



# EASY Abschlussbericht

## Zuwendungsempfänger

Fraunhofer Institut für Entwurfstechnik Mechatronik IEM

## Titel der Förderung

it's OWL: EEmbedded Artificial Intelligence for Production Systems

## Verantwortliche Autoren

Maximilian Bause, Dr.-Ing. Christian Henke, Dr. Simon Michalke, Nour Younes

## Förderkennzeichen

03INT706BA

**„Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 03INT706BA gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor“**

FRAUNHOFER-INSTITUT FÜR ENTWURFSTECHNIK MECHATRONIK IEM

# **EASY** - Embedded **A**rtificial Intelligence for Production **S**ystems

**Maximilian Bause, Dr.-Ing. Christian Henke**

Fraunhofer-Institut für Entwurfstechnik Mechatronik IEM

**Vorhabenlaufzeit: 01.10.2020 – 31.12.2023**

**Berichtszeitraum: 01.10.2020 – 31.12.2023**

Förderkennzeichen: 03INT706BA

Projektpartner: Fraunhofer-Institut für Entwurfstechnik IEM  
düspohl Maschinenbau GmbH,  
encoway GmbH,  
NTWIST Inc.  
University of Victoria

Paderborn, 01.08.2024

GEFÖRDERT VOM



**Bundesministerium  
für Bildung  
und Forschung**

# Inhalt

<b>1</b>	<b>Kurzbericht .....</b>	<b>4</b>
1.1	Einleitung.....	4
1.2	Ziel und wissenschaftlich-technischer Stand der Technik .....	4
1.3	Ablauf des Vorhabens .....	5
1.4	Zentrale Ergebnisse .....	6
<b>2</b>	<b>Eingehende Darstellung .....</b>	<b>7</b>
2.1	Einleitung.....	7
2.2	Arbeitsplanung und Durchführung der Arbeiten .....	9
2.3	Wissenschaftlich-technische Ergebnisse .....	11
2.3.1	Verallgemeinerte Ergebnisse zu den Arbeitspaketen.....	12
2.3.2	Use Case Predictive Maintenance .....	18
2.3.3	Use Case Predictive Quality .....	37
2.3.4	Ergänzende Inhalte .....	52
2.4	Abgeschlossene Meilensteine .....	53
2.5	Zusammenfassung .....	54
2.5.1	Erreichte Resultate .....	54
2.5.2	Zusammenarbeit mit den kanadischen Partnern .....	55
2.5.3	Verwertung der Ergebnisse und Anschlussfähigkeit.....	56
2.6	Literaturverzeichnis .....	58
2.7	Abkürzungsverzeichnis.....	60

# 1 Kurzbericht

## 1.1 Einleitung

Durch die zunehmende Integration von Informations- und Automatisierungstechnik werden technische Produkte eigenständiger, flexibler und handlungsfähiger. Aufgrund erhöhter Funktionsintegration werden die Systeme jedoch immer komplexer und dem Personal fällt es stetig schwerer, die Ausmaße ihres Handelns abzuschätzen oder notwendige Maßnahmen zu identifizieren. Beide Problemfelder haben gemein, dass eine Prädiktion des Maschinenverhaltens notwendig ist. In vielen Fällen wird versucht, dies mithilfe massiver Rechenleistung in der Cloud zu umgehen, jedoch ist für viele Daten eine Übertragung in die Cloud aufgrund einer zu hohen Auflösung technisch und wirtschaftlich nicht sinnvoll sowie aufgrund des Datenschutzes nicht gewünscht. Die Lösung hierfür ist sogenanntes Edge-Computing, welches die Vorteile von Algorithmen wie dem maschinellen Lernen, mit der Maschinennähe einer SPS vereinen soll.

## 1.2 Ziel und wissenschaftlich-technischer Stand der Technik

Im Vorhaben „EASY - Embedded Artificial Intelligence for Production Systems“ wurden dezentral lauffähige Machine Learning (ML) Algorithmen aus dem Bereich der künstlichen Intelligenz (KI) für Embedded und Edge Devices entwickelt. Auswahl und Einsatz von ML-Algorithmen hängen wesentlich vom Ziel, den zur Verfügung stehenden Informationen und damit von der Anwendung ab. Die in diesem Vorhaben betrachteten Use Cases waren die intelligente Qualitätsprädiktion von Profilmantelungsmaschinen und die vorausschauende Wartung von ausfallkritischen Komponenten in den förder-technischen Anwendungen von Hochregallagern. Diese Use Cases bewegen sich thematisch auf Maschinen- bzw. Komponentenebene. Dies erfordert den Einsatz von ML auf begrenzten Ressourcen der Komponente (Embedded Device) bzw. der Maschine (Edge Device), um die ML-Technologien in Produkte und Produktionsmaschinen zu bringen.

Ein wesentliches Ziel dieses Vorhabens bestand in der Entwicklung und Umsetzung von Lernverfahren in verteilten, eingebetteten Systemen unter Berücksichtigung von Ressourcenbeschränkungen. Ausgangspunkt hierfür sind Daten, die aus multimodalen, aktiven Sensorsystemen gewonnen werden und mit eingebetteten, realzeitfähigen Datenanalysemethoden sowie mit ML-Methoden ausgewertet werden. Diese Daten werden lokal in der Maschine oder der Komponente weiterverarbeitet, um damit die Entwicklung einer datenzentrierten Automation des Maschinen- und Anlagenbaus zu ermöglichen. Die entwickelten Technologien wurden anhand von Anwendungen aus der Industrie validiert. Die wesentlichen Ziele des Vorhabens sind nachfolgend zusammengefasst:

- Analyse von ML-Methoden unter Berücksichtigung von Expertenwissen, Interpretierbarkeit von Lernergebnissen, eingebettetes Lernen auf Datenströmen aus aktiven Sensorsystemen in Echtzeit, und deren Übertragbarkeit in praktische Anwendungen
- Entwicklung von laufzeitoptimierten ML-Algorithmen für Embedded- und Edge Devices, die für praxisrelevante Erkenntnisse aufgearbeitet werden
- Transfer der ML-Methoden in zwei konkrete Anwendungsfälle: Predictive Quality in der Ummantelungstechnik, Predictive Maintenance für Antriebs- und Steuerungslösungen in förder-technischen Anwendungen (z.B. Regalbediengeräten)

Zur Adressierung dieser Zielsetzungen wurden moderne Ansätze der Data Science verwendet. Dabei hat sich ML bereits in vielen Anwendungen etabliert, insbesondere in der Vorhersage von mechanischem Verschleiß mittels Vibrationssensorik. Die Herausforderung besteht darin, auf Datenströmen zu lernen, denn Produktionsanlagen weisen fast ausschließlich mechanisch fehlerfreie Daten auf, sodass die Möglichkeit Modelle zu trainieren stark eingeschränkt sind. Da optimale Trainingsdaten initial quasi nie vorhanden sind, bedarf es Strukturen des kontinuierlichen Trainings von ML-Modellen und deren Anwendung auf eine wachsende Datenbasis.

Die Datenverarbeitung und das Training findet oft an einem anderen Ort als der Datenquelle statt, was Risiken in Bezug auf Datensicherheit birgt. Die Integration von ML auf Edge-Geräten kann eine unmittelbare Datenverarbeitung gewährleisten und das Speichern großer Datenmengen vermeiden. Es gibt bereits verschiedene Frameworks, die das Lernen auf Datenströmen ermöglichen. Eine Besonderheit ist die Einbettung der Funktionen auf Edge-Geräten, um Verzögerungen und Unsicherheiten beim Datentransport zu minimieren. Im einfachsten Fall werden ML-Modelle direkt aus den verfügbaren Daten generiert, was große Datenmengen und hohen Trainingsaufwand erfordert. Bei vorhandenem Expertenwissen empfiehlt sich die Anwendung hybrider Modellierungsansätze, die die Nachvollziehbarkeit der Modelle fördern.

### 1.3 Ablauf des Vorhabens

Im Rahmen des Projektes EASY stand die Analyse von ML-Algorithmen hinsichtlich ihrer Umsetzbarkeit auf Embedded und Edge-Geräten, die Erarbeitung ressourceneffizienter Soft- und Hardwarearchitekturen und die Laufzeitoptimierung der Algorithmen im Fokus. Das Vorhaben wurde im Zeitraum vom 01.10.2020 bis zum 31.12.2023 durchgeführt. Gefördert wurde das Vorhaben durch das Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Fördermaßnahme „Internationalisierung von Spitzenclustern, Zukunftsprojekten und vergleichbaren Netzwerken“. Der im Antrag zum vorliegenden Vorhaben beschriebene Arbeitsplan sah eine Unterteilung der durchzuführenden Aufgaben in insgesamt acht Arbeitspakete sowie die parallel verlaufende Projektkoordination vor.

In Arbeitspaket 1 wurden zunächst die zentralen Anforderungen an ML-Verfahren, Daten, notwendige Hardware und zu entwickelnde Infrastruktur erhoben und eine Feinplanung des Vorhabens durchgeführt. Ausgehend von einem Konzept für Machine Learning auf Edge und Embedded Devices wurden in Arbeitspaket 2 geeignete ML-Verfahren für die Use Cases unter der Randbedingung der Umsetzbarkeit auf der Zielhardware identifiziert. In im dritten Arbeitspaket wurden Ansätze zur Wissensrepräsentation sowie die Integration von Expertenwissen in ML-Verfahren untersucht und anhand der Use Cases validiert. Darüber hinaus wurden im Rahmen der Arbeitspakete AP 4 und AP 5 Lösungen zum Lernen auf sich fortlaufend entwickelnden Datenströmen erarbeitet und Methoden zur Laufzeitoptimierung der ML-Algorithmen durch Optimierung von Daten und Modellen untersucht. Abschließend folgte eine Validierung im Rahmen der Use Cases (AP 6 und AP 7). In einem kanadischen Use Case ohne deutsche Beteiligung (AP 8) erfolgte die Entwicklung eines Agenten-basierten intelligenten Manufacturing Execution Systems.

## 1.4 Zentrale Ergebnisse

Im Projekt konnten zentrale wissenschaftliche Ergebnisse in der Konzipierung, Implementierung und Validierung laufzeitoptimierter ML-Verfahren für Embedded und Edge Devices sowie bezüglich der Integration von Expertenwissen gewonnen und in praktischen Use Cases validiert werden. Insbesondere die Analysen der Daten aus den Use Cases haben die Bedeutung der Datenbasis für den Erfolg einer Anwendung von ML-Verfahren aufgezeigt. Zudem wurden neuartige Konzepte zur Integration von Expertenwissen in Lernverfahren erarbeitet und in die Prozesse implementiert. Im Rahmen des EASY-Projekts wurden folgende Hauptergebnisse erzielt:

- Auswahl und Anwendung geeigneter Lernverfahren (ML und neuronale Netze) zur Erkennung mechanischer Fehler sowie zur Vorhersage einer schlechten Produktqualität.
- Eine umfangreiche Anforderungserhebung sowie die Integration von Expertenwissen führten zu spezifischen Verfahren zur Ermittlung mechanischer Schadensfrequenzen sowie zu einer hybriden Architektur für die Verschleißprädiktion von Rollen.
- Demonstration einer sicheren, laufzeitoptimierten und skalierbaren IoT-Architektur, die als Grundlage für zukünftige Projekte dient.
- Laufzeitoptimierte ML-Algorithmen und Möglichkeiten für deren effiziente und performante Übertragung auf Embedded- oder Edge-Devices.
- ML-Verfahren wurden hinsichtlich benötigter Laufzeit und Speicherkapazität optimiert und auf Embedded und Edge-Geräten getestet (z. B. auf einem selbst entwickelten Sensorboard mit zusätzlicher Vibrationssensorik).
- Hybride Architekturen wurden konzipiert, bewertet und ausgewählte Architekturen getestet sowie zum fortlaufenden Training auf Datenströmen verwendet.
- Ergebnisvalidierung und Verwertungsplanung in den Use Cases zur prädiktiven Instandhaltung in der Intralogistik und zur intelligenten Qualitätsvorhersage in der Profilmantelung.

Weiterhin wurde ein Netzwerk mit Fachleuten aus Kanada und weiteren it's OWL-Projekten etabliert. Mit den kanadischen Partnern (University of Victoria, NTWIST) wurden monatliche Webmeetings abgehalten, um technische und wissenschaftliche Erkenntnisse auszutauschen und zu diskutieren. Der wissenschaftliche Diskurs ermöglichte die Entwicklung und Optimierung der Experten- und ML-Verfahren aus einem anderen Blickwinkel zu betrachten und bessere Lösungskonzepte zu erarbeiten. Durch den engen Austausch konnten beide Seiten neue Sichtweisen auf technische Fragestellungen gewinnen und Impulse in das eigene Teilprojekt einbinden. Zudem wurden gemeinsame Workshops durchgeführt und die internationale Partnerschaft im Rahmen eines Symposiums im Mai 2023 in Kanada vertieft.

## 2 Eingehende Darstellung

### 2.1 Einleitung

Technische Produkte werden durch die zunehmende Integration von Informations- und Automatisierungstechnik eigenständiger, flexibler und handlungsfähiger. Durch die erhöhte Funktionsintegration werden die Systeme jedoch immer komplexer und dem Personal fällt es stetig schwerer, die Ausmaße ihrer Handlungen abzuschätzen oder notwendige Handlungen zu identifizieren. Beide Problemfelder haben gemein, dass eine Prädiktion des Maschinenverhaltens notwendig ist. In vielen Fällen wird versucht dies mithilfe massiver Rechenleistung in der Cloud zu umgehen, jedoch ist für viele Daten eine Übertragung in die Cloud aufgrund einer zu hohen Auflösung technisch und wirtschaftlich nicht sinnvoll sowie aufgrund des Datenschutzes nicht gewünscht. Die Lösung hierfür ist sogenanntes Edge-Computing, welches die Vorteile der modernen Algorithmen, wie das maschinelle Lernen, mit der Maschinennähe einer SPS vereinen soll.

Im Vorhaben „EASY - Embedded Artificial Intelligence for Production Systems“ wurden dezentral lauffähige Machine Learning (ML) Algorithmen aus dem Bereich der künstlichen Intelligenz (KI) für Embedded und Edge Devices entwickelt. Die Auswahl und der Einsatz von ML-Algorithmen hängen wesentlich vom Ziel, den zur Verfügung stehenden Informationen und damit von der Anwendung ab. Die in diesem Vorhaben betrachteten Use Cases waren die ganzheitliche Fertigungsplanung im Rahmen eines intelligenten Manufacturing Execution Systems (IMES), die intelligente Qualitätsprädiktion von Profilmantelungsmaschinen und die vorausschauende Wartung von ausfallkritischen Komponenten in den fördertechischen Anwendungen von Hochregallagern. Diese Use Cases bewegen sich thematisch auf Maschinen- bzw. Komponentenebene. Dies erfordert einen Einsatz von ML auf den begrenzten Ressourcen der Komponente (Embedded Device) bzw. der Maschine (Edge Device), um die ML-Technologien in die Produkte und die Produktionsmaschinen zu bringen.

Ein wesentliches Ziel dieses Vorhabens bestand in der Entwicklung und Umsetzung von Lernverfahren in verteilten, eingebetteten Systemen unter Berücksichtigung von Ressourcenbeschränkungen. Ausgangspunkt hierfür sind Daten, die aus multimodalen, aktiven Sensorsystemen gewonnen werden und mit eingebetteten, realzeitfähigen Datenanalysemethoden sowie mit ML-Methoden ausgewertet werden. Diese Daten werden lokal in der Maschine oder der Komponente weiterverarbeitet, um damit die Entwicklung einer datenzentrierten Automation des Maschinen- und Anlagenbaus zu ermöglichen. Die entwickelten Technologien wurden anhand von Anwendungen aus der Industrie validiert. Die wesentlichen Ziele des Vorhabens sind nachfolgend zusammengefasst:

- Analyse von ML-Methoden unter Berücksichtigung von Expertenwissen, Interpretierbarkeit von Lernergebnissen, eingebettetes Lernen auf Datenströmen aus aktiven Sensorsystemen in Echtzeit, und deren Übertragbarkeit in praktische Anwendungen
- Entwicklung von laufzeitoptimierten ML-Algorithmen für Embedded- und Edge Devices, die für praxisrelevante Erkenntnisse aufgearbeitet werden
- Transfer der ML-Methoden in drei konkrete Anwendungsfälle: Predictive Quality in der Ummantelungstechnik, Predictive Maintenance für Antriebs- und Steuerungslösungen in fördertechischen Anwendungen (z.B. Regalbediengeräten) und 3D-Vision im Bereich Robotics

Die in diesem Vorhaben betrachteten Use Cases der deutschen Partner sind die intelligente Qualitätsprädiktion von Profilmantelungsmaschinen und die vorausschauende Wartung von ausfallkritischen Komponenten in fördertechischen Anwendungen, wie z.B. Schlupf von Reibrädern und Antriebsriemen durch Verschleiß oder Schäden an Lagern und Getrieben. Für die Erreichung der Ziele wurden in den Use Cases definierte

Arbeitspakete von der Anforderungsanalyse über die Auswahl von Lernverfahren mit Integration von Expertenwissen über das Lernen auf Datenströmen bis hin zur Laufzeitoptimierung der Algorithmen für Edge- und Embedded Devices bearbeitet.

Die Zielerreichung wurde zudem durch einen regelmäßigen Transfer von Expertise und Erfahrungen zwischen den Projektpartnern während der Projektlaufzeit gestützt. Das Vorhaben wird als Verbundvorhaben zwischen der kanadischen Firma NTWIST (Industrial AI), dem kanadischen Forschungspartner ACIS (Advanced Control & Intelligent Systems Lab der University of Victoria), den deutschen Firmen düspohl und encoway sowie dem Forschungspartner Fraunhofer IEM durchgeführt. Eine Übersicht über Projektpartner, Projektträger und Fördermittelgeber ist in Abb. 1 dargestellt. Dabei zielte insbesondere der internationale Austausch zwischen deutschen und kanadischen Partnern darauf ab, die Kompetenzen der Projektpartner zusammenzuführen. Die einzelnen Partner brachten ihr jeweiliges Wissen in das Gesamtvorhaben ein und bearbeiteten jeweils die einzelnen, in den insgesamt sieben Arbeitspaketen definierten Aufgaben bis hin zu den resultierenden Gesamtlösungen. Das spezifische Knowhow der Firmen düspohl und encoway wird insbesondere um die Methoden und die Erfahrung mit selbstoptimierenden Systemen und Machine Learning des IEM ergänzt. Im Anschluss an die 3-jährige Förderperiode wird eine die Weiterentwicklung der Lösungen zur wirtschaftlichen Verwertung der Qualitätstechnik angestrebt, um die Hochlohnproduktionsstandorte Deutschland und Kanada für die Zukunft zu rüsten.

Die Zielsetzung von EASY weist hohe Praxisrelevanz und ein deutliches Innovationspotential insbesondere für KMU auf. Gleichzeitig wird jedoch eine Skalierbarkeit der Ergebnisse angestrebt, sodass auch große Unternehmen diese nutzen können. Somit soll ein nachhaltiger positiver Einfluss auf die deutsche und kanadische Wirtschaft, vor allem in Bezug auf die Wahrung der eigenen Position im internationalen Markt, erreicht werden. Im folgenden Kapitel erfolgt eine Darstellung der wesentlichen Arbeiten anhand des tatsächlichen Zeitplans. Im anschließenden Kapitel 2.3 werden die aus der Förderung durch das BMBF resultierenden wissenschaftlich-technischen Projektergebnisse aus der Zusammenarbeit des Fraunhofer IEM mit den Projektpartnern düspohl Maschinenbau GmbH und der encoway GmbH anhand der Struktur der Arbeitspakete beschrieben.



Abb. 1: Beteiligte Institutionen (Wissenschaft, Wirtschaft, Projektträger und Fördermittelgeber)

## 2.2 Arbeitsplanung und Durchführung der Arbeiten

Eingehende Darstellung

Das Vorhaben „EASY – Embedded Artificial Intelligence for Production Systems“ wurde im Zeitraum vom 01.10.2020 bis zum 31.12.2023 durchgeführt. Gefördert wurde das Vorhaben durch das Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Fördermaßnahme „Internationalisierung von Spitzenclustern, Zukunftsprojekten und vergleichbaren Netzwerken“. Der im Antrag zum vorliegenden Vorhaben beschriebene Arbeitsplan sah eine Unterteilung der durchzuführenden Aufgaben in insgesamt acht Arbeitspakete sowie die parallel verlaufende Projektkoordination vor. Der Arbeits- und Zeitplan ist in Abb. 2 dargestellt. Die einzelnen Arbeitspakete wurden durch die jeweils beteiligten Partner bearbeitet.

Arbeits- und Zeitplan	1. Projektjahr				2. Projektjahr				3. Projektjahr				4. PJ				Personalaufwand in PM				
	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	dispohl	encoway	IEM	UBC	NTWist
AP 1: Anforderungsanalyse				1													6	6	4	16	10
AP 2: Analyse und Auswahl von Lernverfahren				2													3	4	8	22	2
AP 3: Integration von Expertenwissen in Lernverfahren																	8	4	7	22	10
AP 4: Lernen auf Datenströmen																	-	8	3	24	2
AP 5: Laufzeitoptimierung von ML-Alg. für Edge- und Embedded Devices																	8	9	5	4	-
AP 6: Use Case Predictive Quality																	20	-	12	5	2
AP 7: Use Case Predictive Maintenance																	4	16	6	5	2
AP 8: Intelligentes, multiagentenbasiertes MES (nicht in die kostenneutrale Verlängerung eingeschlossen)																	-	-	-	20	20
Projektkoordination																	1	1	3	5	2
<b>Summe Personenmonate je Projektteilnehmer</b>																	<b>50</b>	<b>48</b>	<b>48</b>	<b>103</b>	<b>50</b>
Meilenstein 1: Anforderungsliste, Feinplanung des Vorhabens und ML-Konzept																					
Meilenstein 2: Auswahl an ML-Verfahren wurde festgelegt, die für beide Anwendungsfälle relevant und auf Embedded und Edge Devices umsetzbar sind																					
Meilenstein 3: Validierte Ansätze zur Wissensrepräsentation																					
Meilenstein 4: Maschinelles Lernen auf fortlaufenden, sich dynamisch entwickelnden Datenströmen																					
Meilenstein 5: Laufzeitoptimierte ML-Algorithmen																					
Meilenstein 6: Validierung in Use-Cases (Erreichen des Meilenstein am am 31.12.2023 statt 30.09.2023)																					

Abb. 2: Zeitliche Einordnung der Arbeitspakete sowie der zugehörigen Meilensteine

Durch den Kontaktbeschränkungen aufgrund der pandemischen Lage im Rahmen der COVID-19 Pandemie wurde Meetings, Austauschrunden und Workshops sowohl auf nationaler als auch auf internationaler Ebene virtuell in Form von Videokonferenzen durchgeführt. Mit Entspannung der pandemischen Lage konnten Arbeitstreffen zunehmend wieder vor Ort und in unmittelbarer Nähe zu den betrachteten technischen Systemen und Anlagen durchgeführt werden. Auch der persönliche Austausch der Partner auf internationaler Ebene konnte in Form eines Symposiums im Mai 2023 an der University of Victoria realisiert werden.

