

Gefördert durch:



Bundesministerium
für Bildung
und Forschung

Anlage zum Verwendungsnachweis

zum Kooperationsprojekt

„Smart Solution for Quality Control of Vehicle Assembly using Ultrasonic Imaging
Sensor with Embedded Artificial Intelligence (SSQC)“

Richtlinie zur Förderung von Zuwendungen für Forschungsprojekte mit Kanada unter
Beteiligung von Wissenschaft und Wirtschaft (2 + 3-Projekte)

Förderkennzeichen: 01DM21004A; 01DM21004B; 1DM21004C

Projektlaufzeit: 01.04.2021 – 31.03.2024

Berichtszeitraum: 01.04.2021 – 31.03.2024

Sachbericht

Werkzeugmaschinenlabor WZL der RWTH Aachen University

IconPro GmbH

Ford-Werke GmbH

Inhaltsverzeichnis

Inhaltsverzeichnis.....	2
1. Einleitung.....	3
2. Stand der Technik.....	3
2.1 Widerstandspunktschweißen.....	4
2.2 Laserlöten.....	5
2.3 Qualitätssicherung beim Schweißen.....	7
2.4 Computer Vision.....	8
3. Stand der Forschung.....	9
3.1 Bestehende KI-Ansätze zur automatisierten Qualitätskontrolle für NDT beim Schweißen.....	9
3.2 Bestehende KI-Ansätze zur Reduktion der Aufwände für die Datenbeschaffung beim Schweißen.....	11
4. Vorgehen im Projekt.....	13
5. Ergebnisse.....	14
5.1 Widerstandspunktschweißen.....	14
5.1.1 Prozess- und Datenverständnis.....	14
5.1.2 Datenaufnahme, -vorverarbeitung und –labeling.....	16
5.1.3 Entwicklung der KI-Software.....	17
5.1.4 Integration und Datenkommunikation.....	23
5.2 Laserlöten.....	24
5.2.1 Prozess- und Datenverständnis.....	24
5.2.2 Datenaufnahme, -vorverarbeitung und –labeling.....	25
5.2.3 Entwicklung der KI-Software.....	25
6. Zusammenfassung & Ausblick.....	27
Kontakt.....	28
Literaturverzeichnis.....	28

1. Einleitung

Einer der essenziellen Fertigungsschritte in der industriellen Produktion bzw. Metallverarbeitung, wie z.B. im Maschinen-, Anlagen- oder Automobilbau, ist das automatisierte Fügen metallischer Strukturen. Ohne den Einsatz von Fügetechnik wären viele industrielle Güter nicht herstellbar. Als Schlüsselbereich der gesamten deutschen Wertschöpfungskette betrug die Wertschöpfung aus Produktion und Anwendung von Fügetechnik im Jahr 2016 24,8 Mrd. € [1]. Eines der wichtigsten Fügeverfahren ist das Schweißen. Der Schweißprozess ist ein hochkomplexes Zusammenwirken verschiedener sowie schwierig zu beschreibender physikalischer und chemischer Einflussparameter [2]. Darüber hinaus bestimmen die Einspannbedingungen, die Geometrie der zu fügenden Bauteile sowie die Umgebungsbedingungen wie die Schutzgaszufuhr die Schweißqualität [3]. In der Automobilproduktion werden primär die Schweißverfahren Widerstandspunktschweißen und Laserlöten eingesetzt, welche robotergeführt im Karosseriebau zum Fügen von Metallblechen verwendet werden. Gleichzeitig muss jedoch im automobilen Karosseriebau ein enorm hoher Qualitätssicherungsaufwand betrieben werden, mit dem unter anderem durch stichprobenartige zerstörende Prüfungen in allen Schweißbaugruppen die Funktionalität der Fügestellen überprüft werden muss [4]. Allein durch die zerstörenden Prüfungen entsteht Ausschuss lediglich zur Qualitätssicherung in allen Aufbaustufen.

Der Hauptschwerpunkt dieses internationalen Vorhabens SSQC lag daher in der Entwicklung eines ultraschallbasierten, nicht-zerstörenden Überwachungs- und Verarbeitungssystemen mit künstlicher Intelligenz (KI) für das Widerstandspunktschweißen (RSW) und das Laserlöten (LB). Diese Sensor-KI-Systeme zielen auf die automatisierte Ultraschallprüfung in Echtzeit während des Fügeprozesses ab, was unter anderem die Konstruktion und Entwicklung neuer Einzelement- oder Phased-Array-Prüfköpfe zur Überwachung jedes Fügeprozesses erforderte. Um eine automatische Interpretation der Ultraschall-daten in Echtzeit zu ermöglichen, wurden modernste Deep Learning Modelle implementiert und trainiert und entsprechende Methoden angewendet. Diese Echtzeit-Prozess-Überwachungssysteme stellen wichtige Schritte in Richtung Null-Fehler-Fertigung, nicht-zerstörender Prozessüberwachung 4.0 und Industrie 4.0 dar, die beträchtliche wirtschaftliche und ökologische Vorteile mit sich bringen.

Das Ziel dieses Projekts lag also in der Entwicklung von intelligenten Schweißüberwachungssystemen für Widerstandspunktschweißen und Laserlöten, welche Echtzeit-Ultraschall-daten nutzen und diese automatisch mit Hilfe von Deep Learning Algorithmen interpretieren, um den aktuellen Fügeprozess zu bewerten. Im vorliegenden Ergebnisbericht zum Forschungsprojekt SSQC werden die relevanten Grundlagen zum Widerstandspunktschweißen, Laserlöten und der nicht-zerstörenden Prüfung sowie Computer Vision beschrieben. In Kapitel 3 wird der Stand der Forschung aufgezeigt. Darauf aufbauend werden in Kapitel 4 das Projektvorgehen beschrieben und die zugehörigen Ergebnisse in Kapitel 5 vorgestellt. Kapitel 6 fasst die wesentlichen Erkenntnisse des Forschungsprojektes zusammen und liefert einen Ausblick auf weitere Forschungsthemen über das Projekt hinaus.

2. Stand der Technik

In den folgenden Abschnitten werden die notwendigen Grundlagen für das Verständnis des Vorgehens im Forschungsprojekt und der zugehörigen Ergebnisse erläutert. Diese umfassen das Verständnis für das Widerstandspunktschweißen und Laserlöten, die die zwei Anwendungsfälle darstellen. Außerdem werden Grundlagen zur nicht-zerstörenden Ultraschallprüfung sowie Computer Vision aufbereitet, welche den Kern der Forschungsaktivitäten darstellen.

2.1 Widerstandspunktschweißen

Das Widerstandspunktschweißen („RSW“) ist eines der wesentlichsten Fügeverfahren in der Automobilindustrie, das seit über einem Jahrhundert verwendet wird. Es wurde erstmals in den frühen 1900er Jahren eingeführt und revolutionierte die Art und Weise, wie Fahrzeuge gebaut wurden. Es wird hauptsächlich verwendet, um dünne Metallbleche (0,5 mm bis 3 mm), die häufig in der Fahrzeugkarosserie verwendet werden, zu verbinden. Aufgrund der kurzen Prozesszeiten und sehr guten Automatisierbarkeit bei hoher Wiederholgenauigkeit ist dieses Verfahren besonders für die Massenproduktion geeignet. Daher werden seit mehreren Fahrzeuggenerationen in modernen Fertigungsstraßen Roboter eingesetzt, um Tausende von Schweißpunkten präzise und schnell zu setzen, was die Produktionskosten senkt und die Effizienz steigert [5].

Das Widerstandspunktschweißen ist eine stoffschlüssige Füge-technik aus der Klasse der Pressschweißverfahren für elektrisch leitfähige Werkstoffe [6]. Der Stoffschluss ist nicht lösbar und erfolgt an der Verbindungsstelle durch einen lokalen Schmelz- und Erstarrungsvorgang von zwei oder mehr Fügepartnern unter Kräfteinfluss [7]. Die benötigte Schweißenergie wird beim Widerstandspunktschweißen als Schweißstrom mit einer typischen Stromstärke im Bereich von mehreren Tausend Ampere durch stiftförmige Elektroden, die Teil der Schweißzange sind, an den Kontaktstellen in die übereinanderliegenden Metallbleche eingeleitet. Die Verbindungsstelle der Bleche erreicht die Schweißtemperatur durch Widerstandserwärmung nach dem Ohm'schen Gesetz. Durch die erzeugte Wärme bildet sich ein Schmelzbad, das als Schweißlinse, im Folgenden auch „Nugget“ genannt, bezeichnet wird. Diese Schweißlinse verbindet die Werkstücke miteinander. Der Stromfluss wird für eine bestimmte Zeit aufrechterhalten, um sicherzustellen, dass die Schweißlinse die richtige Größe und Form erreicht. Der elektrische Strom wird abgeschaltet, während die Anpresskraft F_p aufrechterhalten wird. Dies ermöglicht es, der Schweißlinse abzukühlen und zu erstarren. Nach einer kurzen Abkühlzeit wird der Druck langsam reduziert und die Elektroden werden entfernt. [4]. Die Anpresskraft F_p sorgt für einen guten elektrischen Kontakt und stellt sicher, dass die Wärme gleichmäßig verteilt wird. Der Druck wird kontinuierlich überwacht und angepasst, um ein Herausspritzen, im Folgenden „Expulsion“ genannt, zu verhindern [2]. Ein Zusatzstoff oder Schutzgas ist für das Widerstandspunktschweißen nicht notwendig [8]. Der Aufbau des Widerstandspunktschweißens ist schematisch in Abbildung 1 dargestellt und bildet sowohl die typische Schweißlinsenform als auch den richtungsabhängigen Wärmefluss ab.

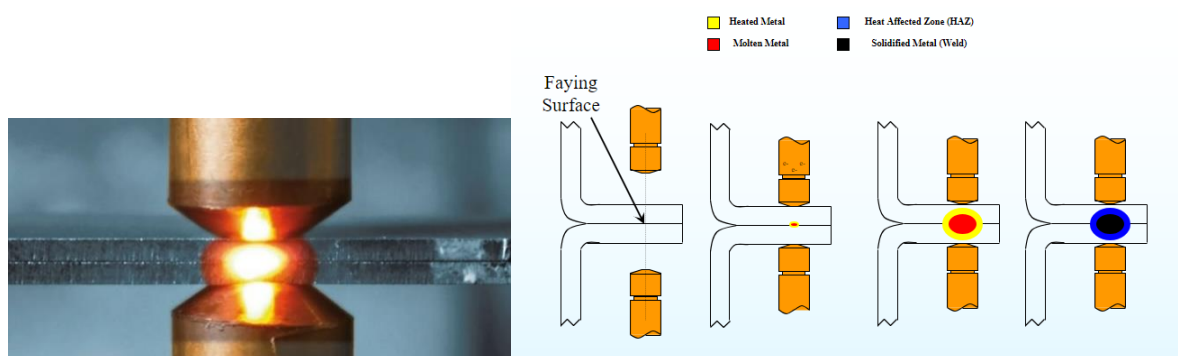


Abbildung 1: Schematische Darstellung des Widerstandspunktschweißprozesses

In der folgenden Tabelle sind die Vor- und Nachteile des Widerstandspunktschweißens gegenübergestellt. Insgesamt ist das Widerstandspunktschweißen ein unverzichtbares Verfahren in der Automobilindustrie, das durch seine Effizienz, Wirtschaftlichkeit und die hohe Qualität der Verbindungen besticht.

Tabelle 1: Übersicht über die Vor- und Nachteile des Widerstandspunktschweißens

Vorteile	Nachteile
<ul style="list-style-type: none"> • Hohe Geschwindigkeit und Effizienz • Leichte Automatisierung • Geringe Betriebskosten und die Fähigkeit, große Stückzahlen zu produzieren, machen das Verfahren kosteneffektiv. • Hohe Festigkeit und Dauerhaltbarkeit • Kein zusätzlicher Materialverbrauch 	<ul style="list-style-type: none"> • Begrenzte Werkstoffkompatibilität: Es kann schwierig sein, verschiedene Metallarten miteinander zu verschweißen, da unterschiedliche Metalle unterschiedliche Widerstandswerte haben. • Hohe anfänglichen Investitionskosten • Eingeschränkte Flexibilität

Das Widerstandspunktschweißen eignet sich für dünne Metallbleche, meist aus Stahl oder Aluminium. Für eine hohe resultierende Qualität müssen die Werkstücke sauber und frei von Verunreinigungen sein, um eine gute elektrische Kontaktierung und Schweißqualität zu gewährleisten. Kupferbasierte Elektroden werden aufgrund ihrer hohen elektrischen und thermischen Leitfähigkeit verwendet. Die Elektroden sind wassergekühlt, um Überhitzung zu vermeiden. Die genaue Positionierung der Metallbleche ist entscheidend, um eine gleichmäßige Schweißnaht zu gewährleisten. Trotz der guten Automatisierbarkeit des Widerstandspunktschweißens ist die Inline-Prozessüberwachung in der Automobilindustrie aufgrund der hohen Prozessgeschwindigkeiten und der schlechten Zugänglichkeit weiterhin herausfordernd. In der Praxis werden daher offline zerstörungsfreie Prüfungen wie Ultraschall oder Röntgen verwendet, um die Integrität der Punktschweißverbindung zu überprüfen. Darüber hinaus finden in regelmäßigen Abständen stichprobenartig zerstörende Prüfungen statt, um die mechanischen Eigenschaften der Schweißverbindung zu bewerten. Ein kosteneffizientes Ultraschallsensor-KI-System für die 100% inline-Prüfung, besonders von kritischen Punktschweißverbindungen, hätte somit erhebliche Mehrwerte für die Automobilindustrie [9].

