

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

Evaluation of the Robustness of Deep Learning Methods in Quality Assurance Using the Example of Friction Stir Welding

## Forschungspraktikum

an der TUM School of Engineering and Design der Technischen Universität München

|                        |  |
|------------------------|--|
| <b>Betreut von</b>     | Prof. Dr.-Ing. Rüdiger Daub  |
| <b>Eingereicht von</b> | Lehrstuhl für Produktionstechnik und Energiespeichersysteme<br>Shaoxiang Tan |
| <b>Eingereicht am</b>  | Clemensstraße 118, 80796 München<br>30. September 2023 in Garching           |



# Aufgabenstellung

## **Titel des Forschungspraktikums:**

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

## **(English Title of Practical Research Course:)**

Evaluation of the Robustness of Deep Learning Methods in Quality Assurance Using the Example of Friction Stir Welding

**Verfasser:in:** Shaoxiang Tan

**Ausgabe:** 01.04.2023

**Betreuer:in:** Johannes Bauer

**Abgabe:** 30.09.2023

## **Ausgangssituation:**

Technologische Entwicklungen im Bereich des maschinellen Lernens in den letzten Jahren eröffnen zahlreiche Anwendungsmöglichkeiten für verschiedene Branchen. So ermöglichen es sogenannte Deep-Learning-Verfahren, basierend auf der Anwendung tiefer neuronaler Netze, auch hochdimensionale Daten, wie z. B. Bilder, effektiv zu verarbeiten. Im Produktionskontext können solche neuronalen Netze zur Qualitätsüberwachung von Rührreibschweißprozessen eingesetzt werden, indem Bilder der erzeugten Schweißnähte hinsichtlich ihrer Qualität klassifiziert werden. Das Potenzial der Verfahren dafür konnte bereits in wissenschaftlicher Literatur gezeigt werden. Dabei basiert die Leistungsfähigkeit der verwendeten Modelle wesentlich auf der Verfügbarkeit einer umfassenden Trainingsdatenbasis. Kommt es im praktischen Einsatz zu Änderungen am Prozess, am Produkt oder der Umgebung, z. B. aufgrund variierter Lichtverhältnisse oder auftretenden Störgrößen, können diese die Leistungsfähigkeit des Modells beeinträchtigen. Die Robustheit von Deep-Learning-Verfahren gegenüber solchen Einflüssen wurde bisher jedoch nicht hinreichend untersucht.

## **Zielsetzung:**

Ziel dieser Arbeit ist die systematische Untersuchung der Robustheit von Deep-Learning-Verfahren bei der Qualitätsüberwachung von Rührreibschweißprozessen. Hierfür sollen zunächst potenzielle Einflussgrößen identifiziert werden. Anschließend ist ein Versuchsstand aufzubauen, der die systematische Variation einzelner Größen ermöglicht. Die Reproduzierbarkeit sowie die Effizienz bei der Generierung der benötigten Bilddaten stellen zentrale Anforderungen dar. Deswegen wird auf ein Roboter-basiertes System, auf der Basis des Frameworks ROS (Robot Operating System), zurückgegriffen. Abschließend sollen geeignete Deep-Learning-Modelle anhand der generierten Bilddaten trainiert und ihre Robustheit gegenüber verschiedenen Einflussgrößen evaluiert werden.

**Vorgehensweise und Arbeitsmethodik:**

- Einarbeitung in folgende Themengebiete
  - Grundlagen der Robotik und Kameratechnik
  - Roboterprogrammierung mit ROS
  - Deep-Learning Verfahren, insbesondere Convolutional Neural Networks
  - Qualitätsüberwachung beim Rührreibschweißen
- Identifikation möglicher Einflussgrößen
- Generierung der Bilddatenbasis
  - Aufbau des Versuchsstands
  - Durchführung der Datenaufnahme
  - Aufbereitung der Bilddaten
- Evaluation der Robustheit
  - Implementierung eines Deep-Learning-basierten Klassifizierungsmodells
  - Evaluation der Robustheit des Modells gegenüber Einflussgrößen anhand der generierten Bilddatenbasis

**Vereinbarung:**

Durch die Betreuung von Herrn B. Sc. Shaoxiang Tan fließt geistiges Eigentum des *iwb* in diese Arbeit ein. Eine Veröffentlichung der Arbeit oder eine Weitergabe an Dritte bedarf der Genehmigung durch den Lehrstuhlinhaber. Der Archivierung der Arbeit in der *iwb*-eigenen und nur für Mitarbeitende des *iwb* zugänglichen Bibliothek als Bestand und in der digitalen Studienarbeitsdatenbank des *iwb* als PDF-Dokument stimme ich zu.

Garching, den 15.03.2023

Prof. Dr.-Ing.  
Rüdiger Daub

B.Sc.  
Shaoxiang Tan

## **Abstract**

The use of deep learning methods demonstrates strong performance in quality monitoring of Friction Stir Welding (FSW) joints. However, the model's performance is often tested solely on independent, identically distributed datasets. What happens when natural variations or influences occur? This study investigates the robustness of deep learning methods in quality assurance of FSW welds. To do this, a robot-based measurement setup was established, and an efficient semi-automatic data generation concept was developed. In total, 576 optimal images without natural influences were collected for training and testing, and 9207 naturally degraded images were gathered to assess robustness. The results of this study indicate that DenseNet-161 exhibits higher robustness compared to other considered architectures like ResNet-50. By applying suitable data augmentation techniques, the model, trained with only 456 image data, achieved significantly increased robustness. A strong overall performance of 91.2 % was achieved for the entire set of 9207 images, considering 16 different natural degradations.

## **Zusammenfassung**

Der Einsatz von Deep-Learning-Verfahren zeigt eine gute Leistungsfähigkeit in der Qualitätsüberwachung der FSW-Schweißverbindung. Allerdings wird die Leistungsfähigkeit des Modells häufig lediglich an unabhängigen, identisch verteilten Datensätzen getestet. Was passiert, wenn es zu natürlichen Veränderungen oder Einflüssen kommt? In dieser Arbeit wird die Robustheit von Deep-Learning-Verfahren in der Qualitätssicherung von FSW untersucht. Dafür wurde ein roboter-basierter Messstand aufgebaut und ein und ein effizientes, halbautomatisches Konzept zur Datengenerierung entwickelt. Insgesamt wurden 576 optimale Bilder ohne natürliche Einflüsse für das Training und Testen sowie 9207 natürliche verschlechterte Bilder zur Bewertung der Robustheit gesammelt. Die Ergebnisse dieser Arbeit zeigen, dass DenseNet-161 im Vergleich zu anderen betrachteten Architekturen wie ResNet-50 eine höhere Robustheit aufweist. Durch die Anwendung geeigneter Data-Augmentation-Techniken wird die Robustheit des Modells, das lediglich mit 456 Bilddaten trainiert wurde, signifikant erhöht. Es wurde eine gute Leistungsfähigkeit von 91,2 % für alle 9207 Bilder erreicht, wobei 16 verschiedene natürliche Verschlechterungen berücksichtigt wurden.



# Inhaltsverzeichnis

|  |           |
|--|-----------|
| <b>Abkürzungsverzeichnis</b>                                   | <b>ix</b> |
| <b>Abbildungsverzeichnis</b>                                   | <b>xi</b> |
| <b>Tabellenverzeichnis</b>                                     | <b>xv</b> |
| <b>1 Einleitung</b>  | <b>1</b>  |
| 1.1 Motivation . . . . .                                       | 1         |
| 1.2 Zielsetzung . . . . .                                      | 2         |
| 1.3 Aufbau der Arbeit . . . . .                                | 3         |
| <b>2 Grundlagen</b>  | <b>5</b>  |
| 2.1 Rührreibschweißen . . . . .                                | 5         |
| 2.1.1 Prozessablauf . . . . .                                  | 6         |
| 2.1.2 Charakteristik von Schweißnähten . . . . .               | 7         |
| 2.2 Bildverarbeitungssystem . . . . .                          | 9         |
| 2.2.1 Beleuchtung . . . . .                                    | 10        |
| 2.2.2 Kameratechnik . . . . .                                  | 12        |
| 2.3 Faltendes neuronales Netzwerk . . . . .                    | 15        |
| 2.3.1 Grundaufbau . . . . .                                    | 15        |
| 2.3.2 Trainingsverfahren . . . . .                             | 17        |
| 2.3.3 Methoden zur Modellselektion . . . . .                   | 21        |
| <b>3 Stand der Forschung und Technik</b>                       | <b>25</b> |
| <b>4 Generierung der Datenbasis</b>                            | <b>29</b> |
| 4.1 Strategie der Datengenerierung . . . . .                   | 29        |
| 4.2 Entwicklung des Roboter-basierten Messstands . . . . .     | 31        |
| 4.2.1 Übersicht des Panda-Roboters . . . . .                   | 31        |
| 4.2.2 Sensorauswahl . . . . .                                  | 33        |
| 4.2.3 Aufbau des Messstands . . . . .                          | 34        |
| 4.3 Umsetzung der Software zur Datengenerierung . . . . .      | 36        |
| 4.3.1 Softwareumgebung . . . . .                               | 36        |
| 4.3.2 Halbautomatisches Konzept der Datengenerierung . . . . . | 38        |
| 4.3.3 Datenaufbereitung . . . . .                              | 42        |
| <b>5 Bewertung der Robustheit</b>                              | <b>45</b> |
| 5.1 Vorgehen zur Bewertung der Robustheit . . . . .            | 45        |
| 5.2 Modellauswahl . . . . .                                    | 46        |
| 5.3 Evaluationsergebnisse . . . . .                            | 48        |
| 5.4 Diskussion . . . . .                                       | 51        |

|  |           |
|--|-----------|
| <b>6 Zusammenfassung und Ausblick</b>              | <b>53</b> |
| 6.1 Zusammenfassung . . . . .                      | 53        |
| 6.2 Ausblick . . . . .                             | 54        |
| <b>A Digitaler Anhang</b>                          | <b>55</b> |
| A.1 Technische Datenblatt der Kamera . . . . .     | 55        |
| A.2 Technische Datenblatt des Objektives . . . . . | 55        |
| A.3 Zeichnungen von Kamerahalterung . . . . .      | 55        |
| <b>Literaturverzeichnis</b>                        | <b>57</b> |

# Abkürzungsverzeichnis

|              |   |
|--------------|---|
| ROS          | Robot Operating System                                    |
| <i>iwb</i>   | Institut für Werkzeugmaschinen und Betriebswissenschaften |
| FSW          | Friction Stir Welding                                     |
| SVM          | Support Vector Machine                                    |
| DL-Verfahren | Deep-Learning-Verfahren                                   |
| DNN          | Deep Neural Network                                       |
| CNN          | Convolutional Neural Network                              |
| GW           | Grundwerkstoff  |
| WEZ          | Wärmeeinflusszone   |
| TMAZ         | Thermomechanically affected zone                          |
| PC           | Computer  |
| LEDs         | light-emitting diode                                      |
| BRDF         | bidirectional reflectance distribution function           |
| MOD          | minimale Objektdistanz                                    |
| FOV          | Field-of-View   |
| ANN          | Artificial neural network                                 |
| GD           | Gradient descent  |
| SGD          | stochastic Gradient Descent                               |
| RMSProp      | Root Mean Square Propagation                              |
| Adam         | Adaptive Moment Estimation                                |
| RIDE         | Integrierte Entwicklungsumgebung                          |
| FCI          | Franka Regelungsschnittstelle                             |
| IDE          | Integrated Development Environment                        |
| OOD          | Out-of-Distribution                                       |
| ID           | In-Distribution   |



# Abbildungsverzeichnis

|      |   |    |
|------|---|----|
| 2.1  | Prozessablauf beim FSW im Stumpfstoß in Anlehnung an (BACHMANN 2021, S. 11) . . . . .   | 6  |
| 2.2  | FSW-Nahtquerschnitt mit charakteristischen Bereichen in Anlehnung an (BACHMANN 2021, S. 15) . . . . .   | 7  |
| 2.3  | Charakteristische Merkmale einer FSW-Naht bei einer Stumpfstoßverbindung in Anlehnung an (RUHSTORFER 2012, S. 14) . . . . .   | 8  |
| 2.4  | Oberflächenqualitäten von Rührreibschweißen in Anlehnung an (HARTL 2021, S. 10) . . . . .   | 8  |
| 2.5  | Komponenten eines Bildverarbeitungssystems in Anlehnung an (STEGER et al. 2018) und (BEYERER et al. 2016, S. 9) . . . . .   | 10 |
| 2.6  | Aufnahmen des Schlagbolzeneindrucks auf einer Patronenhülse zum Zwecke der automatischen Identifikation von Schusswaffen: (a)–(c) gerichtete Beleuchtungen; (d) diffuse Beleuchtung, in Anlehnung an (BEYERER et al. 2016, S. 301) .                                | 11 |
| 2.7  | (a) Prinzip der Hellfeldbeleuchtung; (b) Prinzip der Dunkelfeldbeleuchtung, in Anlehnung an (BEYERER et al. 2016, S. 302, 303) . . . . .  | 12 |
| 2.8  | Standardanschluss für industrielle Objektive in Anlehnung an (STEGER et al. 2018) . . . . .   | 13 |
| 2.9  | Abbildungsgesetze für dicke Linsen in Anlehnung an (BEYERER et al. 2016, S. 136). Die Ausdehnung des abgebildeten Strahlenbündels ist rot dargestellt. Blau eingefärbt sind die ausgezeichneten Strahlen. (1: Mittelpunktsstrahl, 2, 3: Parallelstrahlen) . . . . . | 14 |
| 2.10 | Vergleich von Schärfentiefe in Anlegung an (STEGER et al. 2018). ZK:Zerstreuungskreis (a) Zerstreuungskreis ohne Blende; (b) Zerstreuungskreis mit Blende. . . . .  | 14 |
| 2.11 | Darstellung des verallgemeinerten linearen Modells als einfaches neuronales Netz . . . . .  | 16 |
| 2.12 | Eine schematische Darstellung eines mehrschichtigen Feedforward-Netzwerks mit zwei verborgenen Schichten. . . . .   | 16 |
| 2.13 | Darstellung der Beispielarchitektur eines CNN . . . . .   | 17 |
| 2.14 | Von links nach rechts: Die Auswirkung der Erhöhung des Lernrates in Anlegung an (JULIJA ZAVADLAV 2022) . . . . .  | 18 |
| 2.15 | Darstellung der Konvergenz von Batch-, Online- und Mini-Batch-Algorithmen in Anlegung an (JULIJA ZAVADLAV 2022) . . . . .   | 18 |
| 2.16 | Lokale Minima (links) und Sattelpunkte (rechts) . . . . .   | 19 |

|   |    |
|---|----|
| 2.17 Die Anpassungen von einer Sinus-Funktion in Anlegung an (JULIJA ZAVADLAV 2022) . . . . .   | 21 |
| 2.18 Darstellung für K-Fold-Kreuzvalidierung in Anlegung an (JULIJA ZAVADLAV 2022). . . . .   | 22 |
| 3.1 Methodisches Flussdiagramm für die Bildverarbeitung in Anlegung an (SINHA et al. 2008, S. 485). . . . .   | 25 |
| 3.2 Kombinierte Detektion und Klassifikation angewendet auf eine Schweißnaht in Anlegung an (HARTL et al. 2019). . . . .  | 26 |
| 4.1 Veranschaulichung von Verteilungsverschiebung in der Anlegung von (GOJIĆ et al. 2023, S. 1). Eine im Voraus unbekannte reale Datenverteilung $D$ (dargestellt durch gestrichelte Linien) mit der Quellverteilung $D_s$ und der Zielverteilung $D_t$ (dargestellt durch orange und blaue durchgezogene Linien), die im Vergleich zu $D_s$ verschoben sind. . . . . | 30 |
| 4.2 Schematische Darstellung der Komponente des Robotersystems in Anlegung an (FRANKA EMIKA GMBH 2021, S. 104). . . . .   | 32 |
| 4.3 Achsbezeichnungen mit Achsausrichtung ohne Endeffektor und mit Franka-Hand in der Anlegung von (FRANKA EMIKA GMBH 2021, S. 35) und (FRANKA EMIKA GMBH 2023b, S. 55). . . . .  | 32 |
| 4.4 Darstellung von Sichtfeld der Schweißnaht. . . . .  | 33 |
| 4.5 Verbindung zwischen Roboter und Kamera. . . . .   | 34 |
| 4.6 Aufbau des Messstandes. . . . .   | 35 |
| 4.7 Positionierung der Schweißprobe und Anordnung der Beleuchtung in Bezug auf das Weltkoordinatensystem. . . . .   | 36 |
| 4.8 Grafische Darstellung der Komponenten der Software. . . . .   | 38 |
| 4.9 Schematische Darstellung der optischen Erfassung einer Schweißprobe während der linearen Bewegung des Roboters. . . . .   | 39 |
| 4.10 Darstellung der Sichtfeldabmessungen in einem Bild der Schweißprobe. . . . .   | 39 |
| 4.11 Darstellung von Schweißprobenbildern bei verschiedenen Helligkeitsstufen. . . . .  | 40 |
| 4.12 Darstellung von Schweißprobenbildern bei verschiedenen Drehstufen. . . . .   | 40 |
| 4.13 Darstellung von Bildern unter Verwendung unscharfer Fokussierung. Grüner Pfeil zeigt scharfe Stelle und Roter Pfeil zeigt unscharfe Stelle. . . . .  | 41 |
| 4.14 Flussdiagramm zur Programmierlogik des halbautomatischen Konzepts zur effizienten Datengenerierung. Der Prozess in grüner Farbe wird automatisch durch den Roboter durchgeführt, während der Prozess in roter Farbe manuell ausgeführt wird. . . . .   | 41 |
| 4.15 Darstellung des Bildes während des Überprüfungsprozesses. . . . .  | 42 |
| 4.16 Darstellung der Teilbilder mit entsprechender Bildgröße. . . . .   | 43 |
| 4.17 Darstellung der Gratbildung von zwei Schweißnähten. Das Bild ist in blauer Farbe mit dem Label von Null und in roter Farbe mit dem Label von Eins gekennzeichnet. . . . .  | 43 |

|  |    |
|--|----|
| 4.18 Schematische Darstellung der verwendeten Datensätze für Training und Evaluation. . . . .  | 43 |
| 5.1 Lernverlauf von DenseNet-121 mit verschiedenen Lernraten. . . . .  | 47 |
| 5.2 Der Lernverlauf verschiedener Architekturen bei einer Lernrate von $10^{-6}$ . . . . .   | 47 |
| 5.3 Die durchschnittlichen Validierungsgenauigkeiten von drei verschiedenen Modellen mit entsprechenden Standardabweichungen in der K-Fold-Validierung. . . . .  | 48 |
| 5.4 Genauigkeiten des Referenzmodells DenseNet-161 für die Evaluationsdatensätze der einzelnen Einflussgrößen und alle Evaluationsdaten im Vergleich zur Testgenauigkeit. . . . .                            | 48 |
| 5.5 Genauigkeiten der Modelle für den Testdatensatz, alle Evaluationsdaten und die Evaluationsdatensätze der einzelnen Einflussgrößen im Vergleich zum Referenzmodell. . . . .                               | 50 |
| 5.6 Genauigkeiten der unterschiedlichen Modellarchitekturen für den Testdatensatz, alle Evaluationsdaten und die Evaluationsdatensätze der einzelnen Einflussgrößen im Vergleich zum Referenzmodell. . . . . | 50 |
| 5.7 Genauigkeiten der unterschiedlichen Netztiefen für den Testdatensatz, alle Evaluationsdaten und die Evaluationsdatensätze der einzelnen Einflussgrößen im Vergleich zum Referenzmodell. . . . .          | 51 |



# Tabellenverzeichnis

|     |   |    |
|-----|---|----|
| 4.1 | Tabellarische Darstellung der Versionen des Linux-Betriebssystems und der installierten Pakete . . . . .  | 37 |
| 5.1 | Tabellarische Darstellung der verwendeten Hardware- und Softwareumgebung.   | 46 |
| 5.2 | Robustheitswerte des Referenzmodells DenseNet-161 für die Evaluationsdatensätze der einzelnen Einflussgrößen und alle Evaluationsdaten. (E: alle Evaluationsdaten; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von $\pi/100$ , $\pi/50$ , um X-Achse von $\pi/25$ , U: Unschärfe) . . . . .  | 49 |
| 5.3 | Relative Robustheitswerte der Modelle für den Testdatensatz, alle Evaluationsdaten und verschiedene Evaluationsdatensätze der Einzelnen Einflussgrößen bei der Anwendung von Data-Augmentation-Techniken. (T: Test, E: alle Evaluationsdaten; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von $\pi/100$ , $\pi/50$ , um X-Achse von $\pi/25$ , U: Unschärfe) . . . . . | 49 |
| 5.4 | Relative Robustheitswerte der Modelle für den Testdatensatz, alle Evaluationsdaten und verschiedene Evaluationsdatensätze der Einzelnen Einflussgrößen bei unterschiedlichen Modellarchitekturen. (T: Test, E: alle Evaluationsdaten; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von $\pi/100$ , $\pi/50$ , um X-Achse von $\pi/25$ , U: Unschärfe) . . . . .         | 50 |
| 5.5 | Relative Robustheitswerte der Modelle für den Testdatensatz, alle Evaluationsdaten und verschiedene Evaluationsdatensätze der Einzelnen Einflussgrößen bei unterschiedlichen Netztiefen. (T: Test, A: Alle Einflussgrößen; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von $\pi/100$ , $\pi/50$ , um X-Achse von $\pi/25$ , U: Unschärfe) . . . . .                    | 51 |



# **Kapitel 1**

## **Einleitung**

Im vorliegenden Kapitel werden Forschungsbedarf und Zielsetzung der Arbeit erläutert. In Abschnitt 1.1 werden Forschungshintergrund und Motivation beschrieben. Abschnitt 1.2 betrachtet das Forschungsziel dieser Arbeit. In Abschnitt 1.3 wird der Aufbau der Arbeit präsentiert.

### **1.1 Motivation**

Die Qualität ist eine entscheidende Dimension von Produkten und Prozessen und somit ein Wettbewerbsvorteil für Unternehmen. Die industrielle Qualitätssicherung spielt eine wichtige Rolle im Produktionsprozess, wobei ihr Hauptziel in der Verbesserung der Zuverlässigkeit der Produkte liegt. Die Anwendung von Qualitätssicherungsmaßnahmen trägt zur Reduzierung von Produktionsausschuss bei und steigert die Gesamteffizienz der Produktionsumgebung (ZONNENSHAIN und KENETT 2020). In diesem Zusammenhang hat sich die visuelle Qualitätsüberwachung zu einem integralen Bestandteil entwickelt (ZIPFEL et al. 2023). Die Sichtprüfung ist eine zerstörungsfreie Qualitätsüberwachungsmaßnahme. Sie dient dazu, Defekte während oder nach dem Produktionsprozess zu erkennen und gegebenenfalls zu beseitigen, bevor die Produkte verkauft und ausgeliefert werden (AVOLA et al. 2022; OLIMOV et al. 2022).

In traditionellen Inspektionsmethoden ist menschliche Arbeitskraft erforderlich. Die manuelle Inspektion ist jedoch zeitaufwändig und aufgrund der menschlichen Erschöpfung oft fehleranfällig (CHANDOLA et al. 2009; PANG et al. 2022). Eine fortschrittlichere Möglichkeit besteht in der Anwendung automatischer Sichtprüfungen mittels klassischer industrieller Bildverarbeitungsverfahren. Dabei werden verschiedene Algorithmen auf die Pixel der Bilder angewendet, um eine Methode zur Lösung einer spezifischen Aufgabe zu entwickeln. Mit zunehmender Komplexität der Aufgabenstellung steigt die Herausforderung bei der Entwicklung solcher Methoden (BEYERER et al. 2016; STEGER et al. 2018). Die Entwicklung der künstlichen Intelligenz hat sich in der jüngsten Vergangenheit äußerst dynamisch vollzogen. Die Aufgabe der Inspektion kann in Bezug auf die Klassifizierung als Problem betrachtet werden. Traditionelle maschinelle Lernverfahren wie die Support Vector Machine (SVM) kann zum Einsatz kommen. Dabei erfordert es oft eine manuelle Gestaltung und Extraktion von Merkmalen des Zielpunkts, um eine erfolgreiche Klassifizierung zu erreichen. Die Leistungsfähigkeit hängt in hohem Maße von der sorgfältigen Auswahl oder Extraktion der Merkmale ab (HARTL 2021). In der Regel ist eine Bildvorverarbeitung nach wie vor notwendig, um die Bildqualität zu verbessern. Dies gewährleistet die robuste Anwendung der entwickelten Methode. Alternativ kann die Lösung des Klassifizierungsproblems ebenfalls durch

Deep-Learning-Verfahren (DL-Verfahren) ermöglicht werden. Durch das Training großer Datensätze kann ein Modell entwickelt werden, das präzise Klassifizierungen ermöglicht. Im Vergleich zu anderen Methoden bietet das DL-Verfahren einen erheblichen Vorteil, da keine manuelle Extraktion von Merkmalen erforderlich ist, was den Schwierigkeitsgrad der Methodenentwicklung reduziert. Die Erstellung eines leistungsfähigen Modells erfordert jedoch hochwertige, umfassende und anwendungsspezifische Datensätze. Die Generierung solcher Datensätze stellt eine herausfordernde Aufgabe dar.

Die beschriebenen Verfahren der automatischen Sichtprüfung können bei Schweißtechniken wie dem Rührreibschweißen (engl. Friction Stir Welding, FSW) angewendet werden. FSW hat sich in den 1990er Jahren als umweltfreundliche Alternative zum Schmelzschweißen etabliert. Es ermöglicht das Verbinden von Aluminiumlegierungen und besitzt viele Vorteile im Vergleich zu anderen Schweißarten. Insbesondere in der Luft- und Raumfahrtindustrie sowie in der Automobilindustrie hat es großes Interesse gefunden. Die steigende Verwendung von FSW hat die Nachfrage nach zerstörungsfreien Prüfverfahren erhöht, insbesondere bei sicherheitskritischen Anwendungen wie in der Luft-, Raumfahrt und der Elektromobilität, wo eine effektive Qualitätsüberwachung von entscheidender Bedeutung ist (HARTL 2021). Infolgedessen wurden zahlreiche Forschungsanstrengungen unternommen, um verschiedene zerstörungsfreie Prüfverfahren zu entwickeln. In diesem Kontext haben sich DL-Verfahren als äußerst leistungsfähig bei Inspektionsaufgaben erwiesen, wobei ihre Leistungsfähigkeit der menschlichen Leistung nahekommt (HARTL et al. 2019). In industriellen Anwendungen ist die Berücksichtigung der Robustheit des Systems von entscheidender Bedeutung. Die Leistungsfähigkeit des Modells sollte weitgehend unbeeinflusst von Umgebungsveränderungen bleiben, was eine wesentliche Voraussetzung für erfolgreiche industrielle Anwendungen darstellt. Allerdings gibt es in der Realität viele natürliche Einflüsse, die oft zu erheblichem Leistungsabfall des Modells führen können. Diese Ergebnisse geben Anlass zur Sorge bezüglich der Verlässlichkeit von tiefen neuronalen Netzwerken (engl. Deep Neural Network, DNN), insbesondere wenn sie in Systeme mit steigendem Sicherheitsbedarf und wachsender gesellschaftlicher Relevanz eingebunden werden (DRENKOW et al. 2022). Es ist daher wissenschaftlich sinnvoll, die Auswirkungen unterschiedlicher Umgebungseinflüsse auf die Robustheit zu erforschen und diese Robustheit zu bewerten.

