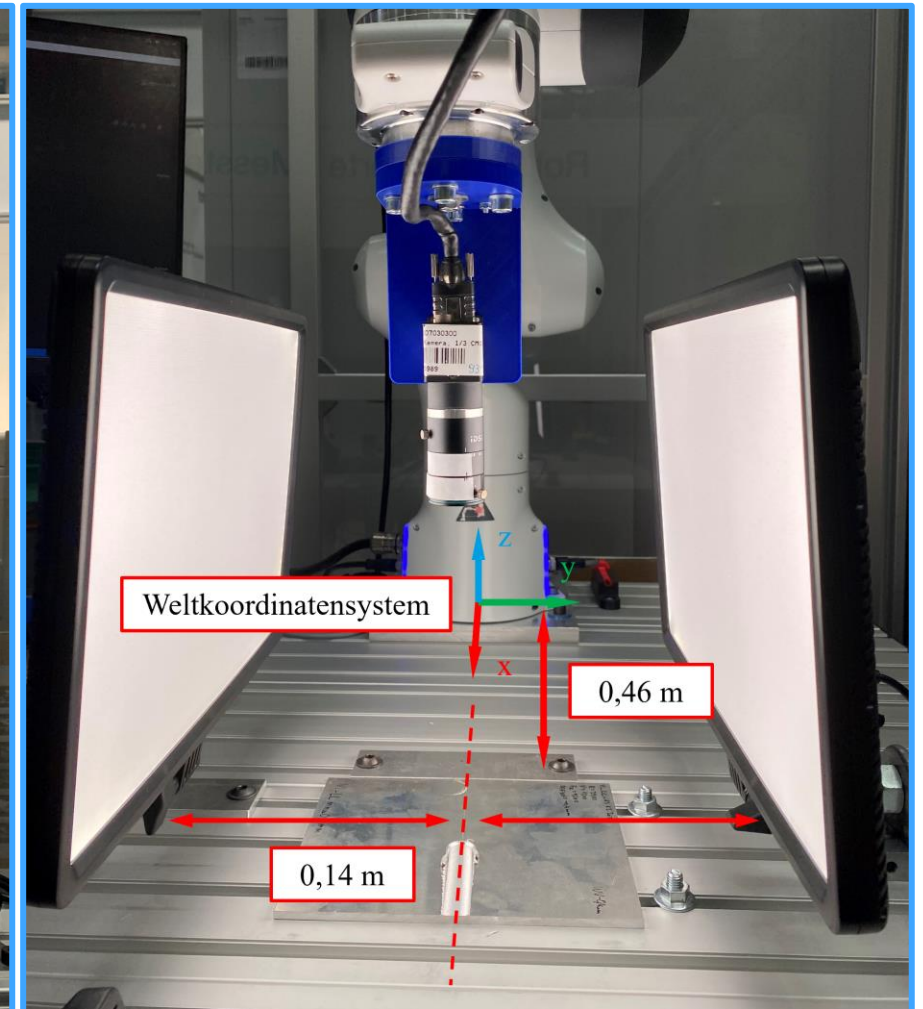
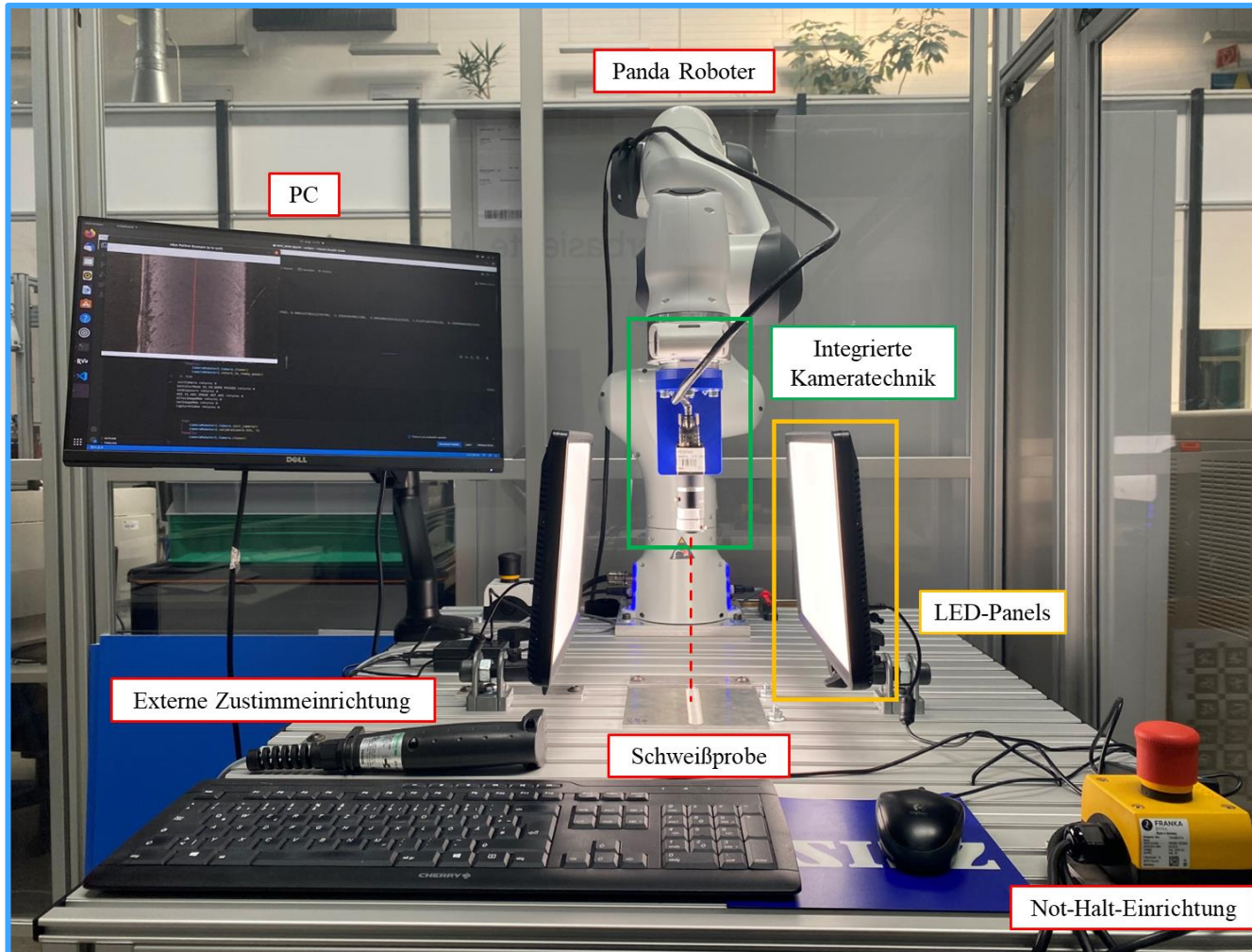
unscharfe Bilder