2.2 Laserlöten

Das Laserlöten ist eine noch relativ neue und fortschrittliche Fügetechnologie, die in der Automobilindustrie erheblich an Bedeutung gewonnen hat. Bei diesem Verfahren wird ein Laser verwendet, um ein Füllmaterial aufzuschmelzen, das anschließend zwei oder mehr Metallkomponenten verbindet. Der Prozess bietet mehrere Vorteile gegenüber herkömmlichen Schweiß- und Lötmethoden, insbesondere in Bezug auf Präzision, Geschwindigkeit und die Qualität der hergestellten Verbindungen. Im automobilen Karosseriebau wird das Laserlöten zum Verbinden verschiedener Karosserieteile, insbesondere für sichtbare Verbindungen wie z.B. an Dächern, Türen und Motorhauben, verwendet, um eine hohe strukturelle Integrität und ästhetische Qualität zu gewährleisten. Im Zuge der Elektromobilitäts-Transformation gewinnt das Laserlöten an Bedeutung, da es eine entscheidende Bedeutung beim Verbinden von Batteriekästen und anderen Komponenten, die hohe Präzision und minimale thermische Einwirkungen erfordern, einnimmt. Neben dem automobilen Karosseriebau findet es vor allem in der Fertigung elektronischer Bauteile Anwendung [9].

Löten im Allgemeinen ist ein stoffschlüssiges, nicht lösbares Fügeverfahren, das eine Wärmezufuhr benötigt. Die Grundwerkstoffe werden unter Zunahme eines nicht artgleichen Zusatzwerkstoffs, dem

Lot, verbunden. Dazu wird ausschließlich der Zusatzwerkstoff aufgeschmolzen. Dies unterscheidet das Löten vom Schweißen, wo auch die Grundwerkstoffe aufschmelzen. Das Lot benetzt im flüssigen Zustand die festen Grundwerkstoffe und dringt aufgrund der Kapillarwirkung in den Lötspalt ein. Sind die Oberflächen der Grundwerkstoffe gut benetzt und wird die Löttemperatur ausreichend lange gehalten, tritt eine Art Legierungsbildung durch Diffusion ein. Die im Grundwerkstoff und Lot ausgebildeten Diffusionszonen erhöhen die Festigkeit der Lötverbindung [10]. Das Laserlöten wird zur Verfahrensvariante „Löten durch Strahl“ zugezählt. Die für den Fügeprozess benötigte Prozessenergie wird durch einen Laserstrahl bereitgestellt [11]. Das Laserlöten unterscheidet sich im Wesentlichen von anderen Lötverfahren durch eine präzise lokale Energieeinbringung, wodurch die Erwärmung der Grundwerkstoffe minimiert wird. Somit wird der Verzug der zu fügenden Bauteile stark verringert. Weitere Vorteile des Laserlötens sind die hohe Prozessgeschwindigkeit, sehr gute Automatisierbarkeit und die einseitige Zugänglichkeit. Da der Energieeintrag präzise und lokal erfolgt, ist eine hohe Positioniergenauigkeit von Laserstrahl und Werkstück notwendig. Trotz seiner diversen Vorteile sind mit dem Laserlöten einige Herausforderungen verbunden, wie die anfängliche Investition in Laserausrüstung und die Notwendigkeit von Fachkräften für die Einstellung des Prozesses und der Qualitätskontrolle. Die Vor- und Nachteile sind in Tabelle 2 gegenübergestellt [12].

Tabelle 2: Übersicht über die Vor- und Nachteile des Laserlötens

Vorteile	Nachteile
<ul style="list-style-type: none"> • <u>Hohe Präzision:</u> Das Laserlöten ermöglicht eine hohe Kontrolle über den Fügeprozess, was zu präzisen und konsistenten Verbindungen führt. Der fokussierte Laserstrahl sorgt für eine minimale Wärmeinbringung in die umliegenden Materialien, wodurch das Risiko von thermischen Verzerrungen verringert wird und die Integrität der Komponenten erhalten bleibt. • <u>Verbesserte ästhetische Anmutung:</u> Ein bemerkenswerter Vorteil des Laserlötens ist die überlegene Oberflächenbeschaffenheit der Verbindungen. Dies reduziert Nacharbeiten für Sichtnähten gegenüber anderen Verfahren massiv. • <u>Erhöhte Effizienz:</u> Die Geschwindigkeit des Laserlötens ist deutlich höher als bei herkömmlichen Methoden, was zu erhöhten Produktionsraten führt. • <u>Materialkompatibilität:</u> Das Laserlöten ist äußerst vielseitig und kann zum Verbinden einer Vielzahl von Materialien verwendet werden, einschließlich verschiedener Stahlsorten und Aluminiumlegierungen. Diese Flexibilität unterstützt den Einsatz von Leichtbaumaterialien. 	<ul style="list-style-type: none"> • Hohe anfängliche Investitionskosten • Aufwändige Qualitätssicherung • Hohe Anforderungen an Geometrie und Positioniergenauigkeit • Hohe Sicherheitsanforderung zur Gewährleistung der Lasersicherheit

Aufgrund des Einsatzes des Laserlötens im Sichtbereich der Karosserie müssen hohe Qualitätsstandards eingehalten werden. Dementsprechend werden hohe Qualitätssicherungsaufwände in der Automobilproduktion betrieben. Trotz dieser Qualitätsstandards können sogenannte Nahtimperfectionen entstehen [12]. Diese Nahtimperfectionen sind in der DIN EN ISO 18279 beschrieben und lassen sich in sechs Gruppen einteilen [13]. In Abbildung 2 sind die typischen Fehler, die in der Automobilproduktion vorkommen, anhand von exemplarischen Bildern aufgeführt [9].

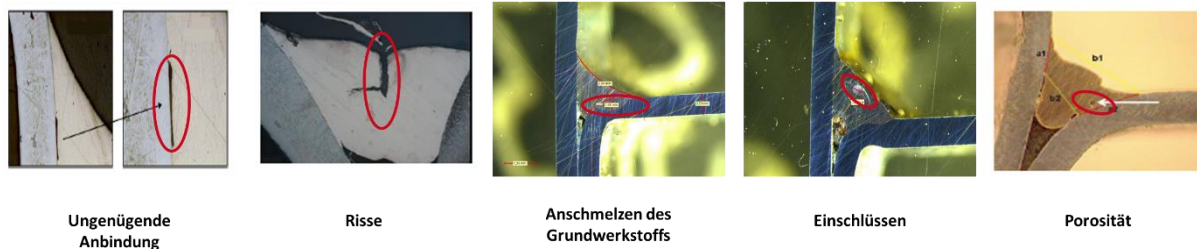


Abbildung 2: Übersicht über typische Fehler beim Laserlöten

Aufgrund der Vielzahl der möglichen Fehler und der schlechten Zugänglichkeit ist eine zerstörungsfreie Prüfung mit herkömmlichen Methoden kaum möglich. Es werden im wesentlichen Sichtprüfungen durchgeführt, die die Nahtqualität hinsichtlich Spritzer etc. anhand des äußeren Eindrucks bewertet. Für eine Analyse des Gefüges werden in regelmäßigen Abständen daher aufwändige Querschnitte an einer Karosserie durchgeführt und entsprechend zerstörend geprüft. Auch für das Laserlöten ist die in-line-Prozesskontrolle eine wesentliche Herausforderung. Daher hätte auch für diesen Anwendungsfall ein kosteneffizientes Ultraschallsensor-KI-System großes Potenzial, um die Art und Komplexität der Fehler in-situ zu beurteilen [9].

2.3 Qualitätssicherung beim Schweißen

Um die Qualität eines Schweißpunktes sicherzustellen, existieren verschiedene Arten von Prüfverfahren. Unterschieden wird zwischen Prüfverfahren für äußere und innere Unregelmäßigkeiten. Die Überprüfung auf äußere Unregelmäßigkeiten erfolgt üblicherweise durch eine Sichtprüfung. Innere Unregelmäßigkeiten einer Fügeverbindung lassen sich durch zerstörende und zerstörungsfreie Prüfverfahren untersuchen [14]. Die derzeit in Automobilwerken verwendeten Inspektionsmethoden zur Sicherstellung der Qualität von Laserlötverbindungen umfassen maßgeblich visuelle Inspektionen zur Erkennung von Oberflächenfehlern und Querschliffe von Verbindungen und deren Betrachtung unter einem optischen Mikroskop zur Messung der Größen von inneren Unvollkommenheiten. Echtzeitfähige Systeme zur Qualitätsprüfung im RSW- und LB-Prozess greifen auf die Überwachung von Prozessparametern wie Strom, Spannung, Laserleistung oder Prozessemissionen wie Kamerabild oder Lichtsignale im sichtbaren und infraroten Bereich zurück. Diese Systeme sind zwar im industriellen Einsatz, allerdings kann in allen Systemen nicht in die Fügeverbindung „hineingeschaut“ werden kann. Die so gewonnenen Qualitätsdaten müssen daher regelmäßig in Stichproben mit zerstörender oder durchleuchtender Prüfung wie der Ultraschallprüfung abgesichert werden [9].

Betrachtungsgegenstand des Forschungsprojekts ist die Ultraschallprüfung. Die Ultraschallprüfung gehört zu den akustischen, zerstörungsfreien Prüfverfahren und ermöglicht die volumetrische Analyse von Werkstücken auf innere Unregelmäßigkeiten. Die Ultraschallprüfung basiert auf Schallwellen, die durch einen Ultraschallprüfkopf in das zu prüfende Material eingeleitet werden und sich dort ausbreiten. Die Schallgeschwindigkeit ist hierbei material- und temperaturabhängig. Somit breiten sich die

eingeleiteten Schallwellen in verschiedenen Materialien unterschiedlich schnell aus und werden an Grenzflächen ungleicher Wellenimpedanz teilweise reflektiert. Eine Reflektion der Schallwellen tritt somit vor allem an Phasenübergängen, Einschlüssen, Rissen oder an der Werkstückoberfläche auf. Die Ultraschallprüfung wird in zwei verschiedenen Verfahren angewandt, dem Impuls-Echo- bzw. Reflexionsschall- und Durchschallungsverfahren. Beim Durchschallungsverfahren werden zwei Prüfköpfe, einer als Sender und einer als Empfänger, am Ultraschallprüfgerät angeschlossen und der Amplitudenabfall der Durchschallung gemessen. Beim Impuls-Echo-Verfahren fungiert der Ultraschallprüfkopf sowohl als Sender, als auch als Empfänger und es wird die Laufzeit zwischen entsendetem Schallimpuls und Reflektion gemessen. Diese reflektierten Signale werden als Amplitude über die Zeit gemessen, was man A-Scan nennt. Durch das horizontale Aneinanderreihen zeitlich aufeinanderfolgende A-Scans kann ein sogenannter M-Scan erzeugt werden und in eine zweidimensionale Bilddarstellung überführt werden. Hierbei ist auf der x-Achse die Zeit und auf der y-Achse die Laufzeit dargestellt. Die Pixelintensitätswerte ergeben sich aus der Amplitude der A-Scans [15]. Diese Art der Darstellung wird für die inline Ultraschallüberwachung beim Widerstandspunktschweißen verwendet, wobei ein Transducer in die Schweißzange ortsfest integriert ist. Da beim Laserlöten nicht in-situ sondern nur nachgelagert geprüft werden kann, besteht das Ultraschallprüfsystem beim Laserlöten aus mehreren Ultraschallwandlern, die unabhängig voneinander betrieben werden und über der Lötnaht positioniert werden. Die aufgenommenen A-Scans werden ebenfalls in ein zweidimensionales Bild überführt, wobei hier die x-Achse nicht der Zeit, sondern der Position entspricht. In diesem Fall spricht man vom B-Scan [15].

2.4 Computer Vision

Computer Vision (CV) ist ein Teilgebiet der Künstlichen Intelligenz (KI), das sich mit der automatischen Extraktion, Analyse und dem Verständnis nützlicher Informationen aus einem einzelnen Bild oder einer Bildfolge befasst. CV umfasst verschiedene Lernaufgaben wie Bildklassifikation, Objekterkennung oder semantische Segmentierung. Letztere ist eine CV-Aufgabe, bei der jedem Pixel in einem Bild ein Klassenlabel zugewiesen wird und dadurch semantisch zusammenhängende Bildregionen identifiziert werden (Referenz).

Da herkömmliche CV-Ansätze auf manuell definierter Merkmalsextraktion beruhen, die stark vom Fine-Tuning vieler Parameter abhängen, haben sich DL-Ansätze, die datenbasiert die zu extrahierenden Merkmale erlernen, für CV-Aufgaben wie die semantische Segmentierung etabliert [16]. Es gibt viele verschiedene konkrete DL-Modelle für die semantische Segmentierung, die hauptsächlich auf zwei verschiedenen Architekturprinzipien beruhen: Convolutional Neural Networks (CNNs) und Vision Transformer (ViTs).