## 1.2 Zielsetzung

Ziel der vorliegenden Arbeit ist die Robustheit von DL-Verfahren in der Qualitätssicherung am Beispiel des FSWs zu untersuchen. Dazu war eine umfassende Literaturrecherche zu den Themen FSW, Kameratechnik und DL-Verfahren erforderlich. Zunächst waren die potenziellen Umgebungseinflussgrößen zu identifizieren. Auf Basis dieser Einflüsse sollte ein roboterbasiertes Messsystem und eine entsprechende Software entwickelt werden, um Bilddaten unter variierenden Umgebungsbedingungen effizient aufzunehmen. Die DL-Modelle sollten durch die generierten Daten trainiert und abschließend die Robustheit der Modelle bewertet werden.

### 1.3 Aufbau der Arbeit

Zum Beginn werden in Kapitel 2 die Grundlagen des FSWs beschrieben. Dabei werden die relevanten Oberflächeneigenschaften von FSW-Nähten zur Sichtprüfung erläutert. Anschließend werden die Grundlagen zu den Themen Bildaufnahmesystem und DL-Verfahren betrachtet. Sie sind für die Auswahl der Messsensoren, den Aufbau des Versuchsstandes und das Training der Modelle erforderlich. Kapitel 3 gibt einen umfassenden Überblick über den Stand der Forschung und Technik. In Kapitel 4 wird die Strategie zur Datengenerierung erläutert. Die Umgebungseinflussgrößen werden anschließend diskutiert und identifiziert. Der Aufbau des roboter-basierten Messstands sowie der gesamte Aufnahmeprozess werden ausführlich dargestellt. Die Aufbereitung der Bilddaten wird zum Ende dieses Kapitels erläutert. In Kapitel 5 wird zunächst das Vorgehen zur Bewertung der Robustheit betrachtet und anschließend werden die Bewertungsergebnisse präsentiert und diskutiert. Abschließend fasst Kapitel 6 die vorliegende Arbeit zusammen und bietet einen Ausblick auf weiterführende Forschungsthemen.



# **Kapitel 2**

## **Grundlagen**

Im vorliegenden Kapitel geht es um die erforderlichen Grundlagen für den wissenschaftlichen Betrachtungsraum dieser Arbeit. Abschnitt 2.1 beschreibt Prozessablauf und Oberflächeneigenschaften des Rührreibschweißens. In Abschnitt 2.2 werden relevante Komponenten eines Bildverarbeitungssystems wie Kamera, Objektiv sowie Beleuchtung beschrieben. Abschnitt 2.3 betrachtet faltende neuronale Netze (engl. Convolutional Neural Network, CNN). Grundaufbau, Trainingsverfahren und Methode zur Modellauswahl werden erläutert.

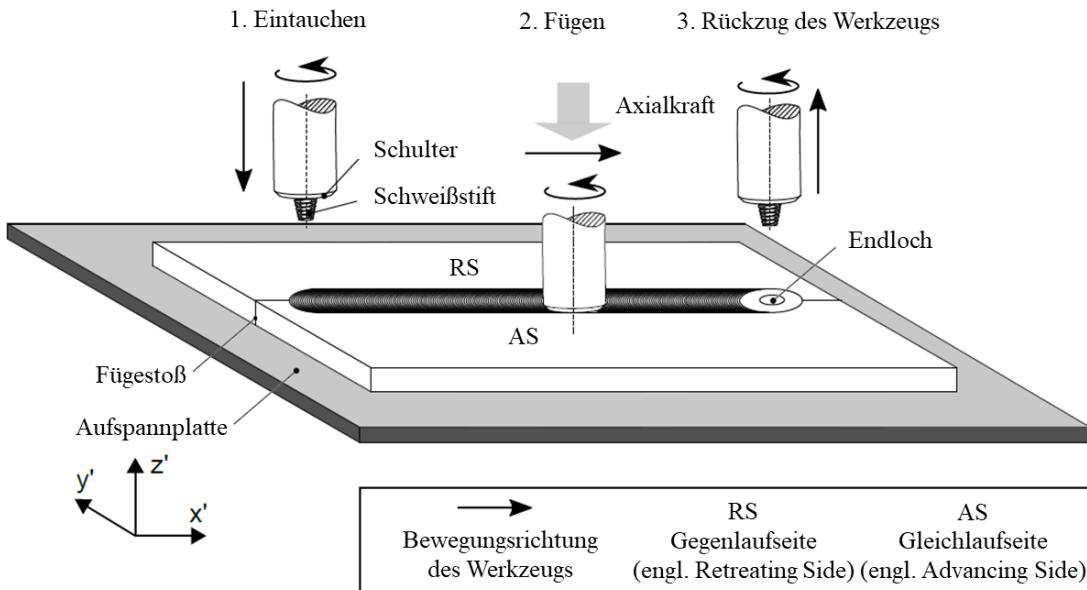
### **2.1 Rührreibschweißen**

Gegenwärtig wird es immer wichtiger, in der Produktionsindustrie, unterschiedliche Materialien einzusetzen und miteinander zu verbinden. Bei herkömmlichen Schweißverfahren treten häufig Probleme aufgrund von Materialinkompatibilitäten auf. Als Lösung wurde das Reibschweißen entwickelt. Dabei werden die erzeugte Hitze in der Bewegung der Materialien und der Druck verwendet, um robuste metallurgische Verbindungen zwischen unterschiedlichen Materialien herzustellen (AKINLABI 2020, S. 2–4).

FSW ist eine weitere Form des Feststoffschweißens und ähnelt dem Reibschweißen, das auf dem Prinzip der Reibung basiert. FSW wurde im Jahr 1991 von Welding Institute erfunden und fand innerhalb kurzer Zeit Anwendung in der Industrie. Mit der Entwicklung der Technologie wird FSW zunehmend in verschiedenen Bereichen eingesetzt. Häufig wird für Mischverbindungen wie bspw. Aluminium und Edelstahl, Aluminium und Kupfer oder für reine Kupferverbindungen verwendet (MUBIAYI et al. 2019, S. 67–68). Das Anwendungsgebiet umfasst die Herstellung von Komponenten für die Luft- und Raumfahrt, die Automobilindustrie, elektronische Gehäuse, Kühler, den Eisenbahnsektor, den Schiffbau, Wärmetauscher und Lagerbehälter für nukleare Abfälle. Beim FSW sind keine Zusatstoffe und kein Schutzgas erforderlich. Es entstehen keine Dämpfe oder giftigen Gase, daher ist FSW emissionsarm und umweltfreundlich. Weitere Vorteile umfassen die hervorragende Schweißnahtqualität, geringe Verformung, geringe Schrumpfung und keine Poren, da keine Schmelze der Grundwerkstoffe erfolgt. (AKINLABI 2020, S. 4–8)

### 2.1.1 Prozessablauf

FSW wird nach DIN EN ISO 25239-1 (BEUTH VERLAG GMBH 2020b) angewendet. Dabei kommt ein rotierendes, nahezu verschleißfestes Werkzeug zum Einsatz, damit die Schweißnaht durch reibungsbedingte Erwärmung und Vermischung des Materials im plastischen Zustand erzeugt wird (BACHMANN 2021; HARTL 2021). Ein herkömmliches Werkzeug zum FSW weist zwei grundlegende funktionale Komponenten auf: Schweißstift und Schulter. Der Schweißstift befindet sich während des Schweißprozesses vollständig in der Schweißzone. Die Schulter begrenzt die Schweißzone in Richtung der Bauteiloberfläche (RUHSTORFER 2012, S. 3). Abb. 2.1 veranschaulicht der Ablauf beim FSW zweier Bleche im Stumpfstoß dargestellt. Der FSW-Prozess umfasst drei aufeinanderfolgende Prozessphasen. Zuerst kommt



**Abbildung 2.1:** Prozessablauf beim FSW im Stumpfstoß in Anlehnung an (BACHMANN 2021, S. 11).

die Eintauchphase, in der das Werkzeug vorbereitet wird. Danach folgt die Hauptphase, in der die Schweißnaht erzeugt wird. Abschließend wird die Beendigungsphase durchgeführt, in der das Werkzeug vom Werkstück zurückgezogen wird. Die drei Phasen werden im Folgenden näher erläutert (COLLIGAN 2010, S. 18):

#### 1. Eintauchphase (Eintauchen)

Am Anfang bewegt sich das Werkzeug in Richtung Eintauchpunkt. Unter kontinuierlicher Rotation um seine eigene Symmetrieachse wird es durch die axiale Anpresskraft in die Fügezone gedrückt. Der Schweißstift dringt in die Fügepartner ein, bis Schulter und Werkstückoberfläche in vollständigem Kontakt sind. Die plastische Deformation zwischen dem rotierenden Werkzeug und den beiden Fügepartnern werden durch starke Reibungswärme erzeugt. Diese lokale Erwärmung ermöglicht das Eindringen des Schweißstifts in den Fügestoß und nach der Abkühlung eine Entfestigung des Werkstoffes in der Umgebung des Werkzeugs. In der Regel ist es notwendig, am Eintauchpunkt eine gewisse Zeit zu verweilen, um eine ausreichend hohe Temperatur im Schweißwerkzeug und im Werkstück zu erreichen. (BACHMANN 2021; COLLIGAN 2010; HARTL 2021; RUHSTORFER 2012)

## 2. Hauptphase (Fügen)

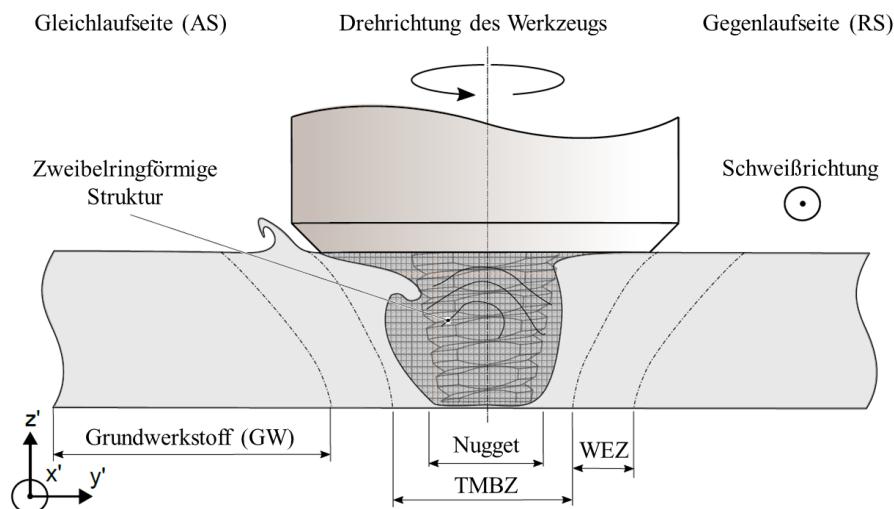
Bei ausreichender Erwärmung bewegt sich das Werkzeug entlang der Fügelinie mit einer definierten Geschwindigkeit (Schweißgeschwindigkeit oder Vorschubgeschwindigkeit). Mit der Bewegung werden lokale Werkstoffe in der Umgebung des Werkzeugs kontinuierlich verschmolzen und vermischt (COLLIGAN 2010, S. 19). Die Schmelztemperatur der Bauteile sollte beim Schweißprozess nicht überschritten werden. Aufgrund der Überlagerung der translatorischen und rotatorischen Bewegung des Schweißwerkzeugs entlang der Fugenlinie entsteht eine asymmetrische Schweißnaht. (BACHMANN 2021; HARTL 2021; RUHSTORFER 2012)

## 3. Beendigungsphase (Rückzug des Werkzeugs)

Wenn das Werkzeug den Endpunkt der Schweißnaht erreicht, wird die Vorwärtsbewegung gestoppt. Das Werkzeug wird vom Werkstück zurückgezogen und ein sog. Endloch verbleibt in der Schweißnaht. (COLLIGAN 2010, S. 19)

### 2.1.2 Charakteristik von Schweißnähten

Aufgrund des spezifischen Schweißprinzips unterscheidet sich das entstehende Schweißnahtgefüge beim FSW signifikant von jenem, das bei anderen Schweißverfahren entsteht. Typischerweise erfolgt die Unterteilung des Querschnitts einer FSW-Schweißnaht in vier Bereiche, wie in Abb. 2.2 dargestellt (COLLIGAN 2010, S. 17):

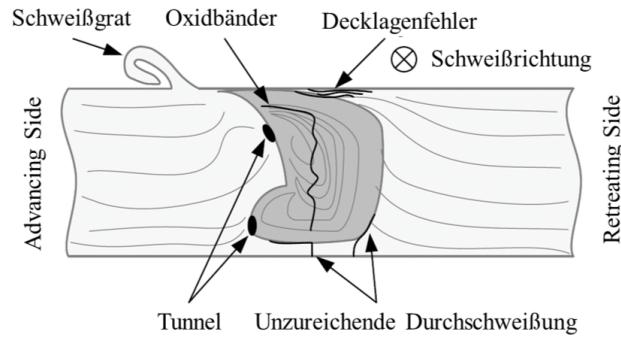


**Abbildung 2.2:** FSW-Nahtquerschnitt mit charakteristischen Bereichen in Anlehnung an (BACHMANN 2021, S. 15).

- **Grundwerkstoff (engl. Base metal, GW):** Das Material ist fern von der Schweißnaht und wird kaum verformt oder durch Hitze beeinflusst. Das Gefüge in diesem Bereich besitzt die gleiche Struktur wie der Grundwerkstoff ohne messbare Veränderungen.
- **Wärmeinflusszone (engl. Heat-affected zone, WEZ):** Die thermische Belastung tritt in der WEZ auf, was Gefügeveränderungen dort verursacht. Dieser Bereich liegt näher am Zentrum der Schweißnaht. Normalerweise verändern sich dort Mikrostruktur und mechanische Eigenschaften. Plastische Verformungen treten in diesem Bereich nicht auf.

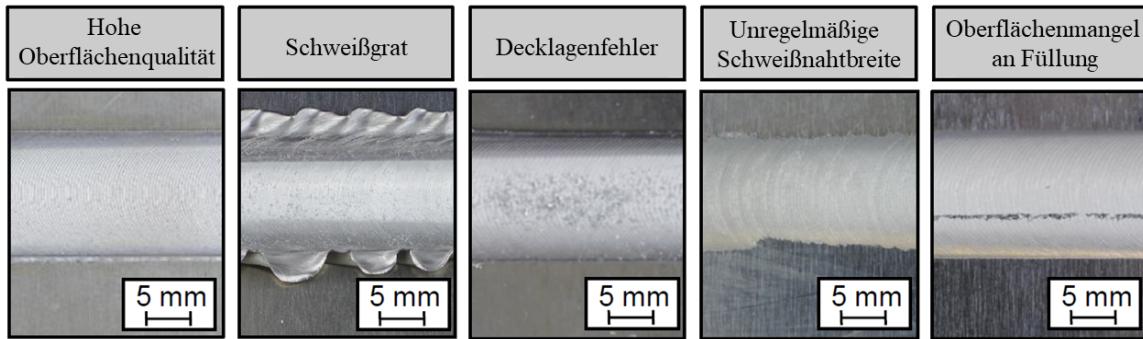
- Thermomechanisch beeinflusste Zone (engl. Thermomechanically affected zone, TMAZ): TMAZ befindet sich im Zentrumsbereich des Nahtquerschnitts. In diesem Bereich wird das Material verformt, wobei die entstehende Wärme eine thermische Beeinflussung auf das Material ausübt. In diesem Zentrumsbereich kann in rekristallisierte und nicht rekristallisierte Zone weiter eingeteilt werden.
- Nugget: Das Nugget beschreibt eine rekristallisierte Zone in TMAZ. Dort findet Rekristallisation statt.

Es gibt viele mögliche Schweißfehler beim FSW. In DIN EN ISO 25239-5 (BEUTH VERLAG GMBH 2020a) werden Unregelmäßigkeiten in Oberflächen-Unregelmäßigkeiten, innere Unregelmäßigkeiten und Schweißwurzelfehler unterteilt. In Abb. 2.3 werden charakteristische Merkmale der FSW-Naht dargestellt. Im folgende werden relevante auftretende Schweißfehler,



**Abbildung 2.3:** Charakteristische Merkmale einer FSW-Naht bei einer Stumpfstoßverbindung in Anlehnung an (RUHSTORFER 2012, S. 14).

ler, die durch Sichtprüfung erfasst werden können, nach HARTL (2021) und RUHSTORFER (2012) zusammengefasst und vorgestellt. Andere Schweißfehler werden in der vorliegenden Arbeit nicht weiter diskutiert. Abb. 2.4 zeigt unterschiedliche Oberflächenqualitäten von FSW.



**Abbildung 2.4:** Oberflächenqualitäten von Rührreibschweißen in Anlehnung an (HARTL 2021, S. 10).

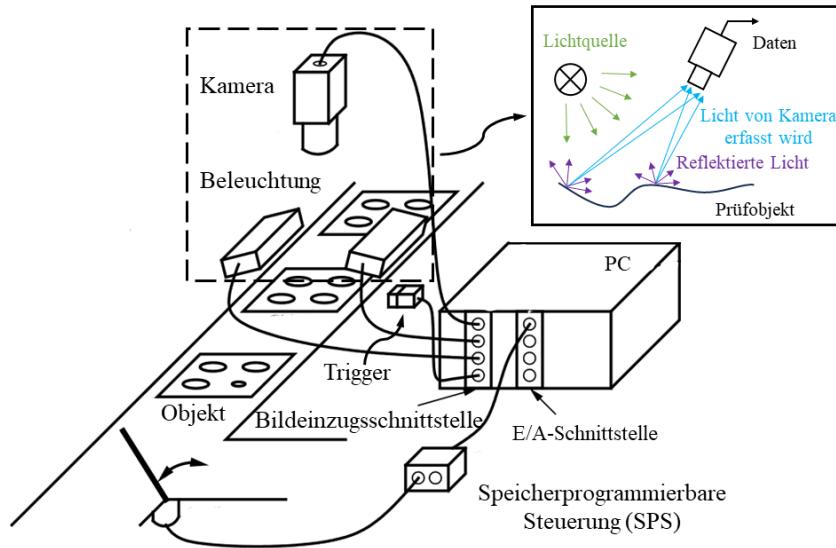
- Schweißgrat: Beim FSW kann ein unregelmäßiger erhöhter Schweißgrat entstehen. Schweißgrat ist ein auftretender Materialauswurf am Rand der Schweißnaht. Aufgrund erheblicher Wärmeenergie infolge z. B. einer überhöhten Werkzeugdrehzahl werden Teile des plastifizierten Werkstoffes vom Zentrumsbereich der Schweißnaht gedrängt und ausgeworfen. Nach der Abkühlung wird der Schweißgrat gebildet, welcher durch

die Anwendung eines nachfolgenden Schweißprozesses wie Fräsen entfernt werden kann. Dies besitzt in der Regel nur einen geringfügigen Einfluss auf die mechanischen Eigenschaften.

- Übermäßige Nahtunterfüllung: Wenn eine Schulter in die Schweißzone eintaucht und Material aus der Schweißnaht verdrängt, tritt häufig Nahtunterfüllung auf. Ein hohe Menge an Unterfüllung wird als Schweißfehler betrachtet. Da die Tiefe der Nahtunterfüllung in der Regel lediglich im Bereich weniger Zehntel Millimeter liegt, lässt sich bei einer Sichtprüfung normalerweise nicht feststellen, ob die Unterfüllung akzeptabel oder zu stark ausgeprägt ist.
- Decklagenfehler: Das Material wird in unmittelbarer Nähe zur Schulter vom restlichen Material abgeschert. Aufgrund der verringerten Axialkraft erfolgt jedoch keine ausreichende Verdichtung.
- Unregelmäßige Schweißnahtbreite: Bei diesem Schweißfehler zeigt die Schweißnaht an ihren Rändern eine unregelmäßige Kontur, wobei die Breite der Schweißnaht oft geringer ist als der Durchmesser der Schulter. Die Hauptursache liegt in einer unzureichenden Axialkraft.
- Oberflächenmangel an Füllung: Der Zustand des Oberflächenmangels, der ebenfalls als Oberflächenhohlraum oder Oberflächenbruch bezeichnet wird, tritt meist als Riss entlang der Schweißrichtung auf. Dieser Schweißfehler resultiert aus einer zu geringen Axialkraft, ungeeigneten Prozessparametern, einer ungeeigneten Werkzeuggeometrie oder Lücken zwischen den Werkstücken.

## 2.2 Bildverarbeitungssystem

Um spezifische Aufgaben in der industriellen Bildverarbeitung erfolgreich zu bewältigen, bedarf es eines geeigneten Bildverarbeitungssystems. Einige der gängigen Aufgabenstellungen in einem Bildverarbeitungssystem umfassen die Objektidentifikation, die Positionserkennung, die Vollständigkeitsprüfung, die Überprüfung von Form und Abmessungen sowie die Oberflächenprüfung (BAUER 2003). Abb. 2.5 veranschaulicht ein Beispiel für ein industrielles Bildverarbeitungssystem. Das Objekt wird auf z. B. durch ein Förderband mechanisch transportiert. Die Bilder des Objektes werden in einer bestimmten Position aufgenommen. Dabei kommt eine geeignete oder speziell entwickelte Beleuchtung zum Einsatz, um optimale Helligkeit und Bildqualität zu gewährleisten. Das Objekt wird durch eine Kamera abgebildet, die ein geeignetes Objektiv verwendet. Die Bilder werden über eine Kamera-Computer-Schnittstelle wie bspw. Frame-Grabber an einem Computer (PC) übertragen. Die Steuerung der Beleuchtung kann ebenfalls über den Frame-Grabber erfolgen. Falls das System keinen Frame-Grabber verwendet, sondern eine Standardschnittstelle wie USB bzw. Ethernet, wird typischerweise ein Trigger direkt mit Kamera und Beleuchtung verbunden oder über speicherprogrammierbare Steuerung gesteuert. Die weiteren Verarbeitungsschritte der aufgenommenen Bilder werden am PC durchgeführt (STEGER et al. 2018, S. 1–4). Die Qualität der aufgenommenen Bilder hat oft einen erheblichen Einfluss auf die Ergebnisse, daher ist die sorgfältige Auswahl der geeigneten Messkomponenten zur Gewährleistung einer hohen Bildqualität unerlässlich.



**Abbildung 2.5:** Komponenten eines Bildverarbeitungssystems in Anlehnung an (STEGER et al. 2018) und (BEYERER et al. 2016, S. 9).

### 2.2.1 Beleuchtung

Die Beleuchtung bildet die Grundvoraussetzung für die Bildabbildung und ist ein Schlüsselfaktor für die Bildqualität. Eine ungünstig ausgewählte oder eingerichtete Lichtquelle erhöht den algorithmischen Aufwand in der Bildaufbereitung. Dies führt dazu, dass Aufgaben oft nicht erledigt werden können (AZAD et al. 2011, S. 14). Der Zweck der Beleuchtung besteht darin, relevante Merkmale hervorzuheben und unerwünschte zu vernachlässigen. Dabei können verschiedene Lichtrichtungen in Bezug auf die Position der Kamera verwendet werden, um die gewünschten Merkmale zu verstärken. Ebenfalls lassen sich unterschiedliche Spektraleigenschaften des Lichts verwenden. In den meisten Anwendungen ist zeitlich konstante Beleuchtung notwendig, wobei ebenfalls der Aspekt der Alterung berücksichtigt werden muss (ERHARDT 2008, S. 21–27).

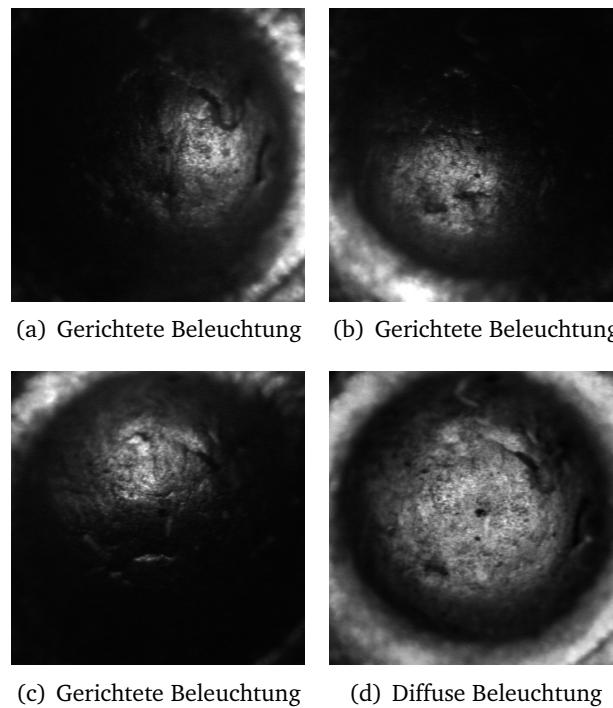
#### Leuchtmittel

Mehrere Arten der Lichtquelle stehen in der Bildverarbeitung zur Verfügung. Die Zusammensetzung der verschiedenen Lichtquellen wie z. B. Hochfrequenz-Leuchtstoffröhren, Halogenlampen sowie Leuchtdioden (engl. light-emitting diode, LEDs) wird von AZAD et al. (2011, S.19) aufgeführt. Eine Betrachtung dieser Technologie erfolgt von BEYERER et al. (2016, S.19). Besonders LEDs weisen viele Vorteile auf, darunter eine lange Lebensdauer, eine hohe Langzeitkonstanz sowie eine kompakte Bauform und eine unkomplizierte Vorschaltelektronik mit geringerer Versorgungsspannung. LEDs sind ebenfalls im Ultravioletten- und Infrarotbereich erhältlich. Ein wesentlicher Nachteil besteht darin, dass die Leistung der LEDs von der Umgebungstemperatur abhängig ist. Bei einer hohen Umgebungstemperatur kann sich die Leistung der LEDs sowie deren Lebensdauer verringern. Dennoch werden LEDs aufgrund ihrer zahlreichen Vorteile häufig in Bildverarbeitungsanwendungen eingesetzt (AZAD et al. 2011; STEGER et al. 2018).

### Verwendung der Richtung des Lichtes

In der industriellen Bildverarbeitung gibt es viele verschiedene Möglichkeiten der Beleuchtung, die nach den vielfältigen Anforderungen der jeweiligen Aufgabenstellungen ausgewählt werden. Die Reflektanz-eigenschaften eines Objekts hängen von der Einstrahlrichtung des Lichtes, der Position der Kamera sowie den Reflexionseigenschaften der Oberfläche des Objekts ab. Diese Beziehung wird durch die bidirektionale Reflektanzverteilungsfunktion (engl. bidirectional reflectance distribution function, BRDF) beschrieben (BEYERER et al. 2016, S. 300–305). Daher lässt sich die Reflektanz-eigenschaften des Objektes durch Änderung der Abstrahlrichtung der Lichtquelle oder der relativen Position von Lichtquelle und Kamera variieren.