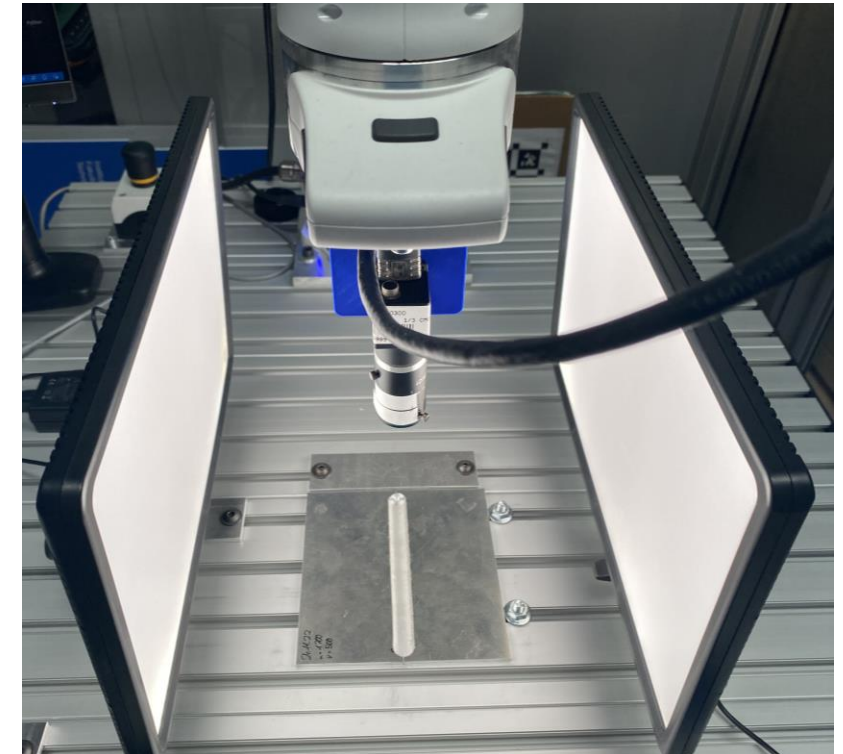
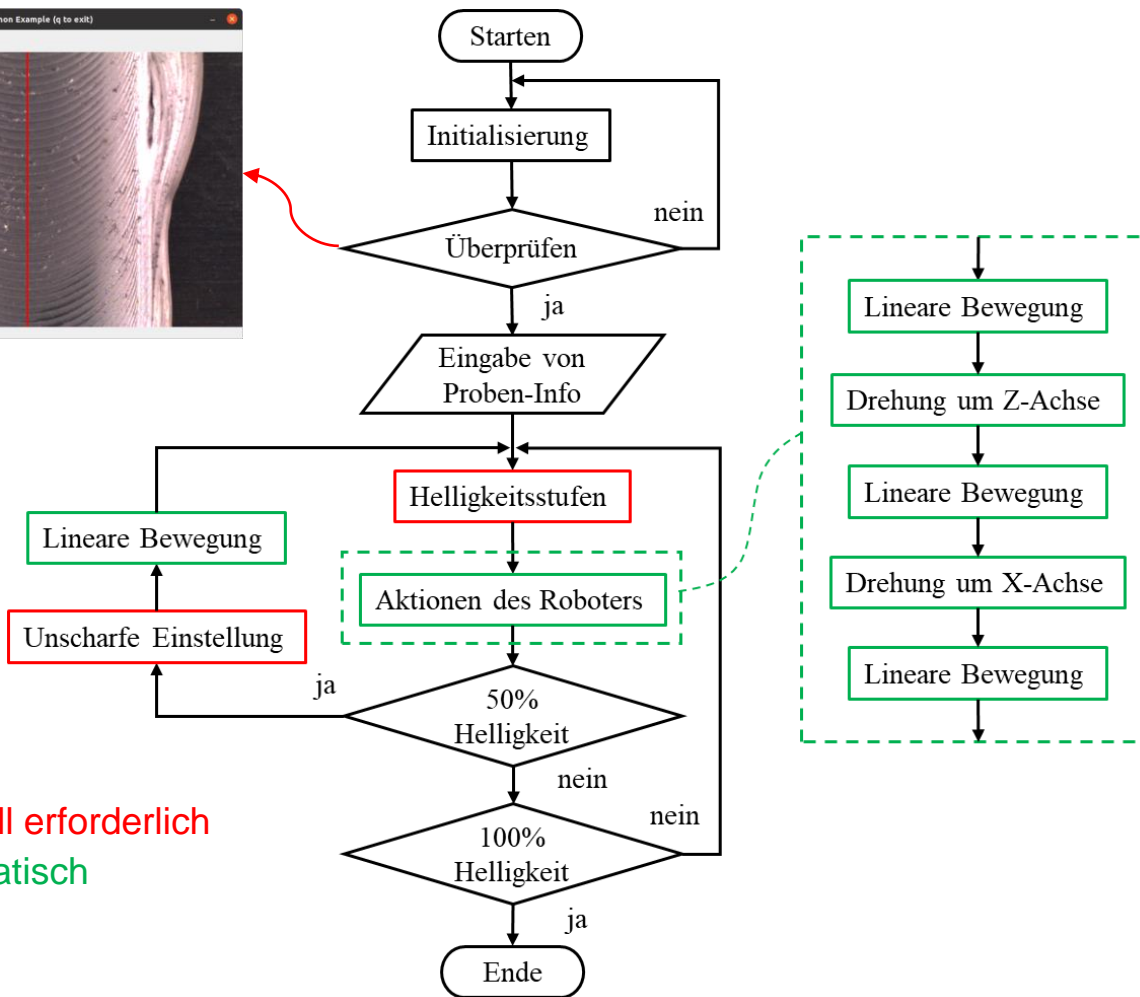
Ein roboter-basierter Messstand wurde aufgebaut, um Trainings- und Evaluationsdaten zu generieren.