Für die erfolgreiche Durchführung des Vorhabens sowie die Validierung der Ergebnisse stellte die Verfügbarkeit der richtigen Daten in der benötigten Qualität und Quantität einen kritischen Faktor dar. Lieferengpässe insbesondere für mikroelektronische Bauteile im Nachgang der COVID-19 Pandemie haben zu Verzögerungen bei der Aufzeichnung von realen Prozessdaten in den Use Cases geführt. Für die Optimierung und Bewertung der entwickelten Lösungen am realen Prozess ist zudem die Erhebung von Daten im produktiven Umfeld unter rauen Bedingungen und harten Anforderungen erforderlich. Die Behebung von Problemen bei der Datenerfassung war dabei mit einem hohen Aufwand verbunden. Aus diesem Grund war die Datenlage für die Validierung der Use Cases im Rahmen der Arbeitspakete AP6 und AP7 bei der Bearbeitung kleiner als erwartet. Daher wurde zur weiteren Aufzeichnung benötigter Daten sowie deren Auswertung eine kostenneutrale Verlängerung der Laufzeit um 3 Monate beantragt. Die vorgelagerten Arbeitspakete blieben davon unberührt. Die entsprechende Anpassung des Arbeits- und Zeitplans ist ebenfalls in Abb. 2 ersichtlich.

Im Hochregallager der Firma Lenze, dass im September 2022 in Betrieb genommen wurde, wurden ab dem 1. März 2023 bedingt durch produktionskritische

Untersuchungen der Regalbediengeräte Motorstromdaten nicht mehr in einer Menge und Qualität übermittelt, welche ein Training von Algorithmen zur prädiktiven Wartung von Regalbediengeräten bzw. der Validierung der Ergebnisse des Forschungsprojekts ermöglichen. Während Lenze als Projektpartner weiterhin die Bereitschaft und Unterstützung des Projekts gewährleistete, war es jedoch technisch nicht möglich die für das Projekt EASY notwendig Daten parallel zu den Untersuchungen der Werksinstandhaltung aufzuzeichnen.

Zunächst sollte sich die Untersuchung nur auf wenige Tage bzw. Wochen beschränken. Die Einschränkungen dauerten aber final bis Anfang September 2023 an, bis die IT-Infrastruktur wieder für die Aufnahme bereitgestellt werden konnte. Danach konnte die Datenaufnahme fortgesetzt werden. Jedoch genügte die verbleibende Projektzeit nicht, um Daten für mehrere Wochen aufzunehmen, ein adäquat großes Sample zu generieren und dieses dann bis Ende September auszuwerten. Die kostenneutrale Verlängerung trug hier maßgeblich zur Erreichung der Projektziele bei, wie in Abschnitt 2.3.2 ersichtlich.

Für den zweiten Use Case wurden im Projektzeitraum laufend Daten aus der Produktion erhoben und durch die Firma düspohl bereitgestellt. Allerdings waren mehrmalige Änderungen in der Datengrundlage erforderlich. Insbesondere die Erhebung von bestimmten Metadaten wie den Geschwindigkeiten und ummantelten Metern für die Ummantelungsanlage RoboWrap sind hier hervorzuheben. Diese Adaptionen in der Datenerhebung erforderten große Aufwände und da die Umsetzung an einer produzierenden Anlage beim Kunden erfolgte, ist die finale Version der Datenerhebung erst später als erwartet im Projekt online gegangen. Herausfordernd war hierbei vor allem, dass historische Daten aus den vorherigen Versionen aufgrund fehlender Informationen nur bedingt nutzbar und für die finalen ML-Modelle aufgrund der fehlenden Metainformationen nicht dienlich nutzbar waren. Zwar stand mit der neuesten Datenversion eine grundlegende Datenmenge zur Verfügung, doch ein zusätzlicher Zeitraum zur Datenerhebung ermöglichte Untersuchungen und Validierungen auf einer repräsentativen Datenmenge. Aufgrund der mangelnden Datenlage vor geplanten Projektabschluss am 30-September 2023 und der Bedeutung der Datengrundlage für die Validierung der Ergebnisse in den Use Cases wurde die Projektlaufzeit kostenneutral bis zum 31.12.2023 verlängert. Im Use Case Predictive Maintenance konnten so weitere wertvolle Daten über Verschleiß und Wartungsvorgänge für die finale Validierung gesammelt werden. Im Use Case Predictive Quality konnte ebenfalls eine verbesserte Datengrundlage aus dem Prozess der Ummantelung aufgezeichnet werden.

## 2.3 Wissenschaftlich-technische Ergebnisse

Das IEM hat im Rahmen des Verbundvorhabens „EASY“ die projektübergreifende Organisation zwischen den beteiligten Partnern durchgeführt. In regelmäßigen Abständen wurden Meetings durchgeführt, in denen nicht nur organisatorische Aspekte geklärt, sondern auch wissenschaftlich technische Ergebnisse und Problemstellungen vorgestellt und anschließend gemeinsam diskutiert wurden. Die dabei entwickelten Anmerkungen und Ideen wurden in die einzelnen Use Cases eingebracht, um bei der Erreichung der Meilensteine und Projektziele aus dem internationalen Austausch der Projektpartner zu profitieren. Aufgrund der sich entspannenden Lage in Bezug auf die Covid-19 Pandemie, wurden Meetings der deutschen Partner zunehmend in Präsenz durchgeführt, der regelmäßige Austausch auf organisatorischer und insbesondere auch fachlicher Ebene mit den kanadischen Partnern wurde weiterhin über regelmäßige Online-Meetings vertieft. Zudem wurden mehrstündige Workshops durchgeführt, in denen die Projektpartner zum Teil auch gemeinsam mit anderen deutsch-kanadischen Forschungsprojekten ihr Projekt mitsamt Ergebnissen und Herausforderungen präsentiert und konstruktiv diskutiert haben. Weiterhin wurde die wissenschaftliche Kooperation auch durch den Austausch von Algorithmen und Daten intensiviert. Der wissenschaftliche Austausch wurde unter anderem auch durch ein Symposium an der University of Victoria gefördert, wo die Projektergebnisse im Mai 2023 mit deutschen und kanadischen Teilnehmern aus Wissenschaft und Industrie sowie mit weiteren Projekten mit ähnlichem thematischem Schwerpunkt geteilt und diskutiert wurden.

Die durch das Fraunhofer IEM im Berichtszeitraum bearbeiteten Arbeitspakete und die dabei erzielten Ergebnisse werden nachfolgend beschrieben und hinsichtlich der Projektziele und Meilensteine bewertet. Die zentralen Ergebnisse der Meilensteine wurden zudem bereits im Rahmen verschiedener Meilensteintreffen gemeinsam durch die Projektpartner präsentiert. Die Berichterstattung des Fraunhofer IEM über die vorliegenden Projektergebnisse wird nachfolgend zunächst auf allgemeiner Ebene und anschließend spezifisch für folgende Use Cases beschrieben:

- **Predictive Maintenance** (encoway GmbH, Fraunhofer IEM)
- **Predictive Quality** (düspohl Maschinenbau GmbH, Fraunhofer IEM)

Ein Überblick zu den Use Cases und deren wesentlichen Forschungszielen ist in Abb. 3 dargestellt. Der Use Case Predictive Maintenance bezieht sich auf die intelligente Instandhaltung in der Intralogistik. Im Fokus stehen dabei im wesentlichen Regalbediengeräte (RBG) als stark verschleißbehaftete Systeme in ausfallkritischen Prozessen der Ein- und Auslagerung von Waren in Hochregallagern. Im Predictive Quality Use Case werden die Auswirkungen von Verschleiß und Prozessführung zur Prädiktion einer kritischen Produktqualität untersucht. Ziel ist die Erkennung von drohenden Ausschussteilen im Vorfeld, um proaktiv die Fertigungsqualität beim Ummanteln sicherzustellen. Gegenstand der Betrachtung ist dabei die RoboWrap, die weltweit erste automatisch rüstende Profilmantelungsanlage. Der Use Case „IMES“ wird durch die kanadischen Partner im Rahmen einer kanadischen Förderung erarbeitet und daher nicht in diesen Bericht inkludiert.

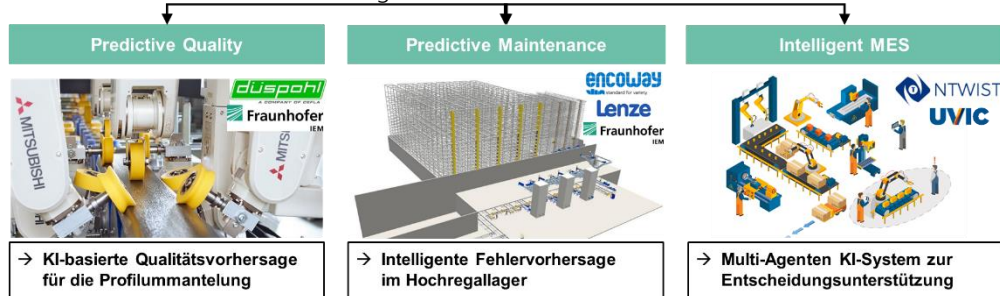


Abb. 3: Überblick über die Use Cases und Projektpartner im Projekt „EASY“

### 2.3.1 Verallgemeinerte Ergebnisse zu den Arbeitspaketen

Nachfolgend werden zunächst die verallgemeinerten wissenschaftlich-technischen Ergebnisse zu den einzelnen Arbeitspaketen dargestellt. Diese sind losgelöst von den Use Cases und für eine Übertragung auf weitere Anwendungen und Use Cases abstrahiert. Der dazu gewählte Lösungsansatz umfasst die Analyse der betrachteten Anwendungen zur Ableitung von Anforderungen und unter Berücksichtigung der Übertragbarkeit der Lösungen. Zudem erfolgt die Spezifikation von Edge-Devices als Zielhardware sowie die Entwicklung von KI-Bausteinen. Weiterhin unterliegen technische Anwendungen und Prozesse oftmals komplexen Wechselwirkungen aus Maschineneinstellungen, Betriebsparametern und Verschleißzuständen. Daher wurde die Integration von Experten untersucht, um die Robustheit der KI-Bausteine auf verfügbaren Daten zu erhöhen. Auf diesem Weg soll die Dateneffizienz deutlich verbessert werden und gezielter auf veränderliches Verhalten (Concept Drift) reagiert werden. Abschließend erfolgt eine Applikation in den Anwendungsfällen in Form von Demonstratoren, die im Nachgang an das Projekt in Serienprodukte überführt werden sollen.

#### **AP 1: Anforderungsanalyse**

Das Ziel der Anforderungsanalyse bestand darin, verbindliche Anforderungen an die Systeme abzuleiten und zu dokumentieren, um anhand dieser die Konzipierung von KI-Bausteinen zu erarbeiten. Dazu wurden zunächst die Use Cases näher beschrieben und darauf aufbauend die Anforderungen definiert. Dazu haben die Projektpartner von Beginn an sehr eng zusammengearbeitet, um die Pain Points und Herausforderungen im realen Betrieb mit den wissenschaftlichen Fragestellungen und Herausforderungen beim Einsatz von ML-Verfahren zusammen zu bringen. Ein wesentlicher Schritt war dabei die Definition der Ziele, aber auch die Identifikation notwendiger Daten aus den Prozessen sowie die Anforderungen an diese Daten. Dabei wurden entscheidende Erkenntnisse für die Anforderungserhebung an die Datenbasis erarbeitet.

So wurden in den ersten Analysen und Workshops zur Anforderungserhebung wesentliche Herausforderungen bei der Entwicklung datenbasierter Verfahren identifiziert, die eine breite, anwendungsunabhängige Relevanz aufweisen. Dazu gehören unter anderem der Umgang mit Datenraumlücken und verrauschten, bzw. fehlerhaften Daten, insbesondere aber auch die Kosten und Quantität der notwendigen Daten in Abhängigkeit der Komplexität der Prädiktionsaufgabe und der zugrundeliegenden Systemzusammenhänge, sowie die Robustheit der Prädiktion. Diese Herausforderungen werden auch in der Wissenschaft regelmäßig benannt [1], [2].

Zentrale Anforderungen an KI-Systeme und deren Datenbasis sind hieraus entstanden, darunter die Anforderung der Beobachtbarkeit des zu präzisierenden Systemverhaltens in den Eingangsdaten der ML-Verfahren. Eine weitere Anforderung ist die Erhöhung der funktionalen Sicherheit und der Dateneffizienz durch die Nutzung von Expertenwissen und alternativen Datenquellen wie Demonstratoren und Simulationen. Diese Anforderungen haben auch allgemein für technische Use Cases im Bereich der Produktion eine hohe Relevanz, da sie zum einen die Machbarkeit der Lösung sicherstellen und zum anderen die Wirtschaftlichkeit durch eine effiziente Einführung in den realen Prozess begünstigen.

Zusätzlich ist in Form einer Label-Tabelle für Prozesse der Zustandsüberwachung ein Tool entstanden, das sich auch auf Prozesse außerhalb der Zustandsüberwachung abstrahieren lässt und Unternehmen und Entwickler bei der Anforderungserhebung im Einklang mit der Zielsetzung bei der Entwicklung maschineller Lernverfahren unterstützt. Diese Tabelle ermöglicht die Reflexion der Zielsetzung bei der Nutzung überwachter Lernverfahren gegen die Anforderungen an das dazu erforderliche Labeling der Daten. Eine wesentliche Fragestellung, die dabei in der Anforderungserhebung geklärt werden kann, adressiert, welche Informationen in Form von Labels den Messdaten zugeordnet werden müssen. Auch die Granularität dieses Labels sowie die Fragestellung der Abbildung des Labels als qualitative oder quantitative Information kann anhand des Tools erarbeitet werden. Dies ermöglicht eine frühzeitige Schärfung der Zielsetzung bei der Entwicklung

von KI-Systemen hinsichtlich der (meist wirtschaftlichen) Realisierbarkeit. Zudem werden vorhandene Schwächen und Bedarfe in der bisherigen Datenerfassung offengelegt. Als Ergebnis der Anwendung des Tools resultieren klare Anforderungen an KI-Systeme und die Datenbasis.

Weiterhin wurden grundlegende Anforderungen an die Robustheit und Recheneffizienz der ML-Verfahren gestellt. Für eine effiziente Ausführbarkeit auf Embedded bzw. Edge Devices wurden daher grundlegende Anforderungen an die Laufzeitoptimierung der Algorithmen beispielsweise durch Quantisierung oder Pruning gestellt.

## AP 2: Analyse und Auswahl von Lernverfahren

Das AP 2 diente der Identifikation und Festlegung von ML-Verfahren für die Use Cases unter Berücksichtigung der Umsetzbarkeit auf Embedded bzw. Edge Devices. Für die Analyse und Auswahl von Lernverfahren hat das IEM eine Taxonomie für die Einordnung verschiedener ML-Verfahren erstellt. Diese ist in Abb. 4 dargestellt und dient als Überblick bei der Analyse und Auswahl von Lernverfahren. Dabei lassen sich die Verfahren auf der unteren Ebene noch weiter aufgliedern in individuelle Architekturen aus der jeweiligen Kategorie von Lernverfahren [3]. Die grundsätzlich relevanten Verfahren, die aufgrund ihrer Performance und Anwendbarkeit im Kontext der Use Cases in Frage kamen, wurden in dieser Darstellung bereits hervorgehoben.

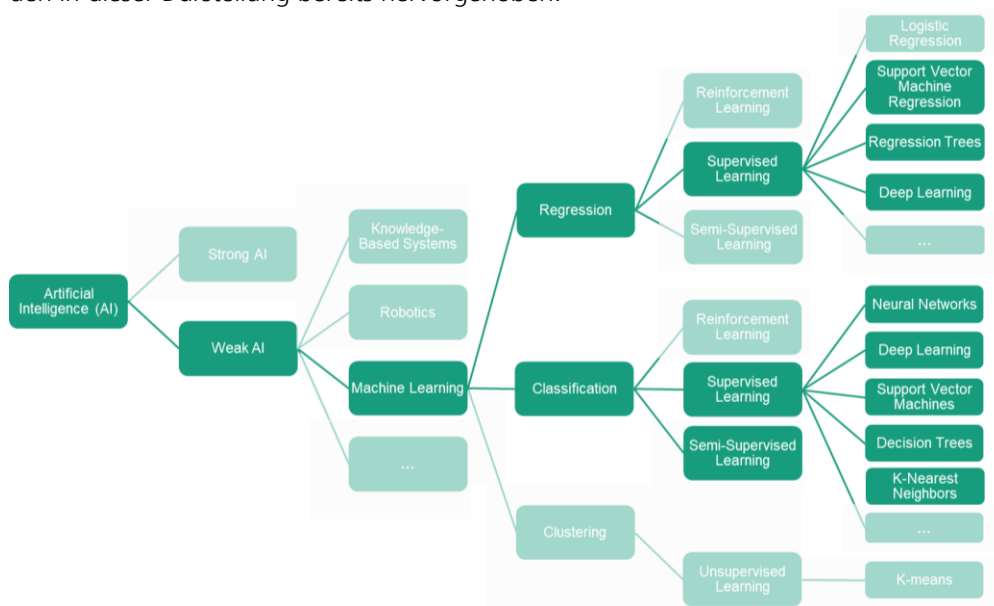


Abb. 4: Taxonomische Einordnung von Verfahren des Machine Learning unter dem Begriff KI

Weiterhin wurden verschiedene ML-Verfahren näher untersucht und hinsichtlich ihrer Vor- und Nachteile bewertet. Die Betrachtung umfasste dabei unter anderem die in Abb. 5 anhand ihrer Funktionsweise dargestellten Lernverfahren: Entscheidungsbäume, k-Nearest Neighbors, Support Vector Machines (SVM) und künstliche neuronale Netze (KNN) [4], [5] [6]. Für die Verarbeitung von Merkmalsdaten wurden insbesondere die KNNs aufgrund ihrer Genauigkeit, Flexibilität und der Möglichkeit zur Parallelisierung von Berechnungen mit Blick auf die Anforderungen an Recheneffizienz ausgewählt. Auch ihre besondere Eignung zur Approximation beliebig komplexer, nichtlinearer Zusammenhänge ist für komplexe technische Prozesse von Vorteil. Für die Verarbeitung von Messreihen wurden zudem Convolutional Neural Networks (CNNs) ausgewählt. Dabei handelt es sich um spezielle KNNs aus dem Bereich des Deep Learnings, die mittels lokal gruppierter Neuronen – sogenannter Kernel oder Filter – eine Extraktion und automatische Fusion von lokalen und komplexen Merkmalen aus zusammenhängenden Datenreihen, wie Zeitserien und geometrischen Informationen, extrahieren können [7], [8].

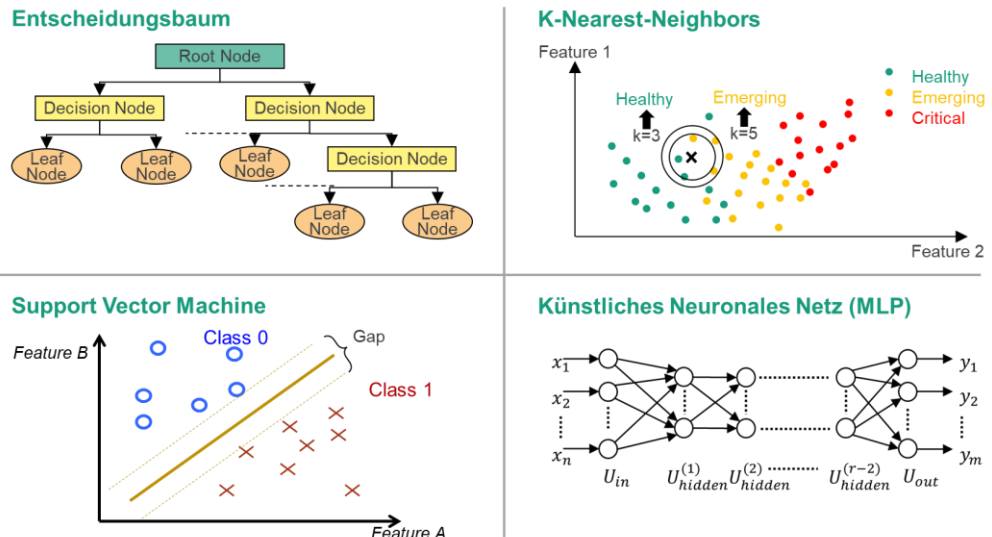


Abb. 5: Darstellung der Funktionsweise ausgewählter maschineller Lernverfahren

### AP 3: Integration von Expertenwissen

Im Rahmen von AP 3 galt es, geeignete formale Modelle zur externalisierten Wissensrepräsentation zu identifizieren und zu validieren und Möglichkeiten zur Kombination mit den datengetriebenen Lernverfahren auszuarbeiten. Für die Integration von Expertenwissen beim ML stellt sich zunächst die Frage, was Expertenwissen bedeutet. In Anlehnung an die aktuelle Literatur zum Thema ist dabei zunächst zu unterscheiden, um welche Form von Expertenwissen es sich handelt. Neben der Quelle des Wissens, die von menschlicher Natur sein kann, aber auch aus wissenschaftlich verfügbarem Wissen bestehen kann, ist insbesondere die Repräsentation des Wissens von zentraler Bedeutung. Expertenwissen kann beispielsweise in Form von algebraischen Gesetzen oder Differentialgleichungen vorliegen. Auch die Ergebnisse von Simulationen, beispielsweise synthetische Daten, stellen eine Repräsentationsform von Expertenwissen dar. Mit abnehmendem Grad der Formalisierung des Wissens erstreckt sich die Gesamtheit der Repräsentationsmöglichkeiten unter anderem über logische Regeln und probabilistische Beziehungen bis hin zum menschlichen Wissen und Feedback, das in einer informellen Form vorliegt [9].