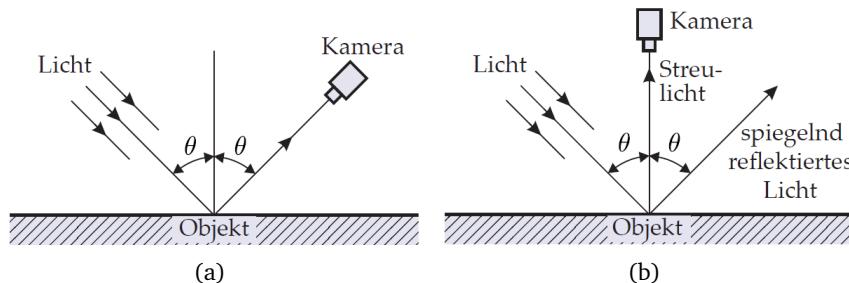
Anhand der Richtung der Lichtstrahlen an der Lichtquelle lässt sich die Beleuchtung in diffuse und gerichtete Beleuchtung unterteilen. Bei diffuser Beleuchtung erfolgt eine theoretisch gleichmäßige Abstrahlung in alle Richtungen, während bei gerichteter Beleuchtung die Lichtquelle in einen engen Richtungsbereich strahlt. Im Grenzfall sind alle Lichtstrahlen parallel angeordnet, was als telezentrische Beleuchtung bezeichnet wird. Im Vergleich zur gerichteten Beleuchtung treten bei diffuser Beleuchtung kaum Schattenwürfe auf. Bei stark reflektierenden Oberflächen wie Metall lässt sich die Stärke von Glanzlichtern vermindern, was zu einer verstärkten Kontrastwirkung führt. In Abb. 2.6 wird ein Beispiel für einen Vergleich zwischen gerichteter und diffuser Beleuchtung in einer Anwendung aus der Kriminaltechnik dargestellt. (BEYERER et al. 2016; STEGER et al. 2018)



**Abbildung 2.6:** Aufnahmen des Schlagbolzeindrucks auf einer Patronenhülse zum Zwecke der automatischen Identifikation von Schusswaffen: (a)–(c) gerichtete Beleuchtungen; (d) diffuse Beleuchtung, in Anlehnung an (BEYERER et al. 2016, S. 301).

Die Position der Lichtquelle relativ zur Kamera und zum Objekt spielt ebenfalls eine wichtige Rolle. Im Allgemeinen lässt sich die Beleuchtung in folgende Kategorien unterteilen: Auflicht-, Durchlicht-, Hellfeld- und Dunkelfeldbeleuchtung. Lichtquelle und Kamera befinden sich auf derselben Seite des Objektes, handelt es sich um eine Auflichtbeleuchtung. Im Gegensatz

dazu wird eine Beleuchtung als Durchlichtbeleuchtung bezeichnet, wenn sich Lichtquelle und Kamera auf unterschiedlichen Seiten des Objektes befinden. Abb. 2.7 veranschaulicht das Prinzip der Hellfeld- und Dunkelfeldbeleuchtung. Bei einer Hellfeldbeleuchtung wird die



**Abbildung 2.7:** (a) Prinzip der Hellfeldbeleuchtung; (b) Prinzip der Dunkelfeldbeleuchtung, in Anlehnung an (BEYERER et al. 2016, S. 302, 303).

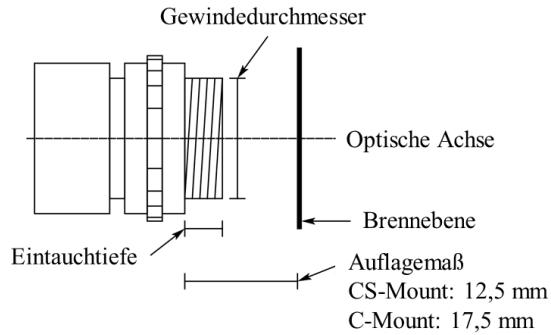
Lichtquelle so angeordnet, dass ein Großteil des Lichtes zur Kamera reflektiert wird. Dadurch erfolgt eine gute Lichtausnutzung. Im Fall einer Dunkelfeldbeleuchtung wird das meiste Licht von der Kamera weg reflektiert. (ERHARDT 2008)

## 2.2.2 Kameratechnik

Die Auswahl einer richtigen Kamera und eines geeigneten Objektives spielt eine entscheidende Rolle in der industriellen Bildverarbeitung. Die Verwendung eines Objektives sammelt einerseits die Lichtstrahlen und erhöht die Helligkeit des Bildes. Andererseits besitzt ein Objektiv die Funktion, das Bild scharf abzubilden und Licht aus einem erweiterten Bereich einzufangen (FORSYTH und PONCE 2003, S. 9). Gegenwärtig werden Kameras hauptsächlich in CCD- und CMOS-Kameras eingeteilt. Die beiden Sensoren basieren auf dem Fotoeffekt. Der Unterschied liegt in der Bauform der Sensoren. Der Vorteil eines CCD-Sensor besteht darin, dass er eine signifikant verbesserte Bildqualität aufweist und ein geringes Rauschen aufzeigt. Im Vergleich zu CMOS-Sensor gleicher Größe besitzt der CCD-Sensor eine höhere Auflösung und Lichtempfindlichkeit. Im Gegensatz dazu hat der CCD-Sensor höheren Stromverbrauch und Produktionskosten (BATCHELOR 2012, S. 358–394). In AZAD et al. (2011, S.48) wird eine ausführlichere Gegenüberstellung gegeben. Im Folgenden werden relevante Objektivparameter und Berechnung praxisgerecht beschrieben und erläutert. In der vorliegenden Arbeit wird zunächst eine geeignete Kamera gemäß den spezifischen Anforderungen der Aufgabe ausgewählt, bevor eine Entscheidung hinsichtlich des Objektives getroffen wird.

### Anschluss

Die Verbindung zwischen Kamera und Objektiv erfolgt hauptsächlich über die C- und CS-Mount-Anschluss. Abb. 2.8 stellt den Standardanschluss für industrielle Objektive dar. Bei einem normierten Objektivanschlussgewinde beträgt der Durchmesser ein Zoll (25,4 mm) und die Gewindesteigung 1/32 Zoll (32 Windungen pro Zoll). Das Auflagemaß beschreibt den Abstand zwischen Objektivauflagefläche und der Bildebene. Es beträgt beim C-Mount-Objektiv 17,5 mm und beim CS-Mount-Objektiv 12,5 mm. Der C/CS-Mount-Adapter ermöglicht die Verbindung zwischen unterschiedlichen Anschlüssen. (STEGER et al. 2018)



**Abbildung 2.8:** Standardanschluss für industrielle Objektive in Anlehnung an (STEGER et al. 2018).

## Format

Das Format definiert die Abmessungen des Bildaufnehmers oder Bildsensors in der Kamera, der für die Erfassung von Bildern verantwortlich ist. Bildsensor-Größen werden nach den Größen der Außendurchmesser der Bildaufnahmeröhren klassifiziert (STEGER et al. 2018). Das Format des Objektives muss mit dem Format der Kamera übereinstimmen und gleich oder größer ausgewählt werden. Ein großformatiges Objektiv mit einem kleineren Bildsensor bewirkt eine Verringerung der Bildverzerrungen.

## Brennweite

Bei der Auswahl des Objektives ist die Berechnung der Brennweite nach Aufgabeanforderungen erforderlich. Die Brennweite  $f$  kann mittels Abbildungsgesetze abgeschätzt werden. Abb. 2.9 stellt die Abbildungsgesetze für dicke Linsen dar. Der Zusammenhang zwischen Gegenstandsgröße  $G$  und Bildgröße  $B$  wird durch Strahlensatz beschrieben:

$$\frac{G}{g} = \frac{B}{b} \quad (2.1)$$

Die Abbildungsgleichung für dicke Linsen lautet:

$$\frac{1}{f} = \frac{1}{b} + \frac{1}{g} \quad (2.2)$$

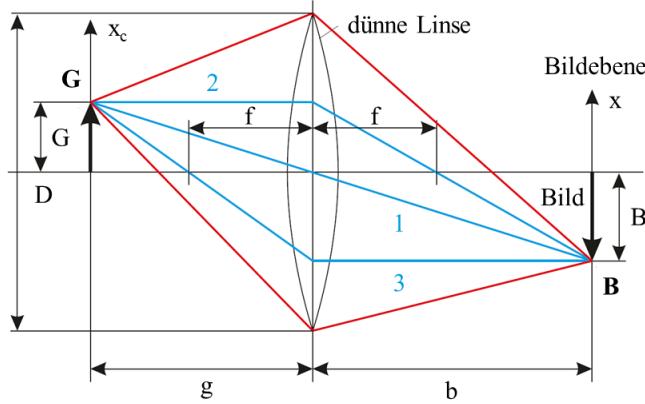
Aus Gl. 2.1 und Gl. 2.2 ergibt sich die Brennweite:

$$f = \frac{g \cdot B}{G + B} \quad (2.3)$$

Wie aus der Gl. 2.3 folgt, muss die Bildweite  $b$  verändert werden, um scharfe Bilder zu erfassen. Mit abnehmender Gegenstandsweite  $g$  vergrößert sich die Bildweite. Gemäß den geforderten Arbeitsabständen und der Größe des Objektes kann die Brennweite bestimmt werden. Abschließend ist es notwendig, die minimale Objektdistanz (MOD) sorgfältig zu überprüfen und den Einfluss des Schärfentiefenbereichs angemessen zu berücksichtigen.

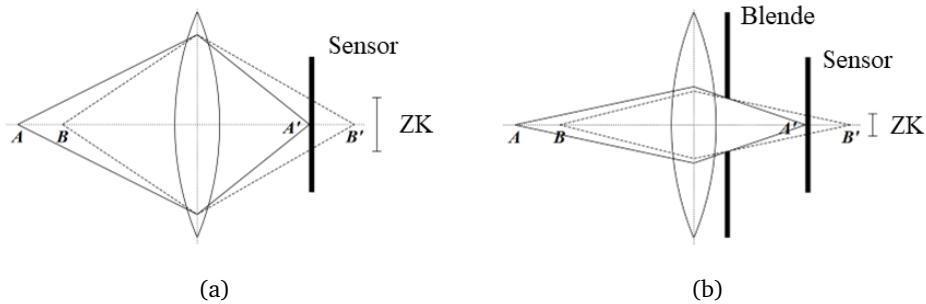
## Minimale Objektdistanz und Schärfentiefe

Die minimale Objektdistanz ist definiert als der kürzeste Abstand, bei dem eine scharfe Fokussierung möglich ist. Die MOD kann mittels Zwischenringen maximal bis zur Brennweite



**Abbildung 2.9:** Abbildungsgesetze für dicke Linsen in Anlehnung an (BEYERER et al. 2016, S. 136). Die Ausdehnung des abgebildeten Strahlenbündels ist rot dargestellt. Blau eingefärbt sind die ausgezeichneten Strahlen. (1: Mittelpunktsstrahl, 2, 3: Parallelstrahlen)

reduziert werden (STEGER et al. 2018). Die Bildweite, bei der ein Objekt exakt scharf abgebildet wird, kann aus der Gegenstandsweite und der Brennweite abgeleitet werden. Entferntere oder nähere Bildpunkte werden auf einem Zerstreuungskreis abgebildet. In einem bestimmten Bereich der Gegenstandsweite ist der Zerstreuungskreis gering, so dass eine messbare Unschärfe nicht erkennbar ist. Dieser Bereich wird als Schärfentiefe bezeichnet. Typische Maximalgröße des Zerstreuungskreises ist die Größe eines Sensorelementes (STEGER et al. 2018). Durch den Einsatz einer Blende kann der Zerstreuungskreis verkleinert und die Schärfentiefe vergrößert werden, wie in Abb. 2.10 dargestellt wird. Bei der Verkleinerung der Blende erfolgt eine längere Belichtungszeit.



**Abbildung 2.10:** Vergleich von Schärfentiefe in Anlegung an (STEGER et al. 2018). ZK:Zerstreuungskreis (a) Zerstreuungskreis ohne Blende; (b) Zerstreuungskreis mit Blende.

## Auflösung

Der Zweck der Auswahl der Auflösung besteht darin, dass Fotos in der Lage sind, die kleinsten Merkmale zu unterscheiden. Der Zusammenhang von Bildaufnahmerauflösung  $R_{Kamera}$  und Ortsauflösung  $R_{Ort}$  wird in folgender Gleichung beschrieben (AZAD et al. 2011, S. 37):

$$R_{Kamera} = \frac{G'}{R_{Ort}} \quad (2.4)$$

Die Größe des Gesichtsfeldes  $G'$  (engl. Field-of-View, FOV) ergibt sich aus der Summe der Gegenstandsgröße  $G$ , der Toleranz der Objektposition und eines Sicherheitsfaktors. Die Orts-

auflösung lässt sich durch folgende Gleichung beschreiben (AZAD et al. 2011, S. 37):

$$R_{Ord} = \frac{s_{Merkmal}}{n_{Pixel}} \quad (2.5)$$

Bei der Betrachtung der pixelgenauen Rechnung beträgt die Pixelanzahl  $n_{Pixel}$  1. Falls die Subpixelmethode verwendet wird, lautet Pixelanzahl  $n_{Pixel}$ :

$$n_{Pixel} = \frac{1}{Subpixelfaktor} \quad (2.6)$$

Durch die Anwendung der oben genannten Gleichungen ergibt sich die Kameraauflösung  $R_{Kamera}$ :

$$R_{Kamera} = G' \cdot \frac{1}{Subpixelfaktor \cdot s_{Merkmal}} \quad (2.7)$$

Üblicherweise erfolgt zunächst die Auswahl einer geeigneten Kamera entsprechend den Anforderungen hinsichtlich der Auflösung, welche durch Gl. 2.7 berechnet wird. Anschließend kann die Brennweite mithilfe von Gl. 2.3 für die spezifische Aufgabe bestimmt werden. Nach Festlegung der Sensorgröße wird das geeignete Objektiv ausgewählt. Bei der Objektivauswahl müssen weitere Aspekte wie bspw. MOD und Schärfentiefe berücksichtigt werden.

## 2.3 Faltendes neuronales Netzwerk

In der heutigen Zeit wird mit dem Fortschritt des Deep-Learning-Verfahrens eine zunehmende Anwendung im Alltag und der wissenschaftlichen Forschung beobachtet. Insbesondere in den Bereichen Regression und Klassifizierung wird es umfangreich eingesetzt. CNN ist eine Art von künstlichen neuronalen Netzwerken, die hauptsächlich zur Verarbeitung von Bildern verwendet werden.

### 2.3.1 Grundaufbau

Das DL-Verfahren ist von menschlichen neuronalen Netzen inspiriert. Das menschliche Gehirn besteht aus Milliarden von Neuronen. Jedes Neuron ist über Synapsen mit zehntausenden anderen Neuronen verbunden. Die Übertragung von Informationen zwischen den Neuronen erfolgt über elektrochemische Signale. Die Eingangssignale gelangen über Dendriten zum Neuron. Wenn die gesammelten Signale einen bestimmten Schwellenwert überschreiten, gibt das Neuron ein Ausgangssignal ab. Das vorgestellte biologische Modell diente als Grundlage für den Perzepron-Algorithmus, der im Jahr 1958 von Rosenblatt entwickelt wurde. (SOMANI et al. 2023, S. 70–74)

Abb. 2.11 stellt ein einfaches neuronales Netz (engl. Artificial neural network, ANN) dar. Ein künstliches Neuron besitzt mehrere Eingänge und einen Ausgang. Die Ausgabe ergibt sich durch die Summe der gewichteten Eingänge, gefolgt von der Anwendung einer Aktivierungsfunktion. Die Gleichung lautet wie folgend:

$$\hat{y} = \sigma \left( \sum_{n=1}^N w_i x_i + w_0 \right) \quad (2.8)$$

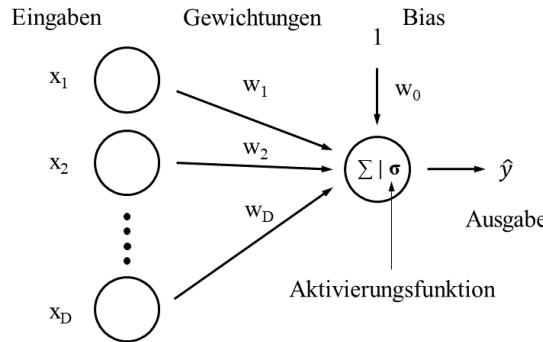


Abbildung 2.11: Darstellung des verallgemeinerten linearen Modells als einfaches neuronales Netz

Es ist offensichtlich, dass das einfache neuronale Netz ein lineares Modell repräsentieren kann. Eine lineare Regression wird durch eine lineare Aktivierungsfunktion durchgeführt, während eine logistische Regression durch eine Sigmoid-Funktion ermöglicht wird. Bei Verwendung einer Stufenfunktion entsteht der Perzeptron-Algorithmus. Es gibt viele unterschiedliche Aktivierungsfunktionen, um spezifische Ausgaben zu erhalten. Um eine Nichtlinearität zu erreichen, wird ein mehrschichtiges Netz eingesetzt. Seine Struktur wird in Abb. 2.12 schematisch dargestellt. Ein mehrschichtiges Netz besteht aus einer Eingabeschicht, mehreren verborgenen Schichten und einer Ausgabeschicht. Ein Netz, wo eine Schicht nur mit der nächsthöheren Schicht verbunden ist, wird als Feedforward-Netz bezeichnet. Die Ausgaben der Schichten werden mittels eines Vorwärtspasses (engl. forward pass) berechnet. (AGGARWAL 2023, S. 5–17)

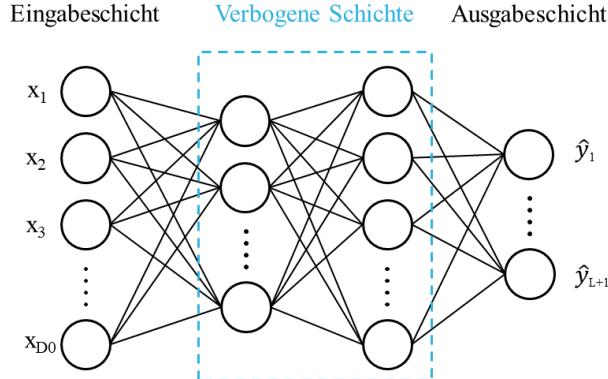
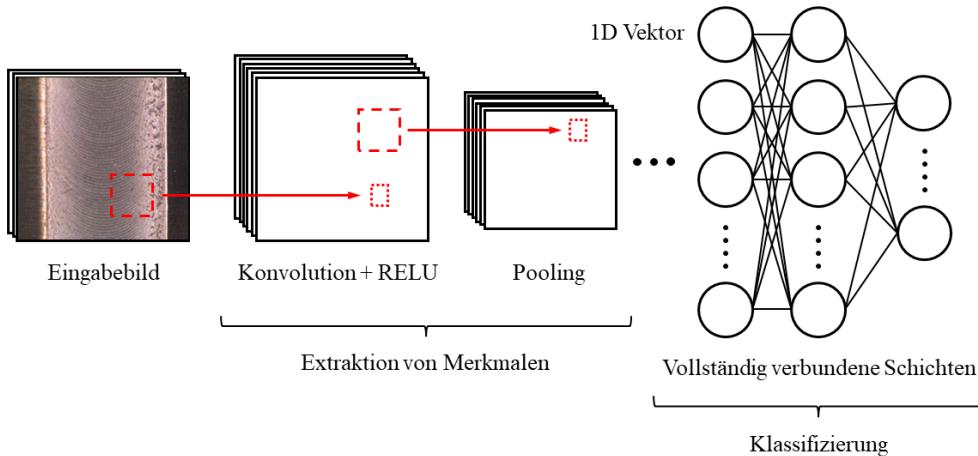


Abbildung 2.12: Eine schematische Darstellung eines mehrschichtigen Feedforward-Netzwerks mit zwei verborgenen Schichten.

Ein ANN ist für die Bildverarbeitung weniger geeignet, da es potenziell eine hohe Anzahl an Gewichtungen aufweisen kann. Dies führt zu erhöhtem Speicherbedarf. Bei der Struktur des ANN gehen räumliche Informationen verloren. Aus diesem Grund wurde eine CNN-Architektur neu entwickelt. Durch den Einsatz der Faltungstechnik mittels Kernels kann die Anzahl der Gewichtungen erheblich reduziert werden. Die Schichten, die durch Faltung berechnet werden, werden als Konvolutionsschichten bezeichnet. Ihre Berechnungen erfolgen nach dem gleichen Prinzip wie beim ANN, nur wird die gewichtete Summe durch Faltungen ersetzt. Die Konvolutionsschichten können als Merkmalsextraktoren betrachtet werden. Mit einer größeren Anzahl von Konvolutionsschichten werden mehr Merkmale tief extrahiert. Nach der Konvolutionsschichten kann eine sog. Pooling-Schicht (ebenfalls Upsampling-



**Abbildung 2.13:** Darstellung der Beispielarchitektur eines CNN

Schicht) verwendet werden, um die extrahierten Merkmale zusammenzufassen und weiter zu reduzieren. Nach der Merkmalsextraktion kommen vollständig verbundene Schichten zur Klassifizierung zum Einsatz (Jo 2023, S. 303–325). Die Beispielarchitektur eines CNN wird in Abb. 2.13 veranschaulicht. Gegenwärtig werden vielfältige Architekturen von CNN entwickelt. Beispiele sind ResNet und DenseNet, welche herausragende Leistungen aufweisen. Eine umfassende Aufzählung sowie Zusammenfassung moderner CNN-Architekturen wird von KHAN et al. (2020, S.5464) beschrieben.

### 2.3.2 Trainingsverfahren

Bisher wird lediglich die Architektur des NN vorgestellt. Es stellt sich die Frage, wie die Gewichtungen in NN bestimmt werden können, um die Trainingsdaten anzupassen. Bei linearen Regressionen oder Klassifizierungsproblemen erfolgt die Aktualisierung der Gewichtungen durch eine analytische Lösung. Allerdings ist bei mehrschichtigen NN-Strukturen oder komplexen Architekturen wie CNN eine analytische Lösung für zahlreiche Gewichtungen nicht realisierbar. Aus diesem Grund wird das Problem durch iterative Methoden wie das Gradientenabstiegsverfahren (engl. Gradient descent, GD) numerisch gelöst.

#### Gradientenabstieg

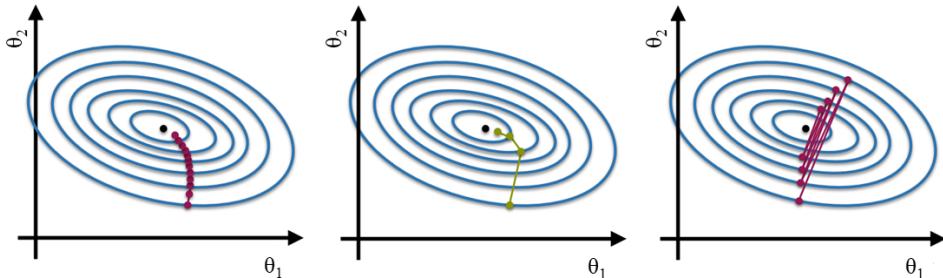
Zur Lösung eines numerischen Optimierungsproblems wird zunächst eine Verlustfunktion definiert. Die Verlustfunktion beschreibt grundsätzlich die Fehler zwischen Modellausgaben und Wahrheitswert. Die Kreuzentropie wird z. B. häufig zur Lösung binärer Klassifizierungsprobleme eingesetzt. Es existieren zudem zahlreiche weitere Verlustfunktionen, die je nach der spezifischen Problemstellung ausgewählt werden können. Das Grundprinzip des Gradientenabstiegsverfahrens liegt darin, die Gewichtungen in Richtung der Minimierung der Verlustfunktion iterativ zu aktualisieren. Die folgende Gl. 2.9 stellt das Gradientenabstiegsverfahren dar.

$$\theta_{i+1} = \theta_i - \alpha_i \nabla E(\theta_i) \quad (2.9)$$

Die negative partielle Ableitung  $-\nabla E(\theta_i)$  charakterisiert die Richtung der maximalen Minimierung der Verlustfunktion in der  $i$ -ten Iteration, während die Lernrate  $\alpha_i$  die Schrittweite

der Aktualisierungen steuert. (SOMANI et al. 2023, S. 145–149)

Die Auswahl einer geeigneten Lernrate ist von grundlegender Bedeutung für das Training. In Abb. 2.14 ist die Auswirkung der Erhöhung der Lernrate veranschaulicht. Bei einer zu geringen Lernrate erfolgt die langsame Konvergenz. Das erfordert mehr Iterationen, um das Minimum zu erreichen. Andererseits kann eine zu hohe Lernrate ebenfalls zu verzögterer Konvergenz oder sogar zur Nichtkonvergenz führen. (SOMANI et al. 2023, S. 149–154)

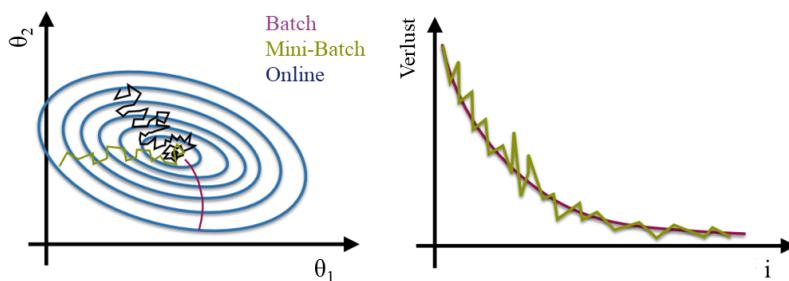


**Abbildung 2.14:** Von links nach rechts: Die Auswirkung der Erhöhung des Lernrates in Anlegung an (JULIJA ZAVADLAV 2022).

### Stochastischer Gradientenabstieg

Die Verlustfunktion wird durch Trainingsdatensatz mit N Datenproben berechnet. In diesem Zusammenhang werden alle Proben verarbeitet, was als eine Epoche bezeichnet wird. Bei einem großen Datensatz ist die Verarbeitung aller Proben in einer Iteration rechenaufwendig. Daher ist es sinnvoll, die Gesamtmenge der Proben auf mehrere Teilmengen aufzuteilen. Auf diesem Prinzip lassen sich im Folgenden verschiedene Arten von numerischen Optimierungsverfahren zuordnen:

- Batch-Algorithmus: der gesamte Datensatz wird in jeder Iteration verarbeitet.
- Online-Algorithmus: in jeder Iteration wird nur eine Datenprobe berücksichtigt. Der Begriff „Online“ wird häufiger von kontinuierlichen Datenströmen verwendet, bei denen keine feste Datensatzgröße vorhanden ist.
- Mini-Batch-Algorithmus: nur ein Teil des Datensatzes ( $K < N$  Proben) wird in Betracht gezogen.

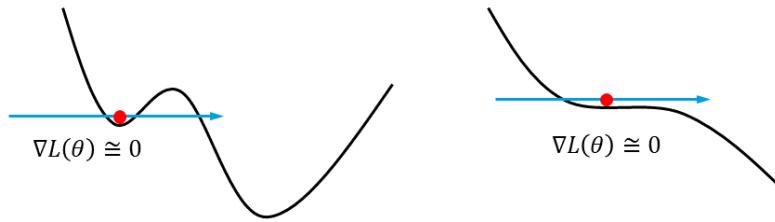


**Abbildung 2.15:** Darstellung der Konvergenz von Batch-, Online- und Mini-Batch-Algorithmen in Anlegung an (JULIJA ZAVADLAV 2022).