CNNs verarbeiten Bilder als zweidimensionale Arrays und sind aufgrund ihrer Fähigkeiten zur hierarchischen Merkmalsextraktion und ihrer lokalen rezeptiven Filter, die in der Lage sind, räumliche Bildinformationen zu berücksichtigen, seit vielen Jahren für CV-Aufgaben etabliert. Inspiriert vom visuellen Kortex des Menschen bestehen CNNs aus mehreren Faltungs- und Pooling-Schichten, die es ermöglichen, eine Hierarchie zunehmend komplexerer Merkmale aus Rohdaten zu lernen. In den Faltungsschichten werden sogenannte „Feature Maps“ durch die Verwendung von Filterkernen erhalten, indem sie sich mit einer bestimmten Schrittweite über das Inputbild bewegen und eine Faltungsoperation auf dem rezeptiven Feld des Filters anwenden. Diese Filterkerne werden während des Modelltrainings mit Hilfe des Backpropagation-Algorithmus aus den Daten gelernt und so für die jeweilige Aufgabe optimiert. Um die räumliche Auflösung und Menge an Informationen zu verringern, werden die Werte benachbarter Pixel in den Pooling-Schichten aggregiert, indem in der Regel der höchste Wert

aus dem vom Pooling-Filter abgedeckten Bereich ausgewählt wird [16]. Encoder-Decoder-Netzwerke sind speziell für die semantische Segmentierungsaufgabe konzipiert und umfassen einen kontrahierenden und einen expansiven Pfad. Im Encoder werden mehrere Faltungsschichten und Pooling-Schichten übereinandergelegt, die die im Inputbild enthaltenen Merkmale kodieren. Die erhaltenen Merkmalskarten werden im expansiven Pfad dekodiert, z. B. durch "Aufwärtsfaltung", wodurch die Segmentierungsmaske für das entsprechende Eingangsbild entsteht [17]. U-Net und DeepLabv3+ sind zwei spezielle CNNs, die die Encoder-Decoder-Architektur implementieren. Während U-Net, das aus einer Encoder-Decoder-Architektur mit Skip-Verbindungen besteht, um räumliche Informationen hochauflösend zu erhalten, in biologischen Bildgebungsanwendungen weit verbreitet ist [17], zielt DeepLabv3+ auf die Erfassung komplexer und vielfältiger Bilddetails durch „atrous spatial pyramid pooling“ ab [18].

In der jüngsten Vergangenheit haben ViTs, die auf dem Attention-Mechanismus beruhen, große Aufmerksamkeit für CV-Aufgaben erreicht [19]. Im Gegensatz zu CNNs modellieren ViTs ein Bild nicht als zweidimensionales Array, sondern als eine Sequenz von Patches. Basierend auf dem Attention-Mechanismus zielen sie darauf ab, zu erlernen, auf welche Bildregionen sie sich konzentrieren sollen und sind somit besonders gut in der Lage, globale Merkmalsabhängigkeiten zu erfassen [20].

In weiteren Verlauf des Forschungsvorhabens werden die zwei repräsentativen CNNs (U-Net und DeepLabv3+) mit einem neuartigen ViT (SegFormer) hinsichtlich der Fähigkeit miteinander verglichen, Ultraschallbilder von Schweißprozessen akkurat zu segmentieren.

3. Stand der Forschung

Aufgrund der Komplexität von Schweißprozessen mit einer Vielzahl an Einflussfaktoren auf die resultierende Qualität, deren Zusammenwirken nicht explizit mathematisch modellierbar ist und den daraus resultierenden hohen Qualitätssicherungsaufwänden in der produzierenden Industrie, zielen bestehende Ansätze in der Forschung auf eine Automatisierung der Qualitätskontrolle durch KI ab. Echtzeitfähige Systeme zur Qualitätsprüfung im Widerstandspunktschweißen und Laserlöten greifen auf die Überwachung von Prozessparametern wie Strom, Spannung, Laserleistung oder Prozessemissionen wie Kamerabild oder Lichtsignale im sichtbaren und infraroten Bereich zurück. Diese Systeme können allerdings nicht in die Fügestelle „hineinschauen“, sodass gerade in der Automobilproduktion regelmäßig in Stichproben durchleuchtende Prüfungen wie die Ultraschallprüfung durchgeführt werden.

Da im Forschungsprojekt die Entwicklung eines Sensor-KI-Systems für die nicht zerstörende Prüfung für das Widerstandspunktschweißen und Laserlöten erforscht wurde, wird im Folgenden eine repräsentative Übersicht über bestehende Ansätze in der Forschung hierzu gegeben.

3.1 Bestehende KI-Ansätze zur automatisierten Qualitätskontrolle für NDT beim Schweißen

Der nachfolgende Abschnitt gibt einen Überblick über den Stand der Forschung im Bereich des Einsatzes von DL zur automatisierten, nicht zerstörenden Qualitätskontrolle (*engl.*: "Non-Destructive Examination" (NDE) oder "Non-Destructive Testing" (NDT)) von Schweißnähten. Hierzu wurde die Literaturliteraturdatenbank Scopus durchsucht. Eine repräsentative Übersicht der als relevant eingestuften Forschungsansätze wird hinsichtlich des konkreten NDT-Verfahrens, der zu überwachenden Qualitätsgröße und KI-Aufgabe sowie verwendeter DL-Architektur klassifiziert und ist in Tabelle 3 zusammengefasst.

Tabelle 3: Klassifizierung der relevanten Literatur über Deep Learning zur Qualitätskontrolle von Schweißnähten mittels zerstörungsfreier Prüfung

NDT-Verfahren	Qualitätsgröße	KI-Aufgabe	Deep Learning Modell	Referenz
X-Rays	Erkennung typischer Defekte (Porosität, Lack of Penetration, Cracks, ...)	Klassifikation	Resnet152 Model	[21]
X-Ray	Porosität	Objektdetektion	Faster RCNN	[22]
X-Ray	Porosität	Segmentierung	Point-REND Res-UNet-Modell	[23]
X-Rays	Nahtgeometrie	Segmentierung	NDD-Net	[24]
Ultraschall	i.O. vs. n.i.O.	Klassifikation	Deep Convolution Neural Network	[25]

Neben den zum Forschungsprojekt entstandenen Veröffentlichungen werden in der Literatur CV-Modelle auf Röntgen- oder Ultraschallbilder angewandt, die mit entsprechenden NDT-Methoden erzeugt wurden, um die innere Qualität der Schweißnaht zu beurteilen. [21] vergleicht beispielsweise 16 verschiedene DL-Modelle hinsichtlich der Klassifizierung von sieben Schweißnahtdefekttypen basierend auf Röntgenbildern. Für das Modelltraining erstellen die Autoren einen Datensatz mit wenigen realen Proben und erweitern diesen durch synthetische Daten. Im Durchschnitt erzielen sie eine Genauigkeit von 96,75 %, wobei das Resnet152-Modell am besten abschneidet. [22] trainieren ein Faster-RCNN-Modell für die Klassifizierung und Lokalisierung von kleineren Porositätsdefekten in Röntgenbildern auf einem selbst generierten Datensatz. Ihr Objekterkennungsalgorithmus erreicht eine mittlere durchschnittliche Präzision von 44,5 %. [23] und [24] trainieren Encoder-Decoder-Netzwerke für die semantische Segmentierung, was eine spezifische CV-Aufgabe zur pixelweisen Klassifizierung von Schweißnahtdefekten darstellt. Während [23] den öffentlichen GDxray-Datensatz verwenden und einen Dice-Score von 83,5 % erreichen, erzielen [24] einen Dice-Score von 58,25 % auf einem kleinen, selbst erstellten Datensatz. [25] verwenden ein CNN zur binären Klassifikation von Phased-Array Ultraschallsignalen in i.O / n.i.O-Teile.

Es lässt sich beobachten, dass DL zur automatischen Schweißnahtqualitätskontrolle von NDT-Anwendungen in der Forschung intensiv untersucht wird, wobei typischerweise CV-Modelle angewendet werden. Der Fokus liegt dabei allerdings auf X-Ray-Datensätzen und weniger auf der automatisierten Ultraschallprüfung. Außerdem wird auf einzelne öffentlich verfügbare Benchmark-Datensätze zurückgegriffen, die nicht unbedingt die realen Produktionsgegebenheiten abbilden. Es besteht also ein Forschungsdefizit in der Entwicklung eines robusten Sensor-KI-Systems zur Automatisierung der Ultra-

schallprüfung beim Widerstandspunktschweißen und Laserlöten unter ganzheitlicher Berücksichtigung industrieller Rahmenbedingungen. Außerdem gibt es nur wenige Arbeiten, die verschiedene hochmoderne DL-Architekturen systematisch evaluieren und vergleichen. Neuartige ViT-Architekturen werden kaum untersucht und mit traditionellen CNN-Modellen verglichen.

3.2 Bestehende KI-Ansätze zur Reduktion der Aufwände für die Datenbeschaffung beim Schweißen

Ein typisches Problem für den Einsatz von DL im Produktionskontext sind die hohen Aufwände für die Erzeugung industrieller Datensätze zum Modelltraining. [26] adressieren diesen Aspekt als "Hidden Technical Debt in Machine Learning Systems" und zeigen auf, dass nur ein Bruchteil eines produktiven ML-Systems das Machine Learning Modell ausmacht und wesentlich höhere Aufwände für die Datenbeschaffung und -labeling anfallen.

Dabei gibt es KI-Ansätze wie GANs [27] und Active Learning [28], die in der Lage sind, die hohen Aufwände für die Erzeugung eines umfangreichen, gelabelten und ausbalancierten Datensatz zu reduzieren. Im Rahmen des Forschungsprojektes wurde daher analysiert, welche dieser Ansätze es in der Forschung für den Anwendungsfall der Deep Learning basierten Qualitätskontrolle beim Schweißen gibt.

Da in der Literaturdatenbank Scopus keine relevanten Veröffentlichungen zum Einsatz von Active Learning für die semantische Segmentierung beim Anwendungsfall der automatisierten Qualitätskontrolle beim Schweißen gefunden wurden, wird im Folgenden eine repräsentative Übersicht über Forschungsansätze zur Deep-Learning-basierten Generierung synthetischer Daten für Anwendungen zur Qualitätskontrolle beim Schweißen gegeben. Die Abfrage der Literaturdatenbank Scopus mit dem Suchbegriff "welding" AND "synthetic data" OR "generative adversarial networks" ergab 59 Veröffentlichungen seit 2018. Es wurde eine systematische Literaturrecherche durchgeführt, bei der zunächst die identifizierte Literatur gesichtet und gefiltert wurde. Lediglich acht der 59 Paper konnten als relevant erachtet werden. Diese Veröffentlichungen wurden dann nach dem Schweißverfahren, den Eingangsdaten, der Lernaufgabe sowie dem detaillierten Modellierungsansatz klassifiziert. Die Ergebnisse der systematischen Literaturrecherche sind in Tabelle 4 zusammengefasst.

Von diesen acht Forschungsansätzen konzentrieren sich vier auf Lichtbogenschweißverfahren und zwei auf das Laserschweißen. Bei zwei weiteren Ansätzen werden keine Angaben zum untersuchten Schweißverfahren gemacht. Mit Ausnahme von [36], die einen Variational-Autoencoder (VAE) trainieren, um Zeitreihendaten zu erzeugen, werden alle anderen Ansätze verwendet, um synthetische Bilddaten entweder für Klassifizierungs- [34,29,35,30], Objekterkennungs- [33,31] oder semantische Segmentierungsaufgaben [32] zu erzeugen.

Von diesen sieben Ansätzen, die sich auf die Erzeugung von Bilddaten konzentrieren, befassen sich drei Ansätze mit Schweißnahtbildern für die nachgelagerte Qualitätsbewertung. [31] verwenden einen CycleGAN-basierten Modellierungsansatz, um synthetische Bilder von Lichtbogenschweißnähten zu erzeugen und anschließend einen Objekterkennungsalgorithmus zur Erkennung und Lokalisierung der Schweißnaht zu trainieren. Sie zeigen, dass das GAN die Leistung des Objekterkennungsalgorithmus verbessert. In ähnlicher Weise zielen [33] darauf ab, synthetische Schweißnahtbilder für die Objekterkennung zu erzeugen. Allerdings wenden sie ein StyleGAN2-basiertes Modell auf Laserschweißprozesse an. Zur Erzeugung synthetischer Bilder von Lichtbogenschweißnähten und entsprechender Segmentierungsmasken verwenden [32] eine spezielle neuronale Netzwerkarchitektur namens "SEAN" und trainieren anschließend ein Modell zur semantischen Segmentierung für die pixelweise Klassifizie-

nung von Schweißnahtbildern. Sie zeigen, dass die Verwendung eines generativen Netzes für die semantische Bildsynthese die Menge und Vielfalt der Trainingsdatensätze erhöhen und damit die Genauigkeit der Segmentierung auf unterrepräsentierten Schweißnahtbildern verbessern kann [32]. Im Gegensatz dazu werden in drei anderen Veröffentlichungen unterschiedliche Modellierungsansätze für die Erzeugung prozessbegleitender Bilder des Schweißbads untersucht, die anschließend für Qualitätsklassifikationen verwendet werden. [29] beispielsweise verwenden ein conditioned GAN-basiertes Bild-zu-Bild-Übersetzungsmodell namens pix2pix, um entauschte Schweißbadbilder zu erzeugen, die anschließend zur Qualitätsklassifizierung verwendet werden. [35] untersuchen die Deep Learning basierte Generierung synthetischer Bilddaten der inneren Gefügestruktur von Schweißnähten, die im Rahmen von zerstörungsfreien Prüfungen entstehen. Sie verwenden CycleGAN auf dem veröffentlichten GDXray-Datensatz [37] und generieren damit synthetische Daten für eine Qualitätsklassifikation. Sie zeigen die Effektivität von CycleGAN für Röntgenbilder von Schweißnähten, indem sie die Genauigkeit des Neuronalen Netzes zur Defektklassifikation, trainiert mit oder ohne synthetischen Daten, vergleichen.