Die Repräsentation von Expertenwissen ist entscheidend dafür, wie dieses Wissen in den ML-Prozess integriert werden kann. Neben mathematische formalisierte Gesetzmäßigkeiten, die im Modellaufbau sowie der Fehlerfunktion des ML-Algorithmus berücksichtigt werden können, besteht auch die Möglichkeit, Simulationsergebnisse als Trainingsdatenbasis zu nutzen. Menschliches Feedback hingegen kann beim überwachten Lernen die bekannte Ausgabe („Ground Truth“) für Lernbeispiele zum ML-Training definieren und findet auch stark Berücksichtigung bei der Entwicklung und Auslegung der Merkmalsextraktion im Rahmen der Datenvorverarbeitung. Die Bedeutung von Expertenwissen für ML-Verfahren ist abhängig von der zur Verfügung stehenden Datenbasis. Je kleiner die Menge an repräsentativen Daten und damit an Lernbeispielen ist, desto weniger Informationen stehen zur erfolgreichen Nutzung von ML im industriellen Kontext zur Verfügung. Daher ist die Bedeutung von Expertenwissen umso höher, je kleiner die Datenmengen sind. Ziel ist es, über die Einbindung von zusätzlichem Wissen vorhandenen Datenmangel zu kompensieren und die Prädiktionsgüte der ML-Verfahren auf spärlichen Datenmengen zu erhöhen [10].

Für die Einbindung von Expertenwissen im Projekt wurden insbesondere hybride Verschaltungen als vielversprechend identifiziert. Zu diesen zählen neben den physics-guided ML-Verfahren, bei denen physikalische Zusammenhänge mathematisch formalisiert in der Architektur des ML-Algorithmus oder in der für das Training

ausschlaggebenden Fehlerfunktion berücksichtigt werden, auch Parameteranordnungen. Von besonderer Bedeutung sind auch die seriellen und parallelen Verschaltungen. Bei den seriellen Verschaltungen wird dem ML-Algorithmus entweder physikalisches Vorwissen zur Informationsanreicherung vorgeschaltet, oder die ML-Ausgabe über mathematisch-physikalische Gesetzmäßigkeiten weiterverarbeitet. Parallele Verschaltungen hingegen ermöglichen eine getrennte Berechnung von Ausgaben durch ML und analytischen Algorithmus, die es anschließen angemessen zu fusionieren gilt. Parallele und serielle Verschaltungen können zudem auch beliebig komplex kombiniert werden. In Abb. 6 sind die hybriden Verschaltungsmöglichkeiten, die auch für die individuellen Use Cases von Bedeutung sind, schematisch dargestellt.

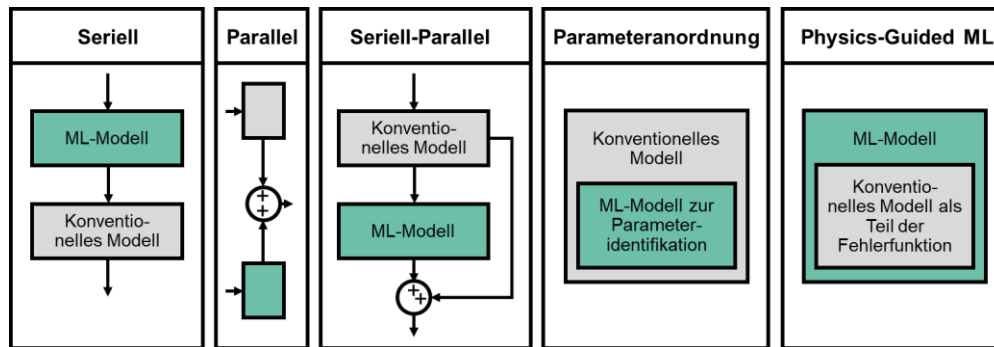


Abb. 6: Verschaltungsszenarien der hybriden Modellbildung

Weiterhin kommen Daten von Demonstratoren oder synthetische Daten auf Basis von mathematisch-physikalischen Simulationsmodellen als Quelle zusätzlichen Wissens für ML in Frage. Diese aus der Simulation entstehenden, anwendungsnahen Daten können genutzt werden, um ML-Modelle im Vorlauf zur Zielanwendung zu trainieren. Dabei kann ein positiver Effekt erzielt werden, indem die Modellparameter bereits beim Training auf Daten angepasst werden, die denen der Zielanwendung ähnlich sind. Anschließend kann das vortrainierte ML-Modell im Zuge eines Transfer Learning (TL) auf den Realdaten angewendet und weiter optimiert werden. Dabei wird erlerntes Wissen über Zusammenhänge in den Daten als eine Form von Vorwissen mit dem ML-Modell auf die reale Anwendung übertragen.

Bei der Adaption des vortrainierten ML-Modells auf die Zielanwendung können sich in Nutzung und fortlaufendem Training auf den Daten verschiedene positive Effekte einstellen, wie in Abb. 7 skizziert. So ist die initiale Modellperformance eines Modells mit TL der eines nicht vorab trainierten Modells in der Regel überlegen. Auch kann durch Training auf Daten der Zielanwendung meist eine höhere maximale Performance erzielt werden, oftmals sogar auf einer kleineren Datenbasis oder unter Nutzung einer geringeren Anzahl von Trainingsepochen [11], [12].

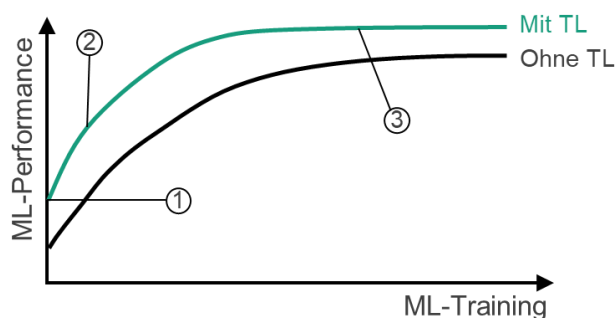


Abb. 7: Performance-Vergleich von ML-Verfahren mit und ohne Transfer Learning, nach [13]

#### AP 4: Lernen auf Datenströmen

Aufbauend auf den vorhergehenden APs zielt AP 4 auf die Konzeptionierung der Sensordatenauswertung und Signalverarbeitung für kontinuierlich anfallende Daten ab. Für sich fortlaufend dynamisch entwickelnde Datenströme resultieren verschiedene Herausforderungen, die im Rahmen dieses Projektes adressiert werden. Wachsende Datenmengen stellen zwar die Erfolgsgrundlage für eine gewinnbringende Anwendung von ML in der Industrie dar, allerdings ergeben sich aus beständig zunehmenden Daten auch Herausforderungen im Umgang mit diesen Daten.

Standardmäßig wird bei der Nutzung von ML-Verfahren davon ausgegangen, dass die durch die Daten abgebildeten Zusammenhänge konstant bleiben. Bei realen Datenströmen ist dies oftmals aber nicht der Fall. Es können zeitabhängige Veränderungen in der Verteilung der Daten auftreten, die auch als Concept Drift bezeichnet werden. Ein solcher Concept Drift kann sowohl allmählich schleichend als auch ruckartig auftreten. Dabei kann zwischen zwei Formen von Concept Drift unterschieden werden:

1. Virtueller Concept Drift: Die Verteilung der Eingangsdaten (z.B. Wertebereich) verändert sich.
2. Realer Concept Drift: Der zugrunde liegende Zusammenhang zwischen Eingangs- und Ausgangsdaten verändert sich.

Zum Lernen auf fortlaufenden Datenströmen können verschiedene Strategien angewendet werden. Die einfachste Form ist das klassische Batch-Lernen auf der Gesamtdatenmenge. Dabei wird ein Modell gänzlich neu oder erneut auf allen Anwendungsdaten trainiert. Da die Gesamtdatenmenge in der Regel aber nicht in dezentralen Geräten verfügbar ist, ist dieses Vorgehen für Embedded oder Edge Devices weniger geeignet. Eine Lösung kann allerdings geschaffen werden, indem ML-Modelle an einer zentralen Stelle, z.B. in einer Cloud-Anwendung, aufwendig trainiert und anschließend in lauffeiert optimierter Form auf die ausführenden Devices portiert werden.

Alternativ können auch Strategien des inkrementellen Lernens genutzt werden, wie in Abb. 8 dargestellt. Beim Online-Lernen wird das aktuelle Modell anhand eines neuen Lernbeispiels einmalig trainiert. Das Mini-Batch-Lernen hingegen zeichnet sich dadurch aus, dass die neuen, fortlaufenden Daten zunächst so lange gesammelt werden, bis eine bestimmte Menge an neuen Lernbeispielen verfügbar ist. Anschließend wird das aktuelle Modell auf dem gesamten Mini-Batch trainiert. Da die Anpassung der Parameter in diesem Fall gemittelt erfolgt, ist dieses Vorgehen im Allgemeinen etwas robuster aber weniger sensitiv. Bei den Verfahren des inkrementellen Lernens ist es wichtig, die fortlaufenden Datenströme zu überwachen, sodass u.a. auf nicht balancierte Daten geeignet reagiert werden kann – z.B. durch gezielte Auswahl von Daten für inkrementelles Lernen.

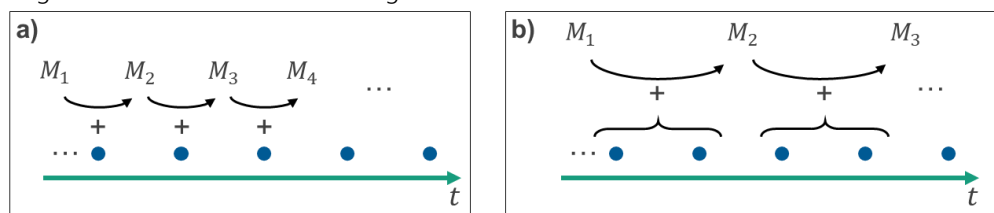
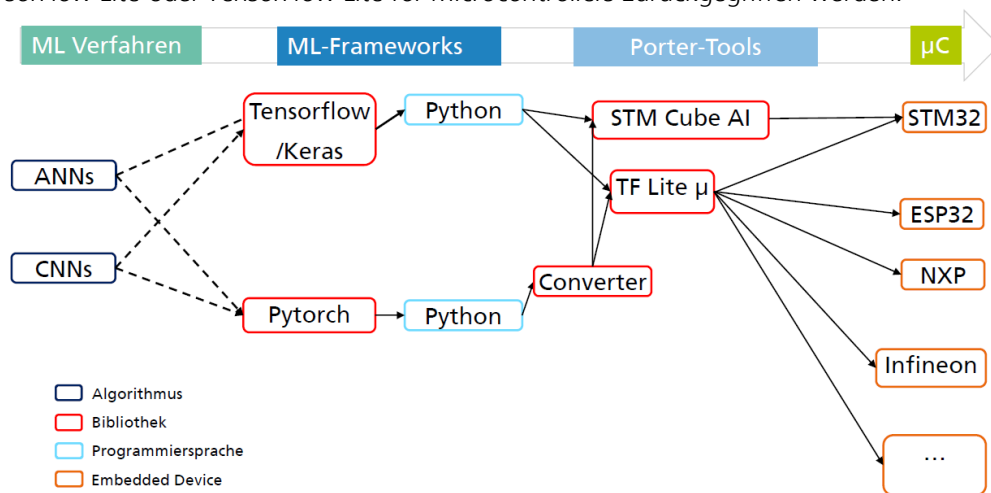


Abb. 8: Strategien zum ML auf Datenströmen: a) Online-Lernen, b) Mini-Batch-Lernen

Bei der Anwendung der Verfahren des inkrementellen Lernens ist stets zu beachten, dass die Stabilität des Trainings nicht zwingend gegeben ist. So können insbesondere bei starkem Concept Drift zuvor gelernte Zusammenhänge durch das Training auf neuen Daten verloren gehen. Zudem ist zu evaluieren, wie stark die Anpassung an neue Daten stattfinden darf. Dieses Verhalten kann unter anderem durch die Lernrate beeinflusst werden und ist anwendungsspezifisch festzulegen. Bei kleineren Lernraten werden bereits gelernte Datenzusammenhänge zwar besser beibehalten, die Anpassung auf einen vorhandenen Concept Drift findet allerdings eher träge statt.

## AP 5: Laufzeitoptimierung

Eine wesentliche Fragestellung im Projekt sowie Anforderung an die entwickelten Verfahren in den Use Cases betrifft die Ausführbarkeit der entwickelten Algorithmen unter begrenzten Rechenressourcen auf Embedded- und Edge Devices. AP 5 fokussierte diesbezüglich die Optimierung von Lernverfahren hinsichtlich Laufzeit zur echtzeitfähigen Datenanalyse. Für die Laufzeitoptimierung der in AP 2 erarbeiteten ML-Verfahren (Neuronale Netze / Convolutional Neural Networks (CNNs)) wurde ein Leitfaden ausgearbeitet, der zeigt, welche Tools eingesetzt werden können, um die mittels *Python* entwickelten Algorithmen laufzeitoptimiert auf einer Zielhardware wie einem Embedded Device aufzuspielen und dort ausführen zu können (vgl. Abb. 9). Für das im Use Case Predictive Maintenance verwendete Sensorboard ist eine Ausführung direkt auf dem MEMS-Sensorboard über das herstellereigene Porter-Tool *Cube AI* der Firma STM möglich. Allgemeiner kann für die Laufzeitoptimierung auf Embedded oder Edge Devices auf *TensorFlow Lite* oder *TensorFlow Lite for Microcontrollers* zurückgegriffen werden.



**Abb. 9: Leitfaden: Laufzeitoptimierten Portierung von ausgewählten künstlichen neuronalen Netzen**

Im Rahmen einer kompakten Studie wurde untersucht, wie sich die Prädiktionsgüte von CNNs durch eine Laufzeitoptimierung verändert. In dem Anwendungsbeispiel konnten der durch das Modell belegte Speicherplatz sowie die benötigte Rechenkapazität zur Ausgabe einer Prädiktion auf Eingangsdaten durch Pruning und Quantisierung um bis zu 30 % reduziert werden. Eine Abnahme der Prädiktionsqualität konnte auf dem betrachteten Datensatz nicht festgestellt werden. Die Modelle konnten zudem auf einem Microcontroller ausgeführt werden. Da in den Use Cases Edge Devices zur dezentralen Datenverarbeitung verwendet werden, die deutlich mehr Rechenkapazitäten als die betrachteten Mikroprozessoren auf MEMS-Basis aufweisen, waren für die Ausführung der Algorithmen in den Use Cases insbesondere nach Optimierung keine Laufzeitprobleme zu erwarten.

### 2.3.2 Use Case Predictive Maintenance

Im Use Case Predictive Maintenance wurde die prädiktive Instandhaltung von Intralogistiksystemen am Beispiel von Hochregallagern untersucht. Der Fokus lag auf der Überwachung von Regalbediengeräten (RBG) zur Ein- und Auslagerung von Waren in Hochregallagern. Dieser Use Case wurde durch das Fraunhofer IEM in Kooperation mit der encoway GmbH bearbeitet. Unterstützend war auch die Firma Lenze SE als Mutterkonzern der encoway GmbH und Betreiber von Hochregallagern sowie Zulieferer für Antriebstechnik in Hochregallagern in das Vorhaben involviert. Im Rahmen des Use Cases wurden laufzeitoptimierte Verfahren des maschinellen Lernens zur Diagnose und Prognose von Systemausfällen und Wartungszeitpunkten entwickelt.

Nachfolgend werden die wissenschaftlich-technischen Ergebnisse der Arbeitspakete für den Use Case Predictive Maintenance sowie die spezifischen Beiträge des Konsortialpartners Fraunhofer IEM und der encoway GmbH dargestellt. In Rahmen der Arbeitspaketen wurden ausgehend von der Analyse von Anforderungen und Datenbasis sowie der Auswahl von Lernverfahren insbesondere Möglichkeiten zur Integration von Expertenwissen erarbeitet. Zudem wurden das Training der Modelle auf fortlaufenden, sich dynamisch entwickelnden Datenströmen umgesetzt und Funktionalitäten zur Laufzeitoptimierung der ML-Verfahren erarbeitet.

#### **AP 1: Anforderungsanalyse**

Für die Anforderungsanalyse wurden im Use Case Predictive Maintenance zunächst kritische Ausfallkomponenten an Regalbediengeräten über Expertenbefragungen und die Auswertung von Wartungsberichten durch die encoway GmbH identifiziert. Daraus ergeben sich potenzielle Fehlerbilder. Diese Fehlerbilder mit ihren Entstehungsmechanismen sind auch an den Regalbediengeräten im Lenze Mechatronics Competence Capus (MCC) zu erwarten, das am Standort Extertal aufgebaut wurde. Dieses wurde während der Projektlaufzeit im Oktober 2022 in Betrieb genommen und konnte in diesem Vorhaben neben einem im Jahr 2021 fertiggestellten Regalbediengerät-Prüfstand als weitere Validierungsmöglichkeit für die zu entwickelnden Technologie-Demonstratoren genutzt werden. Folgende Fehlerarten von Regalbediengeräten wurden für das Projektvorgehen als priorisiert festgelegt:

Für automatisierte Palettenlagergeräte (siehe Abb. 10, links):

- Abnutzung/Materialabtrag an Reibrädern des Fahrtriebs
- Abplattungen an den Reibrädern
- Führungsspiel am Fahrwerk
- Nachlassende Riemenspannung an der Hubeinheit

Für automatisierte Kleinteillagergeräte (siehe Abb. 10, rechts):

- Nachlassende Riemenspannung am Omega-Riemen angetriebenem Fahrwerk
- Nachlassende Riemenspannung an der Hubeinheit

Aus den Analysen konnten verschiedene hardwaretechnische Anforderungen abgeleitet werden. Unter anderem ist die Zielhardware zur Ausführung der Algorithmen von zentraler Bedeutung bei der Auswahl von ML-Verfahren (AP 2). Durch den systemtechnischen Aufbau von Regalbediengeräten sind Recheneinheiten in Form von Edge Devices (hier z.B. Hutschienen PC Arrakis pico Mk3) oftmals bereits im Schaltschrank vorhanden. Aufgrund ihrer Verfügbarkeit im Rahmen der Erprobungsmöglichkeiten beim Projektpartner Lenze sollten diese im Projektverlauf als Zielhardware für die dezentrale Ausführung von ML-Algorithmen während des Betriebs genutzt werden.

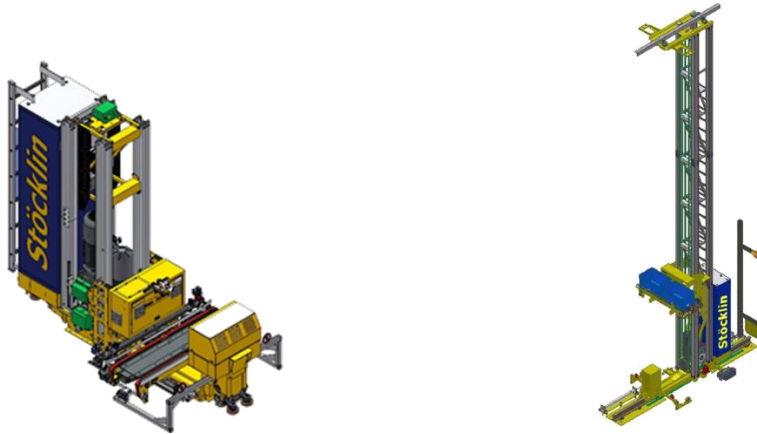


Abb. 10: CAD-Darstellung der MCC-RBG: Palettenlagergerät (links) und Kleinteillagergerät (rechts)

Darüber hinaus konnten aus den identifizierten Fehlerbildern auch Anforderungen an benötigte Messgrößen zur sensorbasierten Zustandsdiagnose erhoben werden. Bereits durch vorhandene Sensorik der Lenze-Umrichter waren Messdaten zu erwarten, mit denen bestimmte technische Defekte insbesondere an den Antriebskomponenten (Motor, Getriebe, etc.) der Regalbediengeräte erkannt werden können. Unter anderem zur Detektion von mechanischen Defekten sind allerdings Vibrationsdaten von zentraler Bedeutung. Daher wurde für den Use Case Predictive Maintenance die Anbindung einer zusätzlichen Sensorik vorgesehen, die zur Messung von Vibrationen/Körperschall an Regalbediengeräten eingesetzt werden kann.

Wichtige Schritte hierbei waren die Validierung der Sensorik hinsichtlich der Betriebsbedingungen und die Integration in das industrielle Umfeld mitsamt Schnittstellen. Abb. 11 zeigt in vereinfachter Darstellung die resultierende Systemarchitektur für den Datenaustausch zwischen verschiedenen Systembestandteilen. Vor dem Hintergrund der stetigen Weiterentwicklung mikroelektronischer Messsysteme wurde ein MEMS-Sensorboard der Firma STMicroelectronics vom Typ STEVAL-STWINKT1B verwendet. Ein solches mikroelektronisches System bietet im Vergleich zu herkömmlichen Messsystemen kostentechnische Vorteile. Diese ermöglichen im Nachgang an das Forschungsprojekt eine geringere Hemmschwelle für interessierte Unternehmen ihre Intralogistiksysteme mit Sensorik nachzurüsten, da die damit verbundenen Anschaffungskosten gering ausfallen.