*Generierung der Datenbasis (2/4) – Aufbau des Roboter-basierten Messstands*



Ein halbautomatisches Messkonzept wurde entworfen, um Bilddaten effizient aufzunehmen.

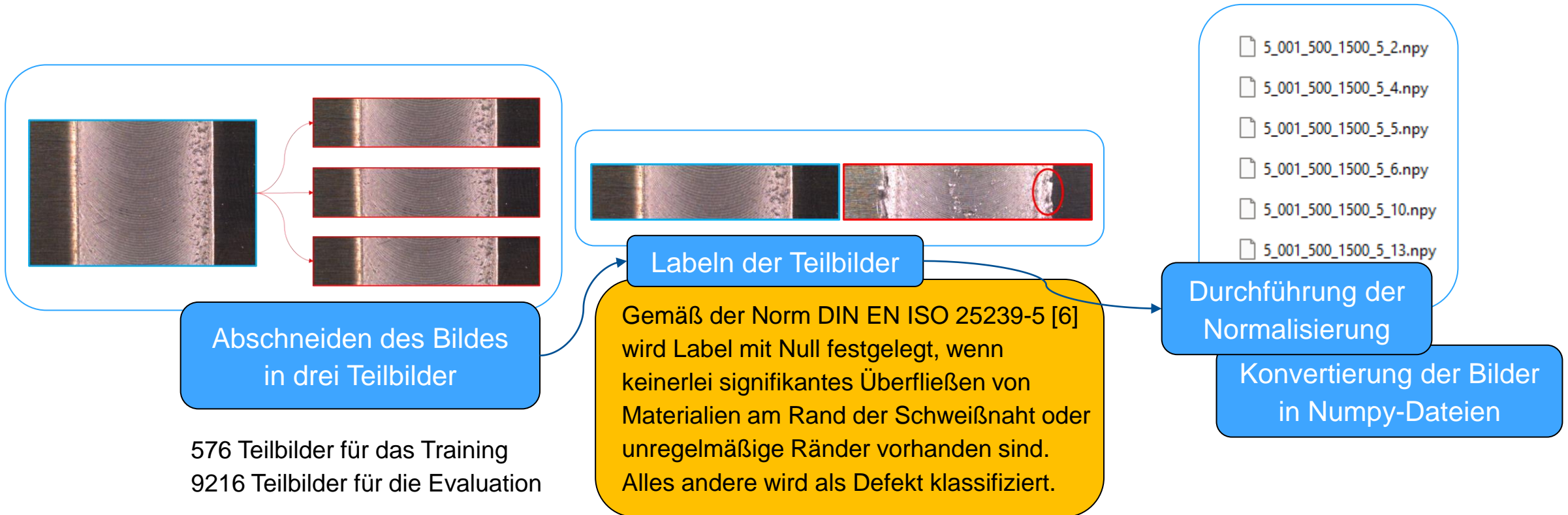
### Generierung der Datenbasis (3/4) – Umsetzung der Software zur Datengenerierung