Tabelle 4: Klassifizierung der relevanten Literatur über Deep Learning basierte synthetische Datengenerierung für die Qualitätskontrolle beim Schweißen

Schweißprozess	Verwendete Daten	KI-Aufgabe	Modellierungsansatz	Referenz
Lichtbogenschweißen	Bilder vom Schweißbad	Klassifikation	Pix2pix	[29]
Lichtbogenschweißen	Bilder vom Schweißbad	Klassifikation	Fusing-and-Filling GAN	[30]
Lichtbogenschweißen	Bilder der Schweißnaht	Objektdetektion	CycleGAN	[31]
Lichtbogenschweißen	Bilder der Schweißnaht	Semantische Segmentierung	SEAN	[32]
Laserschweißen	Bilder der Schweißnaht	Objektdetektion	StyleGAN2	[33]
Laserschweißen	Bilder optischer & thermischer Sensorik	Klassifikation	ACGAN	[34]
-	X-ray Bilder	Klassifikation	CycleGAN	[35]
-	Ultraschallsignale	-	VAE	[36]

Zusammenfassend zeigt sich allerdings, dass das Potenzial der Deep Learning basierten Generierung synthetischer Daten für die weitere Verbesserung von Algorithmen zur automatisierten Qualitätskontrolle von Schweißprozessen in der Forschung kaum untersucht ist. Dies gilt insbesondere für CV-Aufgaben wie die semantische Segmentierung, bei denen die Datenannotation besonders zeitaufwändig ist. Darüber hinaus konzentriert sich der Großteil der identifizierten Literatur auf die Generierung von In-Prozess- oder Post-Prozess-Bildern, die von Kamerasystemen erfasst werden. Nur eine identifizierte

Veröffentlichung untersucht die Anwendung von GANs auf Bilder der inneren Schweißnahtstruktur, die aus der zerstörungsfreien Prüfung stammen, deren Erstellung besonders kostenintensiv ist. Zusammenfassend lässt sich feststellen, dass es an Forschungsarbeiten mangelt, die die Fähigkeit von auf Deep Learning basierenden Ansätzen zur Erzeugung synthetischer Daten für semantische Segmentierungsaufgaben auf Bilddaten aus der zerstörungsfreien Prüfung bei Schweißanwendungen untersuchen. Auch das Potenzial von Active Learning für diesen Anwendungsfall ist in der Forschung nicht untersucht.

4. Vorgehen im Projekt

Im Forschungsprojekt hat das Konsortium bestehend aus Ford, IconPro und dem Werkzeugmaschinenlabor WZL der RWTH Aachen aus Deutschland sowie Tessonics und der University of Waterloo aus Kanada an einem Sensor-KI-System zur automatisierten Qualitätskontrolle der Ultraschallprüfung beim Widerstandspunktschweißen und Laserlöten gearbeitet. Ford agierte im Rahmen des Projektes als Use Case- und Anforderungsgeber und hat zudem Expertise zu den Schweiß- und Qualitätssicherungsprozessen eingebracht. Die University of Waterloo brachte Expertenwissen zu Schweißprozessen und der nicht-zerstörenden Prüfung von Schweißnähten in und hat zudem Labordatensätze für das Laserlöten erzeugt sowie Qualitätsuntersuchungen beim Widerstandspunktschweißen durchgeführt. Tessonics war verantwortlich für die Entwicklung der Ultraschallsensoren und war Teil des KI-Teams, dem auch das WZL und IconPro angehörten. Zudem hat das WZL die Projektkoordination auf deutscher Seite übernommen, die Schnittstelle zwischen KI-Team und Produktionstechnik gebildet und die Datenaufnahme beim Widerstandspunktschweißen unterstützt. In der Projektbearbeitung haben sich entsprechend kleine Subteams wie das KI-Team sowie größer und interdisziplinärer aufgestellte Teams jeweils für das Widerstandspunktschweißen und Laserlöten gebildet, die sich in regelmäßigen Abständen zu den aktuellen Aufgaben besprochen haben. Es wurde agil und hochiterativ in den einzelnen Teams gearbeitet. Für den übergreifenden Austausch fanden monatliche Meetings im gesamten Konsortium statt.

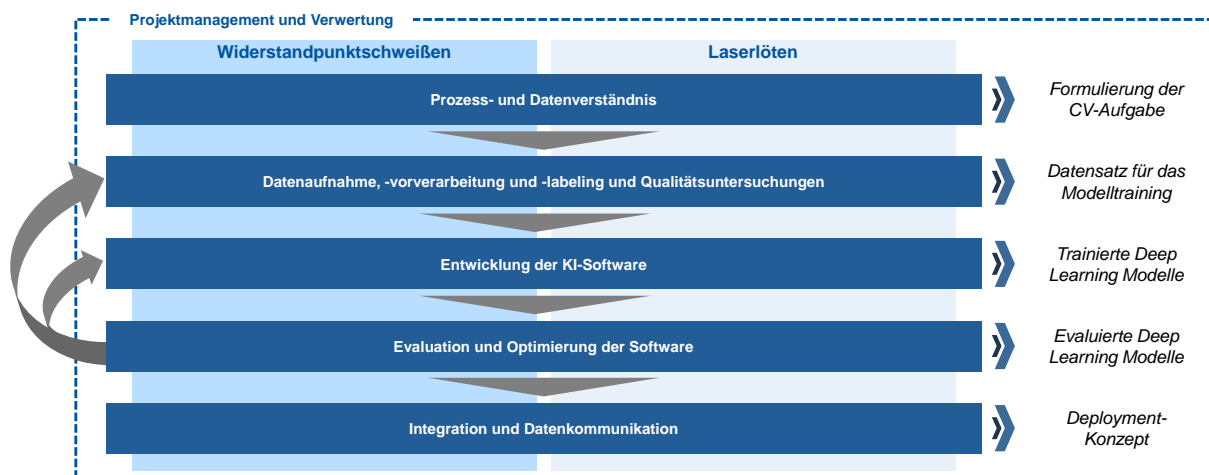


Abbildung 3: Darstellung des Projekttablaufs mit den zwei Workstreams Widerstandspunktschweißen und Laserlöten

Das Vorgehen im Projekt orientierte sich stark am CRISP-DM Framework und ist in Abbildung 3 dargestellt. Zunächst wurde ein Prozess- und Datenverständnis geschaffen und die KI-Aufgabe abgeleitet. Darauf aufbauend wurden die Daten aufgenommen, gelabelt und vorverarbeitet. Außerdem fanden

stichprobenartige Qualitätsuntersuchungen statt, um die Ergebnisse des KI-Modells hiermit korrelieren zu können. Sobald die Datengrundlage geschaffen war, wurden die DL-Modelle trainiert und anschließend evaluiert. Zusätzlich wurde eine Validierung in der laufenden Produktion durchgeführt und ein robustes Deploymentkonzept entwickelt. Diese Schritte wurden für die beiden Workstreams im Projekt, das Widerstandspunktschweißen und Laserlöten, konsekutiv durchlaufen. Aufgrund des höheren Reifegrads des Sensorsystems wurden die obengenannten Schritte für das Widerstandspunktschweißen als erstes durchgeführt. Im letzten Projektjahr waren dann Daten für das Laserlöten vorhanden und die modulare KI-Software wurde hierauf adaptiert und evaluiert.

5. Ergebnisse

In den nachfolgenden Abschnitten werden die Ergebnisse der vorgestellten Arbeitsschritte für die beiden Workstreams Widerstandspunktschweißen und Laserlöten vorgestellt und erläutert.

5.1 Widerstandspunktschweißen

Analog zum konsekutiven Projektverlauf werden zunächst die wesentlichen Ergebnisse des Sensor-KI-Systems für das Widerstandspunktschweißen vorgestellt.

5.1.1 Prozess- und Datenverständnis

Die Grundlage für das Sensor-KI-System bilden die Ultraschallsensoren des Konsortialpartners Tesonics. Diese Sensoren sind, wie in Abbildung 4 zu sehen ist, in den Elektroden der Widerstandspunktschweißzange integriert und senden Ultraschallsignale aus, die an Oberflächen des Bleches reflektiert werden. Während des Schweißprozesses werden diese Reflektionen in bestimmten Zeitabschnitten aufgenommen, um die Veränderungen des Gefüges durch die Bildung der Schweißperle und letztlich die stoffliche Verbindung der zusammengeschweißten Materialien zu verfolgen.

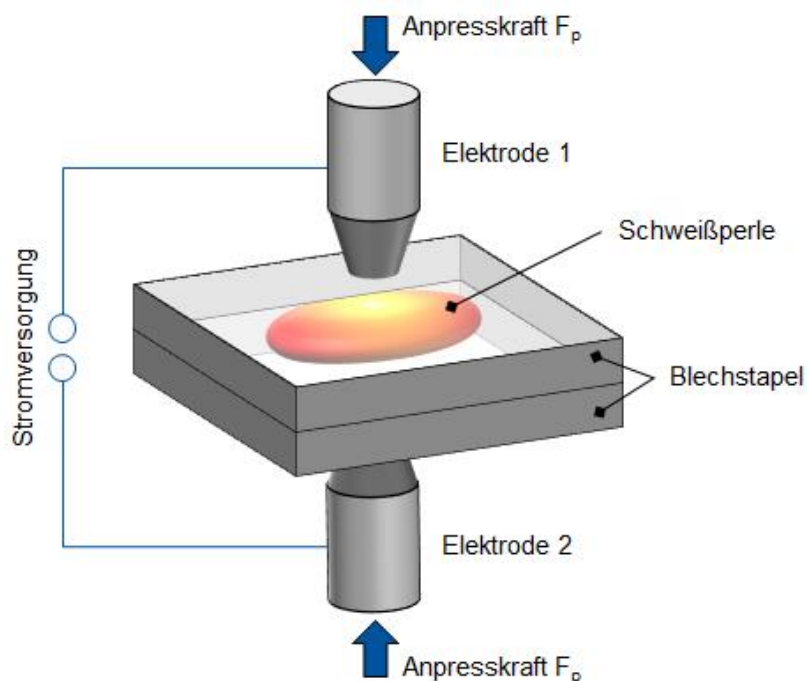


Abbildung 4: Schematische Darstellung des Widerstandspunktschweißens

Wie in Kapitel 2.3 erklärt, werden diese zeitlich segmentierten Aufnahmen der Reflektionen als A-Scans bezeichnet. Wenn die A-Scans horizontal über die Zeit zusammengesetzt werden, ergeben sie den sogenannten M-Scan wie in Abbildung 5 exemplarisch dargestellt, welcher als Input für das KI-System dient.

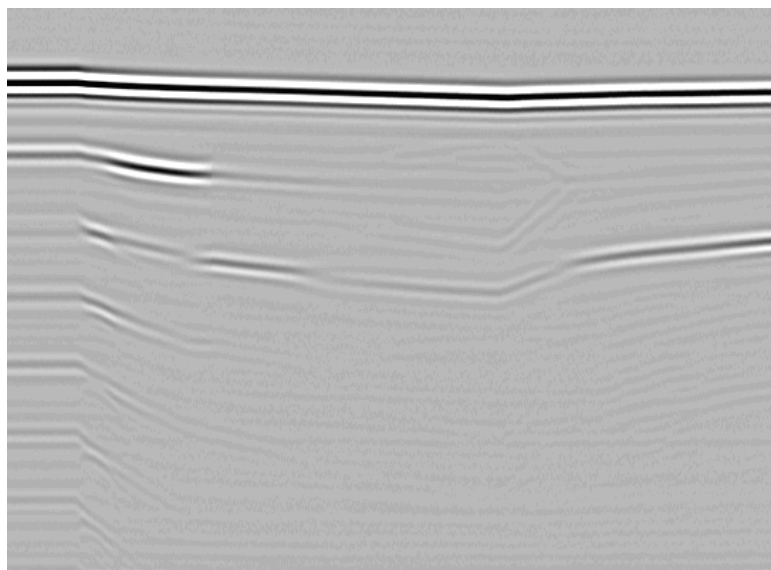


Abbildung 5: Exemplarische Darstellung des M-Scans

Die Interpretation des M-Scans ist schematisch in Abbildung 6 dargestellt. Man erkennt die zeitliche Entwicklung des Schweißprozesses, die für die Qualitätsbeurteilung in Abhängigkeit von den Material- sowie Prozessdaten herangezogen wird. Von besonderer Bedeutung für die Beurteilung der Güte des Schweißprozesses ist hierbei die Entwicklung der Schweißlinse (engl. "Nugget"), insbesondere deren Größe und Form, die auf den M-Scans sichtbar ist.

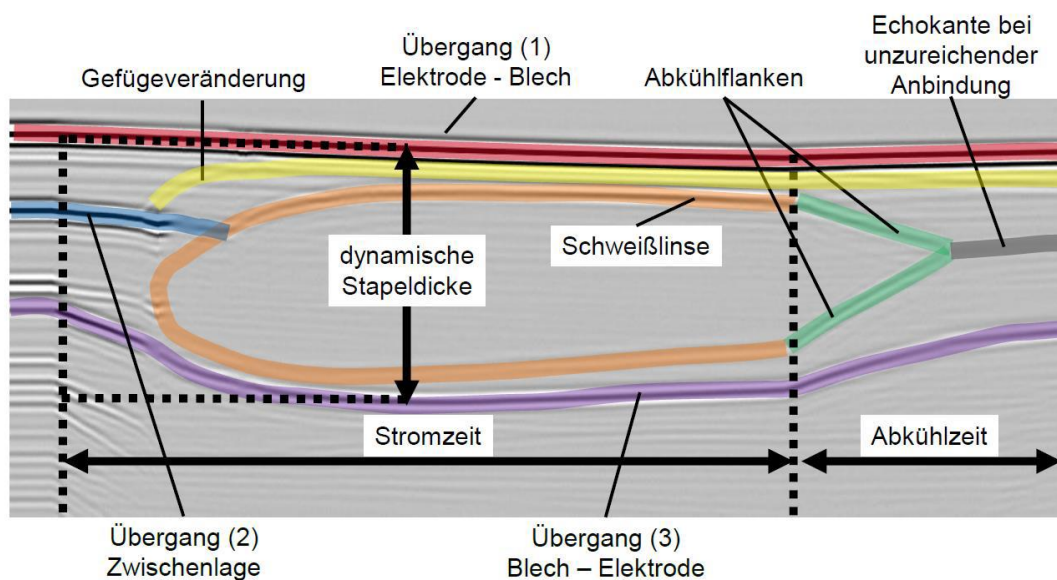


Abbildung 6: Interpretation eines M-Scans vom Widerstandspunktschweißen

Basierend auf dem Prozess- und Datenverständnis lässt sich die KI-Aufgabe ableiten. Da die Größe und Form des Nuggets maßgeblich die Qualität des Widerstandspunktschweißprozesses bestimmen, besteht die zentrale Aufgabe des KI-Systems in der Erkennung des Nuggets auf den M-Scans. Aufgrund der geometrischen Ausprägung des Nuggets und der prozessbedingten Anforderungen ist eine simple Bounding Box nicht ausreichend, sondern eine semantische Segmentierung, also eine pixelweise Klassifikation des Bildes erforderlich. Im Rahmen des Projekts ist die Aufgabe des KI-Systems daher, zu bestimmen, welche Bildpixel der M-Scans der Klasse „Nugget“ zuzuordnen sind. Die durch die Modelle segmentierten M-Scans können anschließend über ein Postprocessing-Skript zur Qualitätsbeurteilung weiterverarbeitet werden, um beispielsweise Größe und Form des Nuggets zu charakterisieren und die Qualität des Schweißprozesses zu bewerten.