Abb. 2.15 stellt die Trajektorie der Konvergenz verschiedener Algorithmen dar. Der Batch-Algorithmus präsentiert eine gleichmäßige Trajektorie. Im Gegensatz dazu veranschaulicht der Mini-Batch-Algorithmus eine verrauchte Trajektorie, während der Online-Algorithmus noch ausgeprägter ist. Allerdings ist bspw. der Mini-Batch-Algorithmus weniger rechenaufwendig und erfordert keinen hohen Speicherbedarf. Ein weiterer Vorteil besteht darin, dass die Konvergenz in lokalen Minima vermieden werden kann. Die Algorithmen der Online- oder Mini-Batch-Algorithmen werden ebenfalls als stochastischer Gradientenabstieg (engl. stochastic Gradient Descent, SGD) bezeichnet. (MONTAVON 2012; TANAKA et al. 2021)

### Effiziente Optimierungsverfahren

Die Abstiegsrichtung kann präziser durch die zweite Ableitung der Verlustfunktion approximiert werden. Dieser Ansatz wird als Newton-Verfahren bezeichnet. Dennoch erfordert die Anwendung der Newton-Methode im Vergleich zur GD einen höheren Rechenaufwand. Im Bereich des Deep-Learning stehen viele weitere Optimierungsverfahren zur Verfügung. Beim GD ergeben sich zwei Herausforderungen, nämlich die Wahl der optimalen Lernrate und der schnellste Weg zum Minimum. Im Sattelpunkt oder im lokalen Minimum wird die Verlustfunktion flach und der Gradient geht an dieser Stelle gegen Null, wie in Abb. 2.16 dargestellt. Daher verlangsamt sich das Lernen. Auf der anderen Seite können verrauchte Gradienten bei



**Abbildung 2.16:** Lokale Minima (links) und Sattelpunkte (rechts)

z.B. ein Mini-Batch-Algorithmus das Lernen ebenfalls verlangsamen. Die folgenden Verfahren werden häufig verwendet, um diese Herausforderungen anzugehen (AGGARWAL 2023; MONTAVON 2012):

- Gradientenabstieg mit Momentum

Statt jeden Schritt unabhängig zu betrachten, werden die aktuellen Gradienten als auch die vorherigen Gradienten berücksichtigt. Die Aktualisierungsregel lautet:

$$\begin{aligned} v_{i+1} &= \beta v_i + \alpha \nabla E(\theta_i) \\ \theta_{i+1} &= \theta_i - v_{i+1} \end{aligned} \tag{2.10}$$

Die Gleichung veranschaulicht, dass der Gradient  $v_{i+1}$  die Summe der exponentiell gewichteten vorherigen Gradienten darstellt. Die Schrittweiten werden in Richtung des Minimums erhöht, während die verrauchte Trajektorie vermindert wird. Normalerweise wird Gewichtung  $\beta = 0,9$  ausgewählt.

- RMSProp-Verfahren (engl. Root Mean Square Propagation, RMSProp)

Beim RMSProp-Verfahren erfolgt eine Skalierung der Summe der quadrierten Gradienten mittels einer Gewichtung  $\beta$ . Anschließend wird der aktuelle Gradient unter Berücksichtigung von  $(1-\beta)$  gewichtet und zum Ergebnis hinzugefügt. Die Aktualisierungsregel lautet:

gel wird im folgenden dargestellt:

$$\begin{aligned} v_{i+1} &= \beta v_i + (1 - \beta) g_i \circ g_i; \quad g_i = \nabla E(\theta_i) \\ \theta_{m,i+1} &= \theta_{m,i} - \frac{\alpha}{\sqrt{v_{m,i+1}} + \epsilon} g_{m,i} \quad m = 1, \dots, M \end{aligned} \quad (2.11)$$

Im RMSProp-Verfahren wird die Lernrate individuell für jeden Gewichtungsparameter optimiert, was zu einer Beschleunigung des Lernprozesses führt.

- Adam-Verfahren (engl. Adaptive Moment Estimation, Adam)

Das Adam-Verfahren (Adam) vereint Momentum und RMSProp-Verfahren. Die Aktualisierungsregel wird im folgende beschrieben:

$$\begin{aligned} m_{i+1} &= \beta_1 m_i + (1 - \beta_1) g_i; \quad \hat{m}_{i+1} = \frac{m_{i+1}}{1 - \beta_1^{i+1}} \\ v_{i+1} &= \beta_2 v_i + (1 - \beta_2) g_i \circ g_i; \quad \hat{v}_{i+1} = \frac{v_{i+1}}{1 - \beta_2^{i+1}} \\ \theta_{j,i+1} &= \theta_{j,i} - \frac{\alpha}{\sqrt{\hat{v}_{m,i+1}} + \epsilon} \hat{m}_{j,i+1} \quad j = 1, \dots, M \end{aligned} \quad (2.12)$$

Der Hyperparameter  $\beta_1$  wird in der Regel auf etwa 0,9 festgelegt, während  $\beta_2$  normalerweise 0,999 beträgt. Für Epsilon wird der Wert 1e-10 gewählt.

## Normalisierung

Im Kontext des Lernprozesses ist es von Bedeutung, eine zügige Konvergenz anzustreben, wobei gleichzeitig die Rechenressourcen effizient eingesetzt werden sollen. Neben der Anwendung geeigneter Optimierungsmethoden stellt die Normalisierung eine weitere Möglichkeit dar. In der praktischen Anwendung werden Datenproben unregelmäßig verteilt, wobei die Mittelwerte von Null abweichen. Dies kann den Lernprozess erschweren. Unterschiedliche Quantitätsniveaus entlang verschiedener Dimensionen führen zu verrauschten Trajektorien und erhöhtem Rechenaufwand. Die Implementierung einer Normalisierung der Eingabedaten mit einem Mittelwert von Null und einer Varianz von eins erweist sich als vorteilhaft für das Training. Durch die Normalisierung wird der Berechnungsaufwand verringert und die zickzackförmige Trajektorie reduziert. Eine erhöhte Lernrate ist daher ermöglicht.

## Fehlerrückführung

Um eine großer Anzahl der Ableitungen effizient zu berechnen, wird eine sog. Fehlerrückführung (engl. Backpropagation) entwickelt. Hierbei wird zunächst ein Vorwärtspass durchgeführt, anschließend wird die Gewichtungen rückwärts aktualisiert. Dieses Verfahren ermöglicht einen geringen Rechenaufwand für die Aktualisierung der Gewichtungen. (SOMANI et al. 2023, S. 143–145)

Für komplexe Aufgaben zeigt ein tieferes NN eine bessere Leistung, da dadurch mehr Merkmale extrahiert werden können. Allerdings können Exploding- oder Vanishing-Gradient bei der Implementierung der Fehlerrückführung auftreten. Beim Vanishing-Gradient werden die Gradienten sehr klein, die durch Multiplikationen entstehen. Beim Exploding-Gradient hingegen entstehen sehr große Gradienten. Beide vorliegenden Probleme haben zur Folge, dass der Trainingsprozess erheblich verlangsamt wird oder das Modell überhaupt nicht in der Lage ist zu lernen. Im Folgenden werden Optimierungsmethoden erörtert, die dazu dienen, diese Schwierigkeiten zu bewältigen.

- Batch-Normalisierung

Die Verteilung der Eingaben jeder Schicht verändert sich während des Trainings, wenn die Gewichtsparameter der vorherigen Schichten sich verändern. Die Normalisierung der Ausgaben jeder Schicht verbessert den Trainingsprozess und unterstützt bei Problemen mit Exploding/Vanishing-Gradient.

- Aktivierungsfunktion

Eine gute Auswahl der Aktivierungsfunktion der NN-Schichten ist notwendig. Zum Beispiel zeigt die ReLU-Funktion eine geringere Anfälligkeit für das Problem des Vanishing-Gradients, daher werden sie meistens in verborgenen Schichten verwendet.

- Initialisierung der Gewichtungen

In der Regel werden die Gewichtungsparameter zufällig mit einem Mittelwert von Null erzeugt, um den Lernprozess zu verbessern und das Problem Exploding-Gradient zu vermeiden. Zum Beispiel wird He-Normalisierung für ReLU-Funktion oder Xavier-Normalisierung für Tanh-Funktion eingesetzt.

### 2.3.3 Methoden zur Modellselektion

Es ist schwer, ein Modell zu suchen, das eine gute Generalisierung aufweist. Über- und Unteranpassung treten beim Training häufig auf. Abb. 2.17 zeigt ein Beispiel von Anpassung einer Sinus-Funktion. Bei der Unteranpassung ist es unmöglich, ein einfaches Modell durch eine Gerade den Datenpunkten anzupassen. Im Gegensatz dazu geht bei der Überanpassung der Fehler gegen Null, allerdings stellt es kein gutes Generalisierungsmodell dar. Daher ist der Fehler aus der Verlustfunktion keine entscheidende Größe für die Modellauswahl. (MONTAVON 2012)

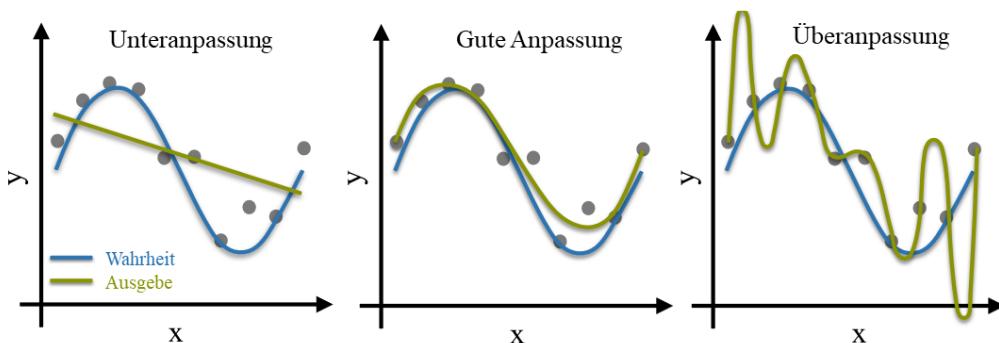


Abbildung 2.17: Die Anpassungen von einer Sinus-Funktion in Anlegung an (JULIJA ZAVADLAV 2022).

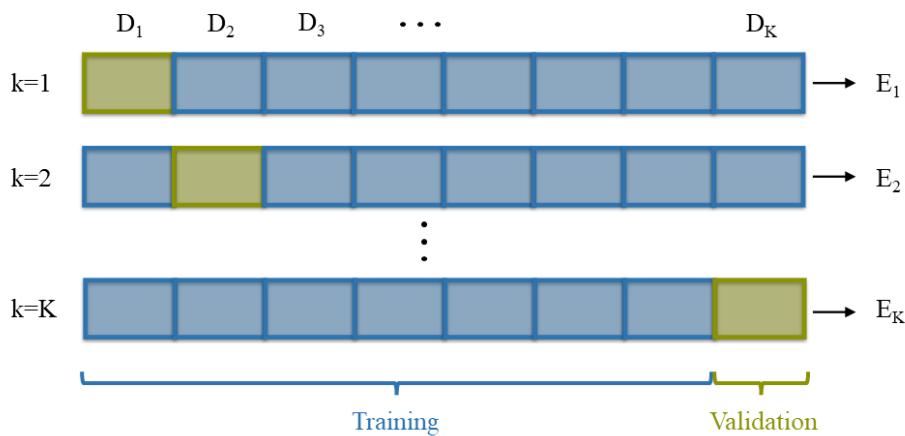
Das Verzerrung-Varianz-Dilemma (engl. bias-variance trade-off) ist ein Grundkonzept für die Modellauswahl, was durch die folgende Gl. 2.13 veranschaulicht. Die Verzerrung beschreibt den Fehler zwischen der Modellausgabe und der Wahrheitswert. Eine hohe Verzerrung führt zu Unteranpassung. Die Varianz zeigt die Empfindlichkeit auf eine kleinere Änderung im Datensatz. Bei hoher Varianz tritt Überanpassung auf. Daher ist es wünschenswert, beide Fehler möglichst gering zu halten. In der Praxis ist der erwartete Verlust nicht erfassbar. Daher werden die Datensätze in Trainings-, Validierungs- und Testdatensätze unterteilt. Der Trainingsdatensatz wird für das Training verwendet. Der Validierungsdatensatz approximiert

den erwarteten Verlust zur Modellauswahl. Die endgültige Leistung des Modells wird durch den Testdatensatz bewertet. (MONTAVON 2012; RASCHKA 2020)

$$\text{Verlust}_{\text{Erwarteter}} = \text{Verzerrung} + \text{Varianz} + \text{Rauschen} \quad (2.13)$$

Ein kleiner Datensatz erschwert die Modellauswahl. Der Validierungsdatensatz sollte durch ausreichende Daten den erwarteten Verlust gut approximieren, während das Modell mit genügend Daten gut trainiert wird. Die K-Fold-Kreuzvalidierung ermöglicht die Ausnutzung der Trainings- und Validierungsdatensätze. Die Konzeption besteht in der mehrfachen Wiederholung der Verwendung von Validierungsdatensätzen, wobei in jeder Wiederholung ein unterschiedlicher Validierungssatz verwendet wird. Bei der K-Fold-Kreuzvalidierung erfolgt der Prozess wie in folgender Abb. 2.18 beschrieben (JULIJA ZAVADLAV 2022):

1. Der Trainingsdatensatz  $D$  wird in  $K$  annähernd gleich große Teilmengen  $D_1, D_2, \dots, D_k$  aufgeteilt.
2. Wiederholung für  $k=1, \dots, K$ :
  - Training durch den Trainingsdatensatz ohne Validierungsdatensatz  $D_k$
  - Berechnung des Validierungsfehlers  $E_k$  des Modells durch Validierungsdatensatz  $D_k$
- Der geschätzte erwartete Verlust wird durch den durchschnittlichen Validierungsfehler berechnet.
3. Das Modell mit dem geringsten erwarteten Verlust wird ausgewählt und mit dem vollständigen Trainingsdatensatz trainiert.



**Abbildung 2.18:** Darstellung für K-Fold-Kreuzvalidierung in Anlegung an (JULIJA ZAVADLAV 2022).

Die K-Fold-Kreuzvalidierung wird häufig in der Optimierung von Hyperparametern eingesetzt. Die Leistungsfähigkeit des Modells mit der ausgewählten Kombination der Hyperparameter wird abschließend durch den Testdatensatz bewertet. (MONTAVON 2012)

Es existieren viele Techniken, die dazu dienen können, Überanpassung zu vermeiden und den Trainingsprozess zu verbessern. Die Überanpassung tritt häufig bei kleinen Datensätzen auf. Dazu gibt es zwei Möglichkeiten: Vergrößerung des Datensatzes und Reduzierung der Modellkomplexität. Im Folgenden werden diese Techniken näher erläutert, um das effiziente Training eines qualitativ hochwertigen Modells zu ermöglichen.

- Data-Augmentation

In der praktischen Anwendung ist es gelegentlich herausfordernd, einen Datensatz durch natürliche Erweiterung zu vergrößern. In solchen Fällen bietet sich die Möglichkeit der künstlichen Datenvermehrung an, wobei Data-Augmentation eine entscheidende Rolle spielt. Zum Beispiel kann ein Datensatz von Bilddaten durch Anwendung von Techniken wie Drehen, Zuschneiden, Anpassung von Kontrast und Helligkeit usw. künstlich erweitert werden. Diese Vorgehensweise zielt darauf ab, die Tendenz zur Überanpassung zu reduzieren und gleichzeitig die Robustheit des Modells zu erhöhen. (GOJIĆ et al. 2023, S. 3–4)

- Regularisierung

Das Optimierungsproblem wird als beschränkten Optimierungsproblem betrachtet. Zum Beispiel erfolgt die Verknüpfung der Verlustfunktion mit der L1-Norm (Lasso) oder der L2-Norm (Ridge). Dadurch werden die Gewichtungsparameter, die mit insignifikanten Merkmalen in Verbindung stehen, auf kleine Werte gedrückt (Ridge) oder genau auf Null gesetzt (Lasso). Bei der Erhöhung des Hyperparameters  $\lambda$  wird die Modellkomplexität verringert. Bei beiden Verfahren werden Modellkomplexität kontrolliert und die Generalisierung verbessert.

- Early-Stopping

Es gibt einen einfachen Weg, die Modellkomplexität zu kontrollieren. Der Validierungsfehler wird während des Trainings beobachtet. Falls sich der Fehler erhöht, wird das Training gestoppt. Dadurch wird die Überanpassung des Modells verhindert.

- Dropout-Verfahren

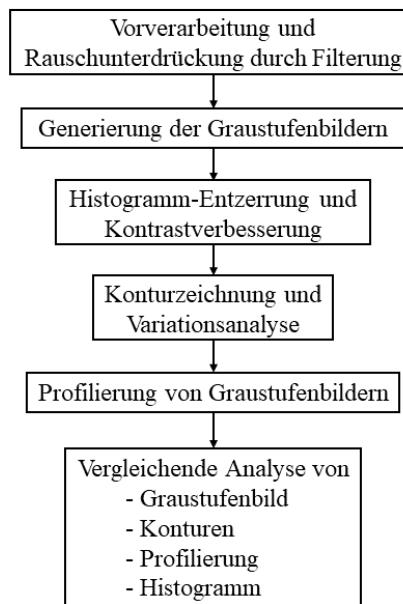
Während des Trainings werden die Verbindungen von Neuronen zufällig ausgeschaltet. In jeder Iteration des GDs werden verschiedene Neuronen ausgewählt und deaktiviert. Nur die Gewichtungen der aktiven Neuronen werden aktualisiert, während die Parameter der deaktivierten Neuronen unverändert bleiben. Typischerweise wird die Wahrscheinlichkeit  $p$  für verborgene Schichten mit 0,5 und für den Eingang mit 0,8 ausgewählt.



# Kapitel 3

## Stand der Forschung und Technik

Frühere Forschungsarbeiten zur automatischen Erkennung von Oberflächenfehlern und Eigenschaften von FSW können in zwei Kategorien unterteilt werden: Die entsprechenden Methoden basieren auf klassischer industrieller Bildverarbeitung und berücksichtigen die Oberflächentopographie der Schweißnähte. Die Mehrheit der veröffentlichten Arbeiten konzentriert sich auf diese erste Kategorie. Darüber hinaus gibt es ebenfalls erste Ansätze zur Bewertung der Schweißnahtoberfläche unter Verwendung von maschinellem Lernen. (HARTL 2021, S. 39)

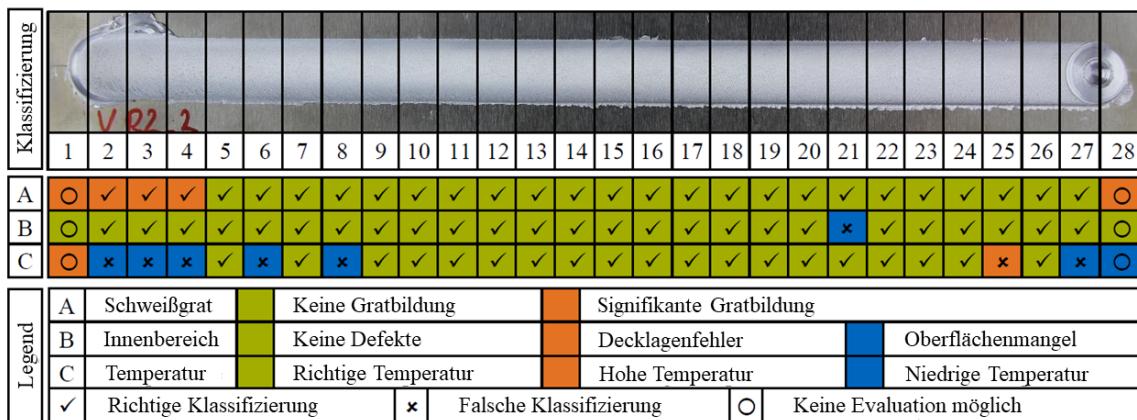


**Abbildung 3.1:** Methodisches Flussdiagramm für die Bildverarbeitung in Anlegung an (SINHA et al. 2008, S. 485).

SINHA et al. (2008) führten eine Untersuchung durch, bei der eine digitale Kamera in eine speziell modifizierte Fräsmaschine integriert wurde, um FSW zu beobachten. Die Kamera befand sich hinter dem Schweißwerkzeug und war mit einem PC verbunden. Dieser PC wiederum war mit einer Maschinensteuereinheit verknüpft. In dieser Studie wurden drei verschiedene Fälle der Schweißnaht betrachtet: intakte Schweißnaht, Schweißnaht, die während eines Probenausfalls entstand, und Schweißnaht, die aufgrund unzureichender Sonden-tiefe erzeugt wurde. Die Bilder wurden offline mithilfe verschiedener Bildverarbeitungstechniken verarbeitet, darunter Vorverarbeitung, Umwandlung von Farbbildern in Graustufenbil-

der, Kontrastverbesserung durch Histogrammausgleich, Kontur- und Linienzeichnung. Durch die Analyse von Histogramm, Kontur- und Linienzeichnung konnten verschiedene Schweißbedingungen und Defekttypen erkannt werden. Das entsprechende methodische Flussdiagramm für die Bildverarbeitung wird in Abb. 3.1 dargestellt. RAJASHEKAR (2013) erweiterte die Methode von SINHA et al. (2008) durch die Durchführung von Röntgenuntersuchungen, um die Schweißnähte in gute und defekte Bereiche zu unterteilen. Zudem wurde eine Methode zur quantitativen Auswertung der Ergebnisse der Linienzeichnung entwickelt, indem sie die Steigung der Linienzeichnung analysierten. SICILAN (2014) setzte die Arbeit von SINHA et al. (2008) fort, indem er die Oberflächenrauheit von Schweißnähten gemessen und deren Beziehung zu den Spitzenwerten in den Histogrammen der Graustufenbilder analysiert hat. RAJASHEKAR und RAJAPRAKASH (2016) verwendeten Bildverarbeitung zur Quantifizierung der Oberflächentextur von Schweißnähten, indem statistische Parameter aus Histogrammen und Grauwertmatrix berechnet wurden. RANJAN et al. (2016) waren die Vorreiter in der Anwendung der Bildverarbeitung, nicht nur zur Identifikation von Oberflächendefekten, sondern auch zu deren Klassifizierung.

BHAT et al. (2015) haben maschinelles Lernen erstmals zur Bewertung der Reibschweißnähte angewendet. Nach der Bildvorverarbeitung wurden verschiedene Merkmale extrahiert. Diese Merkmale wurden in einem SVM verwendet, um die Schweißnähte in die Kategorien gut oder defekt zu klassifizieren. Dabei wurde eine hohe Genauigkeit erzielt. SUDHAGAR et al. (2019) verwendeten eine ähnliche Methode wie in BHAT et al. (2015). Dabei wurde eine andere Methode zur Bildvorverarbeitung und Merkmalsextraktion genutzt.



**Abbildung 3.2:** Kombinierte Detektion und Klassifikation angewendet auf eine Schweißnaht in Anlegung an (HARTL et al. 2019).

In der klassischen Bildverarbeitung und im maschinellen Lernen sind Bildvorverarbeitung und manuelle Merkmalsextraktion erforderlich. Die Berechnungszeit pro Bildern stellt einen entscheidenden Faktor für die Leistungsfähigkeit der Methode dar. Die Studie von HARTL et al. (2019) präsentiert eine fortschrittliche Methode zur Bewertung von Schweißnähten, die mittels DL-Verfahren ermöglicht wird. Dabei entfällt die Notwendigkeit von Bildvorverarbeitung und manueller Merkmalsextraktion. Insgesamt wurden 112 Schweißnähte mit einer Gesamtlänge von 18,4 Metern hergestellt, um CNN-Modelle zu trainieren und zu bewerten. Farbbilder der Schweißnähte wurden unter Verwendung einer digitalen Kamera aufgenommen, während Bilder zur Oberflächenbeschaffenheit der Schweißnähte mithilfe eines dreidimensionalen Profilometers erstellt wurden. Dieser Ansatz umfasste ein zweistufiges Verfahren. In der ersten Stufe ermittelte ein Objektdetektor mithilfe eines CNN die Position der Reibschweißnähte in den Bildern. In der zweiten Stufe erfolgte die Klassifizierung der Oberflächeneigenschaften der Schweißnähte durch ein CNN. Der Objektdetektor erzielte eine be-

achtliche Überschneidung von bis zu 89,5 % bei der Lokalisierung der Reibschweißnähte. Die höchste Genauigkeit bei der Klassifizierung der Oberflächeneigenschaften wurde durch die Verwendung von Topographiebildern erreicht, wobei das CNN-Modell DenseNet-121 eine Klassifizierungsgenauigkeit von 92,1 % erzielte. Abb. 3.2 zeigt die Ergebnisse der Klassifizierungsergebnisse einer Schweißnaht. Dieses Verfahren wurde gemäß der Studie von HARTL (2021) implementiert, um die Überwachung der Oberflächeneigenschaften von Schweißnähten zur Qualitätssicherung zu unterstützen.



# Kapitel 4

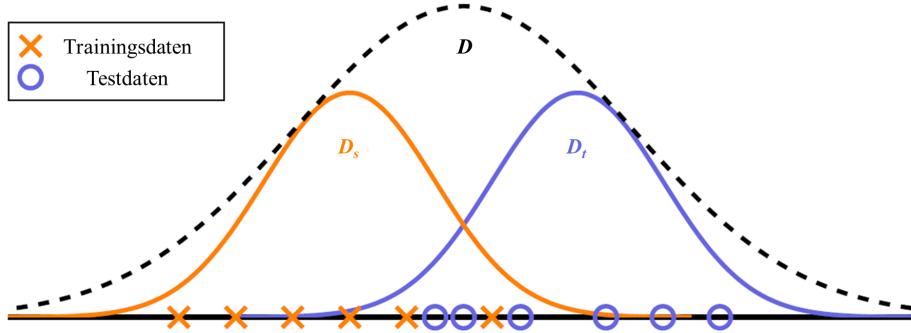
## Generierung der Datenbasis

Das vorliegende Kapitel erläutert die Methode zur Datengenerierung. Abschnitt 4.1 betrachtet die Strategie der Datengenerierung sowie der Identifizierung der zu untersuchenden Einflussgrößen für Bewertung der Robustheit. In Abschnitt 4.2 erfolgt eine ausführliche Vorstellung des roboter-basierten Messstands. Abschnitt 4.3 präsentiert die Umsetzung der Software zur Durchführung der Datengenerierung und die nachfolgende Datenaufbereitung.

### 4.1 Strategie der Datengenerierung

Die Robustheit ist relevante Eigenschaft von DL-Modellen. Dies bezieht sich auf die Fähigkeit des Modells, erwartete Ergebnisse zu erzeugen, wenn die Eingabedaten von den Trainingsdaten abweichen. In der Fachliteratur wird zwischen adversarialer und nicht-adversarialer Robustheit unterschieden. Die adversariale Robustheit beschreibt die Fähigkeit von Modellen, sich gegen feindliche Angriffe zu behaupten. Die Eingabedaten werden absichtlich auf eine Weise verändert, die vom menschlichen Auge nicht erkannt werden kann, um die Modelle zu täuschen. Die nicht-adversariale Robustheit (auch natürliche Robustheit) gewährleistet die Leistung der Modelle trotz natürlicher Veränderungen in der realen Umwelt. Die natürlichen Veränderungen umfassen z. B. unterschiedliche Wetterbedingungen, Verwendung von verschiedenen Sensoren oder andere bekannte und unbekannte Einflüsse (GOJIĆ et al. 2023, S. 1). In der realen Welt sind Datensätze oft natürlichen Einflüssen ausgesetzt. Es kann zu signifikanten Leistungsabfällen von Modellen um 30 % bis 40 % kommen, wenn sie mit natürlichen verschlechterten Daten konfrontiert werden (DRENKOW et al. 2022, S. 1–2). In der vorliegenden Arbeit liegt der Schwerpunkt auf der Untersuchung der natürlichen Robustheit von DL-Verfahren am Beispiel des FSWs.

Im gegenwärtigen Kontext einer standardisierten Maschinenlern-Pipeline wird der Trainingsdatensatz aus einem theoretisch unendlichen, hochdimensionalen Datenspektrum mit einer im Voraus unbekannten Verteilung  $D$  gesammelt, wie in Abb. 4.1 dargestellt. Diese Quellverteilung eines Trainingsdatensatzes bildet die Basis für das Training des Modells. Die Leistungsfähigkeit des Modells wird häufig anhand identisch verteilter Testdaten bewertet. Die Zielverteilung des Testdatensatzes kann allerdings im Vergleich zur Quellverteilung eine Verschiebung aufweisen. Die Ursachen für diese Verteilungsverschiebungen resultieren aus verschiedenen Einflussgrößen in der realen Umwelt. Diese Einflussgrößen können sowohl zeitlicher als auch nicht-zeitlicher Natur sein (GOJIĆ et al. 2023, S. 1). Eine Leistungsverschlechterung kann im Falle einer Verschiebung der Datenverteilung auftreten. Es existieren bereits öffentlich verfügbare Benchmarks zur Bewertung der natürlichen Robustheit. DRENKOW et al.