192 Bilder mit optimaler Bildqualität  
3072 Bilder mit natürlichen Veränderungen  
(16 Szenarien/Evaluationsdatensätze)

# Die Datenaufbereitung für das Training wurde durchgeführt.

## Generierung der Datenbasis (4/4) – Datenaufbereitung



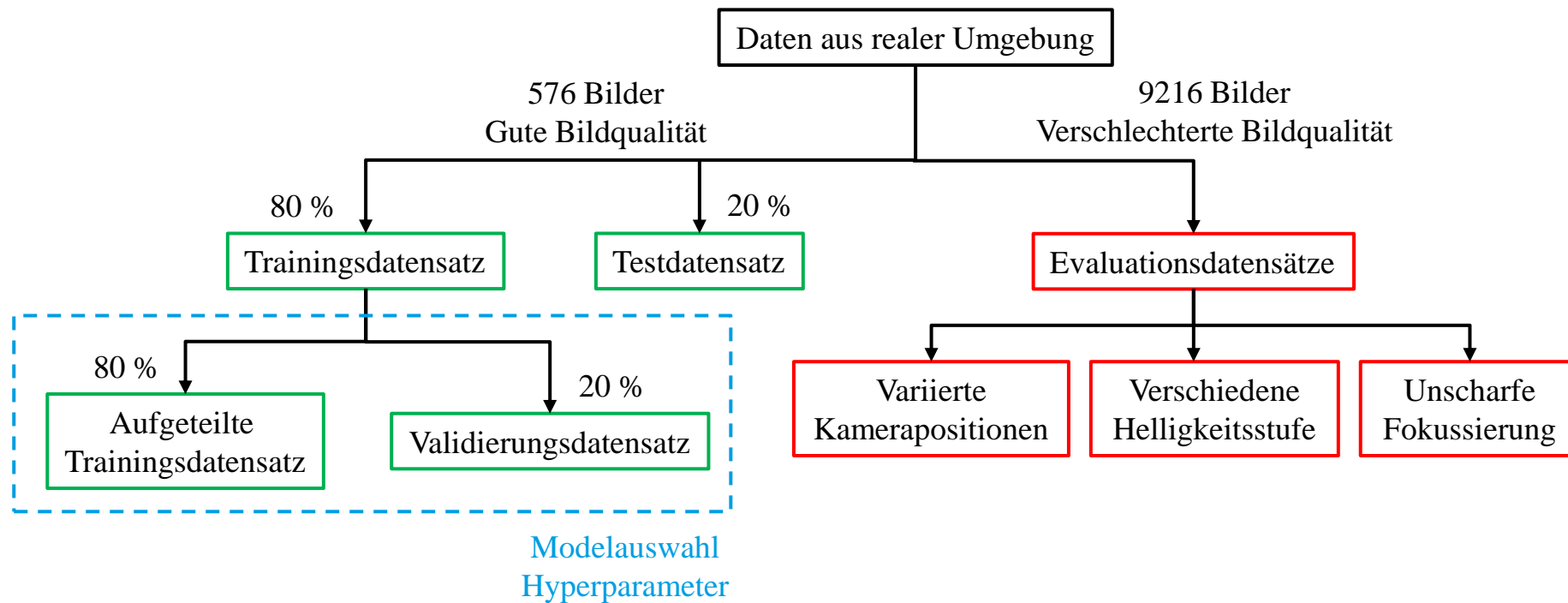


## Bewertung der Robustheit

---

Die gesamten Bilddaten werden in drei Datensätze unterteilt: Trainings-, Test- und Evaluationsdatensatz.

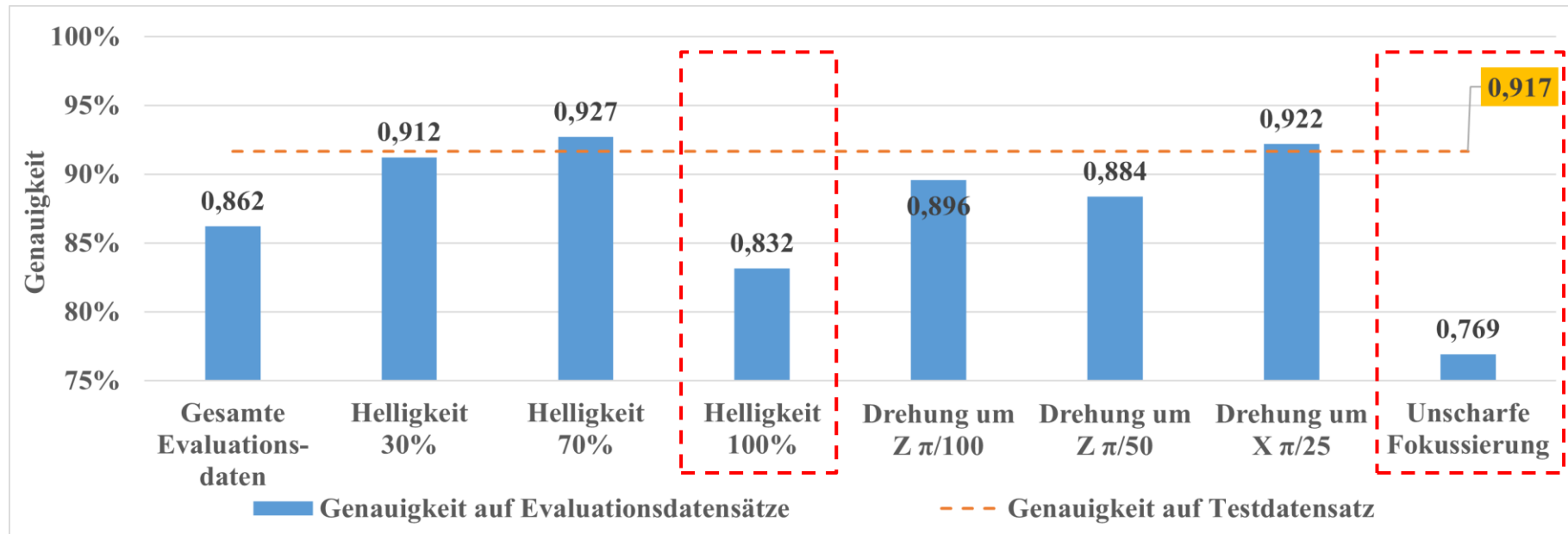
*Bewertung der Robustheit (1/3) – Bewertungsmethodologie*



Ein signifikanter Leistungsabfall des Referenzmodells auf dem Evaluationsdatensatz von 100% Helligkeit und Unscharfer Fokussierung ist zu beobachten.

### Bewertung der Robustheit (2/3) – Evaluationsergebnisse

Evaluationsergebnisse von Referenzmodell DenseNet-161





Die Vergleichsmodelle werden unterschiedlichen Parametern der Data-Augmentation verwendet.