5.1.2 Datenaufnahme, -vorverarbeitung und -labeling

Nach Definition der KI-Aufgabe galt es, die Datengrundlage für die Entwicklung der KI-Software zu schaffen. Ein wichtiger Bestandteil dieses Plans war die Bereitstellung eines „Basisdatensatzes“ aus dem Labor durch die kanadischen Projektpartner, der 1299 Punktschweißungen umfasst. Da es sich bei der definierten KI-Aufgabe zur semantischen Segmentierung von M-Scans um ein überwachtes Lernen handelt, müssen die M-Scans gelabelt werden. Für die spätere pixelweise Klassifikation des Nuggets durch die Deep Learning Modelle war es also erforderlich, das Nugget manuell auf den M-Scans einzuzichnen, um die Pixel zu definieren, die zum Nugget gehören. Dies wurde mit dem Labeling-Tool „SuperAnnotate“ umgesetzt.

Die wesentlichen Anforderungen der Deep Learning Modelle an den Trainingsdatensatz sind, dass dieser eine große Menge an qualitativ hochwertigen gelabelten Bildern umfassen muss. Zusätzlich sollten die im Labor erzeugten Daten den Anwendungsfall in der laufenden Produktion realitätsnah abbilden und die verschiedenen Güteklassen in ähnlicher Häufigkeitsverteilung umfassen. Da sich insbesondere die Menge an Daten, die für ein Modelltraining erforderlich sind, nicht im Vorhinein abschätzen lassen, wurde sich für ein hochiteratives Vorgehen entschieden. Das bedeutet, dass sehr früh die Datenpipeline für das Modelltraining implementiert und bereits mit einem kleinen verfügbaren gelabelten Datensatz trainiert wurde. Neu verfügbare Daten und Label wurden dann sukzessive eingebunden und neue Anforderungen an die Datenaufnahme abgeleitet. Basierend darauf wurde in den Schweißlaboren in Deutschland ein zweiter Labordatensatz mit verschiedenen Materialkombinationen aufgebaut, der auch verstärkt Daten umfasst, die zu fehlerhaften Schweißungen gehören.

In Summe lagen für den Workstream Widerstandspunktschweißen somit die folgenden Datensätze vor:

1. Ein initialer Labordatensatz aus Kanada mit 1299 Datenpunkten.
2. Im Projekt in Deutschland entstandene Labordaten mit 1232 Datenpunkten.

Da ein erfolgreiches KI-Modelltraining eine große Menge an qualitativ hochwertigen gelabelten Daten erfordert, fiel hierfür ein erheblicher Aufwand an. Die Bilder wurden manuell annotiert und die Labelqualität in gemeinsamen Review-Terminen mit den Fachexperten sichergestellt. Aufgrund des hohen Aufwands wurden im Projekt innovative Ansätze für ein semiautomatisiertes Labeling erforscht. Diese Ansätze umfassen Active Learning und Generative Adversarial Networks (GANs).

5.1.3 Entwicklung der KI-Software

Im nachfolgenden Abschnitt werden das Vorgehen beim Modelltraining und der Evaluation sowie die resultierenden Ergebnisse der KI-Software vorgestellt. Dabei wird auf die drei oben genannten verschiedenen Aspekte der semantischen Segmentierung von M-Scans, dem Active Learning Framework zur Reduktion manueller Labeling-Aufwände und den GANs zur Erzeugung neuer, synthetischer M-Scans eingegangen.

Semantische Segmentierung der M-Scans

Zunächst erfolgten eine umfassende Recherche und Vorauswahl geeigneter Deep-Learning-Architekturen für die binäre semantische Segmentierung der M-Scans in "Nugget" und "kein Nugget" basierend auf Literatur und Benchmarks. Es wurden die Architekturen U-Net, DeepLabV3 und SegFormer implementiert:

- **U-Net:** CNN-basierte, symmetrische Encoder-Decoder-Architektur mit Skip-Verbindungen.
- **DeepLabV3:** CNN-basierte Segmentierungsarchitektur mit Atrous Convolution und Atrous Spatial Pyramid Pooling.
- **SegFormer:** Transformer-basierte Segmentierungsarchitektur.

Eine modulare End-to-End-Trainingspipeline wurde entwickelt, die das Einlesen und Vorverarbeiten der Bilder und Masken, die Bildung des train-val-test-Splits, die Implementierung des Batch-Generators mit Augmentierungsfunktionen sowie den Trainings- und Evaluationsprozess umfasst. Ein automatisiertes Bayes'sches Hyperparameter-Tuning optimiert die Dropout-Rate, Modelltiefe und Anzahl der Filter je Layer.

Die Modelle werden nach dem k-fold cross-validation-Prinzip trainiert und evaluiert, um eine robuste Bewertung der Modellleistung zu gewährleisten. Die gesamte Implementierung erfolgt in Python unter Verwendung von Open-Source-Software, insbesondere TensorFlow, um ein effizientes On-Edge-Deployment zu ermöglichen.

Im Projektverlauf wurde hochiterativ trainiert und evaluiert, kleine Bugs gefixt, Optimierungen an der Datenpipeline und -vorverarbeitung vorgenommen, Hyperparameter mit zunehmendem Erfahrungswissen angepasst und sukzessive neu verfügbare Daten und Labels eingebunden. Nachfolgend werden nicht alle Evaluationen der im Projektverlauf durchgeführten Experimente aufgezeigt, sondern nur die finalen Ergebnisse für die verschiedenen Datensätze. Die Interpretation der Ergebnisse erfolgt anhand der mean Intersection over Union (mIoU) und der visuellen Analyse.

Die mIoU ist eine gebräuchliche Kennzahl zur Bestimmung der Genauigkeit bei der semantischen Segmentierung, indem sie den Bereich der Schnittmenge („Intersection“) zwischen der vorhergesagten Segmentierung und der tatsächlichen Segmentierung mit dem Bereich ihrer Vereinigungsmenge („Union“) vergleicht. Die Berechnung erfolgt, indem in einem ersten Schritt für jede zu segmentierende Klasse c die Intersection over Union (IoU) berechnet wird und dann der Durchschnitt der IoU-Werte über alle Klassen C gebildet wird.

$$mIoU = \frac{1}{C} \sum_{c=1}^C IoU_c$$

$$\text{mit } IoU_c = \frac{\text{Intersection}}{\text{Union}} = \frac{\text{Anzahl Pixel in der Prediction und Ground Truth}}{\text{Gesamtzahl Pixel, die entweder in Prediction und Ground Truth liegen}}$$

Es ergeben sich die folgende Performance-Metriken für:

- die initialen Labordaten, die in Kanada erzeugt wurden:

Tabelle 5: Übersicht über die Performance-Metriken des U-Net-Modells gemessen anhand der mIoU

	mIoU
U-Net	97.8 %

- die Labordaten nach Ford-Spezifikationen, die in Deutschland erzeugt wurden:

Tabelle 6: Übersicht über die Performance-Metriken der verschiedenen DL Modelle nach Metallblechkombination und insgesamt

	Anteil	mIoU		
		U-Net	DeepLabv3+	SegFormer
Gesamter Testdatensatz	100 %	83.81 %	81.56 %	81.41 %
Blechmaterialkombination 1	37.97 %	86.52 %	85.47 %	85.95 %
Blechmaterialkombination 2	36.77 %	82.40 %	80.90 %	79.59 %
Blechmaterialkombination 3	25.26 %	79.46 %	56.63 %	66.75 %

Bereits nach den ersten Trainingsiterationen mit etwa 500 M-Scans konnten hohe Genauigkeiten mit einer mIoU von über 90 % beobachtet werden. Im weiteren Projektverlauf wurden teilweise Genauigkeiten von bis zu 97,8 % erzielt. Durch die Einbindung zusätzlicher Daten, insbesondere der in Deutschland generierten Bilder, wurde die Generalisierbarkeit der Modelle weiter gesteigert. Die finalen Ergebnisse der KI-Modelle für das Widerstandspunktschweißen zeichnen sich somit durch sehr hohe Genauigkeiten und eine gute Generalisierbarkeit aus, was anhand des folgenden Bildes verdeutlicht wird.

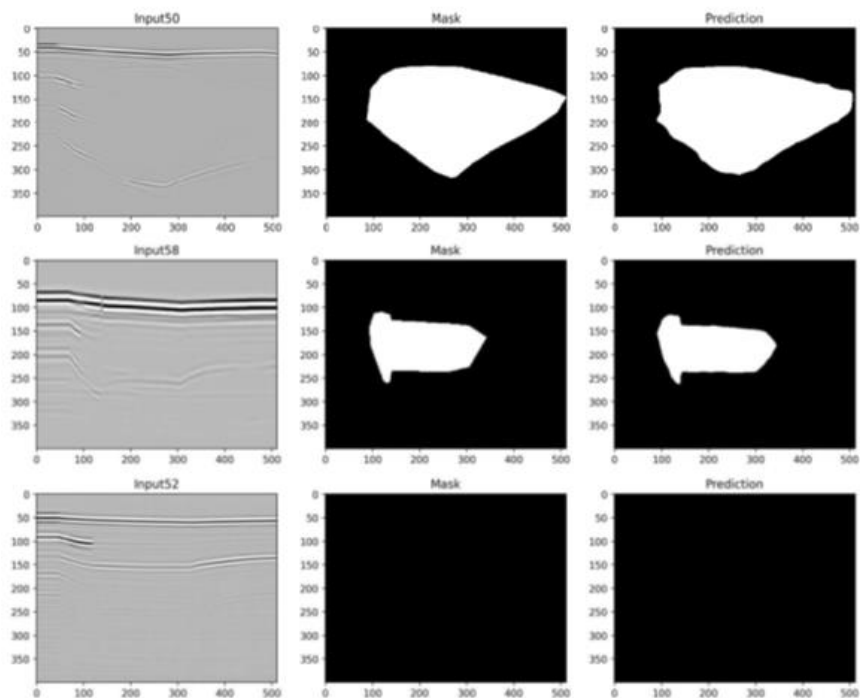


Abbildung 7: Gegenüberstellung von Vorhersagen des U-Nets mit wahrem Label für exemplarische M-Scans vom Widerstandspunktschweißen

Der Modellvergleich und die Auswertungen des Hyperparameter-Tunings zeigen, dass bereits ein U-Net sehr gute Ergebnisse auf den relativ informationsarmen Bildern liefert. Eine zusätzliche Tiefe des Netzes erweist sich als vorteilhaft für die Segmentierung. Zudem hat sich gezeigt, dass nicht zu viele Filter in den einzelnen Schichten des CNNs benötigt werden, um optimale Ergebnisse zu erzielen.

Auf dem zweiten in Deutschland erzeugten Datensatz wurden die drei verschiedenen DL-Modelle hinsichtlich ihrer Fähigkeit zur Generalisierung auf unterschiedliche Metallblechkombinationen untersucht. Die Ergebnisse sind in Tabelle 6 dargestellt. Die Ergebnisse für die spezifisch verwendeten Modell- und Trainingsparameter zeigen, dass das U-Net-Modell das DeepLabv3+ als auch den ViT SegFormer auf dem gesamten Datensatz und für jede Blechkombination übertrifft. Während alle Modelle für die Blechkombinationen 1 und 2 eine relativ hohe Segmentierungsgenauigkeit von etwa 80 % und mehr erreichen, sinkt die Leistung von DeepLabv3+ und SegFormer für Blechkombination 3, die den kleinsten Anteil der verfügbaren Daten ausmacht, signifikant auf 56,63 % bzw. 66,75 %. Das U-Net hingegen kann eine relativ hohe Segmentierungsgenauigkeit von 79,46 % aufrechterhalten und liefert somit nicht nur die genauesten, sondern auch die robustesten M-Scan-Segmentierungen unter den verglichenen Architekturen. Das U-Net scheint für diesen speziellen Anwendungsfall am besten geeignet zu sein, bei dem die Eingangsbilder relativ informationsarm und relevante Merkmale zur Segmentierung des Schweißpunktes im Allgemeinen im Zentrum des M-Scans beschränkt sind. Die Fähigkeit von DeepLabv3+ und SegFormer, globale Merkmalsabhängigkeiten zu erfassen, scheint gegenüber den festen, lokalen rezeptiven Feldern im U-Net keine Vorteile zu bieten. Diese Ergebnisse unterstreichen die Bedeutung der Modellauswahl für die Schweißqualitätskontrolle mittels Ultraschall und Deep Learning. Darüber hinaus zeigt die verschlechterte Leistung von DeepLabv3+ und SegFormer bei M-Scans, die zur Blechkombination 3 gehören, den signifikanten Einfluss, den wechselnde Materialien tatsächlich auf DL-basierte Qualitätskontrolllösungen haben können.