**Abbildung 4.1:** Veranschaulichung von Verteilungsverschiebung in der Anlegung von (GOJIĆ et al. 2023, S. 1). Eine im Voraus unbekannte reale Datenverteilung  $D$  (dargestellt durch gestrichelte Linien) mit der Quellverteilung  $D_s$  und der Zielverteilung  $D_t$  (dargestellt durch orange und blaue durchgezogene Linien), die im Vergleich zu  $D_s$  verschoben sind.

(2022) und GOJIĆ et al. (2023) liefern eine umfassende Übersicht über zahlreiche Studien, die sich mit Common- und Custom-Benchmarks zur Bewertung der Robustheit befassen. Anhand DRENKOW et al. (2022, S.10) werden die zu untersuchten Einflussfaktoren wie folgt zusammengefasst und aufgelistet:

- Umwelt: Wetter (z.B. Regen, Schnee, Nebel oder Dämmerung), Beleuchtung, Blendung, atmosphärische/mediumbedingte Effekte, Verdeckungen
- Sensor: Rauschen (z. B. Gauß-Rauschen, Impulsrauschen, Schussrauschen), Bewegungsunschärfe, Zoom-Unschärfe, Unschärfe durch Fokussierungsfehler, Linsenverzerrung, Farbabweichung, Kontrast
- Rendern: Digitale Artefakte, Filter, JPEG-/Video-Kompression, Skizzen-/Strichzeichnung, Cartoonisierung, Quantisierung

Allerdings handelt es sich bei den meisten um synthetische Benchmarks, die nicht zwangsläufig eine adäquate Abbildung dieser Verschiebung in den realen Verteilungen sicherstellen (GOJIĆ et al. 2023, S. 5–6). Die tatsächliche Datenverteilung ist schwierig einzuschätzen. Die vorliegende Arbeit verfolgt das Forschungsziel, Evaluationsdatensätze unter realen Umweltbedingungen zu generieren. Die Identifizierung der Einflussgrößen erfolgt sowohl anhand vorhandener Benchmarks als auch durch die Betrachtung praktischer Anwendungen in der industriellen Bildverarbeitung.

Wie im Grundlagenteil erläutert wurde, erleichtert eine hohe Bildqualität die Bildverarbeitung. Die Bildqualität wird unmittelbar von der Umgebung beeinflusst. In der klassischen industriellen Bildverarbeitung wird eine Bildvorverarbeitung durchgeführt, um die entwickelte Methode zur Bewältigung spezifischer Aufgaben robust zu gestalten. Die Arbeit von STEGER et al. (2018) zeigt zahlreiche Anwendungsbeispiele, um die unterschiedlichen Einflussgrößen zu veranschaulichen. Zum Beispiel besitzt die Veränderung der Helligkeit oder des Kontrasts der Bilder, welche durch unterschiedliche Beleuchtungsrichtungen, Lichtintensitäten oder variierende reflektierende Eigenschaften der Objekte hervorgerufen werden, einen signifikanten Einfluss auf die Robustheit. Das Sensorrauschen kann die Qualität der Methode ebenso verschlechtern. Die Einflüsse von unerwünschter Rotationen oder Positionsänderungen sind meistens nicht vernachlässigbar. In praktischen Anwendungen ist es gelegentlich erforderlich, bedeutende Bildverzerrungen oder die teilweise Verdeckung von Objekten zu

berücksichtigen. Bei der Implementierung von maschinellem Lernen in die industrielle Bildverarbeitung, bspw. durch die Anwendung von SVM, sind diese Einflussgrößen ebenfalls nicht vernachlässigbar.

Anhand der Betrachtung in der klassischen industriellen Bildverarbeitung und vorhandener Benchmarks zur Bewertung der Robustheit werden die zu untersuchenden Einflussgrößen oder natürlichen Veränderungen unter Berücksichtigung der Realisierbarkeit des Messstandes wie folgt identifiziert:

- Beleuchtung: Mittels der Änderung der Beleuchtung werden helle und dunkle Bilder natürlich erzeugt.
- Kameraposition: Änderung der Kameraposition erzeugt natürlich rotierte und verzerrte Bilder.
- Unscharfe Fokussierung: Die unscharfen Bilder werden durch abweichende Fokussierung generiert.

In der vorliegenden Arbeit werden die optimalen Bilddaten, die nicht durch Einflüssen verändert werden, für das Training der CNN-Modelle verwendet. Die optimalen Bilddaten werden unter günstigen Aufnahmebedingungen ohne natürliche Veränderungen gesammelt, wodurch eine hohe Bildqualität gewährleistet wird. Dabei sind die Texturen der Schweißnähte und andere Merkmale durch das menschliche Auge deutlich erkennbar, ohne Überbelichtungseffekte. Die unterschiedlichen natürlich verschlechterten Bilddaten werden für Bewertung der Robustheit verwendet.

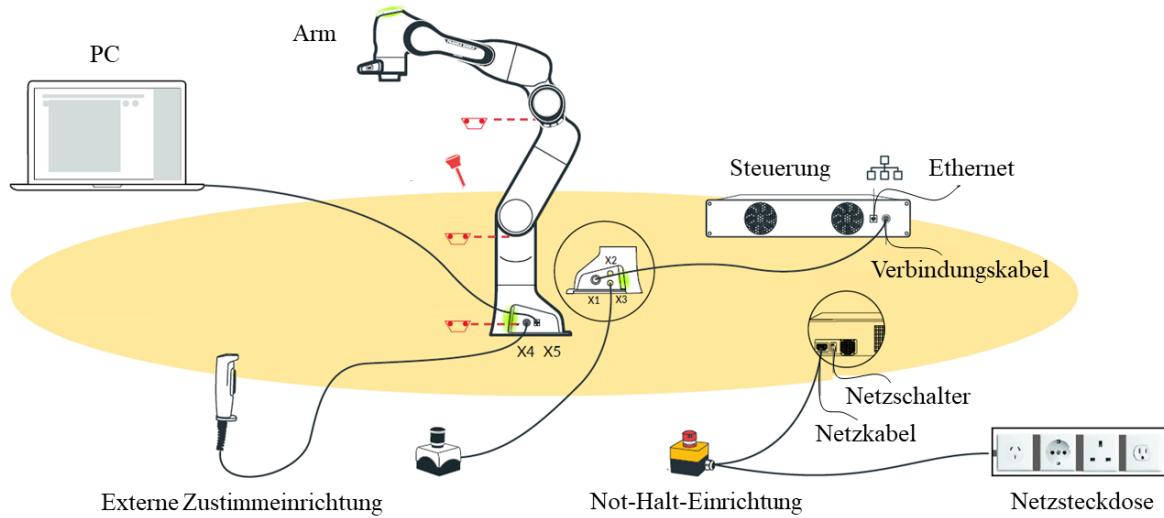
## 4.2 Entwicklung des Roboter-basierten Messstands

Die effiziente Datengenerierung unter variablen Umgebungsbedingungen wird durch den Einsatz eines roboter-basierten Messsystems ermöglicht. Hierbei sind steuerbare Roboter, integrierte Kameratechnik und angemessene Beleuchtungseinrichtungen von entscheidender Bedeutung. Diese Komponenten werden in den folgenden Abschnitten näher erläutert.

### 4.2.1 Übersicht des Panda-Roboters

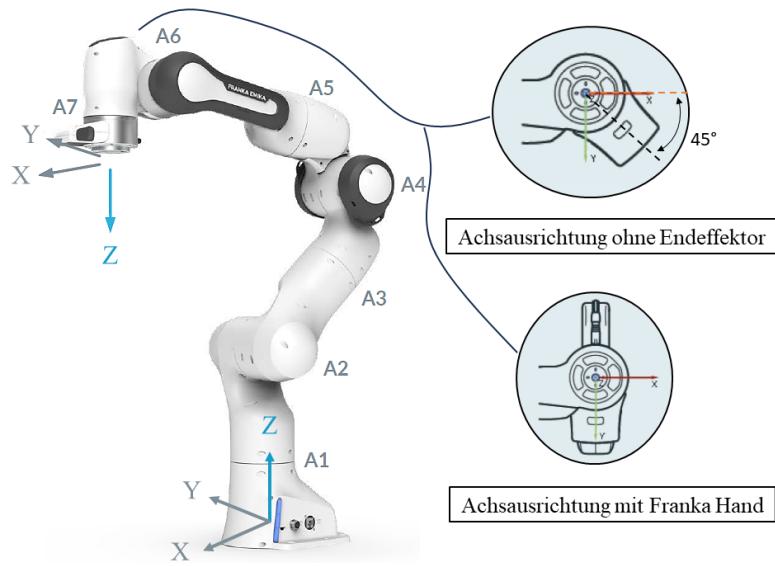
Der Panda-Roboter wurde vom deutschen Deep-Tech-Unternehmen FRANKA EMIKA entwickelt. Der kraftsensitive und bewegliche Arm verfügt über 7 Freiheitsgrade mit Sensoren zur Messung des Drehmoments an jedem Gelenk. Der Roboter erreicht eine industrietaugliche Wiederholgenauigkeit der Position von +/- 0,1 mm und maximale Abweichungen in der Bewegungsbahn von 1,25mm. Die maximale Geschwindigkeit vom Endeffektor kann bis zu 2 m/s erreichen. Der Roboter trägt eine Nutzlast von 3 kg, besitzt eine Reichweite von 855 mm und deckt einen Arbeitsbereich von 94,5 % ab. (FRANKA EMIKA GMBH 2021, S. 34–36)

Es stehen drei verschiedene Programmierschnittstellen für die Bewegungsplanung zur Verfügung: Desk, die integrierte Entwicklungsumgebung für Roboter (RIDE) und die Franka Regelungsschnittstelle (FCI). Desk ist die high-level Schnittstelle des Roboters. Diese erfordert keine Programmierungskenntnisse im Voraus und ermöglicht somit eine schnelle und unkomplizierte Umsetzung von Aufgaben. RIDE stellt das Werkzeug dar, um eigene Desk-Anwendungen zu erstellen und externe Sensoren sowie weitere Funktionalitäten mit dem



**Abbildung 4.2:** Schematische Darstellung der Komponente des Robotersystems in Anlegung an (FRANKA EMIKA GMBH 2021, S. 104).

System zu verbinden. FCI bietet eine direkte und schnelle bidirektionale low-level Verbindung zum Arm und zur Hand des Roboters. Sie gewährt Echtzeitinformationen über den Zustand des Roboters und erlaubt die direkte Steuerung des Drehmoments mit einer Frequenz von 1 kHz. Neben der C++-Schnittstelle *libfranka* besteht ebenfalls die Möglichkeit zur Integration mit den populärsten Ökosystemen ROS, ROS2 und MATLAB-Simulink. Aufgrund dieser direkten Steuerung mit einer 1 kHz-Echtzeitschleife wird FCI als ideale Schnittstelle zur Erforschung neuer Planungs- und Steuerungssysteme verwendet. (FRANKA EMIKA GMBH 2023a)



**Abbildung 4.3:** Achsbezeichnungen mit Achsausrichtung ohne Endeffektor und mit Franka-Hand in der Anlegung von (FRANKA EMIKA GMBH 2021, S. 35) und (FRANKA EMIKA GMBH 2023b, S. 55).

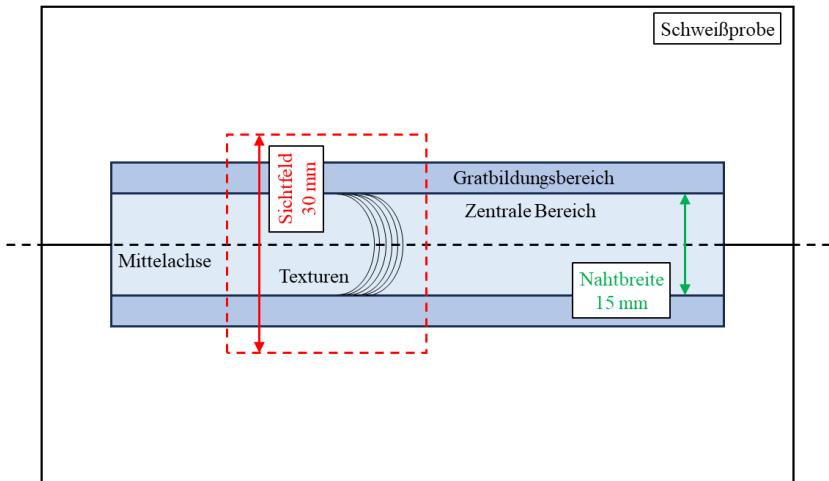
Abb. 4.2 stellt die schematische Darstellung der relevanten Komponente des Robotersystems dar. Während der ersten Implementierung der Programmierung ist externe Zustimmeinrichtung notwendig, die beizubehalten ist, um das Risiko unerwünschter Bewegungen zu mini-

mieren. Durch Betätigung der externen Zustimmeinrichtung wird die Bewegung des Roboters unverzüglich gestoppt. Falls es erforderlich ist, steht zudem eine Not-Halt-Einrichtung zur Verfügung, über die sich die Stromzufuhr des Roboters unterbrechen lässt. Die Sicherheitseinrichtung sollte sich außerhalb des Arbeitsbereichs (gekennzeichnet als gelber Bereich) befinden, um jegliche Gefährdung zu vermeiden.

Bei der Datengenerierung werden die erwünschten Bewegungsplanungen durch ROS und Movelt ermöglicht. Die Datengenerierung erfolgt mittels eines Roboters ohne die Verwendung einer Franka-Hand. Stattdessen wird entsprechende Kameratechnik eingesetzt. Die Achsbezeichnungen sowie die zugehörige Achsausrichtung ohne Endeffektor und mit Franka-Hand werden in Abb. 4.3 veranschaulicht. Bei der Bewegungsplanung ist zu berücksichtigen, dass bei einem Roboter ohne Endeffektor ein Versatz von 45° besteht.

#### 4.2.2 Sensorauswahl

Die Anforderungen für die Kameratechnik bestehen darin, dass bei optimaler Aufnahmebedingung die Texturen der Schweißnähte deutlich erkennbar sind. Die Schweißnaht sollte im Fokus stehen und zentral positioniert sein. In Abb. 4.4 ist ein Sichtfeld der Schweißnaht veranschaulicht. Die Breite der sauberen Schweißnaht beträgt ungefähr 15 mm, während sie



**Abbildung 4.4:** Darstellung von Sichtfeld der Schweißnaht.

mit Gratbildung bis zu 20 mm erreichen kann. Das erforderliche Sichtfeld, das sich vertikal zur Mittelachse der Schweißnaht erstreckt, sollte 30 mm abdecken. Die minimale Größe der Textur beträgt etwa 0,1 mm, was durch mindestens 2 Pixel beschrieben werden sollte. Unter Verwendung von Gl. 2.7 ergibt sich eine minimale Auflösung von 600 Pixel:

$$R_{Kamera} = 2 \text{Pixel} \cdot \frac{30 \text{mm}}{0,1 \text{mm}} = 600 \text{Pixel} \quad (4.1)$$

Die Kamera UI-3220CP-C-HQ (von IDS GmbH) mit einer Auflösung von 752 x 480 Pixel ist verfügbar für die Forschungsanwendung. Die Pixelanzahl von 752 Pixel ist für die aktuelle Aufgabe ausreichend, es sollte berücksichtigt werden, dass durch eine höhere Auflösung eine verbesserte Bildqualität erreicht werden kann. Die Kamera mit einer höheren Auflösung von IDS ist im August nicht bestellbar. Daher wird in der aktuellen Untersuchung die Kamera UI-3220CP-C-HQ eingesetzt. Das detaillierte technische Datenblatt ist im Anhang A.1 dargestellt.

Für die optimale optische Erfassung der Aluminium-Schweißnähte wird die Anwendung einer diffusen Beleuchtung empfohlen. Sie dient dazu, die starken Reflexionen von der Aluminium-Oberfläche zu minimieren sowie Schatten an den Rändern der Schweißnähte zu reduzieren. Die Texturen der Schweißnähte sollten deutlich erkennbar sein. Daher wird eine Hellfeld-Diffuse Auflichtbeleuchtung eingesetzt. Die Lichtintensität sollte variabel eingestellt werden, um den unterschiedlichen Beleuchtungsbedingungen gerecht zu werden. Basierend auf den Anforderungen wurden zwei LED-Panels (LUMIS LED-Panel Bi-Color, von Rollei GmbH & Co. KG) bestellt. Die LED-Panels wurden nah an der Schweißnaht von beiden Seiten angebracht, damit das Störlicht von der Umgebung reduziert werden. LED-Panels bieten eine konstante, flächige und diffuse Beleuchtung. Sie erlauben die Feinjustierung der Farbtemperatur von 3.200 bis 5.600 K in Schritten von 100 K. Ebenso kann die Helligkeit von 0 bis 100 % in Schritten von 1 % präzise reguliert werden. Dadurch wird eine flexible Erzeugung der unterschiedlichen Umgebungsbedingungen ermöglicht.

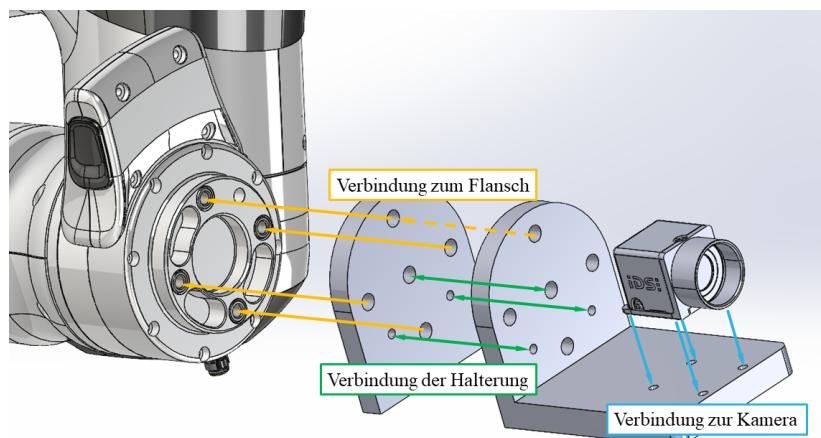
Da das LED-Panel eine Länge von 350 mm und eine Breite von 240 mm aufweist, ist ein Arbeitsabstand für die Bildaufnahme von über 240 mm erforderlich. Das Sichtfeld, wie zuvor dargestellt, sollte ungefähr 30 mm betragen. Unter Verwendung der Gl. 2.3 ergibt sich die Brennweite wie folgt:

$$f = \frac{240\text{mm} \cdot 4.8\text{mm}}{30\text{mm} + 4.8\text{mm}} = 33.1\text{mm} \quad (4.2)$$

Die gewählte Brennweite sollte über 33 mm liegen. Daher eignet sich ein industrielles Objektiv mit einer Brennweite von 35 mm. Bei der Auswahl ist es relevant, den minimalen Arbeitsabstand sowie das Format des Objektives zu berücksichtigen. Die Entscheidung fällt auf das Objektiv IDS-8M118-C3524, welches eine Brennweite von 35 mm und ein Bildformat von 1/1,8" aufweist. Der minimale Arbeitsabstand von 150 mm ist geringer als 240 mm, und das Format des Objektives ist größer als das der Kamera. Detaillierte technische Datenblätter sind in Anhang A.2 aufgeführt.

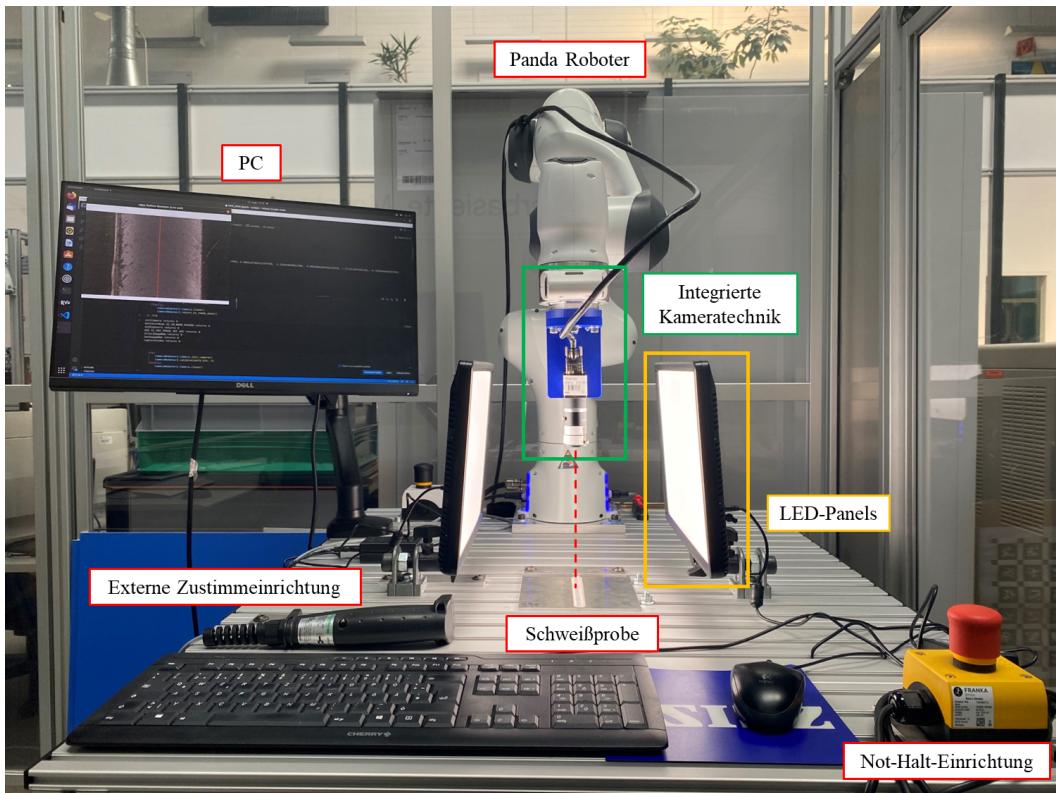
#### 4.2.3 Aufbau des Messstands

Zum Anbringen der Kamera im Roboter wurde eine speziell konzipierte Halterung entwickelt und mithilfe von 3D-Drucktechnologie hergestellt. Die Halterung wird mittels vier M6-Schrauben mit Flansch befestigt. Die große Bohrung in der Mitte mit einem Durchmesser von



**Abbildung 4.5:** Verbindung zwischen Roboter und Kamera.

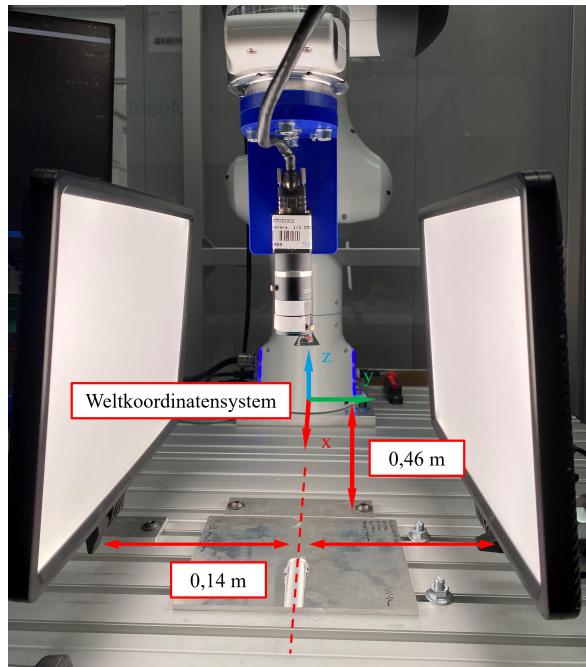
31,5 mm sowie die kleine Bohrung mit einem Durchmesser von 6 mm werden angepasst, um die fixierte Ausrichtung der Halterung sicherzustellen. Da beim 3D-Druck das Bauteil schichtweise von unten nach oben aufgebaut werden muss, sollte die Halterung in zwei Teilen gestaltet werden. Teil A wird mit der Kamera verbunden, während die Bohrungen vom Teil B dem Flansch angepasst wird. Teil B muss ausreichend Platz für die Kabelverbindung zur Kamera bieten. Die minimal erforderlichen Abstände für die C-Mount-Kabelverbindung belaufen sich auf ungefähr 6 cm. Mit Berücksichtigung von Toleranzen werden die Bohrungen etwas größer dimensioniert. Abb. 4.5 stellt die Verbindungen zwischen der Halterung, dem Flansch und der Kamera dar. Die Zeichnungen der Halterung stehen im Anhang A.3 zur Verfügung. Bei der Bewegungsplanung muss beachtet werden, dass die Mittelachse der Kamera nicht mit dem Flansch zentriert ist. Daher sollte diese relative Transformation berücksichtigt werden.



**Abbildung 4.6:** Aufbau des Messstandes.

In Abb. 4.6 wird der Aufbau des Messstands dargestellt. Der Roboter befindet sich am Rand des Labortisches, der die Abmessungen 1,08 m in der Länge, 0,82 m in der Breite und 0,95 m in der Höhe aufweist. Die Schweißnaht wird entlang der X-Achse des Weltkoordinatensystems des Roboters in der Mitte des Tisches positioniert. Dies ermöglicht eine einfache Aufnahme der Schweißprobe durch eine lineare Bewegung des Roboters. Die externe Zustimmeinrichtung sowie die Not-Halt-Einrichtung sind leicht zugänglich, um sicherzustellen, dass unerwünschte Bewegungen sofort gestoppt werden können.

Die Schweißproben werden mithilfe eines Metallblattes und zwei Schrauben sicher positioniert, damit die Schweißproben beim Wechseln immer in derselben Position platziert werden. Das Metallblatt befindet sich in 0,46 m Entfernung von der Weltkoordinate des Roboters. Durch Justierung der zwei Schrauben wird die Schweißnaht entlang der X-Achse der Weltkoordinate des Roboters ausgerichtet. Aufgrund der unregelmäßigen Geometrie der Schweißproben kann es vorkommen, dass eine Probe schief positioniert ist, was durch Justierung



**Abbildung 4.7:** Positionierung der Schweißprobe und Anordnung der Beleuchtung in Bezug auf das Weltkoordinatensystem.

korrigiert werden muss. Die LED-Panels sind ungefähr 0,14 m von der Schweißnaht entfernt angebracht und leicht um 5° nach unten geneigt. Abb. 4.7 veranschaulicht die Positionierung der Schweißprobe und die Anordnung der Beleuchtung in Bezug auf das Weltkoordinatensystem.

### 4.3 Umsetzung der Software zur Datengenerierung

Nach der Betrachtung des Aufbaus des roboter-basierten Messstands wird ein halbautomatisches Konzept zur effiziente Datengenerierung entwickelt. Im Folgenden wird zunächst die verwendete Softwareumgebung beschrieben. Anschließend wird dieses Messkonzept erläutert und abschließend wird die Datenaufbereitung für den Trainings- und Evaluationsprozess betrachtet.