*Bewertung der Robustheit (3/3) – Evaluationsergebnisse*

Leichte Data-Augmentation  
(DA-1)

Farbtonänderung von 0,15  
Sättigungsänderung von 0,15  
Helligkeitsänderung von 0,25  
Drehwinkel bis zu 5 Grad  
Gaußsche Unschärfe mit einem 7x7 Filter und einem Sigma-Wert von 0,1 bis 0,5

Stärkere Gaußsche Unschärfe  
(DA-2)

...  
Gaußsche Unschärfe mit einem 7x7 Filter und einem Sigma-Wert von 0,1 bis 2

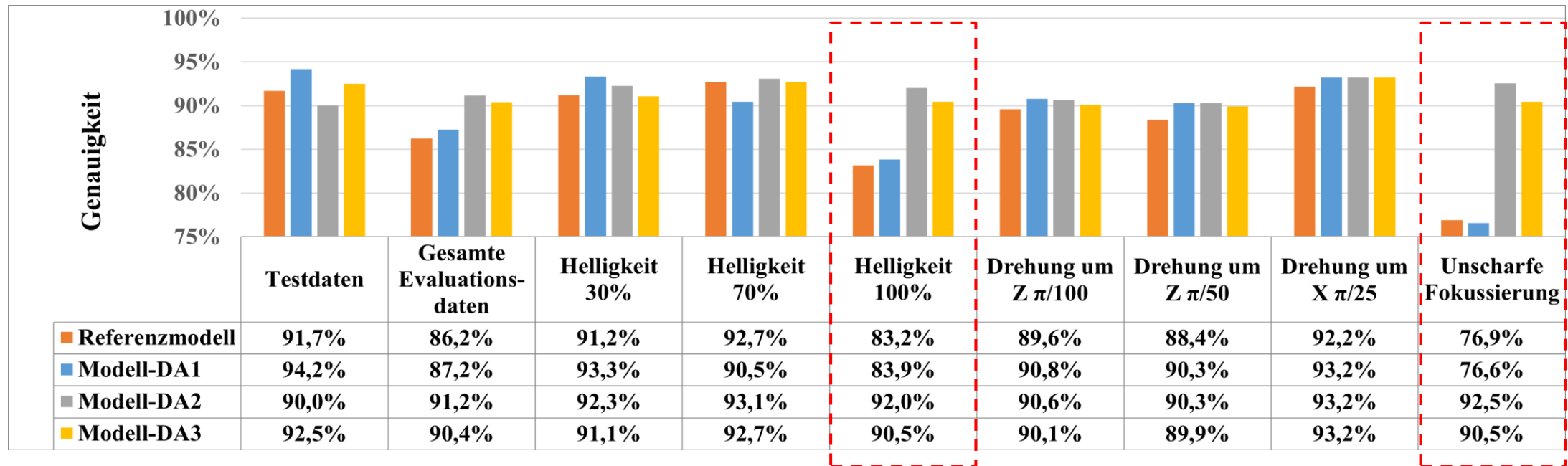
Stärkere Gaußsche Unschärfe  
Stärkere Helligkeitsänderung  
(DA-3)

...  
Helligkeitsänderung von 0,5  
Gaußsche Unschärfe mit einem 7x7 Filter und einem Sigma-Wert von 0,1 bis 2

Mittels ausgeprägter Data-Augmentation kann die Robustheit der Modelle deutlich verbessert werden.

### Bewertung der Robustheit (3/3) – Evaluationsergebnisse

#### Evaluationsergebnisse von DenseNet-161 mit Data-Augmentation



## Zusammenfassung und Ausblick

---



DenseNet-161 mit geeigneter Data-Augmentation zeigt eine gute Robustheit gegenüber den betrachteten Einflüssen. Weitere Einflüsse können weiter erforscht werden.

### *Zusammenfassung und Ausblick*

#### Beitrag der Arbeit

- Entwicklung des Halbautomatischen Messkonzepts zur effizienten Datengenerierung
- Bewertung der Robustheit auf Evaluationsdatensätze

#### Schlüsselergebnisse

- DenseNet-161 zeigt bereits eine gute Robustheit gegenüber leichten Helligkeitsänderungen und variierten Kamerapositionen.
- Durch die Anwendung geeigneter Data-Augmentation kann die Robustheit des Modells signifikant gesteigert werden.

#### Weitere Forschungsbedarf

- Betrachtung weiterer Einflussgrößen wie bspw.. Rauschen, Farbabgleich und Kontrast
- Die Erforschung zusätzlicher Parameter in der Data-Augmentation kann sich positiv auf das Training auswirken und die Robustheit erhöhen

Vielen Dank für Ihre Aufmerksamkeit

---

- [1] HARTL, R., (2021). Monitoring and optimizing the surface quality of friction stir welds using machine learning: Dissertation. Bd. 372. Forschungsberichte IWB. utzverlag GmbH. ISBN: 9783831649495.
- [2] HARTL, R., LANDGRAF, J., SPAHL, J., BACHMANN, A. und ZAEH, M. F., (2019). „Automated visual inspection of friction stir welds: a deep learning approach“. In: Multimodal Sensing: Technologies and Applications. Hrsg. von STELLA, E. Proceedings of SPIE. Bellingham, Washington, USA: SPIE, S. 8. ISBN: 9781510627970. DOI: 10.1117/12.2525947.
- [3] DRENKOW, N., SANI, N., SHPITSER, I. und UNBERATH, M., (2022). A Systematic Review of Robustness in Deep Learning for Computer Vision: Mind the gap?
- [4] GOJIĆ, G., VINCAN, V., KUNDAČINA, O., MIŠKOVIĆ, D. und DRAGAN, D., (2023). Non-adversarial Robustness of Deep Learning Methods for Computer Vision.
- [5] STEGER, C., ULRICH, M. und WIEDEMANN, C., Hrsg., (2018). Machine vision algorithms and applications. Second completely revised and enlarged edition. Weinheim, Germany: Wiley-VCH. ISBN: 9783527812905.
- [6] BEUTH VERLAG GMBH, (2020). DIN EN ISO 25239-5:2020-12, Rührreibschweißen\_- Aluminium\_- Teil\_5: Qualitätsund Prüfungsanforderungen (ISO\_25239-5:2020); Deutsche Fassung EN\_ISO\_25239-5:2020. Berlin. DOI: 10.31030/3161367.
- [7] LAUGROS, A., CAPLIER, A. und OSPICI, M., (2019). Are Adversarial Robustness and Common Perturbation Robustness Independent Attributes?