Active Learning Framework

Im Rahmen der gemeinsamen Forschungsk Kooperation haben das Werkzeugmaschinenlabor WZL, Icon-Pro und Tessonics untersucht, inwiefern Active Learning (AL)-Ansätze die hohen manuellen Aufwände für das Labeling der Bilder reduzieren können. Es wurde ein Human-in-the-Loop Labelingansatz konzipiert und implementiert, der darauf abzielt, den menschlichen Aufwand beim Labeling zu minimieren. Dabei soll der Mensch, wie in Abbildung 8 dargestellt, nur einen Bruchteil der Bilder labeln, indem er diejenigen auswählt, die für das KI-Modell den größten Informationsgewinn liefern.

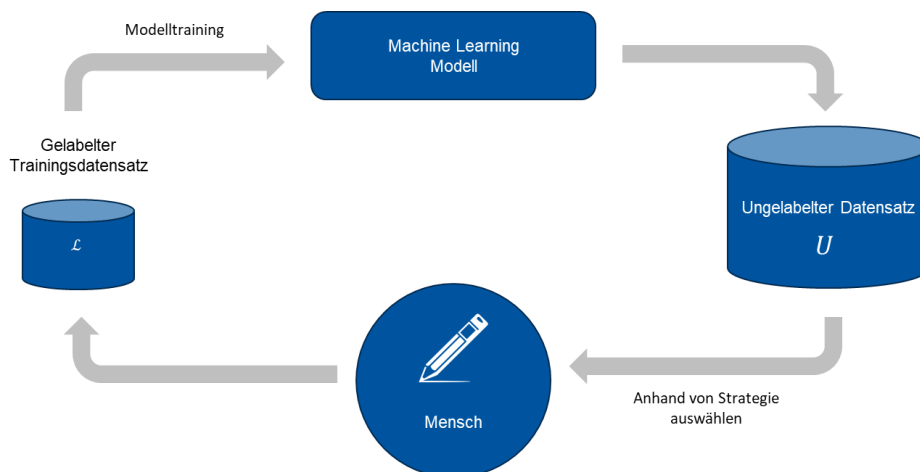


Abbildung 8: Darstellung des Active Learning Prinzips

Das Kernelement des AL-Ansatzes ist die Strategie zur Auswahl der Bilder, die vom Menschen zu labeln sind. Folgende Strategien wurden implementiert und miteinander verglichen:

- **Random Sampling:** Zufällige Auswahl der zu labelnden Bilder.
- **Entropy Sampling:** Auswahl der Bilder anhand der Entropie der Vorhersagen des Modells.
- **Monte Carlo Entropy Sampling:** Auswahl der Bilder basierend auf der mittleren Entropie mehrerer Vorhersagen, indem Dropout während der Inferenz aktiviert wird.
- **Normalized Entropy Sampling:** Normalisierung der Entropie basierend auf der Größe des Bildes.

Im Folgenden wird das Vorgehen zur Beantwortung der Forschungsfragen beschrieben, a) inwiefern sich Active Learning Ansätze zur Reduktion des initialen, manuellen Labeling-Aufwands eignen und b) welche Query-Strategie am wirksamsten ist.

Ausgangspunkt war der oben beschriebene vollständig gelabelte Labordatensatz. Zunächst wurden 10 % der gelabelten Daten für das initiale Modelltraining verwendet, während 30 % als unabhängiger Testdatensatz festgelegt wurden. Die restlichen Daten bildeten einen „unlabeled pool“. Ein Modell wurde auf dem initialen Trainingsdatensatz trainiert und diente als Basis für jede Sampling-Strategie.

Für jede Sampling-Strategie wurden zehn Iterationen durchgeführt. In jeder Iteration wurde das trainierte Modell auf dem unabhängigen Testdatensatz evaluiert und auf die Bilder im „unlabeled pool“ angewendet. Die Metrik zur Auswahl neuer Bilder wurde gemäß der jeweiligen Sampling-Strategie berechnet und entsprechend der Strategie wurden 10% der ungelabelten Daten zum Trainingsdatensatz hinzugefügt. Das Modell wurde dann auf dem erweiterten Trainingsdatensatz erneut trainiert.

Der Vergleich der verschiedenen Sampling-Strategien nach den zehn Iterationen ist in Abbildung 9 dargestellt. Die Visualisierung zeigt die mIoU auf dem unabhängigen Testdatensatz nach jeder Iteration für die verschiedenen Strategien. Man erkennt, dass die Segmentierungsmodelle unabhängig von der gewählten Sampling-Strategie bereits nach wenigen Iterationen sehr hohe Genauigkeiten erzielen. Mit zusätzlichen Daten wird diese Genauigkeit nur noch marginal gesteigert. Zusammenfassend lässt sich festhalten, dass die komplexeren, systematischen Sampling-Strategien keinen signifikanten Mehrwert gegenüber dem Random Sampling liefern.

Die Quantifizierung der Unsicherheit mittels der Entropie-basierten Sampling-Strategien lieferte jedoch wertvolle Erkenntnisse für die Interpretation und Plausibilisierung der Segmentierungsmodelle. So wurde deutlich, dass diejenigen Bilder, bei denen das Modell am unsichersten ist, Punktschweißungen darstellen, die Expulsionen zeigen oder bei denen das Aufschmelzen ungewöhnlich früh einsetzt. Im Gegensatz dazu sind die Bilder, bei denen das Modell hohe Sicherheit aufweist, solche, die zu Gutschweißungen gehören.

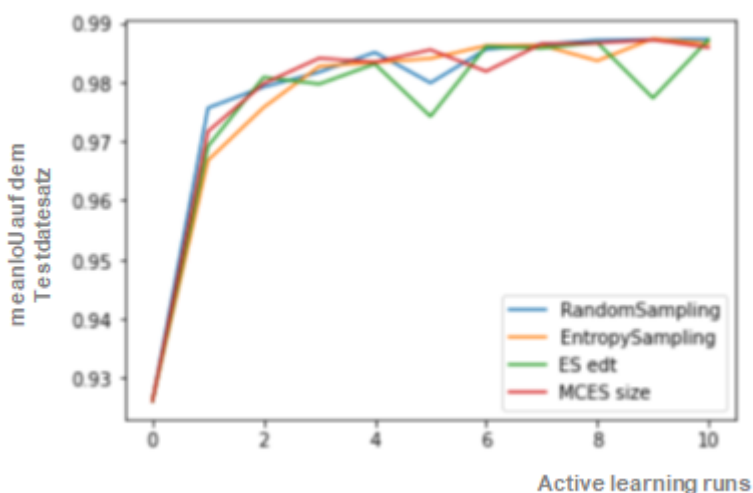


Abbildung 9: Darstellung der mIoU nach jeder AL-Iteration auf dem unabhängigen Testdatensatz für die verschiedenen Sampling-Strategien

Darüber hinaus wurden Heatmaps erstellt, die auf einzelnen Bildern aufzeigen, in welchen Regionen des Bildes das KI-Modell am unsichersten ist. Wie in Abbildung 10 zu sehen ist, sind das die Regionen, die an der Grenze zwischen Nugget und Metallblech liegen.

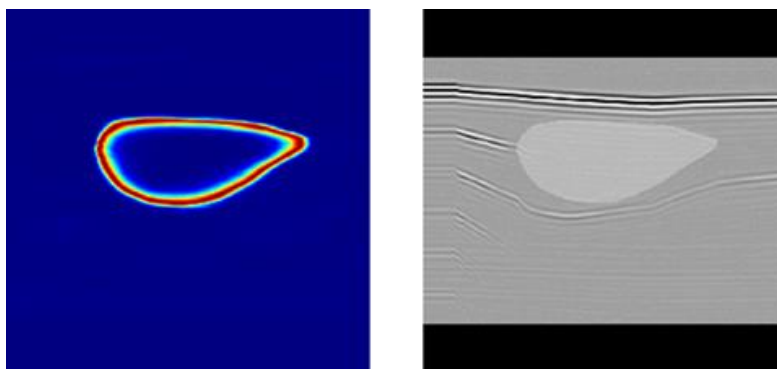


Abbildung 10: Darstellung der Unsichertheitsheatmap (links) für die Vorhersage auf einem M-Scan (rechts)

GANs

Aufbauend auf den Active Learning Erkenntnissen hat das Werkzeugmaschinenlabor WZL untersucht, inwiefern Deep Learning Ansätze zur Erzeugung synthetischer Daten die hohen Aufwände für die Erstellung industrieller Datensätze für die KI-basierte Qualitätskontrolle reduzieren können. Ziel der Untersuchung war es, die Effizienz und Kosteneffektivität bei der Datengenerierung zu verbessern sowie die Genauigkeit und Robustheit der CV-Algorithmen zu steigern. Es wurde ein zweistufiger Forschungsansatz, der in Abbildung 11 dargestellt ist, konzipiert und implementiert, um die Potenziale synthetischer Daten, die mittels Deep Learning erzeugt wurden, für die Verbesserung der semantischen Segmentierung von M-Scans zu evaluieren [38].

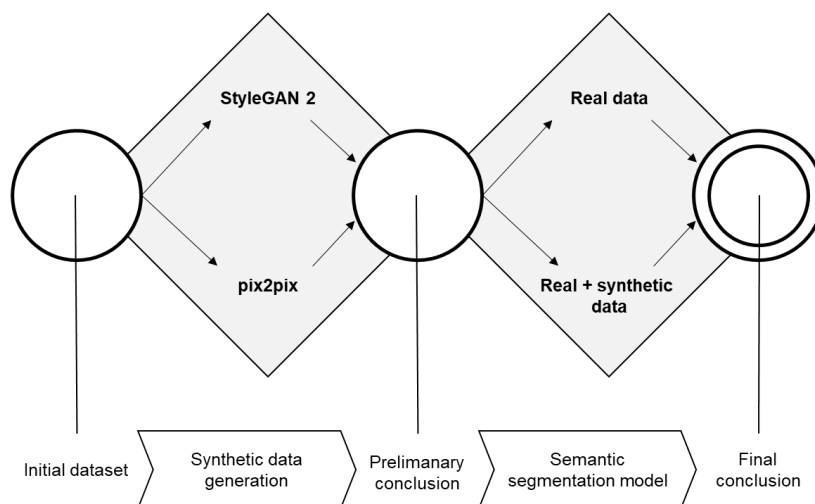


Abbildung 11: Zweistufiger Forschungsansatz zur Untersuchung der GAN-Potenziale

In der ersten Stufe wurden zwei verschiedene Ansätze zur Erzeugung realistischer synthetischer Ultraschallbilder verglichen. Einer dieser Ansätze war die Verwendung eines unconditioned Standard-GANs (StyleGAN2), der simultan Input-Bilder (M-Scans) und die zugehörigen Labels (Nugget-Masken) erzeugt [39]. Der zweite Ansatz war ein conditioned Ansatz namens pix2pix, der basierend auf einer Image-to-Image Translation für ein vorgegebenes Label (Nugget-Maske) das zugehörige Input-Bild (M-Scan) erzeugt [40]. Anhand der Frechet Inception Distance (FID) und einer visuellen Beurteilung wurde der vielversprechendste Ansatz zur Erzeugung synthetischer M-Scans und Nugget-Masken ausgewählt.

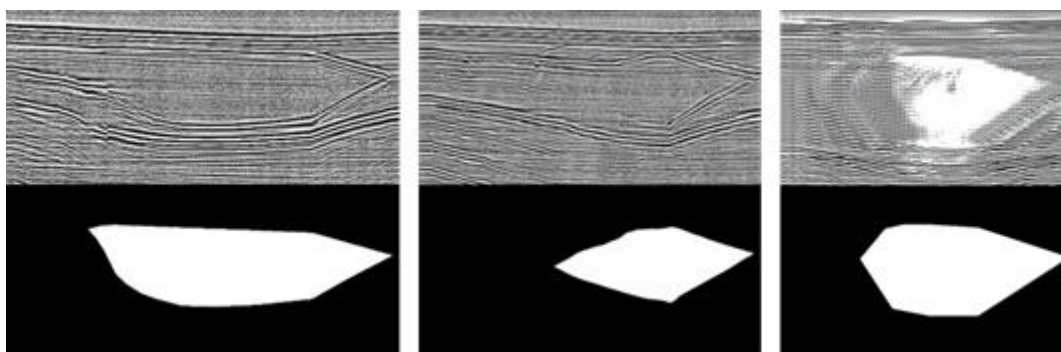


Abbildung 12: M-Scans und zugehörige Nugget-Masken - links: echter M-Scan und manuelle Annotation, Mitte: Mit StyleGAN2 generierter M-Scan und Maske, rechts: von pix2pix generierter M-Scan für eine gegebene Maske

Anhand von Abbildung 12 erkennt man, dass StyleGAN2 bessere synthetische Bilder als pix2pix generiert. Dieser Ansatz wird in der zweiten Stufe zur Evaluation der Mehrwerte der Erzeugung synthetischer Daten für die Genauigkeit der semantischen Segmentierung von M-Scans weiterverfolgt.

Um dies quantifizieren zu können, wurden Modelle auf drei verschiedenen Datensätzen trainiert:

1. einem Trainingsdatensatz ausschließlich mit realen Daten
2. einem Trainingsdatensatz mit realen Daten plus 30% synthetischen Daten
3. einem Trainingsdatensatz mit realen Daten plus 60% synthetischen Daten.