#### 4.3.1 Softwareumgebung

Die Softwareentwicklung erfolgt in Verbindung mit ROS und Movelt. Das ROS ist eine der beliebtesten Robotik-Middleware und wird von Universitäten und Industrien für roboterspezifische Anwendungen verwendet (GANDHINATHAN 2019). ROS ist ein Open-Source-Software-Entwicklungskit für Anwendungen von Robotik. Es bietet Entwicklern aus unterschiedlichen Branchen eine standardisierte Softwareplattform, die sie von der Forschung und Prototypenerstellung bis hin zur Implementierung und Produktion begleitet (STANFORD ARTIFICIAL INTELLIGENCE LABORATORY 2023). Movelt ist eine benutzerfreundliche integrierte Entwicklungsplattform, die aus einer Reihe von Funktionspaketen für mobile Operationen besteht. Diese Pakete umfassen Bewegungsplanung, Betriebssteuerung, 3D-Wahrnehmung, Kinematik

sowie Steuerungs- und Navigationsalgorithmen. Die Plattform bietet eine benutzerfreundliche grafische Benutzeroberfläche, die in Industrie, Handel, Forschung und anderen Bereichen eingesetzt werden kann. Diese Funktionspakete gehören zu den drei am häufigsten genutzten Paketen in der ROS-Gemeinschaft (COLEMAN et al. 2014).

FCI ermöglicht eine schnelle und direkte bidirektionale Verbindung auf niedriger Ebene zum Arm und zur Hand. Sie liefert den aktuellen Status des Roboters und ermöglicht dessen direkte Steuerung mit einem externen PC, der über Ethernet verbunden ist. Durch C++-Schnittstelle *libfranka* werden Echtzeitsteuerungswerte mit einer Frequenz von 1 kHz übertragen und ebenfalls Messwerte des Roboters empfangen. Das Paket *franka\_ros* verbindet den Panda-Roboter mit dem gesamten ROS-Ökosystem. Es integriert *libfranka* in ROS. Zusätzlich enthält es URDF-Modelle und detaillierte 3D-Informationen des Roboters und Endeffektoren, die eine Visualisierung (z. B. in RViz) und kinematische Simulationen ermöglichen. Die Integration von MoveIt erleichtert das Bewegen des Roboters und die Steuerung des Greifers. MoveIt bietet die ROS-Knoten *move\_group* und die entsprechende Schnittstelle aus dem Python-Paket *moveit\_commander* an, um eine kollisionsfreie Bewegungsplanung zu ermöglichen. (FRANKA EMIKA GMBH 2023a)

In der vorliegenden Arbeit wird die entwickelte Software auf einem Linux-Betriebssystem mit Echtzeit-Kernel ausgeführt. Hierfür wird eine Konfiguration auf einem PC mit Ubuntu 20.04 LTS und ROS Neotic verwendet. Das Paket *franka\_ros* wird auf dem Linux-System installiert und *panda\_moveit\_config* wird zur Verfügung gestellt. Die implementierte Software wurde mithilfe von Python erstellt. Die Kamerasteuerung erfolgt mittels des Python-Pakets *pyueye*. Für die Visualisierung und Bearbeitung der Bilder kommt *OpenCV* zum Einsatz. Die nachfolgende Tabelle enthält detaillierte Versionen der Pakete im Kontext der Softwareentwicklung.

|               |  |
|---------------|--|
| Linux         | Ubuntu 20.04 LTS mit Echtzeit-Kernel der Version 5.9.1 |
| ROS-Framework | Noetic Ninjemys  |
| Python        | Version 3.7  |
| Pyueye        | Version 4.96.952                                       |
| OpenCV        | Version 4.8.0.76                                       |

**Tabelle 4.1:** Tabellarische Darstellung der Versionen des Linux-Betriebssystems und der installierten Pakete.

Eine integrierte Entwicklungsumgebung (engl. Integrated Development Environment, IDE) VSCode (Visual Studio Code) wird in der Softwareentwicklung verwendet. VSCode ist eine hochwertige IDE mit einer Vielzahl von erweiterbaren Plugins. Insbesondere eignet sich VSCode für kleine Projekte und Skripte aufgrund seiner Schnelligkeit und Benutzerfreundlichkeit. Ein weiterer Vorteil besteht darin, dass Jupyter-Notebook nahtlos in VSCode integriert werden kann. Die Verwendung von Jupyter-Notebook ermöglicht die Ausführung von Programmcodes in Blockstrukturen, wodurch die Softwareentwicklung effizienter gestaltet wird. In der vorliegenden Arbeit zeigt sich der besondere Vorteil der Verwendung von Jupyter-Notebook für Testzwecke. Es ermöglicht eine schrittweise Steuerung des Roboters in Echtzeit und die Ausführung von Aktionen des Roboters in sequenziellen Blöcken. Dies erhöht die Effizienz der Programmierung.

### 4.3.2 Halbautomatisches Konzept der Datengenerierung

Das Ziel der Entwicklung eines halbautomatischen Konzepts zur Datengenerierung besteht darin, die optische Erfassung von Schweißproben effizient und präzise zu gestalten. Hierbei wird ein Roboter mit integrierter Kameratechnik eingesetzt, um eine vielfältige Erfassung von Bildern aus unterschiedlichen Kamerapositionen zu ermöglichen. Die einstellbare Beleuchtung dient zur Erzeugung von verschiedenen Helligkeitsänderungen. Der Bildaufnahmeprozess erfolgt durch die Ausführung eines schrittweisen Loop-Programms.

Die Software besteht aus drei Python-Dateien und einer Jupyter-Notebook-Datei. Die Funktionalitäten der Dateien sind in Abb. 4.8 veranschaulicht. Die Datei *camera\_roboter.py* stellt

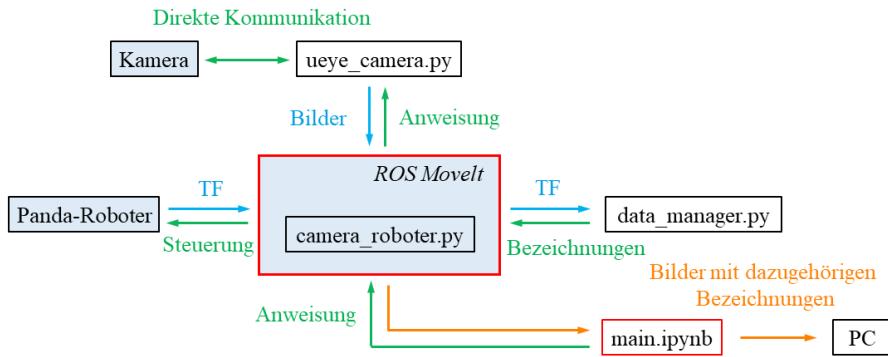
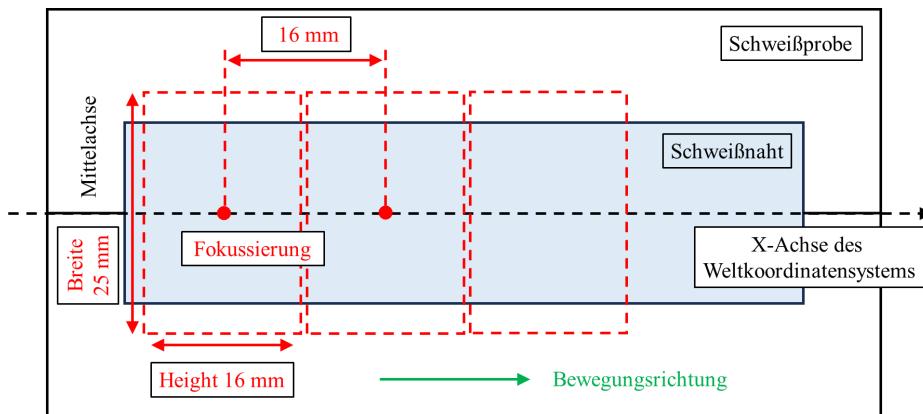


Abbildung 4.8: Grafische Darstellung der Komponenten der Software.

die Funktionen für die Steuerung des Roboters zur Verfügung, wie z. B. eine Durchführung von linearen Bewegungen oder Veränderungen der Kameraposition. Durch diese Funktionen wird ebenfalls die präzise Erfassung der Position, Orientierung sowie der Gelenkwinkel des Roboters ermöglicht. Sämtliche Funktionen sind in der Klasse *CameraRoboter* implementiert. Bei der Instanziierung eines Objekts der Klasse *CameraRoboter* im Hauptprogramm werden zwei weitere Klassen initialisiert: *UeyeCamera* und *DataManager*. Diese Klassen sind in den Dateien *ueye\_camera.py* bzw. *data\_manager.py* implementiert. *UeyeCamera* ermöglicht die Initialisierung der Kamera und die Aufnahme von Bildern. *DataManager* übernimmt die Zuordnung von automatisch generierten Bezeichnungen zu den aufgenommenen Bildern.

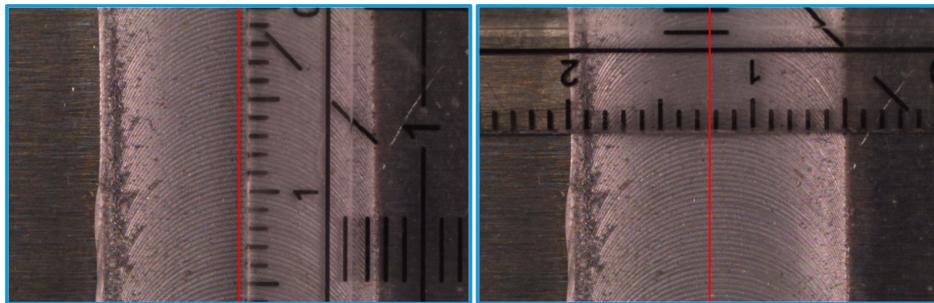
Im Hauptprogramm wird ein ROS-Knoten erstellt, der zur Steuerung verschiedener Aktionen des Roboters dient. Dieser Knoten stellt verschiedene Schnittstelle bereit, über die bspw. Bewegungsaktionen des Roboters, Positions- und Gelenkwinkelerfassung sowie Befehle zur Bildaufnahme ausgeführt werden können. Die Aktionen des Roboters können zunächst gefahrlos in der Simulation Gazebo durchgeführt werden und anschließend im realen Raum umgesetzt werden. Bei der Bewegungsplanung ist es notwendig, die Kollisionsgefahren mit Objekten wie LED-Panels oder Bildschirmen vom PC zu berücksichtigen. Die Visualisierung des Arbeitsraums erfolgt mithilfe von RViz.

Die optische Erfassung einer vollständigen Schweißprobe kann mittels einer linearen Bewegung des Roboters durchgeführt werden. Die Arbeitsdistanz der Kamera wird so eingestellt, dass ein Sichtfeld von 25 mm Breite und 16 mm Höhe erzeugt wird. Je nach Länge der Schweißprobe erfolgt die schrittweise Aufnahme der Probe durch eine lineare Bewegung in Intervallen von jeweils 16 mm. Abb. 4.9 veranschaulicht die optische Erfassung einer gesamten Schweißprobe durch eine lineare Bewegung. Die Kamera wird oberhalb der Schweißprobe positioniert und zentriert auf die Schweißprobe fokussiert. Die Bildebene sollte parallel zur Ebene der Schweißprobe verlaufen. Diese Konfiguration ermöglicht eine exakte Erfas-



**Abbildung 4.9:** Schematische Darstellung der optischen Erfassung einer Schweißprobe während der linearen Bewegung des Roboters.

sung der Merkmale der Schweißnaht und minimiert gleichzeitig jegliche Bildverzerrungen. Die Darstellung der Sichtfeldabmessungen in der vorliegenden Arbeit wird in Abb. 4.10 veranschaulicht. In einer linearen Bewegung erfolgt lediglich eine Veränderung der Kamerapo-

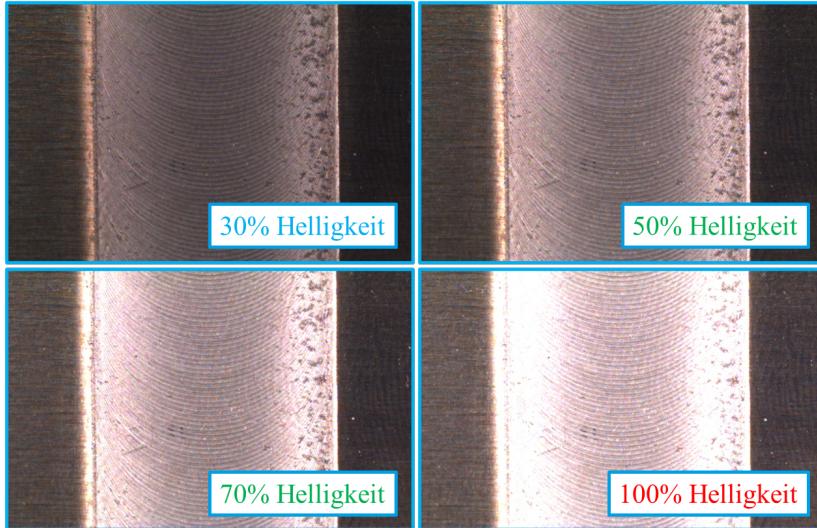


**Abbildung 4.10:** Darstellung der Sichtfeldabmessungen in einem Bild der Schweißprobe.

sition entlang der X-Achse des Weltkoordinatensystems, während Orientierung der Kamera und Position entlang der Y- und Z-Achse unverändert bleiben. Dies bietet den Vorteil, dass die Ausgangsposition angepasst werden kann und während der linearen Bewegung die Pose beibehalten wird, wobei lediglich die Position entlang der X-Achse modifiziert wird. Dies ermöglicht eine effiziente Aufnahme der Schweißnähte mit variierender Kameraposition.

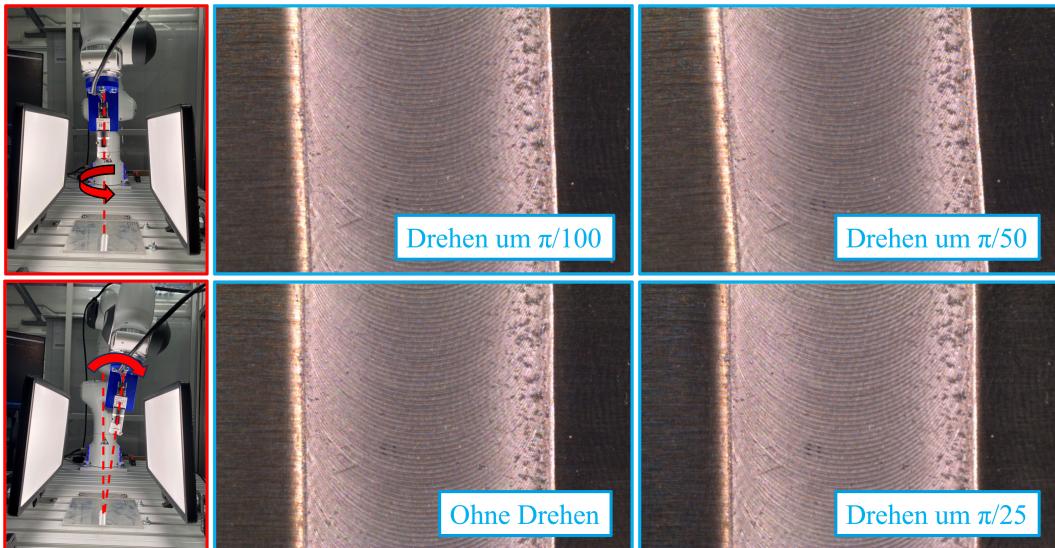
Um die natürliche Helligkeitsänderung zu erzeugen, werden in der vorliegenden Arbeit vier verschiedene Helligkeitsstufen festgelegt. Eine optimale Bildqualität sollte erreicht werden, wenn die Helligkeit der LED-Panels auf 50 % eingestellt ist. Zu diesem Zweck wird die Belichtungszeit auf 10 ms und die Blende auf 1/4 eingestellt. Diese Einstellungen gewährleisten, dass die Texturen der Schweißprobe klar vom menschlichen Auge erfasst werden können und gleichzeitig ebenfalls der Rand oder Gratbildung der Schweißnaht oder andere Details deutlich erkennbar sind, wie in Abb. 4.11 veranschaulicht wird. Bei einer Helligkeitsstufe von 70 % sind die Texturen der Schweißnaht ebenfalls erkennbar. Allerdings gibt es am Rand der Schweißnaht etwas Überbelichtung. Bei den übrigen Helligkeitsstufen geht mehr Detail verloren. Die optimalen Bilddaten werden erzeugt, indem eine lineare Bewegung entlang der Schweißnaht bei einer Helligkeitsstufe von 50 % durchgeführt wird. Bei anderen Helligkeitsstufen werden helle und dunkle Bilder natürlich erzeugt. Belichtungszeit und Blende werden im Verlauf der Untersuchung unverändert beibehalten. Die Modifikation der Helligkeitsstufen erfolgt durch manuelle Einstellung des LED-Panels.

Die natürlich rotierten und verzerrten Bilddaten aus verschiedenen Kamerapositionen wer-



**Abbildung 4.11:** Darstellung von Schweißprobenbildern bei verschiedenen Helligkeitsstufen.

den durch Veränderungen der Pose des Roboterflansches erzeugt. Die Kameradrehung erfolgt durch Änderungen des Drehwinkels um die Z-Achse des Koordinatensystems vom Flansch. In der vorliegenden Arbeit werden zwei Drehstufen in positiver Richtung um die Z-Achse definiert. Abb. 4.12 zeigt die rotierten Bilder nach einer Kameradrehung um einen Drehwinkel von  $\pi/100$  bzw.  $\pi/50$ . Eine zusätzliche Möglichkeit zur Variation der Kameraposition besteht

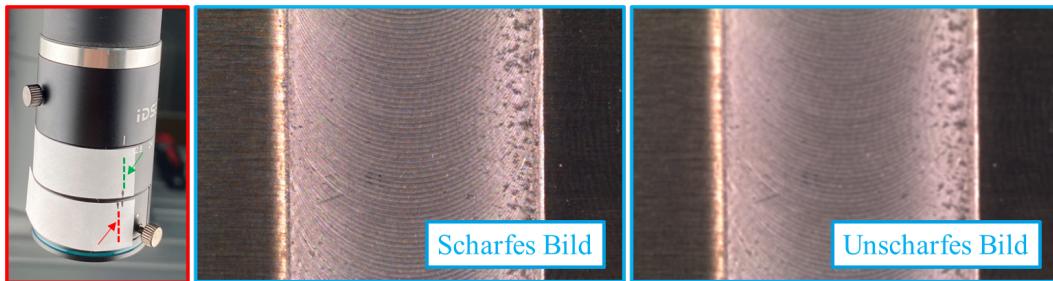


**Abbildung 4.12:** Darstellung von Schweißprobenbildern bei verschiedenen Drehstufen.

in der Rotation um die X-Achse des Weltkoordinatensystems, während der Fokus auf die Mitte der Schweißnaht gerichtet ist. Abb. 4.12 veranschaulicht die verzerrten Bilder nach einer Rotation um die X-Achse von  $\pi/25$ . Die beiden Bilder weisen kaum Unterschiede auf, was bedeutet, dass keine signifikanten Veränderungen vorhanden sind. Dennoch werden diese Daten erfasst, da sich dadurch möglicherweise die Ergebnisse verschlechtern könnten. Bei einer signifikanten Gratbildung sind die Unterschiede aber deutlicher.

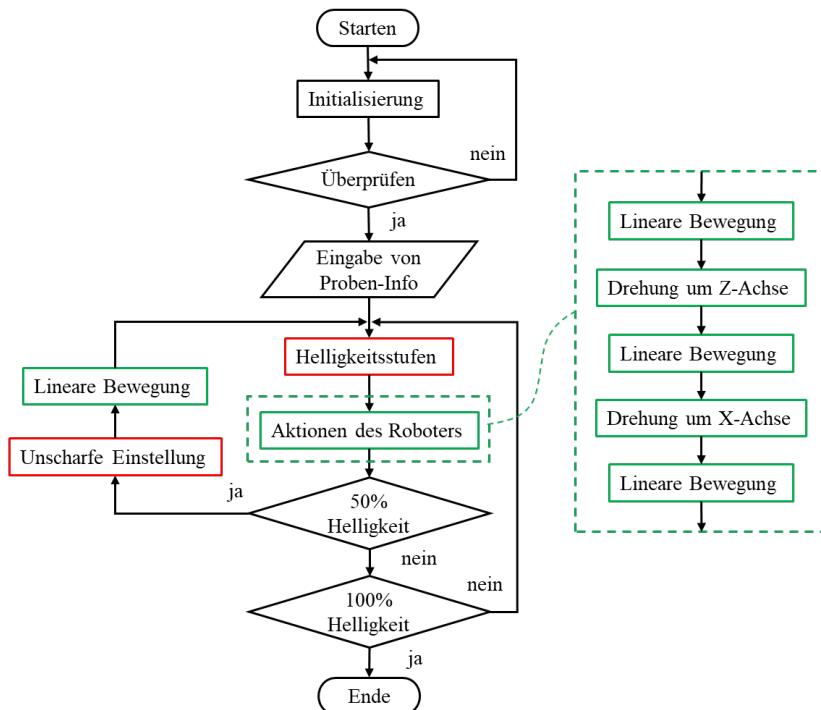
Die unscharfen Bilder werden durch manuelle Einstellung des Objektives erfasst. Die Stelle, die scharfe Bilder erzeugt, wird markiert. Eine unscharfe Stelle wird ebenfalls markiert, sodass eine identische Unschärfe erzeugt werden kann. Abb. 4.13 stellt das unscharfe Bild

dar. Die unscharfe Stelle wird so ausgewählt, dass die Texturen der Schweißnaht weiterhin in gewissem Maße erkennbar bleiben.



**Abbildung 4.13:** Darstellung von Bildern unter Verwendung unscharfer Fokussierung. Grüner Pfeil zeigt scharfe Stelle und Roter Pfeil zeigt unscharfe Stelle.

In der vorliegenden Studie werden nicht nur einzelne Einflussfaktoren untersucht, sondern auch die gleichzeitige Betrachtung mehrerer Einflussfaktoren. Insgesamt werden 16 verschiedene Szenarien betrachtet. Sie können mittels der Roboterbewegung und der Einstellung der Sensoren realisiert werden. Die Bilder z. B. unter guten Aufnahmebedingungen sowie jene bei variiertem Kameraposition werden durch eine automatische Bewegungsplanung ermöglicht. Hierfür sind verschiedene Aktionen des Roboters erforderlich, wie zuvor erläutert: lineare Bewegung, Drehung um die Z-Achse und Drehung um die X-Achse. Die Handhabung von unscharfen Bildern und Helligkeitsänderungen erfolgt manuell. Auf dieser Grundlage kann ein halbautomatisches Konzept zur effizienten Datengenerierung entwickelt werden. Die grafische Darstellung des Prozesses der Datengenerierung ist in Abb. 4.14 veranschaulicht.



**Abbildung 4.14:** Flussdiagramm zur Programmierlogik des halbautomatischen Konzepts zur effizienten Datengenerierung. Der Prozess in grüner Farbe wird automatisch durch den Roboter durchgeführt, während der Prozess in roter Farbe manuell ausgeführt wird.

Die halbautomatische Datengenerierung erfolgt durch das Ausführen des Hauptprogramms

*main.ipynb*. Bevor dieser Vorgang gestartet wird, müssen der Roboter, die Kamera und die ROS-Umgebung bereitgestellt werden. Die Schweißprobe sollte sich an einer markierten Stelle befinden und die Beleuchtung manuell auf 30 % eingestellt werden. Zu Beginn wird der Roboter zur Bereitschaftsposition bewegt, und ein Bild von der Schweißprobe wird auf dem PC angezeigt, wie in Abb. 4.15 dargestellt. Die rote Linie stellt die Mittellinie des Bildes dar.

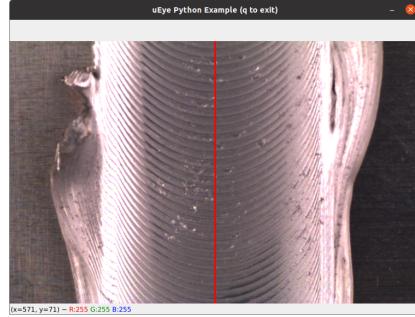


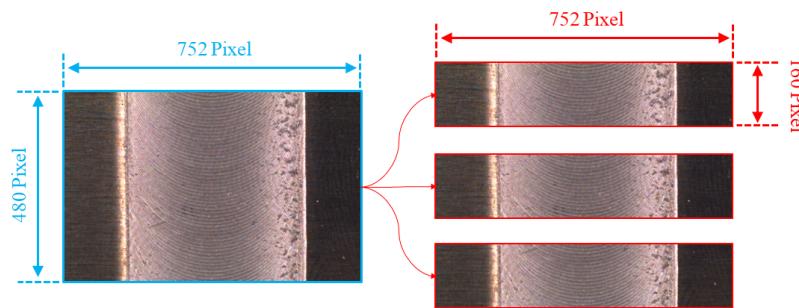
Abbildung 4.15: Darstellung des Bildes während des Überprüfungsprozesses.

Hierbei ist zu überprüfen, ob sich die Schweißprobe in der Mitte des Bildes befindet und ob die Helligkeit korrekt eingestellt ist. Aufgrund der unterschiedlichen Geometrie der Schweißprobe kann es gelegentlich vorkommen, dass die Schweißprobe leicht schief ausgerichtet ist. In solchen Fällen ist eine Neupositionierung der Schweißprobe erforderlich. Dieser Prozess ist notwendig, da die optimalen Bilddaten für das Training nicht von variierenden Kamerapositionen beeinflusst werden sollten. Nach der Überprüfung werden die Informationen der Schweißprobe, wie das Erzeugungsdatum, die Vorschubgeschwindigkeit und die aufgewendeten Kräfte, in die Software eingegeben. Diese Informationen dienen dazu, diejenige Probe zu kennzeichnen, die für das darauf bezogene Bild verwendet wird. Anschließend folgt der Aufnahmeprozess. Hierbei werden für verschiedene Helligkeitsstufen unterschiedliche Posen eingenommen, gefolgt von einer linearen Bewegung zur Aufnahme der Bilder der Schweißprobe. Lediglich bei einer Helligkeit von 50 % werden unscharfe Bilder erzeugt. Die Helligkeitsstufen werden manuell von der niedrigsten zur höchsten Einstellung verändert. Dieses Messkonzept ermöglicht die effiziente Generierung zahlreicher Bilder mit unterschiedlichen Kamerapositionen, variablen Bewegungen sowie vielfältigen Einstellungen der Sensoren und Beleuchtungen.

### 4.3.3 Datenaufbereitung

Im Verlauf dieser Forschung wurden insgesamt 32 Schweißproben verwendet. Hierbei wurden 192 Bilder mit optimaler Bildqualität und 3072 Bilder mit natürlichen Veränderungen erfasst. Ein Bild der Schweißprobe wird in jeweils 3 Teilbilder aufgeteilt, die anschließend gelabelt werden. Die Aufteilung der Bilder wird mithilfe eines Python-Skripts durchgeführt. Die nachfolgende Abb. 4.16 veranschaulicht die aufgeteilten Bilder. Die Bezeichnungen der Teilbilder werden im Format „Probe-Info\_Szenario\_Position“ präsentiert. In der vorliegenden Arbeit sind insgesamt 17 verschiedene Szenarien mit unterschiedlichen Einflussgrößen vorhanden. Ein Beispiel hierfür ist die Berücksichtigung einer Helligkeit von 50 %, ohne Drehung und keine unscharfe Einstellung. Die Bilder unter diesem Szenario bilden den Trainings- und Testdatensatz, während die Bilder in anderen Szenarien den Evaluationsdatensatz bilden.