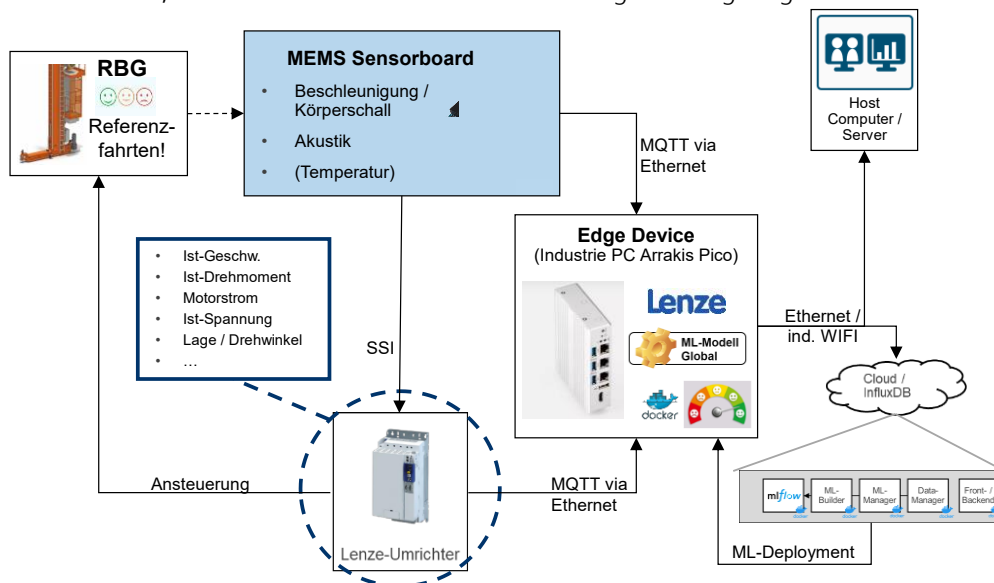


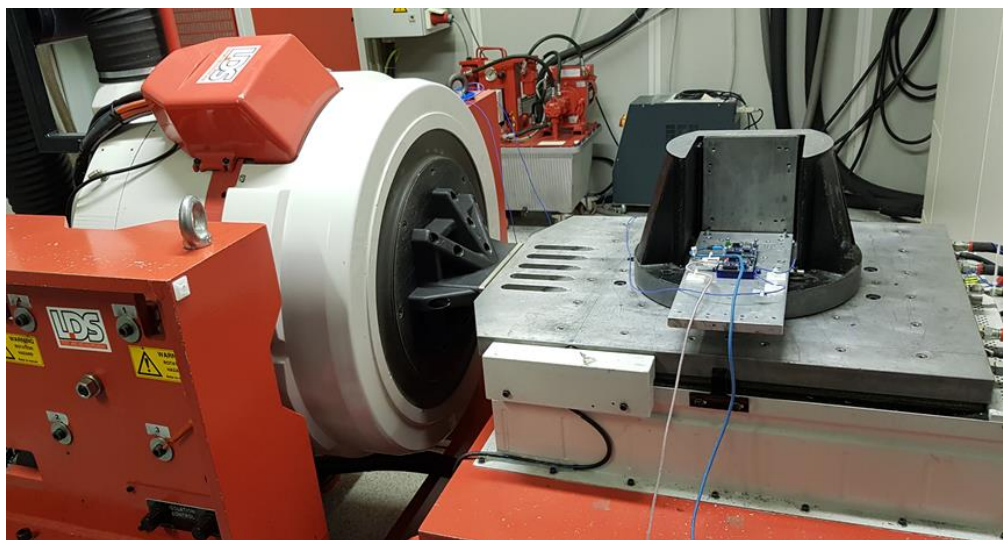
Abb. 11: Übersicht Systemarchitektur zur Zustandsüberwachung eines RGB im Use Case PM

Die aufgezeigte Systemarchitektur zeigt die wesentlichen Schnittstellen und Verbindungen zur Kommunikation von Daten und skizziert damit die angestrebte Funktionsweise des im Projekt entwickelten Technologie-Demonstrators. Um gute Bedingungen für eine zuverlässige Zustandsprädiktion zu schaffen, wurde zunächst festgelegt, dass die Aufzeichnung von Messdaten sowohl an dem vorhandenen Regalbediengerät-Prüfstand als auch im realen Hochregallager für Referenzfahrten mit konstanter Geschwindigkeit und ohne Last erfolgen soll, um die Daten untereinander vergleichbar zu machen. Eine dementsprechende Ansteuerung der Antriebe erfolgt über die verbauten Lenze-Umrichter, deren Daten über Motorstrom und Ist-Spannung ebenso an den Edge-Device übertragen werden, wie Informationen über Drehwinkel, Drehmoment und Drehzahl.

Weiterhin sollen auch die Messdaten des zusätzlichen Sensorboards verfügbar gemacht werden. Zum einen soll ein Transfer der Daten zum Umrichter realisiert werden, um gemessene Beschleunigungen für regelungstechnische Aufgaben nutzen zu können. Für die Verwendung im Rahmen einer prädiktiven Instandhaltung soll zudem eine Weiterleitung der Daten via MQTT per Ethernet an den Edge-Device und damit in das Lenze-System erfolgen. Von dort aus kann eine Übertragung der Daten auf einen externen Server bzw. eine Cloudanwendung erfolgen. Dort wiederum soll die Initialisierung, das Training und die Verwaltung der ML-Algorithmen erfolgen. Darunter fällt auch das Life-Cycle-Management der Algorithmen. Diese Auslagerung des maschinellen Lernens wird angestrebt, da die Umsetzung auf einem Edge Device aufgrund der enormen Datenmengen und Rechenoperationen beim Training von ML-Algorithmen ressourcentechnisch nicht ohne enormen Datenverkehr möglich ist.

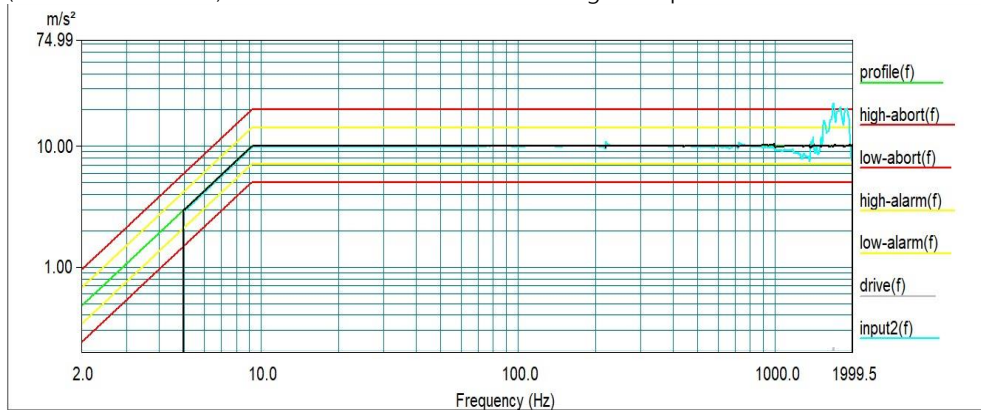
Anschließend können die resultierenden ML-Modelle zur Prädiktion über eine Online-Datenbank verwaltet werden und je nach Bedarf auf den Edge Device übertragen und dort ausgeführt werden. Für eine Portabilität der Algorithmen sollen dazu Docker-Container als abgetrennte Code-Einheiten genutzt werden. Damit wird eine dezentrale Online-Prädiktion auf neuen Messdaten im Betrieb möglich. Die mittels maschinellen Lernens berechneten Ausgaben können im Anschluss anderen Diensten zur Verfügung gestellt werden und so beispielsweise den Service-Mitarbeitern im Hochregallager verfügbar gemacht werden.

Um die Einsatzfähigkeit der Sensorik zu prüfen, wurde das Sensorboard im Rahmen einer Versuchsreihe validiert. Dazu wurde das Sensorboard wie in Abb. 12 dargestellt auf einem Shaker-Prüfstand angebracht und mit Orientierung an der Norm DIN ISO 10816 für mittelgroße Maschinen und Transportnorm IEC/EN 60721-3-2 für Klasse 2M2 mit Schwingungen bei verschiedenen Geschwindigkeiten und Beschleunigungen entlang der verschiedenen Messachsen angeregt.



**Abb. 12: Prüfaufbau zur Validierung des STM-Sensorboards an einem Shaker**

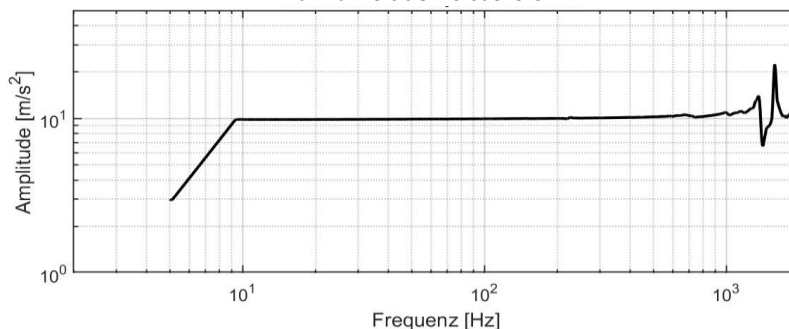
Für diese Versuche zeigt Abb. 13 exemplarisch eine Anregung des Prüfstands, bei der ein sogenannter Sinus-Sweep durchlaufen wurde. Bis zu einer Anregung von ca. 9 Hz wurde die Schwingungsgeschwindigkeit und anschließend von 9 Hz bis 2 kHz die Beschleunigung der Schwingung konstant gehalten. Der hellblaue Verlauf kennzeichnet dabei den Messverlauf eines Referenzsensors, der die Schwingungen der Montageplatte erfasst, auf der sich die Sensorik befand. Dieser Verlauf diente als Referenz bei der Bewertung des mikroelektronischen Sensorboards. Abweichungen von der Sollwertvorgabe (schwarzer Verlauf) resultierten aus auftretenden Eigenfrequenzen des Shakers.



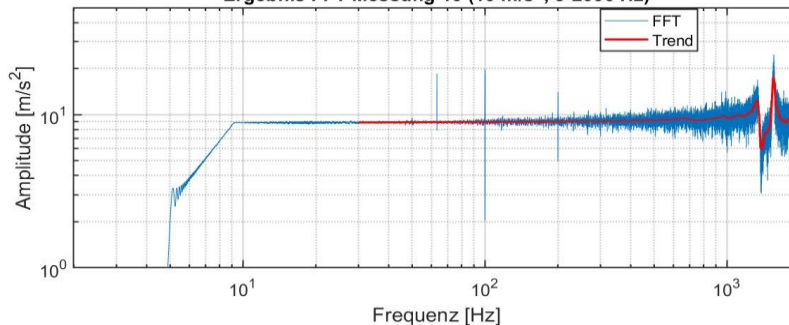
**Abb. 13: Logarithmische Darstellung der Frequenz-Anregung eines exemplarischen Messversuchs**

Die Auswertung der Versuchsdaten erfolgte unter anderem anhand der exemplarischen Verläufe in Abb. 14. Dabei wurde geprüft, ob die angeregten Schwingungen sowohl als Hüllkurven im Zeitbereich als auch über Frequenzanalysen mittels Fourier Transformation aus den Messdaten des Sensorbords reproduzierbar sind. An den in Abb. 14 dargestellten Verläufen wird exemplarisch für die in Abb. 13 gezeigte Anregung deutlich, dass mittels des Sensorboards die angeregten Schwingungen im Vergleich mit den Referenzsensoren in geeigneter Weise erfasst werden können. Dies ergibt sich daraus, dass grundsätzlich sowohl die auftretenden Frequenzen als auch die Amplituden dieser Frequenzen aus den Messungen mittels des zu validierenden Sensorboards ermittelbar sind.

#### Hüllkurve aus Zeitbereich

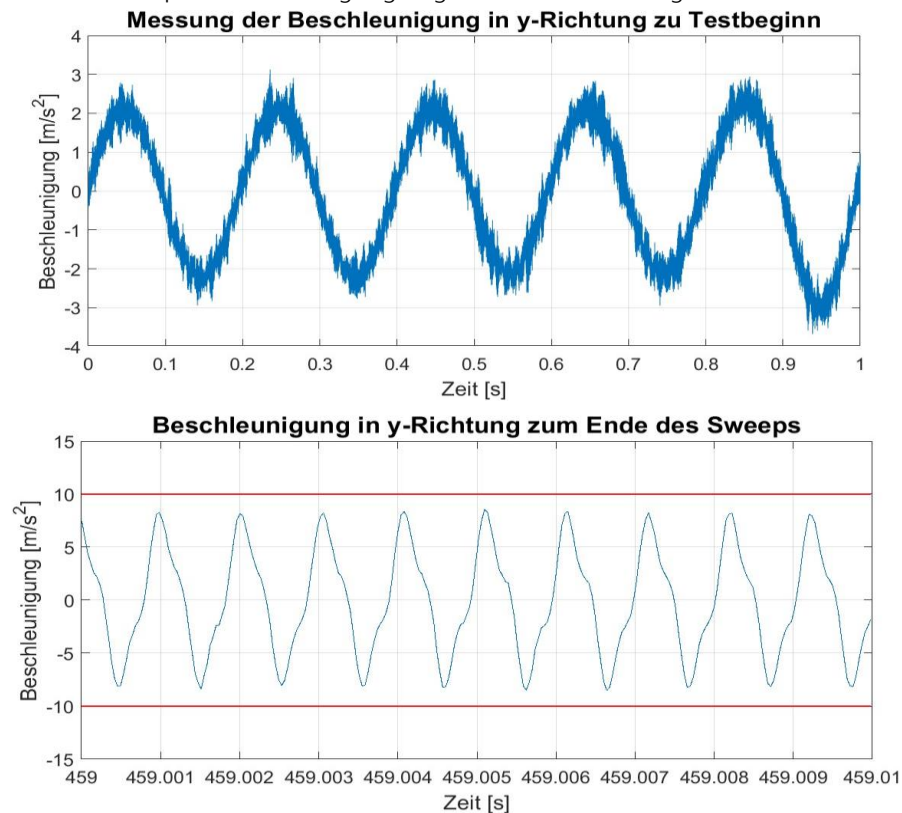


#### Ergebnis FFT Messung 10 (10 m/s<sup>2</sup>, 5-2000 Hz)



**Abb. 14: Beispielhafte Messergebnisse des mikroelektronischen Vibrationssensors**

Auch die detaillierte Betrachtung der Schwingungsverläufe zu Beginn und Ende eines Sweeps bestätigt die Erkenntnis, dass die Sensorik für die Messung von Schwingungen in entsprechenden Frequenzbereichen und für typische Geschwindigkeiten und Beschleunigungen nach Transportnorm geeignet ist, wie in Abb. 15 ersichtlich. Rauschen bei niederfrequenten Schwingungen gilt es dabei zukünftig zu filtern.



**Abb. 15: Signalverlauf einzelner Schwingungen zu Beginn und Ende eines Sweep-Tests**

Insgesamt haben die Untersuchungen gezeigt, dass die relevanten Frequenzen zur Zustandsdiagnose im Use Case mittels des Sensorboards erfasst werden können. Weiterhin wurde ein Robustheitstest durchgeführt. Auch einer Schockprüfung mit einer Beschleunigung von  $300 \text{ m/s}^2$  hat die Sensorik ohne Nullpunktverschiebung standgehalten. Bei der ausstehenden Anbringung der Sensorik an Regalbediengeräten ist lediglich zu berücksichtigen, dass für die z-Achse Eigenschwingungen der Platine bei Frequenzen oberhalb von  $1 \text{ kHz}$  auftreten können.

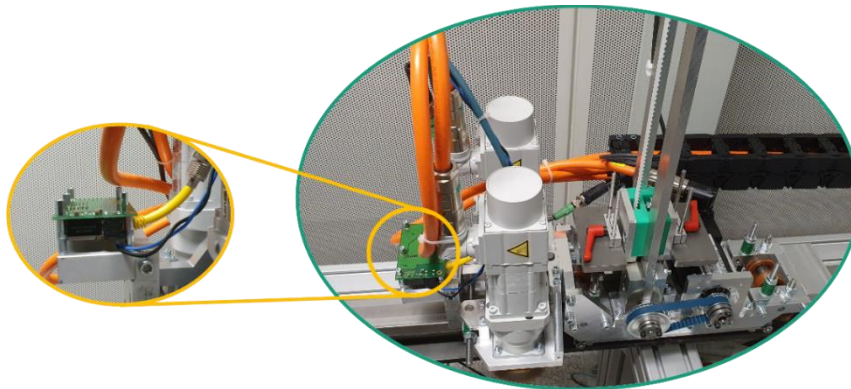
Im Anschluss an die Validierung des Sensorboards wurde gemeinsam mit der weiteren Entwicklung zur Systemintegration (Einhausung, Anbindung Adapterplatinen, Systemintegration) begonnen. In diesem Zuge unterstützte das Fraunhofer IEM die Entwicklung eines Prototyps durch die Firma Lenze hinsichtlich der hardware- und softwaretechnischen Gestaltung der Schnittstellen zwischen Board und den Edge Device. Dabei waren auch die vorhandenen Lenze-Dienste zu berücksichtigen, mit denen das Sensorboard für die Datenakquisition kommunizieren muss. Allgemein waren die Entwicklung der Firmware sowie die Definition und Implementierung der Schnittstellen zum Sensorboard wesentlicher Bestandteil der Arbeit im Projekt im Kalenderjahr 2021.

Anschließend wurde die Systemintegration des entwickelten Sensorboards weiter vorangetrieben. So wurde die Hardware basierend auf den Ergebnissen der Validierung des Sensorboards für den industriellen Einsatz im MCC weiterentwickelt. Dazu wurde unter anderem eine Adapterplatine für die Übertragung der Messdaten via MQTT entwickelt und angeschlossen und ein optimierter mechanischer Aufbau beider Platinen identifiziert. Zudem wurden die schwingungstechnischen Einflüsse einer Einhausung mitsamt angeschlossenem Kabel auf die resultierenden Beschleunigungsmessdaten untersucht.

Die Erprobung des Gesamtaufbaus mittels der im vorherigen dargestellten Shaker-Tests zeigt nur eine leicht veränderte Erfassung von Messfrequenzen, was auf das veränderte Resonanzverhalten des Gesamtaufbaus im Vergleich zur einzelnen Sensorplatine zurückzuführen ist. Die Einflüsse vom Gehäuse konnten als vernachlässigbar identifiziert werden. Einflüsse durch die Kabel zur Stromversorgung sowie das Ethernet-Kabel mitsamt Stecker zur Übertragung der Messdaten konnten nicht festgestellt werden.

Insgesamt ist festzuhalten, dass die MEMS-Sensorik vielversprechende Messdaten für die industrielle Zustandsüberwachung liefert. Zwar reicht die Qualität der Messungen nicht an die von hochwertigen Industriesensoren heran. Allerdings beruht die Zukunftsfähigkeit dieser Lösung zum einen auf der noch immer stark voranschreitenden technischen Weiterentwicklung von Halbleitertechnik und zum anderen auf dem im Vergleich zu Industriesensoren deutlich geringeren Stückkostenpreis von Embedded MEMS-Sensoren, durch den ein breiter Einsatz in der Zustandsüberwachung technischer Anlagen kostentechnisch realisierbar wird.

Ausgehend von den Sensorik-Validierung wurde der Laborprüfstand erweitert. So wurde hier zum einen die Infrastruktur mitsamt der Sensorbox integriert, um einen Demonstrator für die Erprobung und Optimierung der Datenaufzeichnung zu schaffen. Die montierte Platine mitsamt der angeschlossenen Kabelverbindungen ist in Abb. 16 dargestellt. Zudem wurde durch die Firma Lenze ähnlich wie im MCC ein Laser-Absolutwertgeber am Prüfstand angebracht. Die zusätzlichen Messwerte sollten unter Nutzung der ausgewählten Lernverfahren eine Detektion und Isolation von Schlupfeffekten ermöglichen.



**Abb. 16: Demonstrator eines Regalbediengeräts mit finaler Version der Vibrations-Sensorik**

Weiterhin erfolgte auch die Implementierung der gemeinsam entwickelten Sensorik mitsamt Einhausung zur Eröffnung des MCC an einem ersten RBG. Entgegen dem ursprünglichen Plan konnten aufgrund der noch während des Projektes stark beeinträchtigten Lieferketten für Halbleiterelektronik vorerst keine weiteren RBGs mit der MEMS-Sensorik ausgestattet werden. Aus diesem Grund fiel die Datengrundlage geringer aus als vorgesehen. Ein erfolgreicher Abschluss des Use Case Predictive Maintenance durch Erprobung der entwickelten Algorithmen war allerdings auch auf der damit resultierenden, geringeren Datenbasis erwartbar.

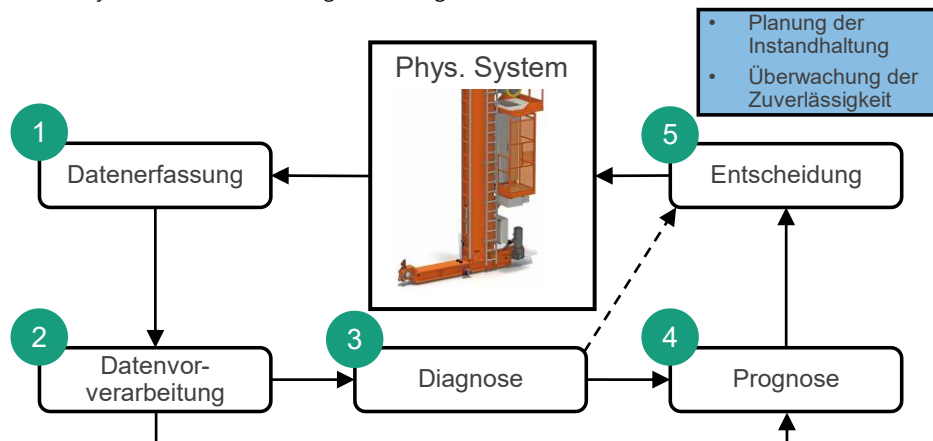
Die Zusammenarbeit sowie die Tätigkeiten der am Use Case Predictive Maintenance im beteiligten Partner Fraunhofer IEM und encoway ist in Tabelle 1 für die Arbeiten im Rahmen des AP1 zur Anforderungserhebung in einer Übersicht dargestellt. Diese Darstellung der Arbeiten wird im auch für die weiteren Arbeitspakete aufgezeigt, um die wesentlichen Anteile der Partner in der engen Zusammenarbeit erkenntlich zu machen.