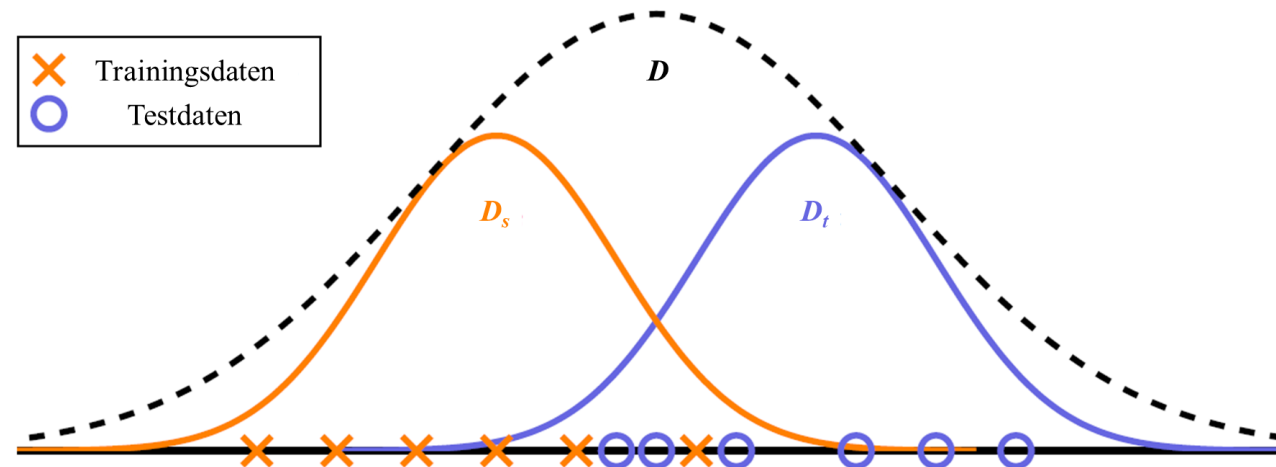


# Anhang

---

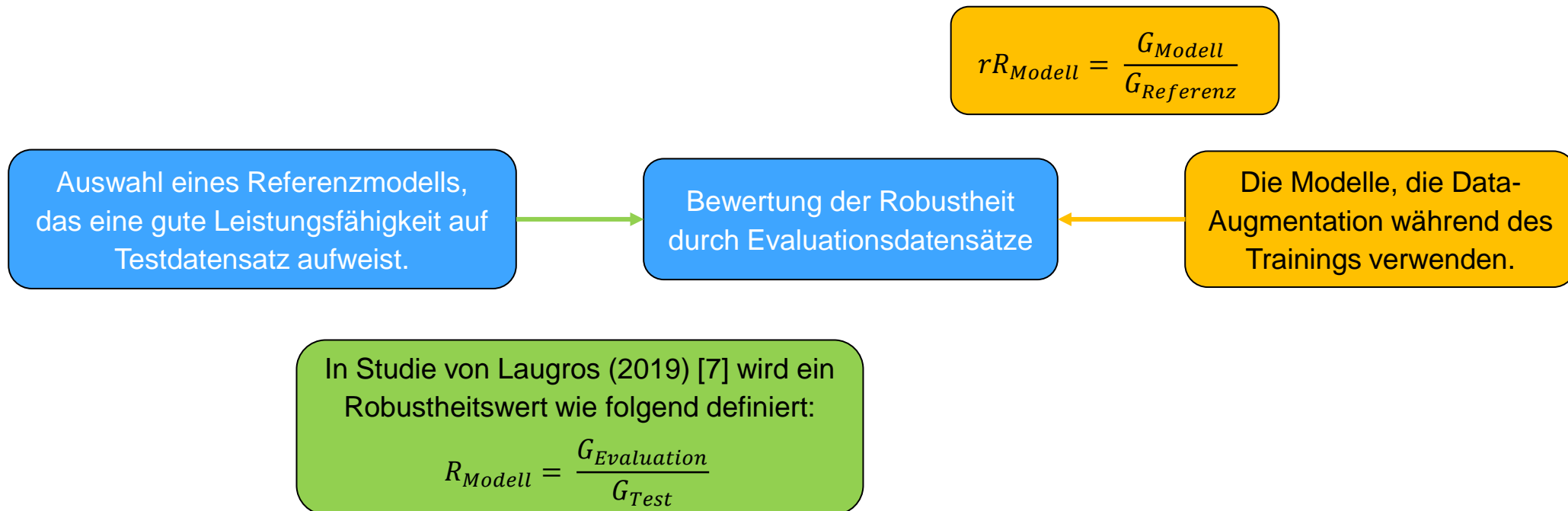


## Generierung der Datenbasis (1/4) – Strategie der Datengenerierung



Die Robustheit der Modelle wird durch einen Robustheitswert bewertet.