In der vorliegenden Arbeit liegt der Fokus auf der Gratbildung des Schweißvorgangs. Die Konzepte zur Gratbildung sind bereits im Grundlagenteil ausführlich erörtert worden. Gemäß der Norm DIN EN ISO 25239-5 (BEUTH VERLAG GMBH 2020b) wird das Label mit Null festgelegt,



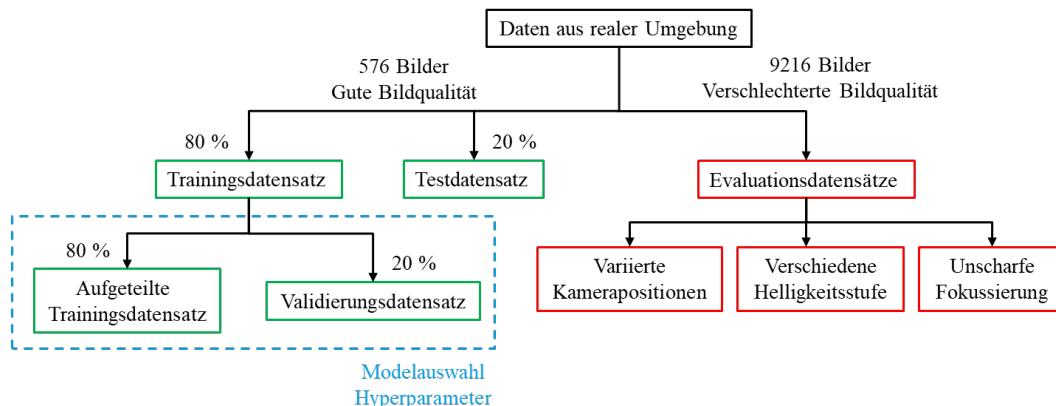
**Abbildung 4.16:** Darstellung der Teilbilder mit entsprechender Bildgröße.

wenn keinerlei signifikantes Überfließen von Materialien am Rand der Schweißnaht oder unregelmäßige Ränder vorhanden sind. Alles andere wird als Defekt klassifiziert. Ein Beispiel von Labels wird in Abb. 4.17 veranschaulicht. Die Labels zu den einzelnen Abschnitten der Schweißnaht werden in entsprechenden JSON-Dateien gespeichert. Insgesamt stehen für den



**Abbildung 4.17:** Darstellung der Gratbildung von zwei Schweißnähten. Das Bild ist in blauer Farbe mit dem Label von Null und in roter Farbe mit dem Label von Eins gekennzeichnet.

Forschungszweck 9792 Bilder mit zugehörigen Labels zur Verfügung. Davon gibt es 576 optimale Bilder und 9216 natürlich verschlechterte Bilder. Abb. 4.18 zeigt die Unterteilung der Bilddaten in unterschiedliche Datensätze für Training und Evaluation. Die optimalen Bilder,



**Abbildung 4.18:** Schematische Darstellung der verwendeten Datensätze für Training und Evaluation.

die nicht natürlichen Einflüssen ausgesetzt waren, werden in Trainings- und Testdatensätze aufgeteilt. Der Trainingsdatensatz wird weiterhin in Trainings- und Validierungsdatensatz unterteilt. Alle verschlechterten Bilder werden nach betrachteten Szenarien in 16 Evaluationsdatensätze aufgeteilt, die zur Bewertung der Robustheit dienen. Alle Bilddaten wurden in NumPy-Arrays konvertiert und normalisiert. Dieser Prozess wurde unter Verwendung eines Python-Skripts durchgeführt. Abschließend wird eine JSON-Datei generiert, die eine Liste von Dateipfaden für die Bilder in Verbindung mit den zugehörigen Labels enthält.



# Kapitel 5

## Bewertung der Robustheit

In diesem Kapitel werden die Ergebnisse zur Bewertung der Robustheit präsentiert und diskutiert. Im Abschnitt 5.1 wird das Vorgehen zur Beurteilung der Robustheit ausführlich erläutert. In den Abschnitten 5.2 und 5.3 werden der Trainings- und Bewertungsprozesse detailliert beschrieben. Abschnitt 5.4 widmet sich der eingehenden Diskussion der erzielten Ergebnisse.

### 5.1 Vorgehen zur Bewertung der Robustheit

Wenn ein Modell bei Instanzen aus einer anderen Verteilung im Vergleich zu derjenigen, auf der es trainiert wurde, gut abschneidet, demonstriert es Robustheit gegenüber Out-of-Distribution (OOD). Wenn das Modell konstant gute Leistungen bei Instanzen aus derselben Verteilung erbringt, demonstriert es Robustheit gegenüber In-Distribution (ID), wodurch eine stabile Leistungsfähigkeit bei unbeobachteten Daten aus derselben Verteilung gewährleistet wird (GOJIĆ et al. 2023, S. 2). Die ID-Robustheit wird zur Auswahl von Modellen verwendet, um sowohl Überanpassung als auch Unteranpassung festzustellen. Die K-Fold-Kreuzvalidierung folgt denselben Prinzipien. Derzeit existieren keine strengen und einheitlichen mathematischen Definitionen für die Robustheit (DRENKOW et al. 2022, S. 8). In LANGROS et al. (2019, S.3) wird ein spezifischer Wert definiert, um die Robustheit zu quantifizieren. In der vorliegenden Arbeit wird die Robustheit ebenfalls auf ähnliche Weise repräsentiert, wie in der folgenden Gleichung dargestellt:

$$R_{Modell} = \frac{G_{Evaluation}}{G_{Test}} \quad (5.1)$$

Der Robustheitswert  $R_{Modell}$  ergibt sich aus dem Verhältnis der Genauigkeit des Modells auf dem Evaluationsdatensatz  $G_{Evaluation}$  zur Genauigkeit auf dem Testdatensatz  $G_{Test}$ . Ein Wert, der nahe bei eins oder größer als eins liegt, deutet auf eine hohe Robustheit hin. In diesem Fall wird die Leistungsfähigkeit des Modells beim Testen auf Daten aus einer verschobenen Datenverteilung kaum beeinflusst oder verbessert sich.

In der vorliegenden Arbeit wird das Testmodell so ausgewählt, dass es eine gute ID-Leistung auf einem Testdatensatz aufweist. Seine OOD-Robustheit gegenüber verschiedene Evaluationsdatensätzen wird analysiert. Zu diesem Zweck wird das Modell zunächst anhand des aufgeteilten Trainingsdatensatzes trainiert und gleichzeitig wird die Validierungsgenauigkeit beobachtet, um geeignete Hyperparameter und Modellarchitekturen auszuwählen. Anschlie-

ßend erfolgt das Training der Modelle mit ausgewählten Hyperparametern anhand des vollständigen Trainingsdatensatzes. Danach wird die ID-Leistungsfähigkeit durch identisch verteilte Testdaten bewertet und abschließend die Robustheit der Modelle durch unterschiedliche Evaluationsdatensätze getestet. Die verwendeten unterschiedlichen Datensätze werden bereits in Abschnitt 4.3.3 beschrieben.

In der vorliegenden Arbeit werden drei Aspekte in der Bewertung der Robustheit betrachtet. Der erste Aspekt bezieht sich auf ein Modell ohne Data-Augmentation im Vergleich zu einem Modell mit Data-Augmentation. Der zweite Aspekt untersucht die Robustheit von verschiedenen Architekturen, während der dritte Aspekt sich mit der Tiefe des Modells befasst. Dazu wird ein Referenzmodell, das eine gute Leistungsfähigkeit auf Testdatensatz aufweist, gesucht. Alle anderen Modelle werden mit diesem Referenzmodell hinsichtlich ihrer Robustheit verglichen. Der relative Robustheitswert  $rR_{Modell}$  ergibt sich aus dem Verhältnis der Genauigkeit des Modells  $G_{Modell}$  zur Genauigkeit des Referenzmodells  $G_{Referenzmodell}$ , wie in folgender Gleichung dargestellt:

$$rR_{Modell} = \frac{G_{Modell}}{G_{Referenzmodell}} \quad (5.2)$$

Ein relativer Robustheitswert größer als eins deutet auf eine überlegene Leistung im Vergleich zum Referenzmodell hin. Mithilfe dieses Wertes verdeutlicht sich, ob sich bei Verwendung von Data-Augmentation-Technik, unterschiedlichen Modellarchitekturen oder tieferen Strukturen, die Robustheit gegenüber verschiedenen Einflussgrößen verbessert.

## 5.2 Modellauswahl

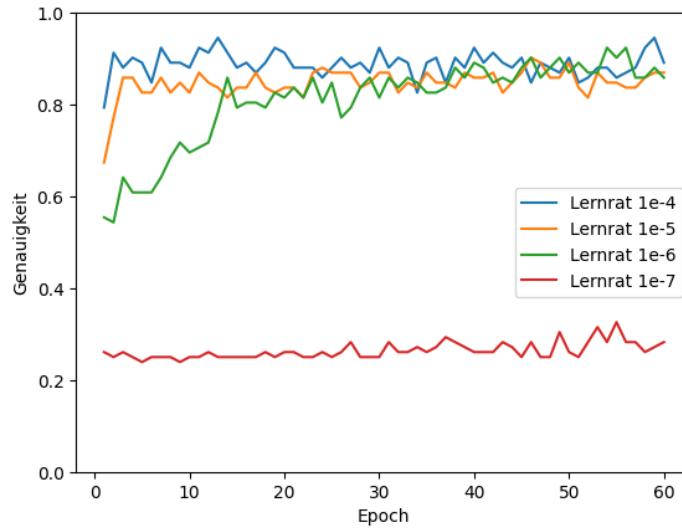
Bei der Auswahl eines Referenzmodells werden verschiedene Modellarchitekturen und unterschiedliche Netztiefen betrachtet und verglichen. Die verwendete Hardware- und Softwareumgebung wird in der folgenden Tabelle detailliert aufgeführt. Aufgrund der begrenzten

|                |                                      |
|----------------|--------------------------------------|
| Betriebssystem | Windows 11                           |
| GPU            | NVIDIA GeForce RTX 3070 Ti           |
| CPU            | 12th Gen Intel(R) Core(TM) i7-12700H |
| pytorch        | Version 2.0.1+cu118                  |
| torchvision    | Version 0.15.2+cu118                 |

**Tabelle 5.1:** Tabellarische Darstellung der verwendeten Hardware- und Softwareumgebung.

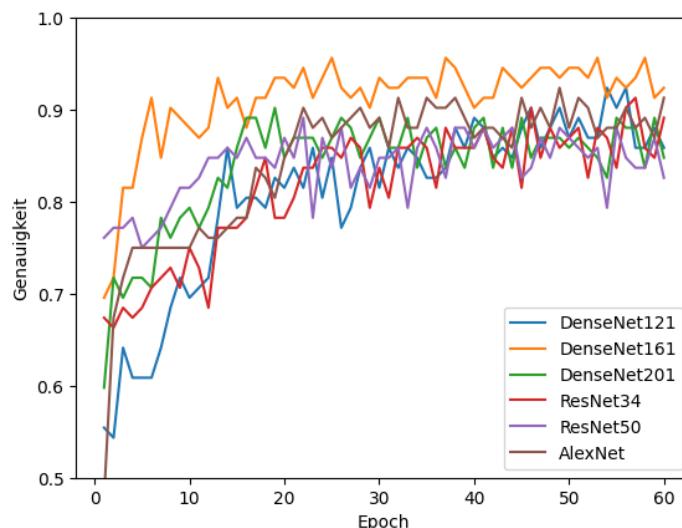
Kapazität der GPU wurden Batch-Größen von 16 beim Training nicht überschritten. Für größere Modelle ist eine Batch-Größe von höchstens 8 einsatzbar. In der vorliegenden Arbeit werden Modelle wie AlexNet, ResNet-34 und DenseNet-121 mit einer Batch-Größe von 16 trainiert, während andere größere Modelle eine Batch-Größe von 8 verwenden.

Um die geeignete Lernrate auszuwählen, wurde ein Testmodell DenseNet-121 über 60 Epochen hinweg mit den Lernraten von  $10^{-4}$ ,  $10^{-5}$ ,  $10^{-6}$  und  $10^{-7}$  trainiert. Insgesamt wurden 365 Bilder für das Training und 91 Bilder für die Validierung verwendet. Der Lernverlauf wird in Abb. 5.1 dargestellt. Das Training mit Lernraten von  $10^{-4}$  und  $10^{-5}$  konvergiert bereits nach 5 Epochen, wobei die Validierungsgenauigkeit stabil bei etwa 0,85 bleibt. Die Lernkurve bei einer Lernrate von  $10^{-6}$  zeigt einen positiven Verlauf und erreicht langsam Konvergenz nach 25 Epochen. Hingegen kann das Modell bei einer Lernrate von  $10^{-7}$  keine Lernfortschritte verzeichnen. Daher wird eine Lernrate von  $10^{-6}$  festgelegt.

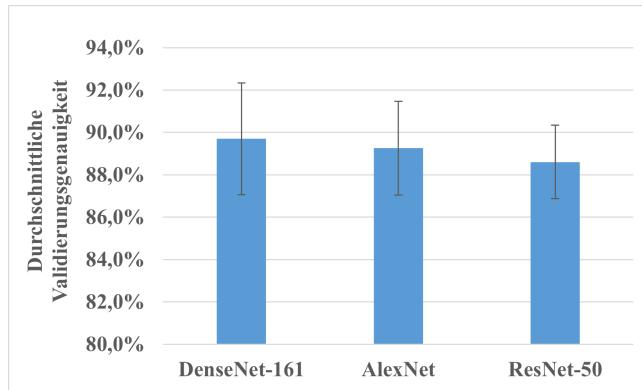


**Abbildung 5.1:** Lernverlauf von DenseNet-121 mit verschiedenen Lernraten.

Nach der Auswahl der Lernraten werden die Modelle mit verschiedenen Architekturen und Tiefen betrachtet. Die Modelle werden über 60 Epochen mit der Lernrate von  $10^{-6}$  trainiert, und der Verlauf des Lernprozesses wird in Abb. 5.2 dargestellt. DenseNet-161 zeigt eine herausragende Leistung. Bereits nach 20 Epochen konvergiert das Modell und die Validierungsgenauigkeit bleibt über 0,9. Andere Modelle erreichen eine vergleichbare Leistung, wobei die Genauigkeiten nach der Konvergenz auf etwa 0,85 stabil bleiben. Besonders hervorzuheben ist, dass AlexNet eine leicht bessere Leistung aufweist. Aufgrund der kleinen Datenmenge für das Training wird die K-Fold-Kreuzvalidierung weiterhin angewendet, um ein robustes Modell auszuwählen. Für die Evaluation werden DenseNet-161, ResNet-50 und AlexNet herangezogen. Die Leistungsfähigkeit dieser Modelle wird mithilfe der k-Fold-Validierung mit  $K$  von 5 untersucht, wobei die Lernrate auf  $10^{-6}$  und die Anzahl der Epochen auf 30 festgelegt wird. Das Ergebnis der K-Fold-Kreuzvalidierung wird in Abb. 5.3 präsentiert. DenseNet-161 zeigt die höchste durchschnittliche Validierungsgenauigkeit, weist jedoch eine etwas größere Standardabweichung auf. Dennoch übertrifft es andere Modelle in Bezug auf die Leistungsfähigkeit.



**Abbildung 5.2:** Der Lernverlauf verschiedener Architekturen bei einer Lernrate von  $10^{-6}$ .

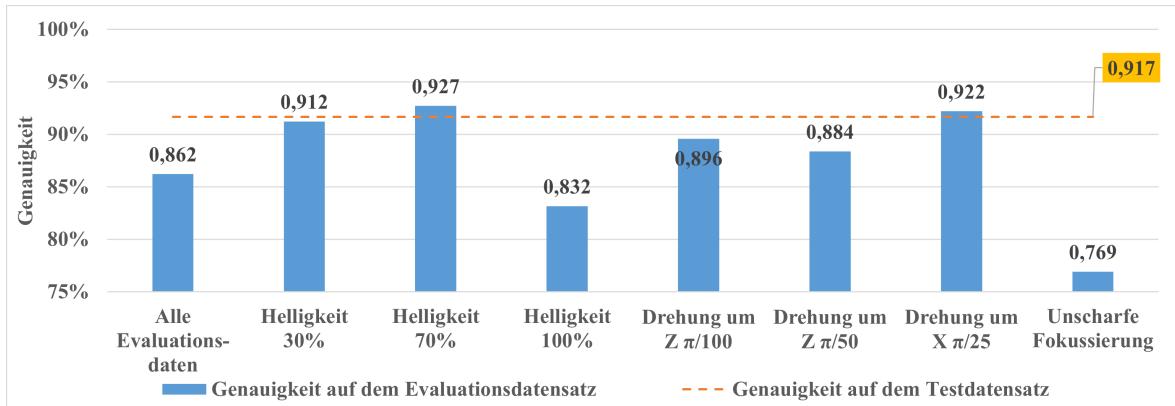


**Abbildung 5.3:** Die durchschnittlichen Validierungsgenauigkeiten von drei verschiedenen Modellen mit entsprechenden Standardabweichungen in der K-Fold-Validierung.

higkeit. AlexNet zeigt ebenfalls eine solide Leistung. ResNet-50 weist die geringste Standardabweichung auf, erreicht jedoch im Vergleich zu anderen Modellen eine etwas schlechtere Validierungsgenauigkeit. Daher wird DenseNet-161 als Referenzmodell herangezogen. Dieses Modell wird mit denselben Hyperparametern mit allen Trainingsdaten erneut trainiert und erzielt eine hohe Genauigkeit von 91,7 % für den Testdatensatz.

### 5.3 Evaluationsergebnisse

Es gibt 16 Evaluationsdatensätze in der Betrachtung der Helligkeiten, der verschiedenen Kameraspositionen und der unscharfen Fokussierung. Insgesamt stehen 9.207 Bilder zur Verfügung, was das 20-fache der Anzahl der Trainingsdaten ausmacht. Die Genauigkeiten des Referenzmodells DenseNet-161 für die Evaluationsdatensätze der einzelnen Einflussgrößen werden in Abb. 5.4 dargestellt. Im vorliegenden Diagramm ist deutlich zu erkennen, dass es



**Abbildung 5.4:** Genauigkeiten des Referenzmodells DenseNet-161 für die Evaluationsdatensätze der einzelnen Einflussgrößen und alle Evaluationsdaten im Vergleich zur Testgenauigkeit.

zu einem signifikanten Leistungsabfall von 8,5 % in der Helligkeit von 100 % und insbesondere in der unscharfen Fokussierung von 14,8 % kommt. Bei der Betrachtung der Drehung um die Z-Achse zeigt sich lediglich eine geringfügige Abnahme der Genauigkeit um etwa 3 % im Vergleich zur Genauigkeit für den Testdatensatz. In Bezug auf andere Einflussfaktoren wie die Drehung um die X-Achse sind kaum Unterschiede festzustellen. Der Leistungsverschlech-

terung für alle Evaluationsdaten beträgt 5,5 %. Dies wird in nachfolgender Tabelle weiter verdeutlicht. Die Robustheitswerte bei variiertem Drehung und Helligkeit von 30 % und 70 %

| Datensätze   | E    | H1   | H2   | H3   | D1   | D2   | D3   | U    |
|--------------|------|------|------|------|------|------|------|------|
| $R_{Modell}$ | 0,94 | 1,00 | 1,01 | 0,91 | 0,98 | 0,96 | 1,01 | 0,84 |

**Tabelle 5.2:** Robustheitswerte des Referenzmodells DenseNet-161 für die Evaluationsdatensätze der einzelnen Einflussgrößen und alle Evaluationsdaten. (E: alle Evaluationsdaten; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von  $\pi/100$ ,  $\pi/50$ , um X-Achse von  $\pi/25$ , U: Unschärfe)

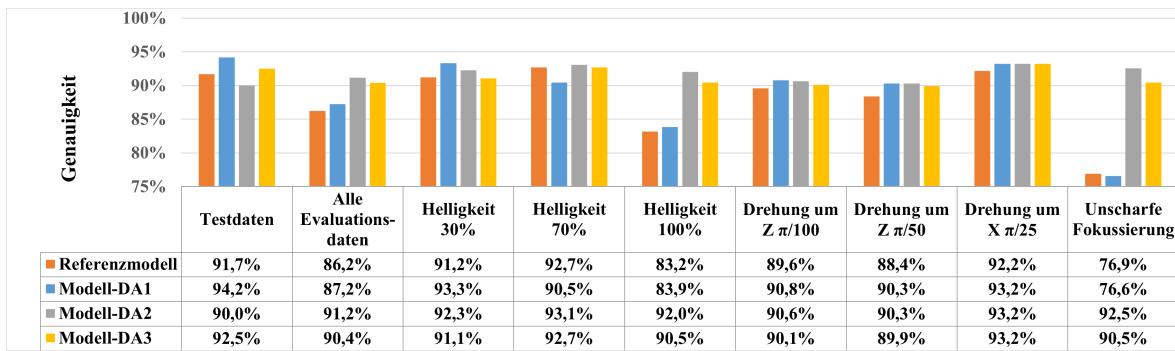
weisen eine hohe Nähe zur Eins auf und erreichen eine signifikante Abnahme bei unscharfen Bildern, nämlich 0,84.

Um den ersten Aspekt zu untersuchen, wurde das Modell DenseNet-161 unter Verwendung von Data-Augmentation erneut trainiert. Zunächst wird eine leichte Data-Augmentation angewendet. Das Modell, das unter Verwendung solcher Data-Augmentation trainiert wurde, wird im Folgenden als Modell-DA1 bezeichnet. Hierbei wird auf die Trainingsbilder zufällig eine Farbveränderung mit einer Farbtonänderung von 0,15, einer Sättigungsänderung von 0,15 und einer Helligkeitsänderung von 0,25 angewendet. Die Bilder werden zudem zufällig um bis zu 5 Grad gedreht. Des Weiteren wird auf die Bilder eine zufällige Gaußsche Unschärfe mit einem 7x7 Filter und einem Sigma-Wert von 0,1 bis 0,5 angewendet. Es wurde eine Gaußsche Unschärfe ausgewählt, die für das menschliche Auge nicht signifikant erkennbar ist. Zur Normalisierung der Bilddaten wird ein Durchschnittswert (0,546, 0,481, 0,491) verwendet. Das Modell, das bei der Anwendung einer Data-Augmentation mit einem erhöhten Gaußschen Sigma-Wert im Bereich von 0,1 bis 2 verwendet wird, wird als Modell-DA2 bezeichnet. Im Vergleich zu Modell-DA2 wird bei Modell-DA3 die Helligkeitsänderung auf einen Wert von 0,5 erhöht. Data-Augmentation wird gezielt angewendet, sodass die bearbeiteten Variationen während des Trainings mit den Einflüssen in der realen Umgebung möglichst in Einklang stehen. Die Ergebnisse von drei Modellen im Vergleichsfall werden in Tabelle 5.3 präsentiert. Die relativen Robustheitswerte von Modell-DA1 sind lediglich bei einer Hellig-

| $rR_{Modell} \setminus$ Datensätze | T    | E    | H1   | H2   | H3   | D1   | D2   | D3   | U    |
|------------------------------------|------|------|------|------|------|------|------|------|------|
| Referenzmodell                     | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| Modell-DA1                         | 1,03 | 1,01 | 1,02 | 0,98 | 1,01 | 1,01 | 1,02 | 1,01 | 1,00 |
| Modell-DA2                         | 0,98 | 1,06 | 1,01 | 1,00 | 1,11 | 1,01 | 1,02 | 1,01 | 1,20 |
| Modell-DA3                         | 1,01 | 1,05 | 1,00 | 1,00 | 1,09 | 1,01 | 1,02 | 1,01 | 1,18 |

**Tabelle 5.3:** Relative Rohustheitswerte der Modelle für den Testdatensatz, alle Evaluationsdaten und verschiedene Evaluationsdatensätze der Einzelnen Einflussgrößen bei der Anwendung von Data-Augmentation-Techniken. (T: Test, E: alle Evaluationsdaten; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von  $\pi/100$ ,  $\pi/50$ , um X-Achse von  $\pi/25$ , U: Unschärfe)

keit von 100 % geringfügig kleiner als Eins, während sie bei anderen Evaluationsdatensätzen Werte größer als Eins aufweisen. Besonders hervorzuheben ist, dass der Robustheitswert bei Unschärfe genau Eins beträgt. Die Robustheitswerte von Modell-DA2 sind bei allen Evaluationsdatensätzen größer als Eins, wobei bei Helligkeit von 100 % und Unschärfe deutlich höhere Werte von 1,11 und 1,20 beobachtet werden. Allerdings liegt der Robustheitswert für den Testdatensatz bei einem etwas geringeren Wert von 0,98. Bei Betrachtung von Modell-DA3 zeigen sich alle Robustheitswerte größer als Eins, insbesondere bei Helligkeit von 100 % und Unschärfe. Die Robustheitswerte sämtlicher Vergleichsmodelle für alle Evaluationsdaten sind durchweg größer als Eins. In Abb. 5.5 wird die Genauigkeiten der entsprechenden Modelle für verschiedene Evaluationsdatensätze anschaulich dargestellt. Es zeigt sich, dass das Modell-DA1 die Genauigkeit für alle Evaluationsdaten nur um 1 % steigert, während die Test-



**Abbildung 5.5:** Genauigkeiten der Modelle für den Testdatensatz, alle Evaluationsdaten und die Evaluationsdatensätze der einzelnen Einflussgrößen im Vergleich zum Referenzmodell.

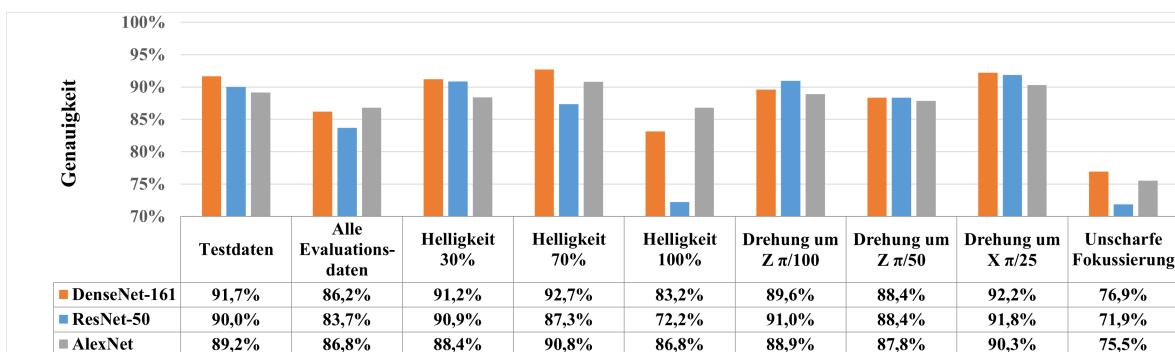
genauigkeit um etwa 3 % erhöht wird. Die Genauigkeit bleibt bei einer Helligkeit von 100 % und bei Unschärfe fast unverändert. Im Gegensatz dazu steigen diese Werte bei Modell-DA2 und Modell-DA3 deutlich an, sowohl bei einer Helligkeit von 100 % als auch bei Unschärfe. Beide Modelle erreichen eine hohe Genauigkeit von über 90 % für alle Evaluationsdaten. Allerdings die Testgenauigkeit von Modell-DA2 verringert sich um etwa 2 %.