**Tabelle 1: Zuordnung der Arbeiten für den Use Case Predictive Maintenance für AP 1**

MEMS-Sensorik					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X		X	X
Encoway/Lenze			X		X
Teststand					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM				X	X
Encoway/Lenze	X	X	X		X
Anforderungserhebung Mechanische Fehler					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM		X		X	X
Encoway/Lenze	X		X	X	X
Anforderungserhebung IT-Architektur					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM		X	X	X	X
Encoway/Lenze	X	X	X	X	X

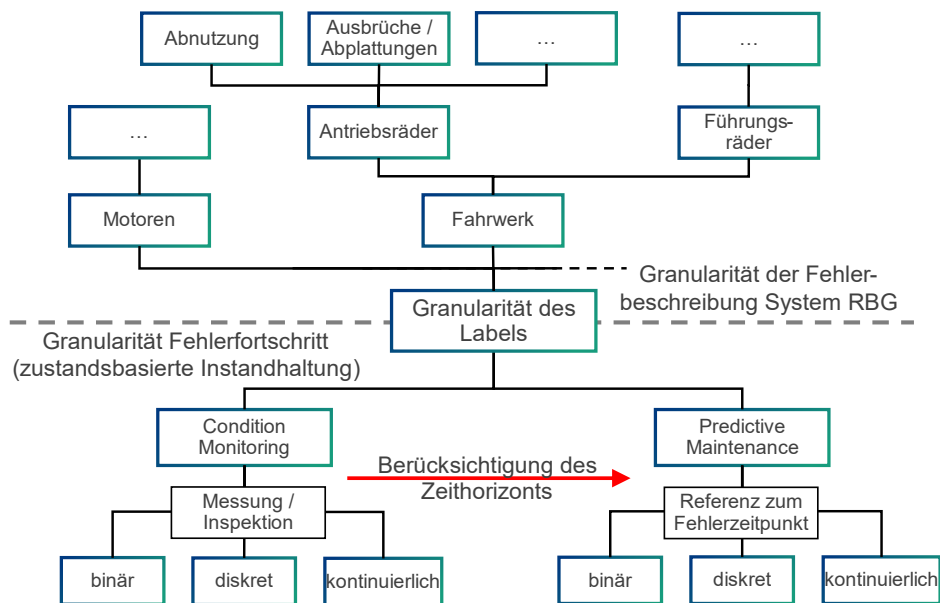
**AP 2: Analyse und Auswahl von Lernverfahren**

Aufbauend auf den Anforderungen hinsichtlich notwendiger Messdaten und der Festlegung auf einen Edge Device konnten vor dem Hintergrund von Arbeitspaket 2 geeignete ML-Algorithmen für die Zustandsprädiktion im Use Case identifiziert werden. Im ersten Schritt wurde dazu das allgemeine ML-Framework nach Abb. 17 beschrieben. Dieses skizziert die wesentlichen Schritte in der zustandsbasierten Diagnose an einem technischen System wie einem Regalbediengerät.



**Abb. 17: Vorgehen in der zustandsbasierten Instandhaltung**

Bei der Anwendung maschineller Lernverfahren zur Fehlerprädiktion wird üblicherweise ein Algorithmus des überwachten Lernens verwendet. Dabei steht für Messdaten ein Feedback über den zugehörigen Zustand, Austauschzeitpunkte oder korrektive Wartungen der zu betrachtenden Komponenten zur Verfügung. Dieses Feedback wird auch als Label bezeichnet. Um im Projekt ein einheitliches Verständnis über das Labeling von Daten zu erreichen, wurde ein Tabelle entworfen, die für unterschiedliche Fehlerarten von Regalbediengeräten eine Festlegung des Labels hinsichtlich Granularität ermöglicht. Dabei kann zum einen bezüglich des Detailierungsgrads der Fehlerbeschreibung von Teilsystemen über Komponenten bis hin zu spezifischen Fehlern und zum anderen bezüglich der Granularität des Fehlerfortschritts unterschieden werden, wie in Abb. 18 ersichtlich.



-----  
Eingehende Darstellung  
-----

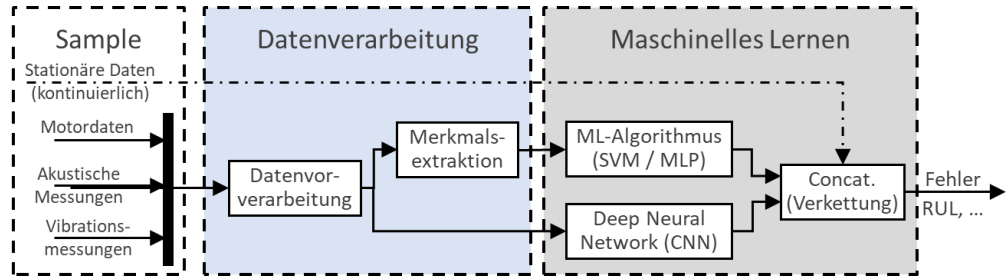
**Abb. 18: Granularität des Labels bezüglich Fehlerfortschritt und Fehlerbeschreibung (RBG)**

Diese Beschreibung des Labels anhand der Tabelle soll dafür sorgen, dass notwendige Zustandsinformationen für die prädiktive Instandhaltung erhoben und in der Datenakquisition berücksichtigt werden. Dieser Vorgang wurde im Projekt abgeschlossen. Anpassungen können allerdings für das MCC jederzeit erfolgen, sobald neue konkrete Informationen über die Durchführung von Wartungen und die damit verbundene Erhebung von Zustandsinformationen als Label vorliegen. Diese werden vom Hersteller der verwendeten Regalbediengeräte zur Einsicht frei gegeben.

Zur Durchführung der Zustandsprädiktion wurden unterschiedliche ML-Algorithmen, wie Support Vector Machines, k-nächste Nachbarn, und Baum-basierte Verfahren, wie Decision Trees und neuronale Netze, im Rahmen einer umfangreichen Literaturrecherche untersucht sowie hinsichtlich ihrer Vor- und Nachteile bewertet. Dabei erschienen insbesondere Support Vector Machines und künstliche neuronale Netze aufgrund ihrer hohen Prädiktionsgenauigkeiten bei ausreichender Datenlage als Zustands-Klassifikatoren für das Vorhaben im Use Case geeignet.

Ein wesentlicher Aspekt in der Datenverarbeitungskette stellt die Datenvorverarbeitung dar, die auch im Bereich des maschinellen Lernens oftmals umfangreiche explorative Datenanalysen erfordert. Zwar gibt es aus dem Bereich des Deep Learnings Algorithmen, die darauf abzielen auch die Merkmalsextraktion zu automatisieren, allerdings sind dazu in der Regel enorme Datenmengen erforderlich. Zudem stellt die manuelle Extraktion von Merkmalen über Signalmodelle bereits einen ersten Schritt zur Integration von Expertenwissen in den Prozess der Diagnose und Prognose dar. Vorteile der Anwendung von Deep Learning Algorithmen ergeben sich aber insbesondere aus der Extraktion versteckter Signalmerkmale auf multi-dimensionalen Datenströme.

Convolutional Neural Networks stechen an dieser Stelle hervor, da sie aufgrund ihrer lokalen Datenverarbeitung über zu Filtern gruppierte Neuronen eine besondere Eignung zur Extraktion von Merkmalen in Daten aufweisen, und das unabhängig von der exakten Lokalität dieser Merkmale. Daher wurde im Weiteren auch die Verschaltung von einem traditionellen maschinellen Lernverfahren mit einem CNN angedacht, falls dies die Datenlage zulässt. Eine solche Anordnung ist in Abb. 19 dargestellt. Bei zu wenig Daten hingegen kann dabei auf das Deep Learning verzichtet werden. Dann ist die Performanz der ML-Prädiktion stärker von der Merkmalsextraktion sowie der Einbringung von Expertenwissen abhängig und weniger von der Quantität der verfügbaren Daten zum Training der ML-Algorithmen.



**Abb. 19: Verschaltung von klassischem ML-Verfahren mit Deep Learning**

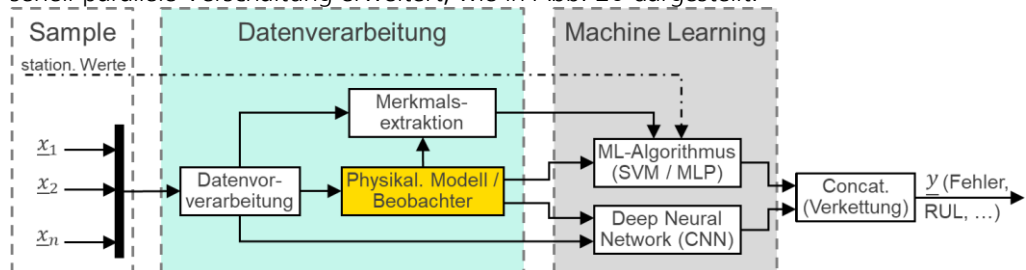
Weiterhin zeigt Tabelle 2 die Rollenverteilung bei der Analyse und Auswahl von lernverfahren, bei der das IEM koordinierend tätig war.

**Tabelle 2: Zuordnung der Arbeiten für den Use Case Predictive Maintenance für AP 2**

Analyse und Auswahl von ML-Verfahren					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
Encoway/Lenze		X	X	X	X

**AP 3: Integration von Expertenwissen**

Für die Integration von Expertenwissen in den ML-Prozess ergeben sich verschiedene Möglichkeiten. Neben der direkten Nutzung von physikalischem Vorwissen und Simulationsmodellen zur Extraktion von zustandscharakterisierenden Features oder zur Generierung zusätzlicher Signale kann Expertenwissen auch aus Modellen oder Demonstratoren resultieren und indirekt beispielsweise in Form von Daten zum Vortrainieren von ML-Algorithmen genutzt werden. In Anlehnung an die in Abschnitt 2.3.1 beschriebenen hybriden Verschaltungsszenarien zur Integration von Expertenwissen in maschinelle Lernverfahren wurde der identifizierte ML-Ansatz aus dem vorhergehenden AP 2 auf eine seriell-parallele Verschaltung erweitert, wie in Abb. 20 dargestellt.



**Abb. 20: Hybride Verschaltung von ML-Algorithmen zur Integration von Expertenwissen**

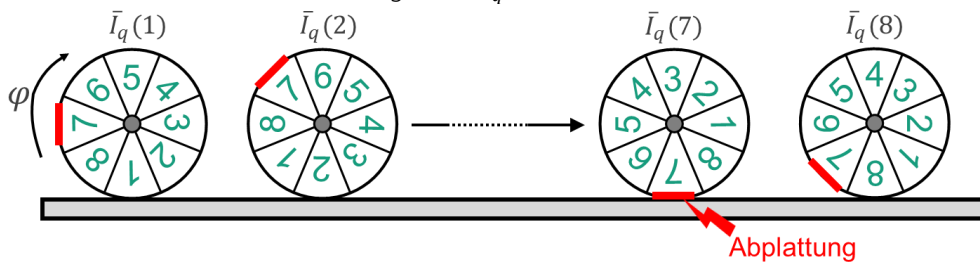
In diesem Szenario werden physikalische Zusammenhänge und Modelle verwendet, um zusätzliche Features und Datenreihen für die ML-Prädiktion zu generieren. Resultierende Datenreihen durchlaufen zusätzlich die Merkmalsextraktion, können aber auch direkt als Eingang für einen CNN-basierten End-to-End Deep Learning Algorithmus dienen.

In der Bearbeitung von AP 3 konnten spezifisch für die Fehlerdiagnose an RBGs Algorithmen identifiziert werden, die Domänenwissen in den Prozess einbringen. Für die Überwachung von Getrieben standen durch die Experten von Lenze und encoway Algorithmen zur Verfügung, die eine auf die Dreh- und Zahneingriffsfrequenzen einzelner Getriebestufen zugeschnittene Berechnung der spektralen Leistungsdichte ermöglichen. Für die charakteristischen Frequenzen, die auf Verschleiß im Getriebe hinweisen, können so gezielt Features berechnet und mittels ML überwacht werden.

Weiterhin wurden in der Anforderungsanalyse ein entstehendes Führungsspiel sowie der Verschleiß an Reibräder von RBGs mit Radantrieb als kritische Systemzustände identifiziert, die durch Verschleiß auftreten. Für eine Überwachung des Reibradverschleißes

zeigt sich die Notwendigkeit zur Einbringung von Expertenwissen besonders deutlich. Während in den reinen Rohdaten der Beschleunigung kaum eine merkmalsbasierte Isolation verschiedener Verschleißformen wie Abplattungen von anderen Effekten an Motor, Getriebe oder Fahrschiene möglich ist, können physikalische Modelle die Entscheidungsfindung durch Aufwertung der Daten um zusätzlichen Informationsgehalt stützen. So kann die Differenz zwischen Absolutwertgeber und der aus der Motorposition berechenbaren idealen Fahrstrecke für einen definierten Messzeitraum als Feature dienen, dass zur Überwachung des Schlupfes genutzt werden kann. Auch der reale Radumfang kann aus der gemeinsamen Vorverarbeitung dieser Signale als eine verschleißcharakterisierende Größe (Health Index) berechnet werden.

Ein Verschleißbild, das in der Bearbeitung des AP 3 für den Use Case Predictive Maintenance besonders im Vordergrund stand, stellen die Abplattung von Antriebsrädern dar. Diese entstehen durch Blockieren der Räder eines RBG beim starken Abbremsen. Einmal induziert, kommt es immer wieder an der gleichen Kontaktstelle des Rades zum Blockieren auf der Schiene, womit sich der Defekt immer weiter verstärkt. Zur Detektion dieses Fehlers können unter anderem Merkmale aus dem Frequenzbereich dienlich sein. Die relevante Fehlerfrequenz einer Abplattung setzt aber eine konstante Drehgeschwindigkeit voraus, die in der Realität im Allgemeinen nicht gegeben ist. Daher wurde ein Verfahren entwickelt, dass basierend auf den Motordaten den  $I_q$ -Strom, der das Motormoment charakterisiert, drehwinkelsynchronisiert überwacht, wie in Abb. 21 schematisch dargestellt. Dem Vorgehen liegt der physikalische Zusammenhang zugrunde, dass beim Überrollen einer Abplattung das mechanische Moment variiert, welches zur Regelung auf die gewünschte Drehgeschwindigkeit benötigt wird. Die resultierende Schwankung wirkt sich somit auch auf den anliegenden  $I_q$ -Strom des Antriebsmotors aus.



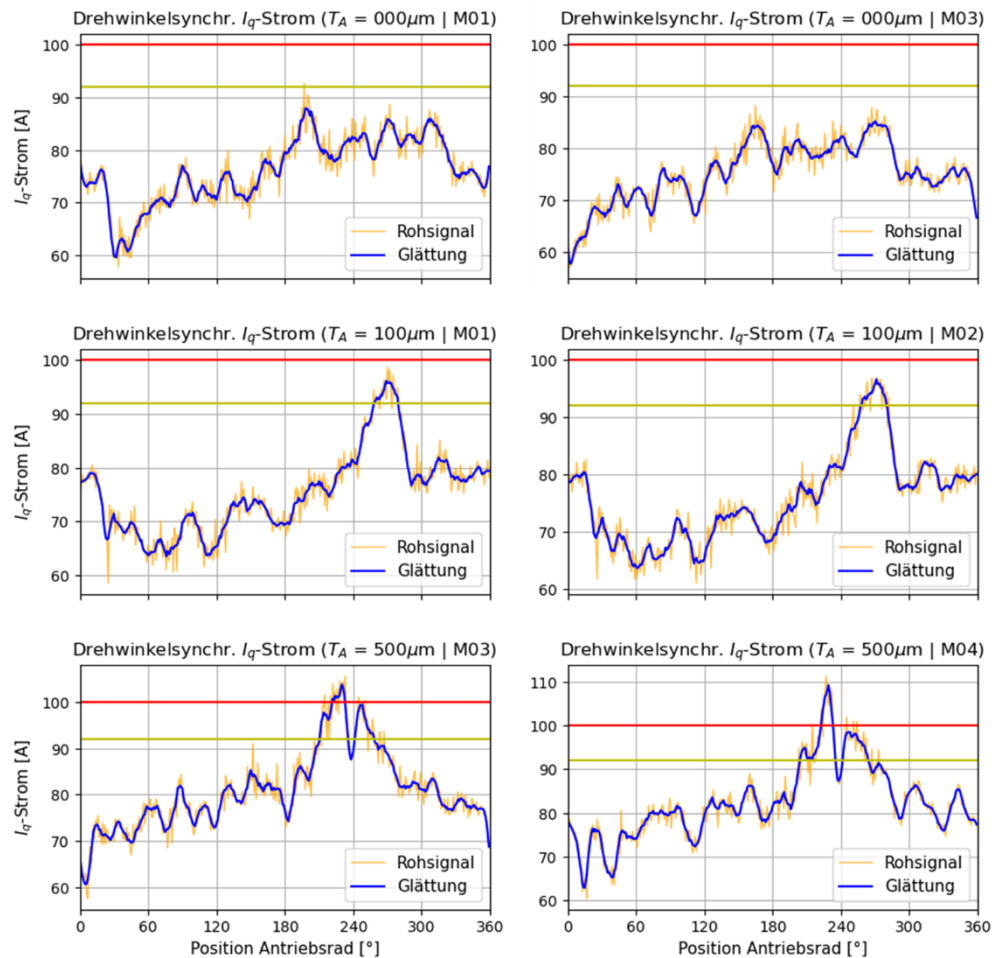
**Abb. 21: Prinzip der auf den Raddrehwinkel synchronisierten Transformation von Motorstromdaten**

Zur Berechnung der auf den Raddrehwinkel synchronisierten  $I_q$ -Ströme wird das Rad in eine definierte Anzahl gleich großer Bereiche aufgeteilt. Die Berechnung des Raddrehwinkels von  $0^\circ$  bis  $360^\circ$  ist aus der vom Umrichter verfügbaren Motorposition unter Nutzung des bekannten Übersetzungsverhältnisses im Getriebe möglich. Aus den jeweils in einem Winkelbereich des Rades auftretenden  $I_q$ -Strömen kann ein Mittelwert berechnet und dem inkrementellen Radbereich zugeordnet werden.

Der Vorteil dieser Methode liegt darin, dass durch Auftragung des  $I_q$ -Stroms über den Drehwinkel des Reibrades rotatorisch auftretende Effekte unabhängig von der Drehzahl abgebildet werden. Somit sind die Ausgaben dieses physikalischen Modells auch repräsentativ für die Daten der Beschleunigungs- und Bremsphasen der RBGs, während derer sich Effekte im  $I_q$ -Strom und damit im Motormoment besonders deutlich äußern.

Abb. 22 zeigt exemplarische Ergebnisse für die drehwinkelsynchronisierte Mittelwertbildung des  $I_q$ -Stroms anhand eines am Laborprüfstand aufgenommen Datensatzes. Dabei wurden Räder mit unterschiedlichem Abplattungsgrad am Prüfstand montiert: zum einen ein gesundes Rad und zum anderen ein Rad mit einer großen Abplattung von  $500\ \mu\text{m}$  Tiefe. Zudem wurden auch Daten für die Nutzung eines Rades mit einer sehr kleinen Abplattung  $100\ \mu\text{m}$  Tiefe, welche die einsetzende Entstehung des Fehlerbildes simuliert, aufgezeichnet. Für jeden Fehlerfall sind exemplarisch je zwei Ergebnisse der drehwinkelsynchronisierten  $I_q$ -Ströme dargestellt. Das Rohsignal ergibt sich dabei aus den Ergebnissen für eine Aufteilung des Rades in 360 Winkelbereiche mit einer Größe

von je 1°. Weiterhin ist ein geglätteter Verlauf visualisiert, der aufgrund seiner höheren Robustheit gegen Ausreißer eine bessere Grundlage zur Berechnung von Features zur Charakterisierung von Abplattungen an Antriebsrädern darstellt.



**Abb. 22: Verläufe der drehwinkelsynchronisierten  $I_q$ -Ströme für unterschiedliche Stufen der Reibradabplattungen am RBG-Demonstrator**

Die visualisierten Ergebnisse deuten klar an, dass für die Prüfstanddaten unter Zuhilfenahme des Wissens über die Systemzusammenhänge zwischen Motor, Getriebe und Antriebsrad eine Detektion von Abplattungseffekten durchgeführt werden kann. Dies ist durch die gelben und roten Linien angedeutet, die suggerieren, dass eine schwellwertbasierte Fehlerklassifikation möglich ist. Dabei lassen sich auch kleine Abplattungen eindeutig von großen Abplattungen sowie gesunden Antriebsrädern unterscheiden. Durch die Berechnung von Features wie Minima, Maxima, Crest-Faktor oder Spitze-zu-Spitze-Wert kann die Datengrundlage für die Zustandsüberwachung mittels ML somit gezielt um wertvolles Expertenwissen angereichert werden. Für den realen Anwendungsfall ist aufgrund zahlreicher Wechselwirkungen und Einflussfaktoren anzunehmen, dass einfache Schwellwerte von Features nicht mehr ausreichen.

Mit wachsender Datenbasis kommt an dieser Stelle die Stärke von ML-Verfahren zum Tragen, auch versteckte Muster in Daten zu erkennen, die weit über die Komplexität schwellwertbasierter Ansätze hinausgehen. Dabei stellt der Ansatz der für das Antriebsrad drehwinkelsynchronisierten Mittelwerte des  $I_q$ -Stroms eine vielversprechende Vorverarbeitungsmethode dar, um den ML-Algorithmen eine Informationsgrundlage mit optimiertem Informationsgehalt bereitzustellen. Die resultierenden Features werden zusätzlich um Features aus Frequenzbereichsanalysen der Motorstromdaten und MEMS-

Vibrationsmessdaten angereichert, um für repräsentative Datenmengen die Prädiktionsgenauigkeit noch weiter zu verbessern.

Neben den Verschleißeffekten an Antriebsrädern entsteht bei RBGs mit der Zeit ein wachsendes Führungsspiel. Auch hierfür wurde mit der Berechnung der Wahrscheinlichkeitsdichte der horizontalen Beschleunigungsamplituden senkrecht zur Fahrtrichtung ein geeignetes Verfahren identifiziert. Mit diesem können Messdaten spezifisch für die Identifikation des Fehlers aufbereitet werden. Grundlage des Verfahrens sind Beschleunigungsdaten mit ausreichend hochfrequenter Abtastung, wie sie durch das entwickelte MEMS-Sensorboard zur Verfügung stehen. Zunächst wurde dieser Ansatz nicht weiterverfolgt, da eine Validierung durch Einstellung eines Führungsspiels am Demonstrator aufgrund der anders gearteten Lagerung nicht möglich war, ebenso wie eine Validierung durch Implementierung des Fehlerbildes im operativen Betrieb im MCC. Daher lag der Fokus im Projektverlauf auf der Prädiktion von Fehlern an den zuvor beschriebenen, drehzahlabhängigen Komponenten wie Motor, Getriebe und Antriebsrädern.