*Bewertung der Robustheit (1/3) – Bewertungsmethodologie*

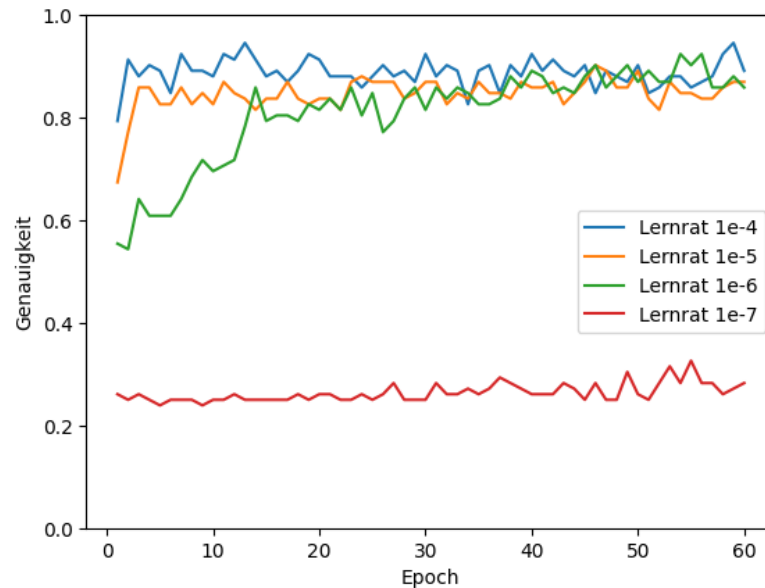


- 1) R: Robustheitswert
- 2) rR: relativer Robustheitswert
- 3) G: Klassifizierungsgenauigkeit

Im Vergleich zu anderen betrachteten Modellen weist DenseNet-161 eine höhere Leistungsfähigkeit auf.

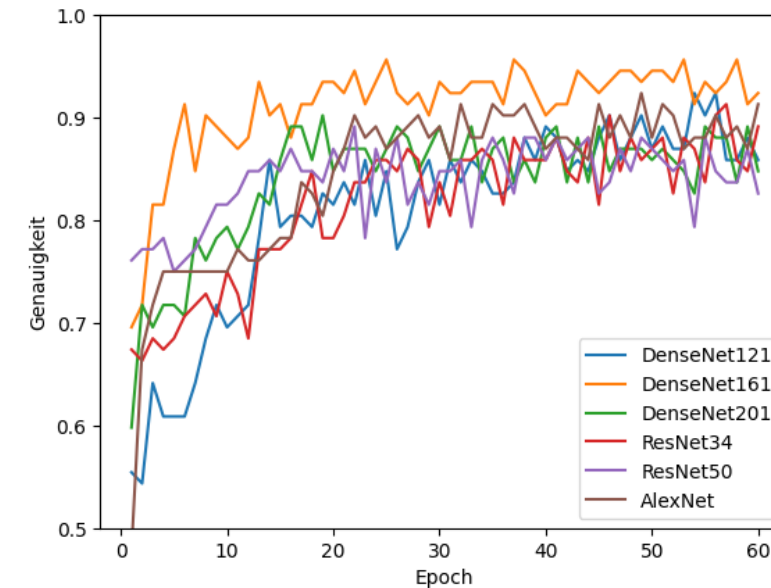
### Bewertung der Robustheit (2/3) – Training der CNN-Modelle

Darstellung der Lernkurven von DenseNet-121



Lernratenwerts von 1e-6

Darstellung der Lernkurve unterschiedlicher Architekturen



DenseNet-161

DenseNet-161:  
Lernratenwerts von 1e-6  
Epoche von 30

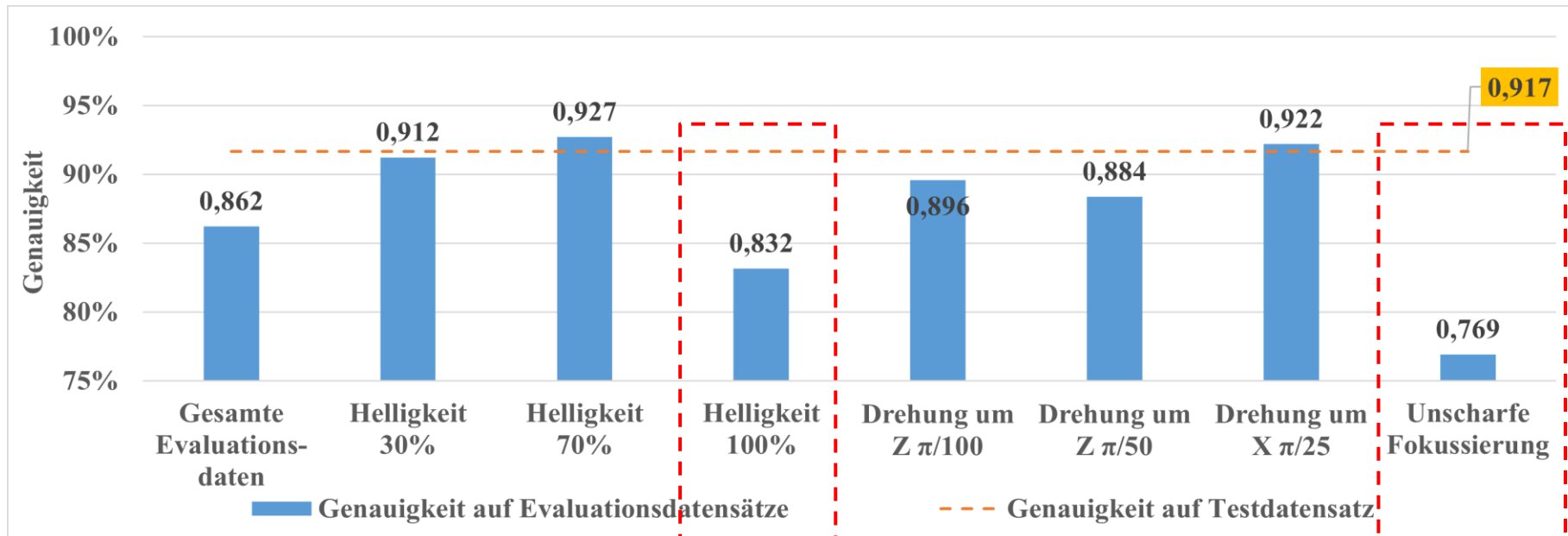
Testgenauigkeit von 91,7%



Ein signifikanter Leistungsabfall des Referenzmodells auf dem Evaluationsdatensatz von 100% Helligkeit und Unscharfer Fokussierung ist zu beobachten.