Anschließend wurden die Modelle auf einem Testdatensatz, der ausschließlich aus realen Daten bestand, anhand der mIoU evaluiert. Die Ergebnisse sind in Tabelle 7 dargestellt.

Tabelle 7: mIoU auf dem Testdatensatz von U-Net-Modellen, die mit verschiedenen Datensätzen trainiert wurden

	Reale Trainingsdaten	Reale Trainingsdaten + 30 % synthetische Daten	Reale Trainingsdaten + 60 % synthetische Daten
mIoU	81.34 %	85.17 %	84.20 %

Die höchste Genauigkeit wurde mit dem Modell erreicht, das auf einem Trainingsdatensatz mit 30% zusätzlichen synthetischen Daten trainiert wurde. Die Ergebnisse zeigen, dass a) StyleGAN2 in der Lage ist, realistische, synthetische M-Scans und zugehörige Masken zu generieren und b) die Integration von synthetischen Daten in den Trainingsprozess die Genauigkeit der semantischen Segmentierung von M-Scans beim Widerstandspunktschweißen verbessern kann. Zusammenfassend lässt sich festhalten, dass die Aufwände für die Erzeugung industrieller Datensätze durch den Einsatz von Generativen Adversarial Networks (GANs) signifikant reduziert werden können.

5.1.4 Integration und Datenkommunikation

Im Rahmen des Projekts wurde das Sensor-KI-System für das RSW für einen konkreten Anwendungsfall in der laufenden Produktion beim Forschungspartner Ford getestet. Dafür wurde es in der Tür-Schweiß-Station im Kölner Werk im Karosseriebau eingebaut. Hierbei handelt es sich um fünf zu überwachende Schweißpunkte mit Zweiblechkombinationen an der Fahrzeugseitentür (Innenblech der Tür, Türquerträger, untere und obere Scharnierverstärkung). Zur Auslegung des Systems wurden einige Anforderungen wie Materialien, Materialdicken, Schweißzangen, Elektrotyp etc. näher spezifiziert. Außerdem waren für die Integration des Sensor-KI-Systems einige Umbauten wie Anpassung des Roboterprogramms, Modifikation der Netzwerkkommunikation oder Wechseln der Elektrodenkappen erforderlich. Der kanadische Partner nahm das Deployment des DL-Modells auf einem Edge-Gerät vor. Das mitlaufende System wurde in der laufenden Produktion unter Serienbedingungen validiert, intern mit Schweißfachexperten intensiv diskutiert sowie mit Labormessungen korreliert. Insgesamt waren die Ergebnisse sehr zufriedenstellend und liefern eine weitere Bestätigung für das Potenzial von KI in der Qualitätskontrolle.

Daneben arbeitete IconPro an einem Deployment-Konzept, das die stabile und effiziente Verwendung von Docker-Containern zur Bereitstellung der DL-Modelle vorsieht. Die Implementierung basiert auf Python 3.9 und Flask. Es wurde eine Multi-container-Docker-Anwendung entwickelt, wobei separate Container für das Training des RSW-KI-Modells und des LB-KI-Modells sowie ein Inferenzcontainer für jedes trainierte Modell eingerichtet wurde. Die Konfiguration der Multicontainer-Anwendung wurde

in einer YAML-Datei mit Docker Compose definiert. Ein separater Minio-Server wurde eingerichtet, um S3-Speicher für die On-Premise- und die Cloud-Version zu ermöglichen. Die Zugriffsdaten können nach der Konfiguration des Minio-Servers aufgezeichnet werden. Für den reibungslosen Betrieb der KI-Segmentierung ist eine effiziente Datenkommunikation erforderlich. Dabei sind zwei Aspekte zu berücksichtigen: Der erste betrifft das Training und der zweite die Inferenz mit dem trainierten Modell. Es wurde eine REST-API entwickelt, um Echtzeitdaten in die Datenbank hochzuladen. Die Kommunikation mit den Docker-Volumes für die Trainingsdaten erfolgt über API-Endpunkte, die nach dem REST-API-Konzept erstellt wurden. Zusätzliche Endpunkte wurden für den Abruf von Validierungsmetriken eingerichtet. Die Software-Architektur für die Implementierung ist in Abbildung 13 dargestellt. Metriken und Nachbearbeitungsergebnisse wurden in einer JSON-Datei im S3-Bucket gespeichert, was ihren Abruf vereinfachte. Die Implementierung wurde ausgiebig für alle möglichen REST-Anfragen getestet. Die Integration und Datenkommunikation wurde nur von den deutschen Partnern konzeptioniert.

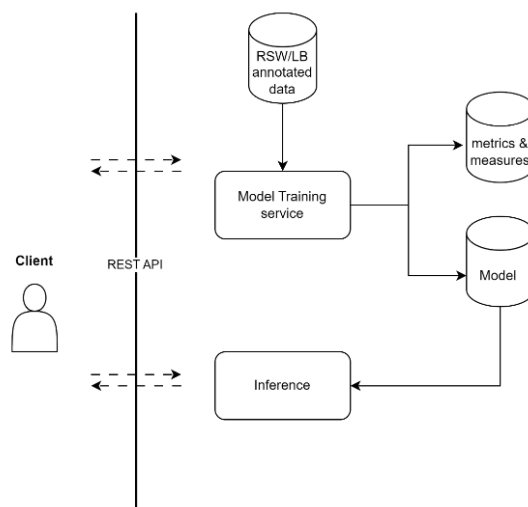


Abbildung 13: Softwarearchitektur mit verschiedenen Docker-Containern und ihren Datenbanken.

5.2 Laserlöten

Nachdem die Ergebnisse des Sensor-KI-Systems für die Ultraschallprüfung beim Widerstandspunktschweißen ausführlich dargestellt wurden, werden nun die Ergebnisse für den zweiten untersuchten Fertigungsprozess Laserlöten vorgestellt. Wie in Kapitel 4 dargestellt, wurde an den beiden Workstreams konsekutiv gearbeitet, sodass dank einer modularen Codestruktur wesentliche Teile vom Widerstandspunktschweißen für das Laserlöten wiederverwendbar waren. In den folgenden Abschnitten wird daher primär auf die Besonderheiten und Unterschiede zwischen den beiden Workstreams eingegangen.

5.2.1 Prozess- und Datenverständnis

Teil des Forschungsprojektes für die kanadischen Partner war die hardwareseitige Weiterentwicklung der Ultraschallsensoren für das Laserlöten, was parallel zum Widerstandspunktschweißen erfolgte. Im Unterschied zum Widerstandspunktschweißen ist eine Ultraschallprüfung in-situ, also während des Fügeprozesses, beim Laserlöten nicht möglich. Diese erfolgt stattdessen nachgelagert, sodass man vom sogenannten B-Scan spricht. Beim Laserlöten liefern CT-Scans qualitativ hochwertige und gut interpretierbare Bilder für die Qualitätskontrolle, da man neben der Löt- und Blechgeometrie außerdem verschiedene Fehler wie Poren, Anbindungsfehler, Einschlüsse und Risse erkennt, sie sind jedoch sehr teuer in der Anfertigung. Daher besteht ein Ziel des Forschungsprojektes darin, ein kosteneffizientes

Ultraschallsensor-KI-System für die automatisierte Qualitätskontrolle beim Laserlöten zu entwickeln. Nach Rücksprache mit den Prozessexperten zur Definition der KI-Aufgabe ist das Ziel weiterhin, eine semantische Segmentierung mit dem primären Fokus auf die Lötgeometrie durchzuführen. In einem ersten Schritt soll also untersucht werden, inwiefern sich Deep Learning Modelle für die semantische Segmentierung der Metallbleche und Lotgeometrie in Ultraschall-B-Scans eignen. Darauf aufbauend soll in einem zweiten Schritt untersucht werden, inwiefern sich zusätzlich zur Lötgeometrie die Defekte durch Deep Learning Modelle segmentieren lassen.

Ein wesentlicher Unterschied für die KI-Aufgabe im Vergleich zum Widerstandspunktschweißen besteht also darin, dass statt einer binären semantischen Segmentierung eine "multi-class" Segmentierung im Anwendungsfall Laserlöten vorgenommen werden muss. Der zweite wesentliche prozessbedingte Unterschied zum Widerstandspunktschweißen, den es für die KI-Softwareentwicklung zu beachten gilt, besteht darin, dass das Laserlöten ein Bahnfügeprozess ist. Das bedeutet, dass statt exakt einem M-Scan je Punktschweißung eine Vielzahl von Ultraschallbildern je Lötnaht erzeugt wird.

5.2.2 Datenaufnahme, -vorverarbeitung und -labeling

Analog zum Widerstandspunktschweißen wurde die Datenaufnahme im Labor der kanadischen Partner gemäß Spezifikationen der Prozessexperten von Ford vorgenommen. In Summe lagen während des Projektes 28 Lötnahte vor, die mit den Ultraschallsensoren nachgelagert geprüft wurden. Gemäß der Auflösung der Sensoren und der Länge der Lötnahte ergaben sich für den Workstream Laserlöten ein Labordatensatz, der je Lötnaht ca. 1.200 Bilder umfasst. Die resultierenden Ultraschallbilder sind schwer manuell zu labeln, wohingegen die verschiedenen zu detektierenden Segmente in den CT-Scans klar erkennbar sind. Daher wurde im Forschungsprojekt ein innovativer, skalierender Ansatz zur Erzeugung gelabelter Ultraschalldaten für lasergelötete Proben entwickelt, der auf einem Domänen-transfer der Label von CT-Scans auf die Ultraschallbilder basiert.

Die Grundidee basiert darauf, für die im Labor erstellten Laserlötproben nicht nur Ultraschall- sondern auch CT-Scans zu erstellen. Mittels unüberwachter Lernverfahren sollen die CT-Scans automatisiert in die Segmente Lot, Metallblech, Pore und Anbindungsfehler geclustered werden. Die so automatisiert gelabelten CT-Scans sollen dann mittels eines regelbasierten Scriptes unter Verwendung von Bildverarbeitungs-funktionen aggregiert und ausgerichtet in die Ultraschalldomäne übertragen werden, so dass vollautomatisiert gelabelte Ultraschallbilder für das Laserlöten entstehen.

Es wurden verschiedene unüberwachte Clustering- und Hyperparametertuningansätze zur automatisierten Segmentierung der CT-Scans implementiert und miteinander verglichen. Am geeignetsten war ein k-means Clustering, wobei k die Anzahl an Segmenten darstellt. Aufgrund der hohen Bildqualität der CT-Scans lassen sich automatisiert sehr akkurate Label erzeugen. Diese Label in der CT-Domäne galt es unter Berücksichtigung der unterschiedlichen Auflösung und Orientierung der CT- und Ultraschallscans in die Ultraschalldomäne zu übertragen. Hierzu wurde auf Basis des Inputs von Prozessexperten ein regelbasiertes Script erstellt. Nach mehreren Entwicklungsiterationen konnte ein vollautomatisiert gelabelter Datensatz von Ultraschall-B-Scans erzeugt werden, der für das Training der Segmentierungsmodelle verwendet werden kann.

5.2.3 Entwicklung der KI-Software

Bei der Entwicklung der KI-Software für das Training und die Evaluation beim Laserlöten wurde ein analoges Vorgehen wie für das Widerstandspunktschweißen gewählt und ein Großteil der für das Widerstandspunktschweißen entwickelten Software konnte aufgrund der Modularität wiederverwendet werden. Prozessbedingt (vgl. 5.2.1) waren folgende zwei Anpassungen an der Software nötig:

1. Anpassung des train-val-test Splits, um Data Leakage zu verhindern, da je Lötnaht mehrere Ultraschallbilder generiert werden.
2. Anpassung von Modelloutput und Loss-Funktion, um die Segmentierung mehrerer Klassen vornehmen zu können.

Im Sinne einer kurzzyklischen, inkrementellen Entwicklung und Evaluation gemäß der mit den Prozessexperten definierten Priorität gliederte sich die KI-Softwareentwicklung beim Laserlöten in zwei Stufen. Zunächst wurde ein KI-Modell für die Segmentierung der Lötgeometrie erstellt. Darauf aufbauend wurden dann die Defekte inkludiert. Im Folgenden werden die Ergebnisse zu den beiden Schritten vorgestellt.

Segmentierung der Geometrie

In einem ersten Schritt wurde eine Segmentierung der Ultraschallbilder von lasergelöteten Proben in Blech und Lötnaht vorgenommen. Die für ein U-Net Modell resultierende Genauigkeit, gemessen anhand der mIoU, ist in Tabelle 8 dargestellt.

Tabelle 8: Übersicht über die Performance-Metriken des U-Net Modells gemessen anhand der mIoU

	U-Net
mIoU	78,32 %

Anhand der mIoU von rund 80 % lässt sich ableiten, dass das KI-Modell grundsätzlich in der Lage ist, die Segmentierung der Lötgeometrie zu erlernen. Allerdings ist ein signifikanter Performanceunterschied im Vergleich zum Widerstandspunktschweißen zu beobachten. Zusammenfassend liefern diese Ergebnisse jedoch einen Proof-of-Concept für die Fähigkeit des Sensor-KI-Systems, die Qualitätsüberwachung der Lötgeometrie vornehmen zu können.

Segmentierung der Geometrie und Defekte

Aufbauend auf dem erstellten Proof-of-Concept für die Segmentierung der Lötgeometrie wurden zusätzlich die Label für die Qualitätsdefekte Poren und Anbindungsfehler integriert. Anhand der mIoU in Tabelle 9 erkennt man, dass dies das Segmentierungsmodell vor Herausforderungen stellt.