Grundsätzlich wurde auch der Demonstrator selbst als Quelle für Expertenwissen angesehen. So wächst durch die Simulation von Fehlern am Prüfstand das vorhandene Expertenwissen über die Charakteristiken, mit denen sich Fehler in Motor- und Beschleunigungsmessdaten äußern. Dieses Wissen wiederum konnte in die Algorithmen zur Datenvorverarbeitung und Merkmalsextraktion einfließen. Auch konnten am Prüfstand Fehlerdaten für potenzielle Defekte erzeugt werden, noch bevor diese erstmalig im realen Hochregallager auftreten. Somit stand unter reduzierten Kosten und Aufwand eine zusätzliche Wissensquelle in Form von Daten zur Verfügung. Neben der Erprobung von Algorithmen wurde hierdurch insbesondere das initiale Trainieren der Algorithmen auf einer künstlichen Datenbasis möglich. Damit wurde in Anlehnung an die Ausführungen zum Transfer Learning in Abschnitt 2.3.1 eine höhere Prädiktionsgüte von ML-Verfahren speziell auf geringen Datenmengen erwartbar. Tabelle 3 zeigt wiederum die Beteiligung der Partner an den Arbeiten.

**Tabelle 3: Zuordnung der Arbeiten für den Use Case Predictive Maintenance für AP 3**

Integration von Expertenwissen					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
Encoway/Lenze		X	X	X	X

#### **AP 4: Lernen auf Datenströmen**

Das fortlaufende Lernen auf Datenströmen erfordert eine beständig wachsende Datengrundlage. Da die reale Anwendung im Hochregallager MCC erst im Oktober 2022 in Betrieb genommen wurde und sich anfänglich noch kleine Probleme bei der Datenakquise zeigten, befand sich diese Datengrundlage länger als erwartet im Aufbau. Auch fehlte es aufgrund des geringen Alters der RBGs an Verschleißdaten. Hinzu kam die auch zu Beginn des letzten Projektdrittels noch immer schwierige Situation auf dem internationalen Halbleitermarkt.

Erste Daten aus dem MCC zeigten, dass die erhobenen Anforderungen sowie die Verfahren zur Berücksichtigung von Expertenwissen eine erfolgreiche Analyse begünstigen. So konnten für die Daten aus dem MCC unter anderem charakteristische Frequenzen des Getriebes in der spektralen Leistungsdichte der Motorstromdaten erkannt werden. Diese Frequenzen konnten fortlaufend bereits ohne Ausfalldaten durch einen unüberwachten ML-Algorithmus der encoway GmbH zur Anomalie-Detektion getrackt werden. Für die Validierung überwachter Lernverfahren standen die Prüfstanddaten zu variierenden Abplattungen an den Antriebsrädern zur Verfügung. Für die Klassifikation der Prüfstanddaten hinsichtlich des Fehlerbildes wurde ein einfaches neuronales Netz in Form eines MLPs als Feature-basierter Klassifikator implementiert und trainiert. Dabei wurden 10 % der Prüfstanddaten als Testdatensatz zurückgehalten. Basierend auf 5 Features des auf den Drehwinkel des Antriebsrades synchronisierten  $I_q$ -Stroms wurde das neuronale Netz auf 90 % der Daten trainiert wobei von diesen 90 % weitere 20 % als

Validierungsdatensatz verwendet wurden. Das optimierte neuronale Netz zeigte nicht nur auf den Trainings- und Validierungsdaten eine Klassifikationsgenauigkeit von 100 %, sondern auch auf dem unabhängigen Testdatensatz. Auch wenn es sich hierbei um einen Datensatz handelt, der unter Laborbedingungen aufgenommen wurde, zeigen die Ergebnisse sowohl den Wert des integrierten Expertenwissens als auch die Vorteilhaftigkeit der verwendeten hybriden Verschaltung für die erfolgreiche Anwendung von ML im Use Case Predictive Maintenance. Ohne Verwendung des Expertenwissens konnten auf Features unmittelbar aus den Motorstromdaten keine vergleichbar guten Ergebnisse zur Detektion bzw. Klassifikation von Abplattungen am Antriebsrad des Labor-RBG erzielt werden.

Die Architektur zur Einbeziehung von Expertenwissen wurde im Projektverlauf weiter optimiert. Zudem werden auf Daten des Demonstrators vortrainierten Algorithmen auf den echten Daten der Anwendung im Sinne eines Transfer Learnings weiter trainiert. Durch ein Mini-Batch-Training, wie in Abschnitt 2.3.1 beschrieben, erfolgt eine fortlaufende Anpassung der ML-Verfahren an aktuelle Daten auch auf der Zielhardware.

**Tabelle 4: Zuordnung der Arbeiten für den Use Case Predictive Maintenance für AP 4**

Auswahl und Bewertung von Lernverfahren					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
Encoway/Lenze		X	X	X	X
Implementierung der Datenaufnahme und -ablage Motorstromdaten					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM				X	X
Encoway/Lenze	X	X	X	X	X
Implementierung der Datenaufnahme und -ablage Vibrationsdaten					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X		X	X
Encoway/Lenze			X		
Trainieren von Modellen					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM		X		X	X
Encoway/Lenze	X	X	X	X	X

### **AP 5: Laufzeitoptimierung**

Im Use Case Predictive Maintenance stand für die Verarbeitung der Sensordaten ein Arakis Pico MK3 der Firma Welotec als Edge Device in den Schaltschränken der RBGs sowie am Laborprüfstand zur Verfügung. Dieser verfügt über ausreichende Rechenkapazitäten zur Verarbeitung der Daten mittels ML-Modellen sowie zum fortlaufenden Training der Modelle. Mittels der in Abschnitt 2.3.1 beschriebenen Tools und Verfahren zur Laufzeitoptimierung maschineller Lernverfahren über beispielsweise *TensorFlow Lite* als Porter-Tool konnte eine Optimierung der Algorithmen unter anderem durch Quantisierung, Pruning und Clustering erreicht werden. Diese Laufzeitoptimierungen gehen so weit, dass einfache Modelle wie ein CNN sogar direkt auf dem Microcontroller des MEMS-Sensorboards der Firma STM, also auf einem eingebetteten (Embedded) Device ausgeführt werden konnten.

Da die Laufzeitoptimierung, welche einen Trade-Off zwischen Prädiktionsgüte und benötigten Rechenkapazitäten darstellt, aufgrund der verfügbaren Rechenressourcen für die Anwendung der Lernverfahren im Use Case unkritisch war, wurde keine Laufzeitoptimierung umgesetzt. Dies liegt unter anderem daran, dass die Daten- und Rechenlast der ML-Verfahren durch die gezielte Integration von Expertenwissen u.a. über Features und domänenspezifische Transformationen im Vergleich zur herkömmlichen Anwendung von End-to-End Deep Learning Verfahren deutlich reduziert wurde. Auch wurde in enger Abstimmung mit encoway und der Lenze SE ein Verfahren integriert, durch das

nur Fahrten von RBGs erfasst werden, die eine definierte Charakteristik aufweisen und damit untereinander vergleichbar sind. Dadurch konnte die anfallende und zu verarbeitende Datenlast nochmals erheblich reduziert werden.

Es wurden abseits des MEMS-Sensorboard vier weitere Optionen für die Anomaliedetektion als Benchmark durchgeführt. Der Vergleich dieser ist in Tabelle 5 dargestellt

**Tabelle 5: Vergleich von Embedded-Devices zur optimierten Ausführung von ML-Modellen**

	<i>CPU</i>	<i>Architecture</i>	<i>RAM</i>	<i>Storage</i>	<i>Price (€)</i>
<b>ESP32</b>	240 MHz	32 Bit LX6	520 KB	4 MB	~2
<b>Raspberry Pi 3B</b>	1.2 GHz Quadcore	64 Bit ARM	1 GB	32 GB	~45
<b>Nvidia Jetson Nano</b>	1.43 GHz Quadcore	64 Bit ARM	4 GB	2 TB	~150
<b>Think Pad E14 Gen 4</b>	4.5 GHz Octacore	64 Bit x86	40 GB	-	~650

Erste Auswertungen mit verschiedenen Anomaliedetektionsverfahren ergaben verschiedene Ergebnisse der Performance, die in Tabelle 6 aufgezeigt sind.

**Tabelle 6: Ergebnisse der Laufzeit-Performance erprobter ML-Modelle auf verschiedenen Devices**

<i>MODEL</i>	<i>SIZE</i>	<i>ESP32</i>	<i>RASP-BERRY</i>	<i>JETSON NANO</i>	<i>THINK PAD</i>
<b>MAHALANOBIS DISTANCE .NPZ</b>	210 KB	154 ms	19 ms	12 ms	16 ms
<b>ISOLATION FOREST</b>	827 KB	-	419 ms	172 ms	26 ms
<b>AUTOENCODER .H5</b>	79 KB	-	16.31 s	10.5 s	2.74 s
<b>AUTOENCODER DEFAULT .TFLITE</b>	16 KB	61 ms	80 ms	34 ms	16 ms
<b>AUTOENCODER OPTIMIZED .TFLITE</b>	7 KB	-	82 ms	33 ms	16 ms
<b>LSTM AUTOENCODER .H5</b>	77 KB	-	-	6.69 s	1.44 s
<b>LSTM AUTOENCODER .TFLITE</b>	73 KB	-	-	142 ms	55 ms

Langfristig steht mit den in Abschnitt 2.3.1 beschriebenen Grundlagen der Laufzeitoptimierung das notwendige Handwerkszeug zur Verfügung, die ML-Verfahren nach Bedarf hinsichtlich ihrer Laufzeit zu optimieren. Dies ist insbesondere von Bedeutung, wenn im Anschluss an das Projekt im Rahmen der Verwertung die Menge an auszuführenden Modellen und Algorithmen sowie die Datenlast auf den Edge Devices sukzessive ansteigt.

**Tabelle 7: Zuordnung der Arbeiten für den Use Case Predictive Maintenance für AP 5**

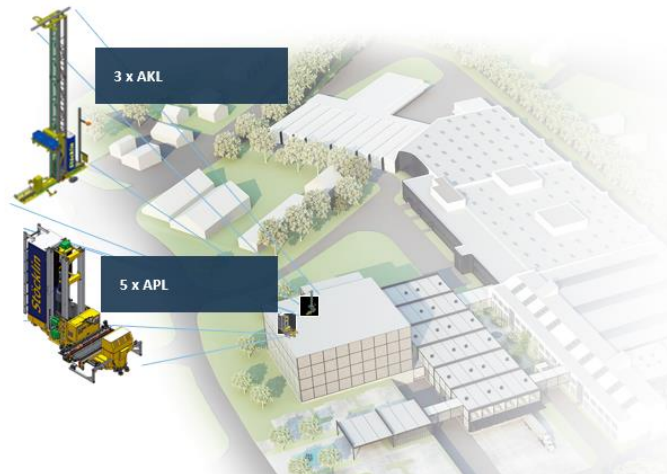
Entwicklung und Erprobung von Laufzeitoptimierten ML-Verfahren					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
Encoway/Lenze		X	X	X	X

### **AP 7: Use Case Predictive Maintenance**

Basierend auf den Anforderungen der Intralogistikbetreiber im MCC, die Abteilung MAR von Lenze, sollten im Rahmen von AP 7: Use Case Predictive Maintenance die Prädiktion von Verschleiß und Wartungsbedarfe in verschiedenen Teilen der automatischen Palettenlager und automatischen Kleinteillager basierend auf Motorstromanalysen und Vibrationsmessungen erfolgen. Die Schadensbilder Riemenspannung im Hub- und Fahrwerk, Reibradverschleiß, Führungsradverschleiß sowie Lager-, Verzahnungs-, Bremsen- und Resolverfehler wurden aufgrund der Ausfallrelevanz für die vorrauschauende Überwachung definiert. Im Ziel sollten durch die Datenaufnahme und intelligente Überwachung Transparenz und Verfügbarkeit der Anlage gesteigert sowie Aufwände in der Wartung und durch geplante und ungeplante Stillstände minimiert werden.

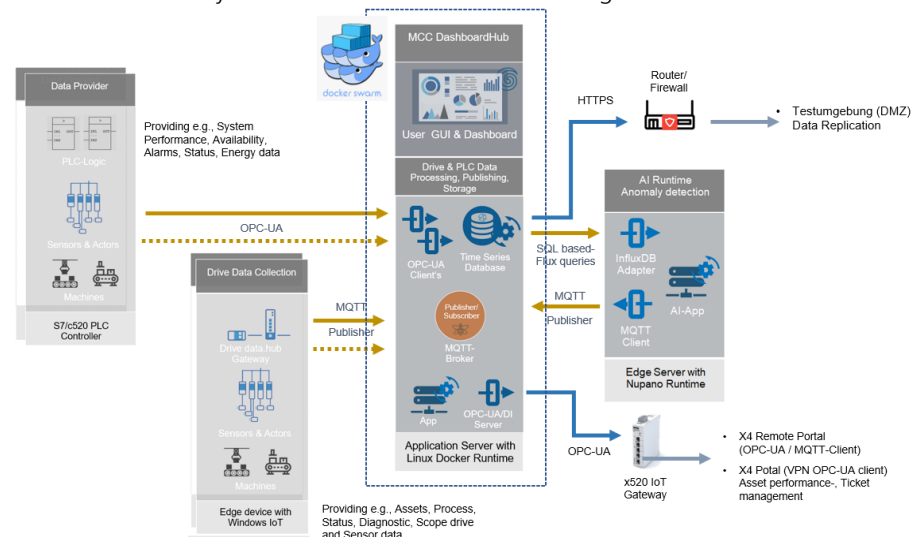
Im Verlauf des Forschungsprojektes wurden diverse Hochregallager besucht, da die Inbetriebnahme des Hochregallagers im Mechatronics Competence Campus (MCC) (Siehe Abb. 23), Extertal erst im Herbst 2022 erfolgte. Bis dahin wurden beschädigte Teile aus

dem Lenze Hochregallager in Aerzen sowie Hochregallagerdefekte von verschiedenen Standorten dritter Parteien gesichtet.



**Abb. 23: Darstellung MCC und Regalbediengeräte**

Für die Datenaufnahme wurde basierend auf den Ergebnissen von AP 1, 2, 3 und 5 eine Docker-Service-basierte Datenaufnahme- und Analyseinfrastruktur konzeptioniert und implementiert. Die Datenaufnahmearchitektur erfasst via OPCUA sowohl Prozess-Daten aus Lenze- als auch von Siemenssteuerungen. Die Motorstromdaten werden mit einer Auflösung von bis zu 16.000 Hz via Oszilloskop im Umrichter oder über Vibrations-sensorik (siehe MEMS-Sensorboard) mittels der Lenze Software „Drives DataHub“ zu bestimmten Zeitpunkten aufgezeichnet. Dabei werden nur gut vergleichbare Bewegungen der Regalbediengeräte erfasst. Via MQTT werden die Daten anschließend von Edge Devices, die im Regalbediengerät verbaut sind, verschiedenen dockerisierten Anwendungen bereitgestellt. Im MCC DashboardHub werden Motorstrom und Prozessdaten visualisiert und verschiedenen Auswertungen unterzogen. Die Daten werden in einer InfluxDB Time-series-Datenbank und einer relationalen Posgres-Datenbank gespeichert, um Anforderungen in der performanten Darstellung und Berechnung zu bedienen. Mit Hilfe der Lenze Softwareverwaltung „NUPANO“ wurden im Forschungsprojekt durch das IEM sowie durch die encoway GmbH entwickelte Überwachungsmodelle im MCC betrieben.

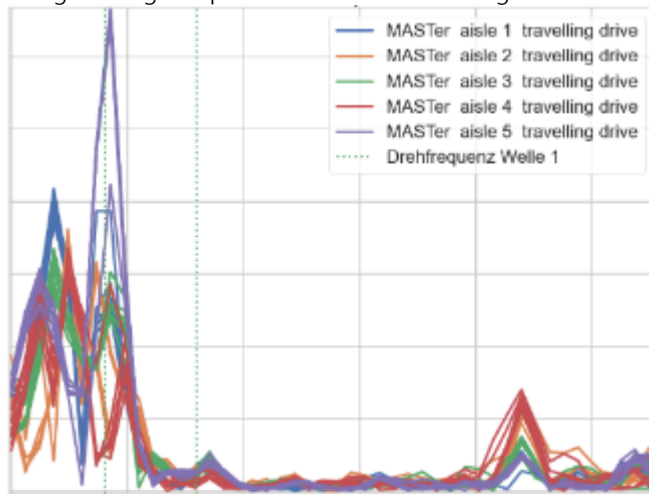


**Abb. 24: Prototypische Dateninfrastruktur für den Use Case Predictive Maintenance**

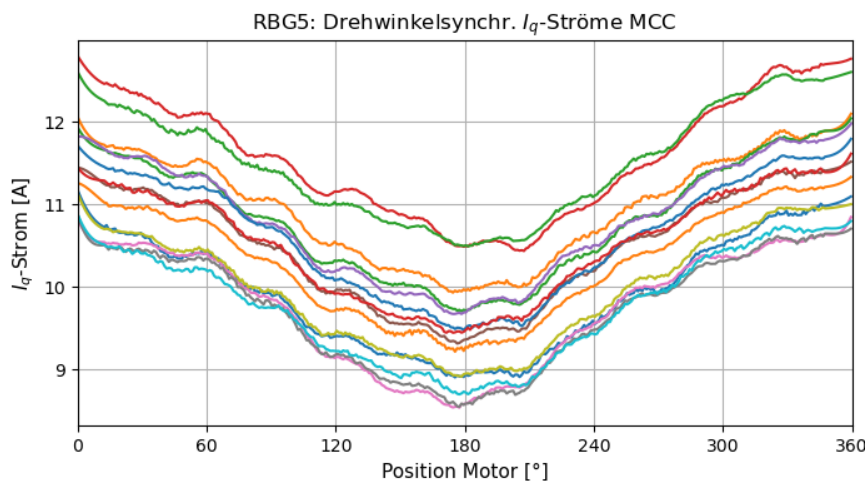
Mit der Eröffnung des MCCs im Oktober 2022 konnten auch erste Daten erhoben und das MEMS-Sensorboard installiert werden, um Referenzdaten neben den Motorstromdaten zu erheben. Erste Aufnahmen und Auswertungen mittels in AP 2 und 3 definierten

Verfahren zeigten höhere Schwingungen aufgrund einer leichten Exzentrizität im Abtrieb eines Regalbediengeräts (Siehe Abb. 25 und Abb. 26). Weitere Prüfungen und eine genauere Untersuchung des Trends ergaben, dass zwar eine Exzentrizität der Welle vorliegt, diese jedoch das System auf absehbare nicht negativ beeinflusst. Dies wurde durch eine intensive und regelmäßige Inspektion des Geräts sichergestellt.

-----  
Eingehende Darstellung  
-----

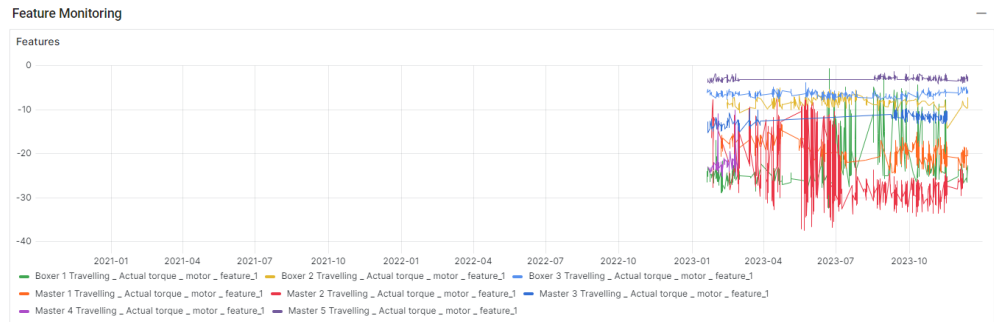


**Abb. 25: Exemplarische Frequenzverläufe der Paletten-Regalbediengeräte im MCC**



**Abb. 26: Charakteristische Signalverläufe für eine Motor-Exzentrizität anhand eines Experten-basierten Verfahrens zur Merkmalsextraktion**

Abseits der definierten Fehlerbilder kam es während der Inbetriebnahme des MCC-Hochregallagers zu Optimierungsprozessen, die eine sukzessive die Datenaufnahme an verschiedenen Regalbediengeräten einschränkten, da zu diesem Zeitpunkt nur eine dauerhafte Datenaufnahmestrategie durchgeführt werden konnte. Daraus resultierend konnten über längere Zeiten von vielen Regalbediengeräten keine Daten aufgenommen werden. Während das Projektteam in dieser Zeit vermehrt Daten mit dem RGB-Demonstrator aufnehmen und validieren konnte, war die Möglichkeit realen Verschleiß zu analysieren, durch die Neuheit des Lagers sowie die eingeschränkte Datenaufnahme erschwert. Die umfassende Aufzeichnung der Daten konnte erst im Herbst 2023 wieder aufgenommen werden und war somit für knapp ein halbes Jahr durch andere Tests verhindert.