Bei der Betrachtung des zweiten Aspekts werden während des Trainings zwei weitere Architekturen, nämlich ResNet-50 und AlexNet, ohne Data-Augmentation verwendet. Die folgende Tabelle zeigt Evaluationsergebnisse. Die vorliegende Tabelle veranschaulicht, dass die relati-

| $rR_{Modell} \setminus$ Datensätze | T    | E    | H1   | H2   | H3   | D1   | D2   | D3   | U    |
|------------------------------------|------|------|------|------|------|------|------|------|------|
| DenseNet-161                       | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| ResNet-50                          | 0,98 | 0,97 | 1,00 | 0,94 | 0,87 | 1,02 | 1,00 | 1,00 | 0,93 |
| AlexNet                            | 0,97 | 1,01 | 0,97 | 0,98 | 1,04 | 0,99 | 0,99 | 0,98 | 0,98 |

**Tabelle 5.4:** Relative Robustheitswerte der Modelle für den Testdatensatz, alle Evaluationsdaten und verschiedene Evaluationsdatensätze der Einzelnen Einflussgrößen bei unterschiedlichen Modellarchitekturen. (T: Test, E: alle Evaluationsdaten; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von  $\pi/100$ ,  $\pi/50$ , um X-Achse von  $\pi/25$ , U: Unschärfe)

ven Robustheitswerte von ResNet-50 entweder gleich oder kleiner als Eins sind. Bei einer Helligkeit von 100 % beträgt der Robustheitswert von ResNet-50 lediglich 0,87 und bei Unschärfe 0,93. Im Gegensatz dazu weisen die Robustheitswerte von AlexNet meist nur geringfügige Abweichungen von Eins auf, außer bei einer Helligkeit von 100 %, wo dieser Wert größer als Eins ist. Der Robustheitswert für alle Evaluationsdaten ist bei AlexNet etwas größer als Eins,



**Abbildung 5.6:** Genauigkeiten der unterschiedlichen Modellarchitekturen für den Testdatensatz, alle Evaluationsdaten und die Evaluationsdatensätze der einzelnen Einflussgrößen im Vergleich zum Referenzmodell.

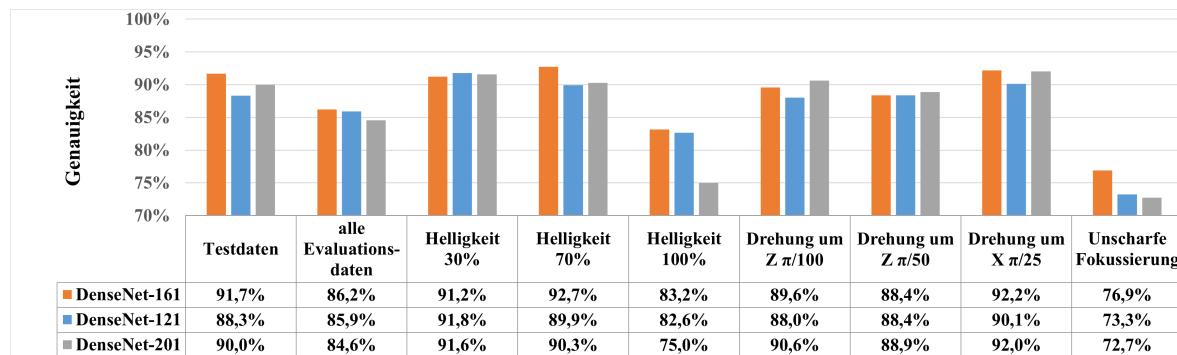
während er bei ResNet-50 kleiner als Eins ist. Insbesondere sind die Robustheitswerte des Testdatensatzes bei zwei Modellen kleiner als eins. Die detaillierten Leistungsfähigkeiten der Modelle im Vergleich zum Referenzmodell werden in Abb. 5.6 veranschaulicht. Die Genauigkeiten von AlexNet weisen geringfügige Unterschiede im Vergleich zum Referenzmodell auf. AlexNet zeigt eine etwa höhere Genauigkeit für alle Evaluationsdaten. Bei einer Helligkeit von 100 % zeigt ResNet-50 einen signifikanten Leistungsabfall von 11 % und bei Unschärfe eine Leistungsabnahme von 5 %. Die Genauigkeit für alle Evaluationsdaten beträgt nur 83,7 %.

Die Modelle DenseNet-121 und DenseNet-201 wurden trainiert, um den dritten Aspekt zu analysieren. Während des Trainings wird wie zuvor keine Data-Augmentation angewendet. In Tabelle 5.5 werden die relativen Robustheitswerte des DenseNet-121 dargestellt. Diese Werte liegen alle unter oder gleich Eins, mit Ausnahme von geringfügig höheren Werten bei Helligkeit von 30 %. Bei DenseNet-201 zeigen sich deutlich kleinere Werte bei einer Hel-

| $rR_{Modell} \setminus \text{Datensätze}$ | T    | A    | H1   | H2   | H3   | D1   | D2   | D3   | U    |
|---|------|------|------|------|------|------|------|------|------|
| DenseNet-161                              | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| DenseNet-121                              | 0,96 | 1,00 | 1,01 | 0,97 | 0,99 | 0,98 | 1,00 | 0,98 | 0,95 |
| DenseNet-201                              | 0,98 | 0,98 | 1,00 | 0,97 | 0,90 | 1,01 | 1,01 | 1,00 | 0,95 |

**Tabelle 5.5:** Relative Robustheitswerte der Modelle für den Testdatensatz, alle Evaluationsdaten und verschiedene Evaluationsdatensätze der einzelnen Einflussgrößen bei unterschiedlichen Netztiefen. (T: Test, A: Alle Einflussgrößen; H1,2,3: Helligkeit von 30%, 70%, 100%; D1,2,3: Drehwinkel um Z-Achse von  $\pi/100$ ,  $\pi/50$ , um X-Achse von  $\pi/25$ , U: Unschärfe)

ligkeit von 100 % im Vergleich zu DenseNet-121, wobei andere Werte nur leicht um eins abweichen. In Abb. 5.7 sind die Genauigkeiten der Modelle für verschiedene Evaluationsdatensätze im Vergleich zum Referenzmodell deutlich zu erkennen. Es wird geringfügige Ab-



**Abbildung 5.7:** Genauigkeiten der unterschiedlichen Netztiefen für den Testdatensatz, alle Evaluationsdaten und die Evaluationsdatensätze der einzelnen Einflussgrößen im Vergleich zum Referenzmodell.

weichungen festgestellt. Bei Helligkeit von 100 % und Unschärfe zeigt sich eine geringfügige erhöhte Abweichung. Die Genauigkeit bei DenseNet-201 weicht etwas mehr ab, während bei DenseNet-121 lediglich eine geringfügige, nicht signifikante Abweichung auftritt.

## 5.4 Diskussion

Die Forschungsergebnisse zeigen, dass das Referenzmodell DenseNet-161 bei unveränderten Bedingungen und leicht natürlicher Helligkeits- und Kamerapositionsänderung eine ausrei-

chende Robustheit aufweist. Es wurde eine gute Testgenauigkeit von 91,7 % erreicht. Bei einer Helligkeit von 30 % und 70 % sowie bei verschiedenen Drehstufen weichen die Genauigkeiten im Vergleich zur Testgenauigkeit leicht ab. Ein signifikanter Leistungsabfall tritt bei extremer Helligkeit und Unschärfe auf, insbesondere bei Unschärfe verschlechtert sich die Genauigkeit um ungefähr 15 %. Bei der Untersuchung anderer Modellarchitekturen zeigt sich in diesem Fall ebenfalls ein signifikanter Leistungsabfall. Diese erhebliche Leistungsabnahme für alle Evaluationsdaten wird hauptsächlich durch die deutlich ungünstigeren Aufnahmebedingungen verursacht. Es ist von besonderer Bedeutung, die Auswirkungen der Helligkeit von 100 % und der Unschärfe weiter zu betrachten.

Das Modell-DA1 zeigt im Vergleich zum Referenzmodell DenseNet-161 eine geringfügige Steigerung der Robustheit durch die Anwendung von leichter Data-Augmentation. Die Anwendung der Parameter von Data-Augmentation wie bspw. einer Helligkeitsänderung von 0,25 und einer leichten Unschärfe mit Gaußschen Sigma-Werten im Bereich von 0,1 bis 0,5, zeigt keine signifikanten Auswirkungen auf die Steigerung der Robustheit. Stattdessen wird die Testgenauigkeit durch diese geringfügigen Änderungen deutlich gesteigert und erreicht einen Wert von 94,2 %. Dies zeigt eine erhöhte ID-Leistungsfähigkeit. Bei der Anwendung von stärkeren Gaußschen Sigma-Werten im Bereich von 0,1 bis 2 in der Data-Augmentation wird eine signifikante Steigerung der Robustheit von Modell-DA2 bei ungünstigen Aufnahmebedingungen festgestellt. Insbesondere steigert sich die Genauigkeit bei Unschärfe um 16 % im Vergleich zum Referenzmodell. Die Genauigkeiten aller Evaluationsdatensätze liegen über 90 %. Es ist lediglich ein leichter Leistungsabfall für den Testdatensatz auf 90 % zu verzeichnen. Dies hat zur Erhöhung der OOD-Robustheit geführt, jedoch gleichzeitig die ID-Leistungsfähigkeit leicht verschlechtert. Das Modell-DA3 verwendet während des Trainings eine noch intensivere Helligkeitsänderung von 0,5 in der Data-Augmentation. Die Genauigkeiten aller Evaluationsdatensätze liegt ebenfalls über 90 %. Im Vergleich zu Modell-DA2 ist die OOD-Robustheit etwas geringer, jedoch ist die ID-Leistungsfähigkeit besser. In der Untersuchung der Data-Augmentation zeigt sich, dass die Anwendung dieser Technik die Robustheit gegenüber betrachteten natürlichen Veränderungen signifikant erhöhen kann. Dabei spielt der Gaußscher Sigma-Wert eine wichtig Rolle. Es ist jedoch zu beachten, dass eine Verschlechterung der ID-Leistungsfähigkeit während der Steigerung der OOD-Robustheit auftreten kann. Durch die systematische Untersuchung verschiedener Parameter der Data-Augmentation kann die natürliche Robustheit des Modells verbessert werden, ohne dabei die ID-Leistungsfähigkeit zu beeinträchtigen. Im Rahmen dieser Untersuchung wurde das Modell mit 456 Bildern trainiert. Durch die Anwendung von Data-Augmentation erreichte es eine herausragende Klassifizierungsleistung für insgesamt 9207 Bilder.

DenseNet-161 zeigt im Vergleich zu anderen Architekturen wie ResNet-50 und AlexNet eine bessere ID-Leistungsfähigkeit. AlexNet und DenseNet-161 zeigen eine höhere Robustheit im Vergleich zu ResNet-50. Bei der Betrachtung der Tiefe von NN hat sich gezeigt, dass eine tiefere Architektur wie DenseNet-201 die Robustheit nicht weiter verbessert. Die Ergebnisse deuten darauf hin, dass DenseNet-161 ist etwas robuster als andere betrachtete Modellarchitekturen.

# **Kapitel 6**

## **Zusammenfassung und Ausblick**

Im abschließenden Kapitel werden die Forschungsziele erläutert und die relevanten Ergebnisse sowie die Schlussfolgerungen zusammengefasst. Darüber hinaus werden Aspekte, die in der Arbeit nicht erfasst wurden, im Ausblick präsentiert.

### **6.1 Zusammenfassung**

Aufgrund seiner großen Vorteile gegenüber anderen Schweißverfahren wird FSW gegenwärtig in vielen Branchen eingesetzt. Die zunehmende Verwendung der FSW-Technologie erhöht die Nachfrage nach effizienten zerstörungsfreien Prüfverfahren. Das DL-Verfahren hat sich als sehr geeignet in der automatischen Qualitätsüberwachung erwiesen. Allerdings handelt es sich bei DL-Verfahren um ein datenbasiertes Verfahren, dessen Klassifikationsleistung deutlich abnimmt, wenn die vorliegenden Daten stark von den Trainingsdaten abweichen. In realen Umgebungen treten viele verschiedene Einflüsse auf, wie Helligkeits-, Positionsänderung und andere bekannte sowie unbekannte natürliche Veränderungen. Diese natürlichen Veränderungen in den Daten beeinträchtigen die Leistungsfähigkeit des Klassifikationsmodells erheblich. Das Ziel der vorliegenden Arbeit besteht darin, die Robustheit unter verschiedenen Umgebungsparametern zu untersuchen. Da viele Studien gezeigt haben, dass vorhandene synthetische Benchmarks zur Bewertung der Robustheit die reale Welt nicht ausreichend repräsentieren. Daher kann die Robustheit nicht angemessen bewertet werden. Zur Lösung dieses Problems wurden in der vorliegenden Arbeit Daten mithilfe eines roboter-basierten Messstands aus der realen Umwelt gesammelt.

Es wurde ein halbautomatisches Konzept zur Datengenerierung entwickelt, um Trainings- und Evaluationsbilder effizient zu generieren. Hierfür wurde ein roboter-basierter Messstand aufgebaut. Die Evaluationsbilder in verschiedenen Umgebungen wurden durch Anpassung der Beleuchtung, des Objektives und der Bewegungsplanung des Roboters generiert. Insgesamt wurden 576 Bilder für das Training und 9207 Bilder für die Evaluation gesammelt. Dieses effiziente Konzept ist anpassbar für weitere Anforderungen der Datengenerierung. Die Daten wurden gelabelt und für den Bewertungsprozess aufbereitet.

Die Ergebnisse der Arbeit zeigen, dass DenseNet-161 im Vergleich zu anderen betrachteten Architekturen wie ResNet-50 eine geeignete Architektur zur Klassifizierung der Gratbildungsfehler von Schweißnähten ist. DenseNet-161 zeichnet sich bereits durch eine gute Robustheit gegenüber leichten Helligkeitsänderungen und Positionsvariationen aus. Ein signifikanter Leistungsabfall tritt jedoch unter ungünstigen Aufnahmebedingungen auf, wie zu hellen

oder unscharfen Bildern. Durch die Anwendung geeigneter Data-Augmentation während des Trainings wird die Robustheit des Modells deutlich erhöht. Die Genauigkeiten liegen nicht nur für den Testdatensatz, sondern auch für allen anderen Evaluationsdatensätzen über 90 %. Die Robustheit des Modells, das lediglich mit 456 Trainingsbildern trainiert wurde, kann durch geeignete Data-Augmentation erheblich verbessert werden, ohne die Leistungsfähigkeit des Testdatensatzes zu verschlechtern. Es wurde hohe Genauigkeit für den kompletten Datensatz von insgesamt 9792 Bildern erzielt, wobei 16 verschiedene Umgebungsveränderungen berücksichtigt wurden.

## 6.2 Ausblick

In der vorliegende Arbeit wird die Robustheit gegenüber natürlichen Einflussgrößen im Kontext von Helligkeit, Kameraposition und unscharfer Fokussierung untersucht. Die Einflüsse von Sensoren wie Rauschen, Farbabgleich und Kontrast wurden nicht in dem Maße betrachtet. Bildverzerrungen und Objektverdeckungen stellen ebenfalls einen interessanten Aspekt dar. Es besteht Potenzial für zukünftige Forschungsarbeiten, um die Robustheit gegenüber weiteren Einflussfaktoren zu erforschen. Eine Generierung weiterer Daten kann mithilfe des entwickelten roboter-basierten Messkonzepts effizient umgesetzt werden.

Die Ergebnisse der Arbeit veranschaulichen die erhebliche Verbesserung der Robustheit des Modells durch die Anwendung geeigneter Data-Augmentation während des Trainings. Dennoch bleibt die Suche nach einer geeigneten Strategie von Data-Augmentation eine komplexe Fragestellung. Es bleibt unklar, ob sich andere Parameter in der Data-Augmentation positiv auf das Training auswirken und die Robustheit steigern können. Die Modelle wurden unter Verwendung derselben Lernraten und Anzahl von Epochen trainiert. Aufgrund von Hardwarebeschränkungen konnte nur eine Batch-Größe von maximal 16 verwendet werden. Es besteht die Möglichkeit, dass eine Erhöhung der Anzahl der Epochen oder die Anpassung von Hyperparametern die Robustheit beeinflussen könnten. Es ist ebenfalls anzumerken, dass in der vorliegenden Arbeit nur 456 Bilder zur Verfügung standen. Ein tieferes DNN-Modell wie DenseNet-201 könnte bei Verwendung eines umfangreicheren Trainingsdatensatzes möglicherweise eine bessere Leistungsfähigkeit und Robustheit aufweisen.

Bei der automatischen Qualitätsüberwachung von Schweißnähten beschränkt sich die bisherige Betrachtung lediglich auf die Analyse von Gratbildung. Es besteht jedoch die Möglichkeit, weitere Aspekte zu berücksichtigen, wie bspw. Decklagenfehler im Innenbereich der Schweißnaht, unregelmäßige Schweißnahtbereite oder Oberflächenmängel in der Füllung. Dies erfordert eine eingehende Untersuchung der Robustheit des komplexen Multi-Klassifikators, insbesondere hinsichtlich seiner Robustheit gegenüber verschiedenen Einflussfaktoren bei unterschiedlichen Arten von Fehlern.

# **Anhang A**

## **Digitaler Anhang**

**A.1 Technische Datenblatt der Kamera**

**A.2 Technische Datenblatt des Objektives**

**A.3 Zeichnungen von Kamerahalterung**



# Literaturverzeichnis

- AGGARWAL, C. C., (2023). *Neural networks and deep learning: A textbook*. Second Edition. Cham: Springer International Publishing AG. ISBN: 9783031296413.
- AKINLABI, E. T., (2020). *Solid-State Welding*. Mechanical Engineering Ser. Cham: Springer International Publishing AG. ISBN: 9783030370152.
- AVOLA, D., CASCIO, M., CINQUE, L., FAGIOLI, A., FORESTI, G. L., MARINI, M. R. und ROSSI, F., (2022). „Real-time deep learning method for automated detection and localization of structural defects in manufactured products“. In: *Computers & Industrial Engineering* 172, S. 108512. ISSN: 0360-8352. DOI: 10.1016/j.cie.2022.108512.
- AZAD, P., GOCKEL, T. und DILLMANN, R., (2011). *Computer Vision: Das Praxisbuch*. 3. Aufl. Aachen: Elektor-Verl. ISBN: 9783895761652.
- BACHMANN, A., (2021). *Regelung der Temperatur beim Rührreibschweißen*. Bd. v.362. Forschungsberichte IWB. München: utzverlag. ISBN: 9783831676651.
- BATCHELOR, B. G., Hrsg., (2012). *Machine Vision Handbook*. SpringerLink Bücher. London: Springer London. ISBN: 9781849961691. DOI: 10.1007/978-1-84996-169-1.
- BAUER, N., (2003). *Guideline for industrial image processing*. Bd. 1. Vision. Erlangen und Stuttgart: Fraunhofer-Allianz Vision und Fraunhofer IRB-Verl. ISBN: 3816762964.
- BEUTH VERLAG GMBH, (2020a). *DIN EN ISO 25239-1:2020-12, Rührreibschweißen - Aluminium - Teil\_1: Begriffe (ISO\_25239-1:2020); Deutsche Fassung EN\_ISO\_25239-1:2020*. Berlin. DOI: 10.31030/3161337.
- (2020b). *DIN EN ISO 25239-5:2020-12, Rührreibschweißen - Aluminium - Teil\_5: Qualitäts- und Prüfungsanforderungen (ISO\_25239-5:2020); Deutsche Fassung EN\_ISO\_25239-5:2020*. Berlin. DOI: 10.31030/3161367.
- BEYERER, J., PUENTE LEÓN, F., FRESE, C. und MEYER, J., (2016). *Automatische Sichtprüfung: Grundlagen, Methoden und Praxis der Bildgewinnung und Bildauswertung*. 2. Aufl. 2016. Berlin, Heidelberg: Springer Berlin Heidelberg. ISBN: 9783662477861.
- BHAT, N. N., KUMARI, K., DUTTA, S., PAL, S. K. und PAL, S., (2015). „Friction stir weld classification by applying wavelet analysis and support vector machine on weld surface images“. In: *Journal of Manufacturing Processes* 20, S. 274–281. ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2015.07.002.
- CHANDOLA, V., BANERJEE, A. und KUMAR, V., (2009). „Anomaly detection“. In: *ACM Computing Surveys* 41.3, S. 1–58. ISSN: 0360-0300. DOI: 10.1145/1541880.1541882.
- COLEMAN, D. T., SUCAN, I. A., CHITTA, S. und CORRELL, N., (2014). *Reducing the Barrier to Entry of Complex Robotic Software: a MoveIt! Case Study*. DOI: 10.6092/JOSER{\textunderscore}2014{\textunderscore}05{\textunderscore}01{\textunderscore}P3.
- COLLIGAN, K. J., (2010). *The friction stir welding process: an overview*. ISBN: 978-1-4398-0211-3.
- DRENKOW, N., SANI, N., SHPITSER, I. und UNBERATH, M., (2022). *A Systematic Review of Robustness in Deep Learning for Computer Vision: Mind the gap?*

- ERHARDT, A., (2008). *Einführung in die Digitale Bildverarbeitung: Grundlagen, Systeme und Anwendungen*. Springer eBook Collection Computer Science & Engineering. Wiesbaden: Vieweg+Teubner. ISBN: 9783834895189. DOI: 10.1007/978-3-8348-9518-9.
- FORSYTH, D. und PONCE, J., (2003). *Computer vision: A modern approach*. Reprinted with corr. Prentice Hall series in artifical intelligence. Upper Saddle River, NJ: Prentice Hall. ISBN: 0130851981.
- FRANKA EMIKA GMBH, (2021). *Franka Emika Robot's Instruction Handbook*. München.
- (2023a). *FRANKA EMIKA Roboter für Forschung Produkthandbuch*. München.
- (2023b). *Offizielle Webweite von FRANKA EMIKA GmbH*. URL: <https://www.franka.de/de>.
- GANDHINATHAN, R., (2019). *ROS ROBOTICS PROJECTS: BUILD AND CONTROL ROBOTS POWERED BY THE ROBOT OPERATING SYSTEM, MACHINE LEARNING, AND VIRTUAL REALITY*. 2nd edition. S.l.: PACKT PUBLISHING. ISBN: 1838645195. URL: <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=2332842>.
- GOJIĆ, G., VINCAN, V., KUNDAČINA, O., MIŠKOVIĆ, D. und DRAGAN, D., (2023). *Non-adversarial Robustness of Deep Learning Methods for Computer Vision*. URL: <https://arxiv.org/pdf/2305.14986>.
- 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.
- 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.
- JO, T., (2023). *Deep Learning Foundations*. 1st ed. 2023. Cham: Springer International Publishing und Imprint Springer. ISBN: 9783031328794. DOI: 10.1007/978-3-031-32879-4.
- JULIJA ZAVADLAV, (2022). „Physics-Informed Machine Learning SoSe 2022“. Diss. München: Technical University of Munich.
- KHAN, A., SOHAIL, A., ZAHORA, U. und QURESHI, A. S., (2020). „A survey of the recent architectures of deep convolutional neural networks“. In: *Artificial Intelligence Review* 53.8, S. 5455–5516. ISSN: 1573-7462. DOI: 10.1007/s10462-020-09825-6.
- LAUGROS, A., CAPLIER, A. und OSPICI, M., (2019). *Are Adversarial Robustness and Common Perturbation Robustness Independent Attributes ?* URL: <https://arxiv.org/pdf/1909.02436>.
- MONTAVON, G., (2012). *Neural Networks*. 2nd ed. Bd. v.7700. Lecture Notes in Computer Science Ser. Berlin, Heidelberg: Springer Berlin / Heidelberg. ISBN: 9783642352898.
- MUBIAYI, M. P., AKINLABI, E. T. und MAKHATHA, M. E., (2019). *Current trends in friction stir welding (FSW) and friction stir spot welding (FSSW): An overview and vase studies*. Bd. 6. Structural integrity. Cham: Springer. ISBN: 9783319927497.
- OLIMOV, B. A. u., VELUVOLU, K. C., PAUL, A. und KIM, J., (2022). „UzADL: Anomaly detection and localization using graph Laplacian matrix-based unsupervised learning method“. In: *Computers & Industrial Engineering* 171, S. 108313. ISSN: 0360-8352. DOI: 10.1016/j.cie.2022.108313.
- PANG, G., SHEN, C., CAO, L. und VAN HENGEL, A. den, (2022). „Deep Learning for Anomaly Detection“. In: *ACM Computing Surveys* 54.2, S. 1–38. ISSN: 0360-0300. DOI: 10.1145/3439950.
- RAJASHEKAR, R. und RAJAPRAKASH, B. M., (2016). „Development of a model for friction stir weld quality assessment using machine vision and acoustic emission techniques“. In: *Journal of Materials Processing Technology* 229, S. 265–274. ISSN: 0924-0136. DOI: 10.1016/j.jmatprotec.2015.09.030.

- RAJASHEKAR, R., (2013). „Digital Image Processing Of Friction Stir Weld Bead Surface Using Profiling and Contouring For Weld Quality Assessment“. In: *Journal of Mechanical and Civil Engineering* vol. 5 5, S. 54–64.
- RANJAN, R., KHAN, A. R., PARIKH, C., JAIN, R., MAHTO, R. P., PAL, S., PAL, S. K. und CHAKRABARTY, D., (2016). „Classification and identification of surface defects in friction stir welding: An image processing approach“. In: *Journal of Manufacturing Processes* 22, S. 237–253. ISSN: 1526-6125. DOI: 10.1016/j.jmapro.2016.03.009.
- RASCHKA, S., (2020). *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*.
- RUHSTORFER, M., (2012). *Rührreibschweißen von Rohren: Zugl.: München, Techn. Univ., Diss., 2012*. Bd. 262. Forschungsberichte / IWB. München: Utz. ISBN: 3831641978.
- SICILAN, K., (2014). *Analysis of Surface Quality of Friction Stir Welding Joints using Image Processing Techniques*.
- SINHA, P., MUTHUKUMARAN, S., SIVAKUMAR, R. und MUKHERJEE, S. K., (2008). „Condition monitoring of first mode of metal transfer in friction stir welding by image processing techniques“. In: *The International Journal of Advanced Manufacturing Technology* 36.5-6, S. 484–489. ISSN: 1433-3015. DOI: 10.1007/s00170-006-0854-2.
- SOMANI, A., HORSCH, A. und PRASAD, D. K., (2023). *Interpretability in Deep Learning*. 1st ed. 2023. Cham: Springer International Publishing und Imprint Springer. ISBN: 9783031206399. DOI: 10.1007/978-3-031-20639-9.
- STANFORD ARTIFICIAL INTELLIGENCE LABORATORY, (2023). *Robotic Operating System*. URL: <https://www.ros.org>.
- 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.
- SUDHAGAR, S., SAKTHIVEL, M. und GANESHKUMAR, P., (2019). „Monitoring of friction stir welding based on vision system coupled with Machine learning algorithm“. In: *Measurement* 144, S. 135–143. ISSN: 0263-2241. DOI: 10.1016/j.measurement.2019.05.018.
- TANAKA, A., TOMIYA, A. und HASHIMOTO, K., (2021). *Deep learning and physics*. Mathematical physics studies. Singapore: Springer. ISBN: 9789813361072. DOI: 10.1007/978-981-33-6108-9.
- ZIPFEL, J., VERWORNER, F., FISCHER, M., WIELAND, U., KRAUS, M. und ZSCHECH, P., (2023). „Anomaly detection for industrial quality assurance: A comparative evaluation of unsupervised deep learning models“. In: *Computers & Industrial Engineering* 177, S. 109045. ISSN: 0360-8352. DOI: 10.1016/j.cie.2023.109045.
- ZONNENSHAIN, A. und KENETT, R. S., (2020). „Quality 4.0—the challenging future of quality engineering“. In: *Quality Engineering* 32.4, S. 614–626. ISSN: 0898-2112. DOI: 10.1080/08982112.2019.1706744.



# **Erklärung**

Ich versichere hiermit, dass ich die von mir eingereichte Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Garching, 30. September 2023

---

(Unterschrift)