### Bewertung der Robustheit (3/3) – Evaluationsergebnisse

Evaluationsergebnisse von Referenzmodell DenseNet-161

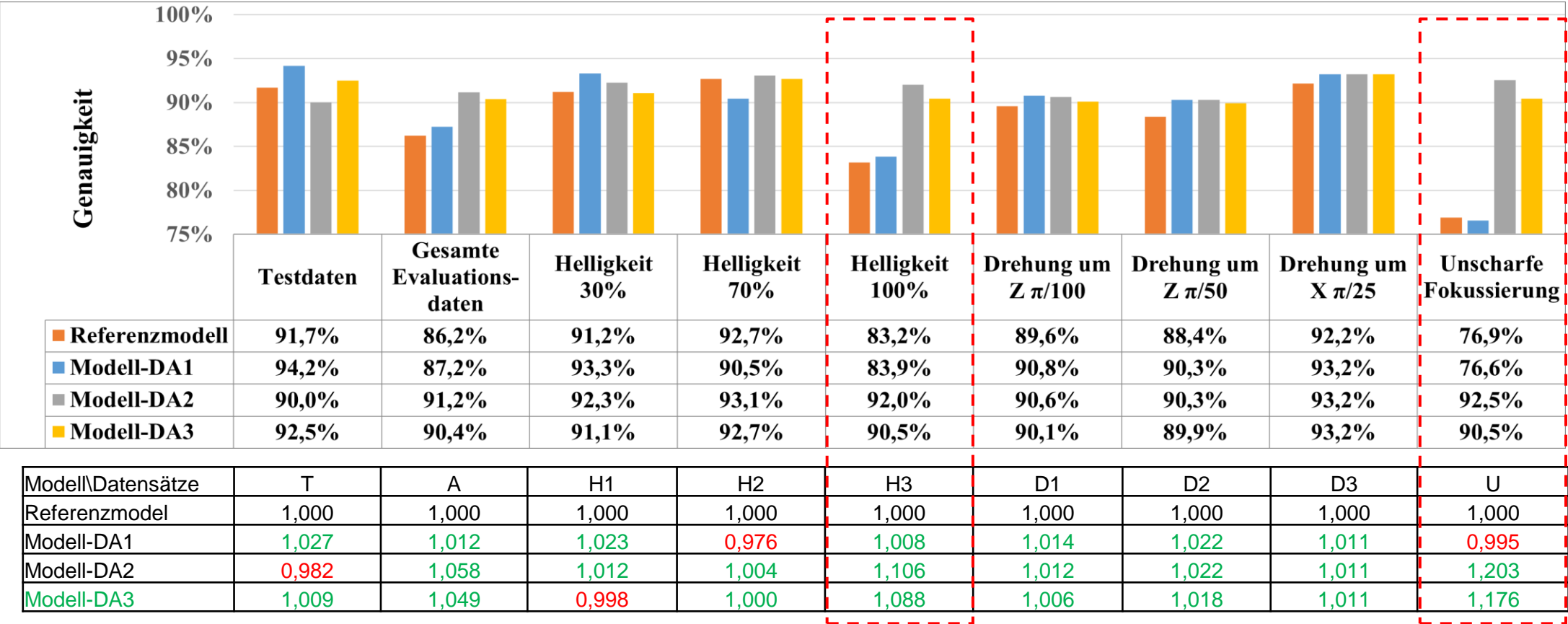


Datensätze	A	H1	H2	H3	D1	D2	D3	U
Robustheitswert	0,941	0,995	1,011	0,907	0,977	0,964	1,006	0,839

Mittels ausgeprägter Data-Augmentation kann die Robustheit der Modelle deutlich verbessert werden.

Bewertung der Robustheit (3/3) – Evaluationsergebnisse

Evaluationsergebnisse von DenseNet-161 mit Data-Augmentation



- 1) R: Robustheitswert
- 2) rR: relativer Robustheitswert
- 3) G: Klassifizierungsgenauigkeit

DenseNet-161 zeigt im Vergleich zu Modellen mit anderen Architekturen und Netztiefen eine höhere Robustheit.

*Bewertung der Robustheit (3/3) – Evaluationsergebnisse*

Weitere Betrachtungen

Modell\Datensätze	T	A	H1	H2	H3	D1	D2	D3	U
DenseNet-161	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
ResNet-50	0,982	0,971	0,996	0,942	0,868	1,016	1,000	0,996	0,935
AlexNet	0,973	1,006	0,969	0,979	1,044	0,992	0,994	0,979	0,982

Modell\Datensätze	T	A	H1	H2	H3	D1	D2	D3	U
DenseNet-161	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
DenseNet-121	0,964	0,996	1,006	0,970	0,994	0,983	1,000	0,977	0,953
DenseNet-201	0,982	0,981	1,004	0,974	0,902	1,012	1,006	0,998	0,946





## Shaoxiang Tan

---

Institut für Werkzeugmaschinen und Betriebswissenschaften  
Technische Universität München  
Boltzmannstraße 15, 85748 Garching



+49 89 289



+49 89 289 15555