**Abb. 27: Datenaufnahme in Realumgebung (MCC) während der Projektlaufzeit**

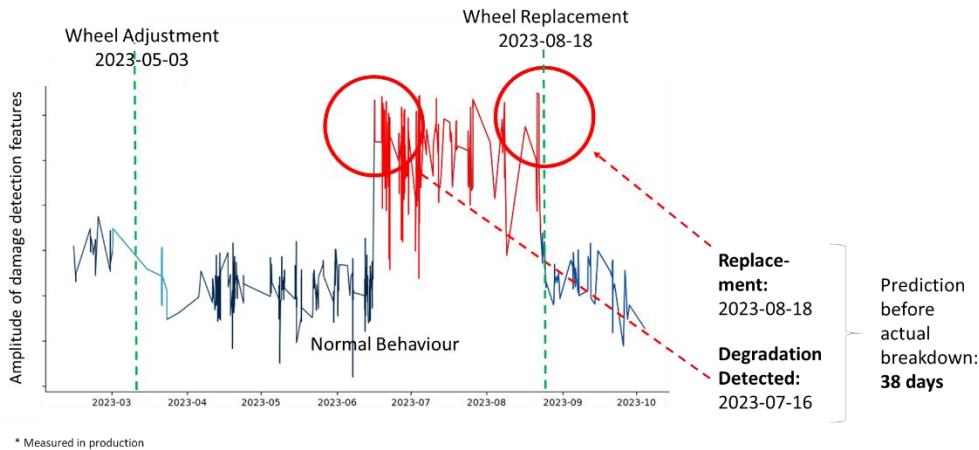
Begünstigt durch eine kostenneutrale Projektverlängerung um drei Monate, konnten weitere Daten erhoben und im Rahmen des Projekts ausgewertet werden. In dieser Zeit konnten zwei reale Schadenfälle beobachtet, analysiert und für das Training von Überwachungsmodellen genutzt werden.

Bei dem ersten Schadensfall handelte es sich um eine Abplattung eines Führungsrads an einem automatischen Kleinteillager (s. Abb. 27 und Abb. 28). Das Führungsrad stabilisiert das RBG während der Fahrt durch das passive Mitlaufen an der Schiene. Das Kleinteillager verfügt über je zwei Vulkollanräder vor und hinter dem Fahrtrieb sowie Führungsrädern an der oberen Führungsschiene. Die Führungsräder am Fahrwerk haben einen Durchmesser von 11 cm und sind bei einer Abplattung von einem Millimeter defekt.



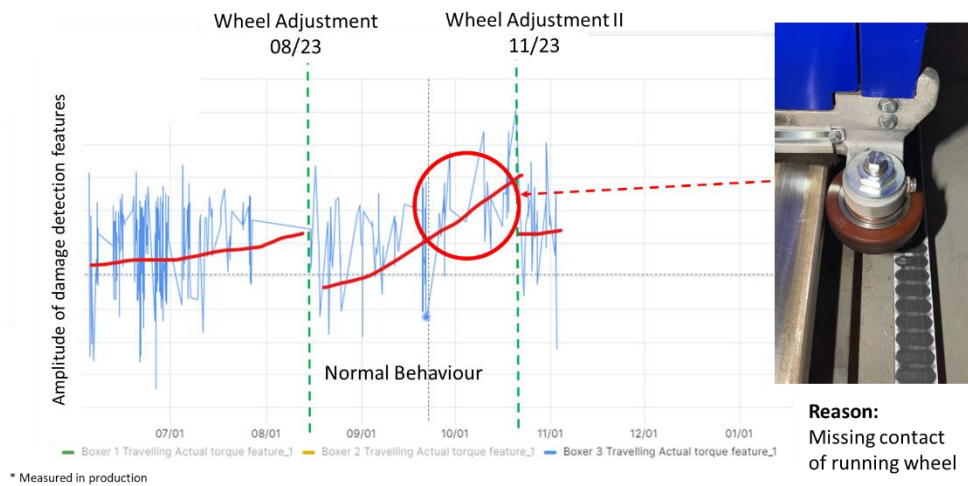
**Abb. 28: Führungsrاد am Kleinteile-RBG: Aufnahme aus Realumgebung (links) und CAD (rechts)**

Eine Analyse der relevanten Frequenzen im Motorstrom des Fahrwerks zeigt, dass sich nicht nur im Labor (RBG-Demonstrator), sondern auch im produktiven Kontext ein Verschleiß eindrücklich in den Daten zeigt. Wie in Abb. 29 erkennbar, lassen sich nicht nur der Schadensfall (ab 16. Juli), sondern auch Wartungsarbeiten (5. März) eindeutig durch Expertenwissen nutzende Verfahren ermitteln. Hierdurch gewinnt die Überwachung an Interpretier- und Übertragbarkeit, da nun durch performante, aber normalerweise nicht fehlerbeschreibende, ML-Verfahren, eine Aussage zum Zustand und der beschädigten Komponente getätigt werden kann. Außerdem erlaubt eine Auswertung des Trends nach der augenscheinlichen Behebung des Fehlers (18. August 2023), die Validierung, ob die Maßnahme den ausschlaggebenden Fehler behoben hat. Durch das angewendete Verfahren hätte der Fehler 38 Tage vor dem ungeplanten Austausch der Führungsrolle identifiziert und behoben werden können. Die Überwachung von Komponenten, die nicht direkt vom Motor betrieben werden, hat sich so als mit Motorstromdaten aus dem Umrichter als machbar herausgestellt.



**Abb. 29: Schadensdetektion und prädiktive Wartung anhand Expertenwissen-basierter Merkmale**

In einem weiteren Fehlerfall an einem anderen Kleinteillagergerät konnte bereits das Lösen des Führungsrad und, dadurch bedingt, ein fehlender Kontakt mit einem Teil der Schiene identifiziert werden (siehe Abb. 30).



**Abb. 30: Frühzeitige Erkennung von einem gelösten Führungsrad anhand von Motorstromdaten**

Basierend auf diese Fehlerfällen wurden, unterstützt durch synthetische Fehlerdaten, ML-Verfahren und Neuronale Netze trainiert, welche eine weitere, automatisierte Ermittlung von Verschleiß sicherstellen. Die Integration der Algorithmen erfolgte fortwährend im MCC. Durch das erfolgreiche Erkennen der Anomalien am Demonstrator RBG (TRL 4) und im MCC konnte im Rahmen des Projektes der Technologie Readiness Level 7 demonstriert werden. Während dies ein großer Erfolg für das Projektteam ist, gilt es nun weitere Fehlerfälle im MCC und in verschiedene Applikation bei Kunden zu validieren, um eine Marktfähigkeit, bzw. Übertragbarkeit sicherzustellen bevor eine Produktifizierung der Ergebnisse erfolgt.

**Tabelle 8: Zuordnung der Arbeiten für den Use Case Predictive Maintenance für AP 7**

Konzeptionierung der IoT-Architektur					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM				X	X
Encoway/Lenze	X	X	X	X	X
Implementierung der IoT-Architektur					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM					
Encoway/Lenze	X	X	X	X	X
Erfassung und Auswertung der Schadensbilder					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM		X		X	X
Encoway/Lenze	X	X	X		
Implementierung von physikalischen Modellen zur Daten Überwachung					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM		X		X	X
Encoway/Lenze	X	X	X	X	X

### 2.3.3 Use Case Predictive Quality

Eingehende Darstellung

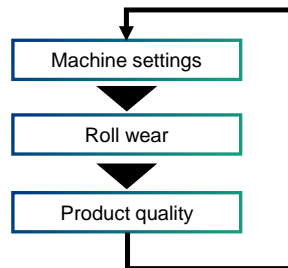
Nachfolgend werden die wissenschaftlich-technischen Ergebnisse der Arbeitspakete für den Use Case Predictive Quality sowie die spezifischen Beiträge des Konsortialpartners Fraunhofer IEM dargestellt. Bereits im vorhergehenden Berichtszeitraum wurden die Ergebnisse der Arbeitspakete AP 1 und AP 2 dargelegt und die zugehörigen Meilensteine abgeschlossen.

#### **AP 1: Anforderungsanalyse**

Die Anforderungen an eine Predictive Quality Anwendung bei den Kunden der düspohl Maschinenbau GmbH wurden gemeinsam erarbeitet. Dabei wurden im Wesentlichen die folgenden Aspekte festgehalten:

- Eine Predictive Quality Anwendung kann und muss ohne zusätzliche Messtechnik im Prozess auskommen, da Qualitätsfehler nicht durch konventionelle Messtechnik erfasst und neue Messtechnik im Projektrahmen nicht entwickelt werden kann.
- Qualitätsverluste werden wesentlich durch das Zusammenspiel von Verschleißerscheinungen auf den Ummantelungswerkzeugen (Rollen) und den Maschineneinstellungen (Auswahl und Positionierung der Ummantelungsrollen) bedingt. Daher wurde die Anforderung formuliert, dass die Qualität durch eine Beobachtung und Anpassung dieser Bereiche überwacht und optimiert wird.
- Es sollen alle Berechnungen zur Prädiktion des Verschleißes vor Ort im Edge-Device der RoboWrap durchgeführt werden. Es sollen keine zusätzlichen Cloud-Dienste zur Berechnung und nicht einmal für das Training von Datenmodellen notwendig sein.

Zur Lösung der Aufgabe wurde ein Vorgehen nach dem folgenden Schema verfolgt:



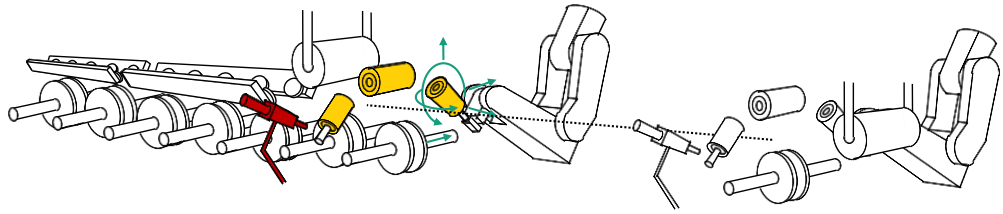
**Abb. 31: Vorgehen zur Predictive Quality Anwendung der RoboWrap von düspohl**

Der erste Schritt umfasst die Prädiktion von Rollenverschleiß in Abhängigkeit der Maschineneinstellungen. Zur Lösung dieses Unterproblems wurden folgende Schritte und Ergebnisse verfolgt und erreicht:

#### **Formalisierung der Maschineneinstellungen**

##### **1. Beschreibung des Aktionsraums der RoboWrap**

Um Maschineneinstellungen als Grundlage zur Beschreibung der Rollennutzung zu verwenden, mussten diese zunächst formalisiert und als Aktionsraum begriffen werden. Dazu wurde dieser Raum mathematisch beschrieben und die Dimensionalität bestimmt. Dadurch ist es möglich, die Größe des Eingangsraums für die Rollenverschleißprädiktion realistisch einzuschätzen. Zur Illustration dient die folgende Abbildung:



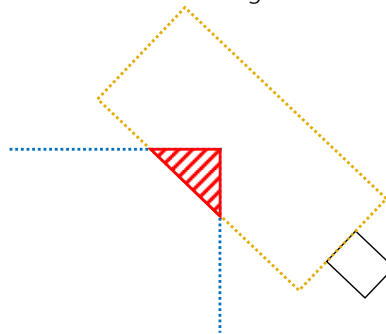
**Abb. 32: Illustration der Bestimmung des Aktionsraums der RoboWrap**

Die mathematische Definition des Aktionsraums wird aufgrund einer zu hohen technischen Tiefe an dieser Stelle nicht dargestellt. Stattdessen sind nachfolgend zusammenfassende Ergebnisse aufgelistet:

- Der Aktionsraum ist sehr groß und bewegt sich im Bereich  $10^{100}$ . Das bedeutet, dass äußerst viele unterschiedliche Einstellungen der RoboWrap möglich sind und auch der Rollenverschleiß auf unterschiedlichste Weise beeinflusst werden kann.
- Aus diesem Grund wird vereinfachend angenommen, dass die auf eine Rolle wirkende Kraft bzw. der auf Rolleninkremente wirkende Druck nur durch die relative Positionierung der Rolle zum Profil bestimmt wird.
- Dieser Druck kann durch geometrische Kalkulationen mithilfe von Materialkennwerten bestimmt werden (s.u.). Durch die Reduktion der Einflüsse auf den Rollenverschleiß auf den Druck verringert sich der Aktionsraum für Verschleiß pro Rolle auf eine Dimensionalität von ca.  $10^3$ .

## 2. Druckberechnung aus Datensätzen aus der Produktion

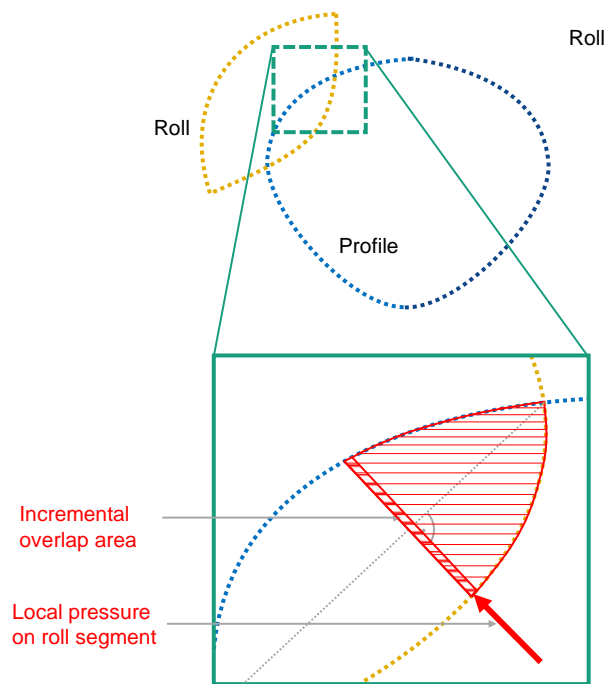
Zur Druckberechnung liegen lediglich Punktwolken im zweidimensionalen Raum vor, die die Geometrie der Rollen und das Profils repräsentieren. Diese geometrischen Informationen können kombiniert mit Materialkennwerten jedoch genutzt werden, um vereinfacht die wirkenden Drücke zwischen Rolle und Profil bezogen auf einzelne Rolleninkremente zu bestimmen. Dazu dient die Bestimmung des Schnittbereichs als Ausgangsbasis.



**Abb. 33: Schnittbereich zwischen 2D-Querschnittspunktwolken zur Druckbestimmung**

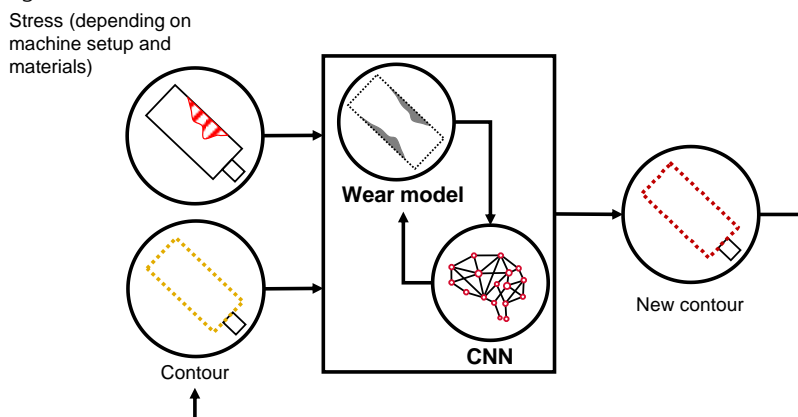
Um den Schnittbereich zwischen Rollen und Profil zu bestimmen, wurden folgende Schritte durchgeführt:

- Bestimmung der Schnittpunkte zwischen Rollen und dem Profil
- Bildung von Schnittpolygonen zur Flächenberechnung der Überlappung
- Umrechnung der Gesamtflächen in inkrementelle Abschnitte auf einem einzelnen Rollensegment
- Multiplikation der inkrementellen Querschnittsflächen mit Materialkonstanten zur Berechnung der resultierenden lokalen Druckkräfte



**Abb. 34: Detailvisualisierung der Druckberechnung**

Nach der Druckberechnung konnte der Inkrementaldruck gemeinsam mit der Rollenkontur vor Einsatz genutzt werden, um als Input für ein Modul zur Vorhersage der Konturänderung bzw. der neuen Kontur zu dienen:



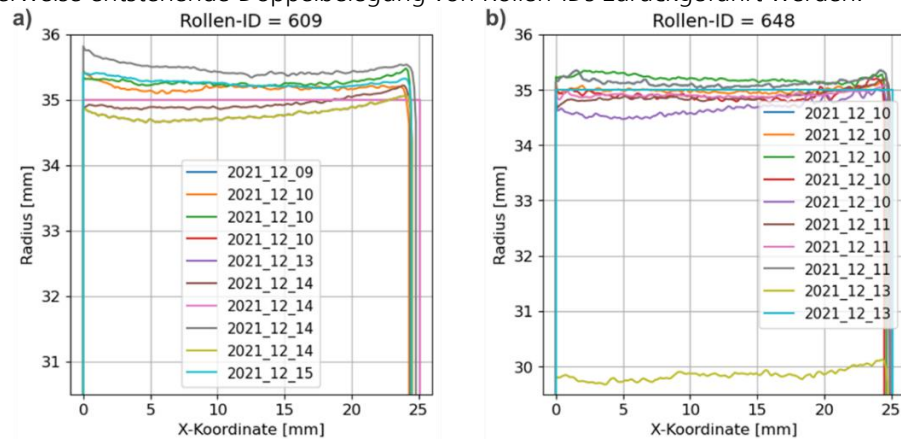
**Abb. 35: Hybrider Ansatz zur Prädiktion des Verschleißes von Ummantelungsrollen**

Das Vorhersagemodul ist hybrid und besteht aus einem daten- und einem physikbasierten Teilmodul. Das Datenmodell besteht im Wesentlichen aus einem Convolutional Neural Network, das die Druck- und Konturinformationen kombiniert verarbeitet und mit einer Encoder-Decoder-Struktur die nächste Kontur der Rolle bei gegebener Nutzung prädiziert. Wie genau die Struktur des Vorhersagemoduls gestaltet ist, wird im nächsten Abschnitt behandelt.

Die Anforderungsanalyse wurde zunächst dokumentiert. Wie bereits in der Vorhabensbeschreibung angedeutet, entstanden im weiteren Projektverlauf neue Erkenntnisse und Anforderungen für den Use Case Predictive Quality. Diese betreffen insbesondere die Datenakquise und sind nachfolgend beschrieben.

Ausgehend von den definierten Anforderungen wurde eine erste Datenbasis aufgezeichnet, die eine fortlaufende zeitliche Veränderung von Ummantelungsrollen durch Abnutzung erfassen kann. In ersten Analysen konnten aber Probleme in der Datenbasis identifiziert werden (vgl. Abb. 36). So traten in den Daten gelegentlich Punktwolken von Rollen

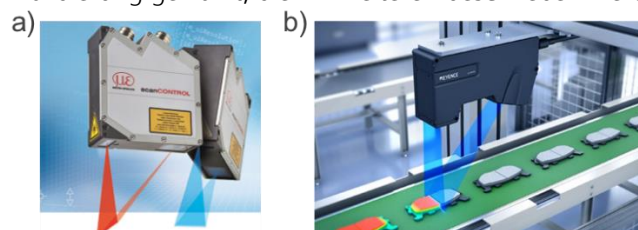
auf, die eine exakt rechteckige Geometrie aufweisen. Diese Geometrie war aufgrund von Fertigungstoleranzen und Messrauschen beim Scan ausgeschlossen und resultierte daraus, dass für die interne Verrechnung in der Steuereinheit der *RoboWrap* in bestimmten Situationen auf die idealisierten CAD-Daten der Ummantelungsrollen zurückgegriffen wurde. Weiterhin traten gelegentlich auch Scan-Daten für eine spezifische ID einer Rolle auf, die einen gänzlich anderen Rollenradius aufweisen. Dies konnte u.a. auf eine fälschlicherweise entstehende Doppelbelegung von Rollen-IDs zurückgeführt werden.



**Abb. 36: Widersprüchliche Scan-Daten von anhand von Barcodes getrackten Ummantelungsrollen**

Die für eine Anwendung von maschinellen Lernverfahren kritischen Widersprüchlichkeiten in den Daten konnten im engen Austausch mit dem Konsortialpartner düspohl für neu entstehende Daten bereinigt werden. So wurde die Datenbasis um die Rohdaten der Scans erweitert, um diese unabhängig von der internen Verarbeitung als Datenbasis für das maschinelle Lernen verwenden zu können. Zudem wurden Metainformationen über die Vorschubgeschwindigkeit der Profile beim Ummanteln sowie die zurückgelegte Strecke in Metern ergänzt.

Eine weitere iterative Änderung ergab sich hinsichtlich der verwendeten Messhardware. Aus diversen Gründen werden die in der Profillummantelungsanlage verwendete Scanner der Firma Micro-Epsilon zukünftig durch präzisere und robustere Scanner der Firma Keyence ersetzt (vgl. Abb. 37). Auch die Vorverarbeitung der Messdaten für die Analyse mittels ML-Verfahren wird in diesem Zuge verbessert. Die bisher aufgezeichnete Datenmenge war damit nicht mehr für das Training eines unmittelbar im Betrieb nutzbaren ML-Modells geeignet, da sich im Sinne der Darstellungen aus Abschnitt 2.3.1 ein spontaner, virtueller Concept Drift ergibt. Die Untersuchungen im Rahmen des Forschungsprojektes wurden dadurch allerdings nicht eingeschränkt. Zudem wurden im Rahmen von AP 4 Möglichkeiten erarbeitet, wie mit spontanen, prozessbedingten Änderungen in der Datenbasis umgegangen werden kann. Da die Identifikation der Anforderung an die Verbesserung der Qualität der Scan-Messdaten iterativ über den Projektverlauf erfolgte, und der Scanner-Austausch mit langen und aufwendigen Adaptionen an der *RoboWrap* verbunden ist, konnten entsprechende Daten im Projektverlauf nicht mehr über einen hinsichtlich Rollenverschleiß auswertbaren Zeitraum erfolgen. Daher wurden alternative Wege zur Validierung gewählt, die im Weiteren beschrieben werden.