Tabelle 9: Übersicht über die Performance-Metriken des U-Net Modells beim Laserlöten inkl. Defekte gemessen anhand der mIoU

	U-Net
mIoU	61,47 %

Die visuelle Betrachtung exemplarischer Bilder bestätigt die Erkenntnis und zeigt zum einen auf, dass die Defekte kaum zu erkennen sind und zum anderen, dass die Genauigkeit der Segmentierung der Lötgeometrie durch die Defektintegration deutlich verschlechtert wird. In Workshops und mittels explorativer Datenanalysen wurden potenzielle Ursachen für den signifikanten Performanceverlust gegenüber dem Widerstandspunktschweißen und besonders durch die Defektintegration ermittelt. Eine mögliche Erklärung kann sein, dass die Defekte in zweierlei Hinsicht im Datensatz unterrepräsentiert sind. Zum einen finden sich im gesamten Datensatz nur sehr wenige Bilder mit Defekten. Zum anderen

sind die Defekte auf den einzelnen Bildern unterrepräsentiert, da es sich um kleine Segmente handelt. Diese Unausgewogenheit in den Daten stellt Deep Learning Ansätze bekanntermaßen vor Herausforderungen. Darüber hinaus besteht Verbesserungspotenzial durch die Verbesserung der Inputdaten- und Labelqualität. Detailanalyse zeigen einige Ausreißer als Ergebnis der vollautomatisierten Labeling-Pipeline. Das Entfernen dieser Ausreißer und die Optimierung der Labeling-Pipeline stellt einen potenziell großen Verbesserungshebel dar. Wie die Ergebnisse zeigen, sind die Deep Learning Modelle in der Lage, Muster in den Daten zu erkennen und eine Segmentierung vorzunehmen. Sollten die einzelnen zu segmentierenden Objekte in den Ultraschallbildern allerdings noch deutlicher sichtbar sein, ist ein sehr großer Anstieg der Modellperformance zu erwarten.

6. Zusammenfassung & Ausblick

Im Projekt wurden durch die Kollaboration der deutschen Partner Ford, IconPro und dem WZL mit den kanadischen Partnern Tessonics und University of Waterloo für die Anwendungsfälle Widerstandspunktschweißen und Laserlöten Ultraschallsensor-KI-Systeme entwickelt. Die Sensorentwicklung und Korrelation der KI-Ergebnisse mit metallurgischen Qualitätsuntersuchungen wurden primär durch die kanadischen Partner erarbeitet, während durch die deutschen Partner das Prozess- und Domänenwissen in die Datenaufnahme und KI-Entwicklung einfluss. Die KI-Entwicklung war gemeinsame Aufgabe der Partner Tessonics, WZL und IconPro.

Die Ergebnisse des Sensor-KI-Systems für das Widerstandspunktschweißen sehen enorm vielversprechend aus und erreichen bei ausreichend vorhandener Datenqualität eine Genauigkeit gemessen in der mIoU von über 95 %. Eine Validierung in laufender Serienproduktion bestätigte die positiven Eindrücke. Weitere Optimierungspotenziale über das Forschungsprojekt hinaus werden zum einen in der Steigerung der Akzeptanz der Shopfloor-Mitarbeitenden als Anwendende des KI-Systems durch Schaffung von Transparenz gesehen. Hierzu können sich Methoden der ExplainableAI und Unsicherheitsmessung sowie der Aufbau einer anwenderzentrierten graphischen Oberfläche eignen. Ein Beispiel für solch eine Unsicherheitsmessung wurde im Zusammenhang mit dem AI-System in Form der Heatmaps umgesetzt. Konkret geeignete Explainable AI Methoden sind sogenannte Feature Activation Maps, die aufzeigen, welche Bildbereiche den KI-Output bestimmen. Eine konkrete Methode zur Berechnung hierfür ist SHAP (SHapley Additive exPlanations) [41].

Da die semantische Segmentierung der Geometrie beim Laserlöten ohne Berücksichtigung der Defekte gute Genauigkeiten erreicht, besteht ein Proof-of-Concept für die Eignung des Sensor-KI-Systems und des innovativen Datenlabelingansatzes. Bei der Integration der Defekte in die KI-Modelle wird jedoch zum einen deutlich, dass die Defekte nicht zuverlässig in hoher Genauigkeit segmentiert werden können. Zum anderen sinkt zusätzlich die Performance der Segmentierung der Lotgeometrie. Die mangelnde Genauigkeit bei der Defektsegmentierung kann in der Datengrundlage begründet sein. Zum einen sind Bilder mit Defekten im Datensatz stark unterrepräsentiert. Zum anderen sind die Defekte auf den einzelnen Bildern stark unterrepräsentiert, da es sich um sehr kleine Objekte handelt. Diese Imbalance im Datensatz kann die Herausforderungen für die KI-Modelle erklären. Weitere Ursachen können einzelne Ausreißer im Script für das Alignment zwischen CT- und Ultraschallbilder sein. Über das Projekt hinaus werden daher weitere Optimierungspotenziale für den Anwendungsfall Laserlöten gesehen. Software-seitige Anpassungen können die systematische Ausreißer-Filterung, die Anwendung von Sampling-Strategien während des Trainings für eine häufigere Auswahl von Bildern mit Defekten oder die Adaption der Loss-Funktion sein, sodass während des Trainings eine stärkere Gewichtung auf kleinen Bildobjekten liegt. Hardware-seitig stellen eine Erweiterung des Datensatzes mit einer

bewussten Erzeugung von Defekten sowie die Weiterentwicklung des Ultraschallsensors Stellhebel dar, um die Genauigkeit des Sensor-KI-Systems beim Laserlöten weiter zu optimieren.

Alles in allem zeigt das Forschungsprojekt jedoch eindrucksvoll die Potenziale von KI für die Qualitätssicherung in Fertigungsprozessen auf. Die deutschen Projektpartner haben wertvolle Erkenntnisse gewonnen, die über das Projekt hinaus in die Lehre sowie in Aktivitäten zur Überführung von KI-Systemen aus einem Forschungsstatus in eine für den Serienbetrieb geeignete Lösung einfließen werden.

Kontakt



Lukas Behnen (*1996, l.behnen@wzl.rwth-aachen.de) ist seit 2021 wissenschaftlicher Mitarbeiter am Werkzeugmaschinenlabor WZL der RWTH Aachen University. Er studierte Wirtschaftsingenieurwesen am Karlsruher Institut für Technologie und schloss dies mit einem Master of Science ab. Er ist nun Doktorand am Lehrstuhl für Produktionssystematik.

Literaturverzeichnis

- [1] Moos, W., Kersting, M. Gesamtwirtschaftliche und sektorale Wertschöpfung aus der Produktion und Anwendung von Fügetechnik in Deutschland, ausgewählten Länder Europas sowie der EU insgesamt, in: , Schweißen und Schneiden, vol. 69, pp. 566–581.
- [2] Dilthey, U., 2006. Schweißtechnische Fertigungsverfahren 1. Springer-Verlag, Berlin/Heidelberg.
- [3] Poprawe, R., 2011. Laser application technology. Springer, Berlin, 605 pp.
- [4] Fahrenwaldt, H.J., Schuler, V., Twrdek, J., 2014. Praxiswissen Schweißtechnik. Springer Fachmedien Wiesbaden, Wiesbaden.
- [5] Schmidt, M., Spieth, H., Bauer, J., Haubach, C. (Eds.), 2017. 100 Betriebe für Ressourceneffizienz – Band 1. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [6] DIN 14610, 2005. Schweißen und verwandte Prozesse: Begriffe für Metallschweißprozesse.
- [7] Ostermann, F., 2014. Anwendungstechnologie Aluminium. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [8] Schiebold, K., 2021. Schweiß- und Lötverbindungen und ihre Prüfung. Springer Fachmedien Wiesbaden, Wiesbaden.
- [9] Advanced Manufacturing Specialist, 2024. Experteninterview.
- [10] Schuler, V., Twrdek, J., 2019. Löten, in: Schuler, V., Twrdek, J. (Eds.), Praxiswissen Schweißtechnik. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 157–177.
- [11] DIN ISO, 2027. Schweißen und verwandte Prozesse - Begriffe - Teil 2: Weichlöten, Hartlöten und verwandte Begriffe.
- [12] Marco Heitmanek, 2014. Reduzierung von Nahtimperfectionen beim Laserstrahlhartlöten.
- [13] DIN ISO. Hartlöten – Unregelmäßigkeiten in hartgelöteten Verbindungen.
- [14] Schuler, V., Twrdek, J., 2019. Qualitätssicherung, in: Schuler, V., Twrdek, J. (Eds.), Praxiswissen Schweißtechnik. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 701–777.
- [15] Schiebold, K., 2015. Zerstörungsfreie Werkstoffprüfung - Ultraschallprüfung. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [16] Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep learning. The MIT Press, Cambridge, Massachusetts, 775 pp.
- [17] Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. <http://arxiv.org/pdf/1505.04597v1>.

- [18] Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.
- [19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need.
- [20] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M., Luo, P., 2021. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers.
- [21] Li, L., Wang, P., Ren, J., Lü, Z., Li, X., Gao, H., Di, R., 2024. Synthetic data augmentation for high-resolution X-ray welding defect detection and classification based on a small number of real samples. *Engineering Applications of Artificial Intelligence* 133, 108379.
- [22] Ajmi, C., Zapata, J., Elferchichi, S., Laabidi, K., 2024. Advanced Faster-RCNN Model for Automated Recognition and Detection of Weld Defects on Limited X-Ray Image Dataset. *J Nondestruct Eval* 43 (1).
- [23] Yang, L., Fan, J., Huo, B., Li, E., Liu, Y., 2022. A nondestructive automatic defect detection method with pixelwise segmentation. *Knowledge-Based Systems* 242, 108338.
- [24] Zhang, B., Wang, X., Cui, J., Yu, X., 2024. Automated Welding Defect Detection using Point-Rend ResUNet. *J Nondestruct Eval* 43 (1).
- [25] Jayasudha, J.C., Lalithakumari, S., 2022. Phased array ultrasonic test signal enhancement and classification using Empirical Wavelet Transform and Deep Convolution Neural Network. *Concurrent Engineering* 30 (3), 229–236.
- [26] Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., Dennison, D., 2015. Hidden Technical Debt in Machine Learning Systems, in: *Advances in Neural Information Processing Systems*. Curran Associates, Inc.
- [27] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2020. Generative adversarial networks. *Commun. ACM* 63 (11), 139–144.
- [28] Cohn, D., Atlas, L., Ladner, R., 1994. Improving generalization with active learning. *Mach Learn* 15 (2), 201–221.
- [29] Feng, Y., Chen, Z., Wang, D., Chen, J., Feng, Z., 2020. DeepWelding: A Deep Learning Enhanced Approach to GTAW Using Multisource Sensing Images. *IEEE Trans. Ind. Inf.* 16 (1), 465–474.
- [30] Zhu, K., Chen, W., Hou, Z., Wang, Q., Chen, H., 2023. Modified Fusing-and-Filling Generative Adversarial Network-based few-shot image generation for GMAW defect detection using multi-sensor monitoring system. *Int J Adv Manuf Technol* 128 (5-6), 2753–2762.
- [31] Yang, L., Liu, Y., Peng, J., 2019. An Automatic Detection and Identification Method of Welded Joints Based on Deep Neural Network. *IEEE Access* 7, 164952–164961.
- [32] Mitri, O. de, Frommknecht, A., Huber, M.F., Müller-Graf, F., Distante, C., 2023 - 2023. Synthetic data generation for improvement of machine learning-based optical quality control: a practical approach in the welding context, in: *Multimodal Sensing and Artificial Intelligence: Technologies and Applications III*. Multimodal Sensing and Artificial Intelligence: Technologies and Applications III, Munich, Germany. 26.06.2023 - 30.06.2023. SPIE, p. 46.
- [33] Shirazi, M., Schmitz, M., Janssen, S., Thies, A., Safronov, G., Rizk, A., Mayr, P., Engelhardt, P., 2021. Verifying the Applicability of Synthetic Image Generation for Object Detection in Industrial Quality Inspection, in: *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), Pasadena, CA, USA. 13.12.2021 - 16.12.2021. IEEE, pp. 1365–1372.
- [34] Fan, K., Peng, P., Zhou, H., Wang, L., Guo, Z., 2021. Real-Time High-Performance Laser Welding Defect Detection by Combining ACGAN-Based Data Enhancement and Multi-Model Fusion. *Sensors (Basel, Switzerland)* 21 (21).

- [35] Jiangsha, A., Tian, L., Bai, L., Zhang, J., 2022. Data augmentation by a CycleGAN-based extra-supervised model for nondestructive testing. *Meas. Sci. Technol.* 33 (4), 45017.
- [36] Ramatlo, D.A., Wilke, D.N., Loveday, P.W., 2023. A data-driven hybrid approach to generate synthetic data for unavailable damage scenarios in welded rails for ultrasonic guided wave monitoring. *Structural Health Monitoring*.
- [37] Mery, D., Riffo, V., Zscherpel, U., Mondragón, G., Lillo, I., Zuccar, I., Lobel, H., Carrasco, M., 2015. GDxray: The Database of X-ray Images for Nondestructive Testing. *J Nondestruct Eval* 34 (4).
- [38] Riesener, M., Keuper, A., Behnen, L., Schuh, G., Scott, R., Chertov, A., Maev, R.G., 2024. Synthetic Data-Enhanced Deep Learning For Quality Control Of Automated Welding Processes. Hannover : publish-Ing.
- [39] Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T., 2019. Analyzing and Improving the Image Quality of StyleGAN. <http://arxiv.org/pdf/1912.04958.pdf>.
- [40] Isola, P., Zhu, J.-Y., Zhou, T., Efros, A.A., 2016. Image-to-Image Translation with Conditional Adversarial Networks. <http://arxiv.org/pdf/1611.07004.pdf>.
- [41] Lundberg, S., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions. <http://arxiv.org/pdf/1705.07874>.