**Abb. 37: Scanner zur Messung der Rollenoberfläche: a) bisheriger Scanner scanCONTROL 25xx (Micro Epsilon), b) neuer Scanner LJ-X8400 (Keyence)**

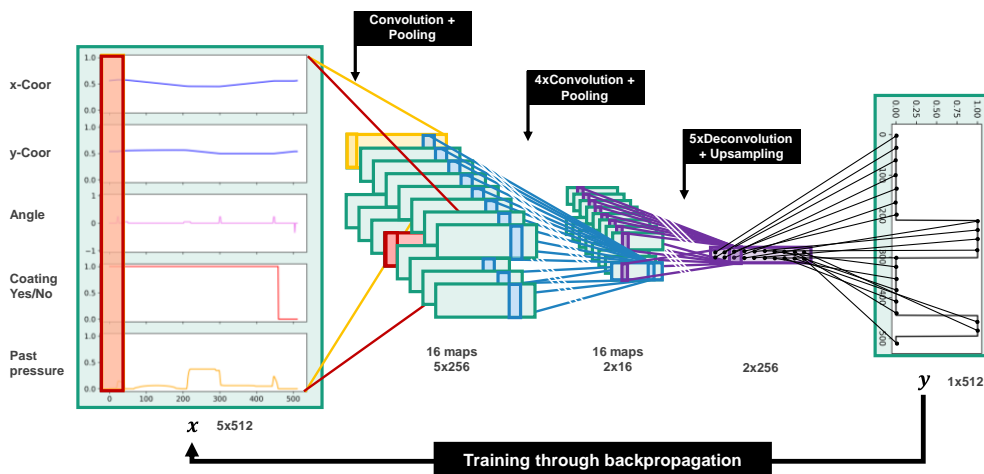
**Tabelle 9: Zuordnung der Arbeiten für den Use Case Predictive Quality für AP 1**

-----  
Eingehende Darstellung  
-----

Implementierung der Datenaufnahme und Datenablage					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM				X	X
düspohl	X	X	X		
Anforderungsanalysen anhand von Daten und Prozess					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl			X	X	X
Formalisierung der Maschineneinstellungen					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl				X	X

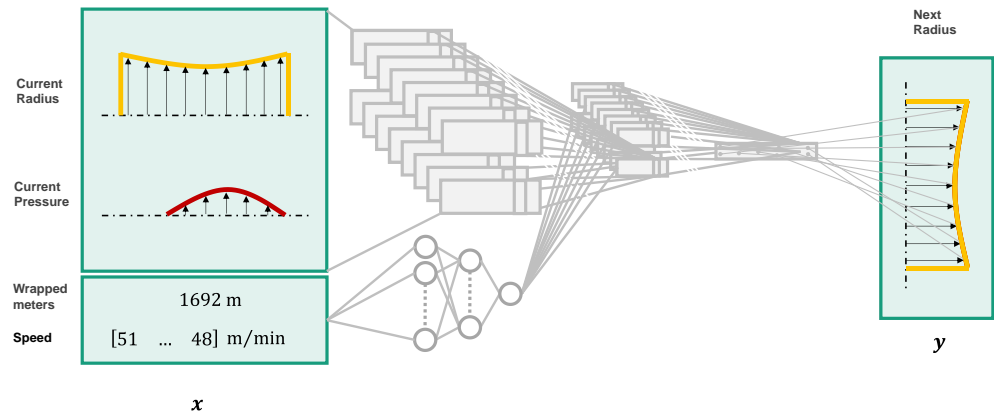
**AP 2: Analyse und Auswahl von Lernverfahren**

Nachdem die Anforderungen für den Use Case Predictive Quality im ersten Arbeitspaket festgelegt wurden, wurde in AP 2 potenzielle Lernverfahren untersucht. Wie im vorherigen Arbeitspaket bereits angeschnitten, wurde ein hybrider Ansatz mit einem Convolutional Neural Network (im Folgenden: CNN) als Ergebnis der Recherche ausgewählt. Als Basis diente ein vorhandenes Systemkonzept aus einem vergangenen Projekt, das nicht Rollengeometrien, sondern einzustellende Drücke am Profil vorhersagen sollte. Die Architektur dieses vergangenen Projektes konnte gute Ergebnisse erzielen und stellte hier den Stand der Technik dar. Die Architektur des Ansatzes ist im Folgenden skizziert:



**Abb. 38: Erfolgreicher Encoder-Decoder-CNN-Ansatz aus vorherigem Projekt**

Die aufgezeigte Architektur verarbeitet eine 5x512 dimensionierte Matrix, sodass 5xn dimensionale Filter des CNN genutzt werden können, um kontextuiert Daten zu verarbeiten und keine zusammenhängenden Informationen unnötigerweise zu trennen. Dieses Vorgehen wurde im Projekt EASY adaptiert und an die gegebenen Inputdaten angepasst. Somit können der Rollenradius (Kontur) und der Druck auf die Rolle gemeinsam verarbeitet werden. Zusätzlich ist es entscheidend, die Anzahl der gefertigten Meter an Produkt und auch ein zugehöriges Geschwindigkeitsprofil zu berücksichtigen, da diese einen wesentlichen Effekt auf die Auswirkung des Druckes auf die Rolle haben. Die adaptierte Version des CNN ist nachfolgend dargestellt:



**Abb. 39: Encoder-Decoder-CNN-Ansatz mit verkettetem MLP in EASY**

Wie in der Darstellung erkennbar, sind die Zusatzinformationen der ummantelten Meter und der Geschwindigkeit durch ein Multi-Layer-Perzeptron (MLP) an das CNN gekettet. Diese „Concatenation“ ermöglicht ein simultanes Training beider Netze, obwohl der Input stark heterogen ist. Der Ansatz vereint die Vorteile eines lokal kombiniert verarbeitenden CNN mit einem Zusatz-Informationen oder Feature-verarbeitenden MLP. Im Gegensatz zum vorherigen Projekt wurde in EASY die nächste Rollenkontur in Form des Radius der Rolle vorhergesagt.

Wie zuvor beschrieben, wurde das Prädiktionsmodul hybrid gestaltet, d.h. ein wesentlicher Bestandteil ist die Verwendung von physikalischem Vorwissen. Dazu folgen im nächsten Abschnitt weitere Erläuterungen.

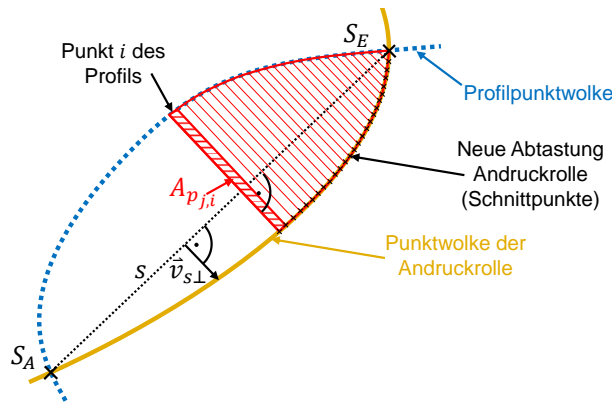
**Tabelle 10: Zuordnung der Arbeiten für den Use Case Predictive Quality für AP 2**

Analyse und Auswahl von ML-Verfahren					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl		X		X	X

### **AP 3: Integration von Expertenwissen**

Expertenwissen kann die zum Erlernen eine Aufgabe notwendige Realdatenmenge bedeutend reduzieren. Dazu wurde im Use Case Predictive Quality in EASY auf Abtrag- bzw. Verschleißmodelle zurückgegriffen. Diese wurden herangezogen, um in Abhängigkeit der zu erwartenden Drücke und Radii in Verbindung mit den ummantelten Metern und dem Geschwindigkeitsprofil modellbasiert die Konturänderung zu berechnen. Weiterhin wurde auch für die Vorverarbeitung der Daten physikalisches Wissen eingebracht, um die Daten sowie ihren Informationsgehalt zielgerichtet für die Vorhersage von Rollenverschleiß als Grundlage für eine prädiktive Qualitätsbewertung aufzubereiten. Diese Herangehensweise ermöglicht, dem neuronalen Netz eine besser vorverarbeitete Informationsgrundlage mit Vorwissen zur Verfügung zu stellen und reduziert den Realdatenaufwand erheblich.

Eine erste Form der Integration von Expertenwissen ist durch die gezielte Vorverarbeitung der Maschinendaten unter Berücksichtigung von Domänenwissen gegeben. Entgegen der üblichen Herangehensweise bei der Anwendung maschineller Lernverfahren, direkt auf Daten oder statistischen Merkmalen zu präzisieren, wurde für den vorhandenen Use Case eine Vorverarbeitung entwickelt, die physikalische Hintergründe des Prozesses abbildet. Um die Kontaktsituation zwischen abnutzenden Ummantelungsrollen und Profilen in komprimierter Form und unter Beibehaltung des maximalen Informationsgehalts darzustellen, erfolgt eine Berechnung des Drucks für die aus den Daten hervorgehenden Überschneidungen von Profil- und Rollenpunktwolken. Dieses Vorgehen ist in Abb. 40 ersichtlich.



**Abb. 40: Druckberechnung aus der geometrischen Überschneidung von Rolle und Profil**

Für die Umrechnung der Information in Druckäquivalente entlang der Rollenoberfläche wurde Formel (1) in Anlehnung an die Theorie der elastischen Fundamente entwickelt [14]. Diese beschreibt den punktwisen Druck  $\sigma_i$  in der Kontaktzone  $j$  von Rolle und Profil basierend auf den am betrachteten Punkt anliegenden Flächen des Schnittbereichs inkrementell. Dieser Druck stellt eine maßgebliche Größe für die aus dem Prozess resultierende Abnutzung der Rollen dar. Dabei werden die vergleichsweise harten Profile als starr angenommen, wohingegen die Materialhärte  $k_h$  der Rollen in die Berechnung einfließt. Somit werden auch die Materialeigenschaft an dieser Stelle in den Daten für die Anwendung von Lernverfahren berücksichtigt.

$$\sigma_i = k_h \cdot \sqrt{\frac{A_{p_{j,i-1}} + A_{p_{j,i}}}{2}} \quad (1)$$

Aufbauend auf der Anforderungsanalyse aus AP 1 existiert der in Abb. 35 dargestellte Lösungsansatz für ein hybrides Vorhersagemodul mit einem datenbasierten und einem physikalischen Teilmodul. Als physikalisches Teilmodul soll dabei ein Verschleißmodell dienen, das den nutzungsabhängigen Rollenverschleiß auf Basis physikalischer Zusammenhänge beschreibt.

Im Rahmen der Integration von Expertenwissen in AP 3 wurde das Verschleißmodell in Anlehnung an existierende Literatur ([15], [16]) und in Zusammenarbeit mit Domänen-Experten der Firma düspohl entwickelt. Bei der Abnutzung der Ummantelungsrollen wirken verschiedene Effekte. Mit einem höheren Druck  $\sigma(x)$  geht auch ein höherer Verschleiß an der lokalen Stelle  $x$  einher. Die Auswirkung des Rollenradius  $R(x)$  auf den Verschleiß wird hingegen invers quadratisch angenommen, da mit abnehmendem Material die Fähigkeit der Rollen zur elastischen Verformung abnimmt und damit auch die Materialabnutzung. Die resultierende Formel ist in Gleichung (2) dargestellt. Der Faktor  $k$  stellt einen Skalierungsfaktor dar, der den Rollenabtrag  $\Delta r(x)$  für die Nutzung im datenbasierten Teilmodul anhand realer Messwerte auf den Wertebereich  $[0,1]$  skaliert.

$$\Delta r(x) = k \cdot \frac{\sigma(x)}{R(x)^2} \quad (2)$$

Weitere Einflussfaktoren auf den Verschleiß stellen die Geschwindigkeit und die bei der Ummantelung zurückgelegte Strecke dar. Da die Akquise dieser Informationen bei der Bearbeitung des APs aufgrund der im Rahmen von AP1 genannten Gründe noch nicht für eine repräsentative Datenmenge erfolgte, wurden sie im aktuellen Stadium näherungsweise als konstant je Ummantelungsvorgang angenommen. Für die spätere Nutzung bei Aufzeichnung dieser Messdaten kann das Verschleißmodell entsprechend erweitert werden, wie in Gleichung (3) ersichtlich. Dabei ist auch unter Berücksichtigung des formalen Expertenwissens der Domänen-Experten anzunehmen, dass der Verschleiß

linear mit der zurückgelegten Ummantelungsstrecke  $d$  einhergeht, aber exponentiell mit zunehmender Ummantelungsgeschwindigkeit  $v$  ansteigt.

$$\Delta r(x) = k_1 \cdot \frac{\sigma(x)}{R(x)^2} \cdot \frac{d}{1000} * e^{k_2 \cdot v} \quad (3)$$

Eine weitere Möglichkeit, Expertenwissen für den Use Case nutzbar zu machen, bestand in der Generierung synthetischer Daten. Dies brachte im Use Case zwei entscheidende Vorteile. Zum einen entstand eine beliebig erweiterbare Menge an synthetischen Daten, anhand derer die Algorithmen aus AP 2 erprobt werden konnten. Weiterhin beinhalten diese Daten bereits implizites Wissen über den Verschleiß von Ummantelungsrollen ähnlich den Zusammenhängen, die in den realen Maschinendaten zu erwarten sind. Somit konnte die synthetische Datenbasis zusätzlich im Sinne eines Transfer Learning Ansatzes genutzt werden, um das CNN vor der Übertragung auf die realen Daten vorab zu trainieren und so das initiale Wissen im datenbasierten Teilmodul zu erhöhen.

Für diesen Ansatz wurde das bereits beschriebene Verschleißmodell verwendet, um anhand von verschiedenen, zufällig initialisierten, aber für den Prozess typischen Kontaktsituationen eine Abnutzung von Ummantelungsrollen zu simulieren. Die einzelnen Teilfunktionen zur Generierung der synthetischen Daten sind in Abb. 41 ersichtlich und werden nachfolgend kurz erläutert.

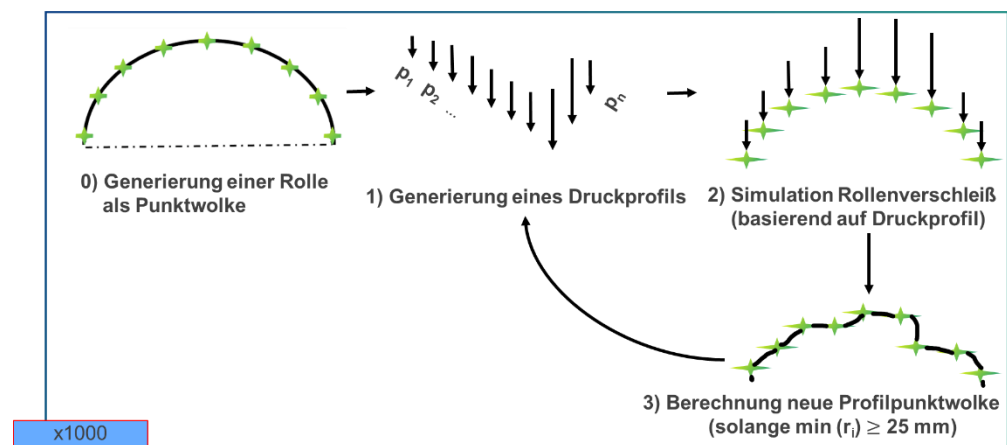


Abb. 41: Vorgehen zur Generierung synthetische Ummantelungsrollen-Verschleißdaten

- 0) Generierung von Rollen: Initial werden verschiedene Rollen generiert. Diese weisen typische Geometrien (linear, kegelförmig, parabelförmig, etc.) auf. Die Ausprägung dieser Geometrien, beispielsweise der initiale Radius und Steigungskoeffizienten der Konturen werden innerhalb vorgegebener Grenzen nach dem Zufallsprinzip gewählt.
- 1) Generierung von Druckprofilen: Die Generierung von Druckprofilen orientiert sich an realistischen Kontaktsituationen zwischen Rollen und Profilen. So entsprechen dreiecksförmige Druckverläufe beispielsweise der Kontaktsituation zwischen einer flachen Rolle an einer Kante. Auch hier erfolgt die Parametrierung bei der Generierung der Druckprofile zufällig, um eine möglichst breite und repräsentative Datengrundlage zu erzeugen.
- 2) Modellbasierte Simulation des Rollenverschleiß: Der Rollenverschleiß wird anhand der Rollen- und Druckinformationen unter Verwendung des entwickelten Verschleißmodells simuliert. Zusätzlich wird dem Ergebnis ein Messrauschen aufgeschaltet, um Rauscheinflüsse der Realdaten zu simulieren.
- 3) Berechnung der resultierenden Rollengeometrie: Basierend auf dem berechneten Rollenverschleiß als Simulationsergebnis wird eine neue Rollenkontur nach einer Abnutzungsiteration berechnet. Diese steht nun für weitere Simulationen zur Verfügung.

- 4) Wiederholung der Schritte 2) bis 4): Die Schritte zur Simulation der durch Verschleiß unter variierenden Druckprofilen entstehenden Geometrien von Rollen werden solange wiederholt, bis der Radius einer Rolle unter 21 mm fällt. Dieser Wert orientiert sich an dem realen Materialvorrat der Rollen, die in der Mitte einen Stahlkern mit integriertem Lager aufweisen. Durch die sukzessive Abnutzung der Rollen in der Simulation soll der Eingangsdatenraum breit abgetastet werden, um ein umfangreiches Vorwissen in Form von Daten bereitzustellen.

Die Generierung synthetischer Daten kann theoretisch für unendlich viele Rollen erfolgen, sodass für Untersuchungen auf den synthetischen Daten eine beliebig große Datenmenge erzeugt werden konnte, wie in Abb. 41 exemplarisch für 1000 Rollen angedeutet. Die entstehende Basis an synthetischen Daten wurde genutzt, um die Fähigkeit der in AP2 identifizierte Algorithmik des maschinellen Lernens im nächsten Abschnitt hinsichtlich ihrer grundlegenden Eignung zur Prädiktion der Abnutzung von Ummantelungsrollen und damit auch zur Abschätzung von Einflüssen auf die resultierende Ummantelungsqualität zu bewerten.

**Tabelle 11: Zuordnung der Arbeiten für den Use Case Predictive Quality für AP 1**

Integration von Expertenwissen in Datenaufbereitung und Datenvorverarbeitung					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl			X		X
Entwicklung eines Verschleißmodells mit Expertenwissen					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM			X	X	X
düspohl	X	X		X	X
Generierung und Validierung synthetischer Daten					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl				X	X

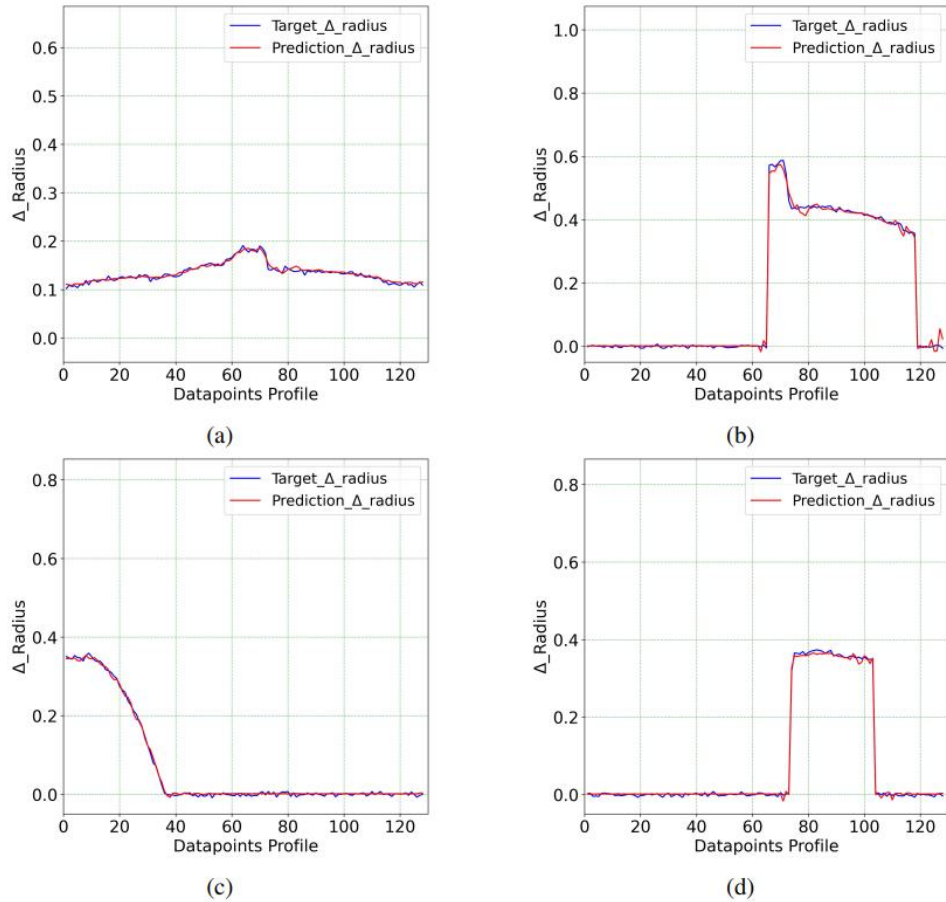
#### **AP 4: Lernen auf Datenströmen**

Für das fortlaufende Lernen auf Datenströmen ist eine beständig wachsende Datenbasis die Grundvoraussetzung. Von zentraler Bedeutung ist neben der Berücksichtigung von Unsicherheiten in Sensorsignalen auch die Überprüfung der Plausibilität von Daten. Wie unter AP 1 beschrieben, wurde die Datenakquise sukzessive verbessert. Da erst von diesem Zeitpunkt an Datenströme aufgezeichnet werden konnten, die unmittelbar für maschinelles Lernen geeignet sind, war die verfügbare Datengrundlage während der Projektlaufzeit für eine repräsentative Nutzung des entwickelten Algorithmus zu klein. Daher wurde das Lernen auf fortlaufenden Datenströmen anhand der synthetischen Daten umgesetzt. Dabei wurde für das Training des datenbasierten Teilmoduls eine Hyperparameteroptimierung implementiert, die unter statistischer Signifikanz eine optimierte CNN-Autoencoder-Architektur mit geeigneter Anzahl trainierbarer Modellparameter identifiziert.

Anhand der Rollen-IDs wurde eine Unterteilung in Trainings-, Validierungs- und Testdatensatz durchgeführt. Während das CNN auf den Validierungsdaten nicht trainiert wurde, diese aber zur Auswahl einer Architektur im Rahmen der Hyperparameteroptimierung genutzt wurden, dienten die Testdaten zur unabhängigen Bewertung der datenbasierten Verschleißprädiktion. Exemplarische Ergebnisse für 4 Verschleißzyklen von Rollen aus dem Testdatensatz sind in Abb. 42 ersichtlich. Die blaue Linie zeigt die mit Messrauschen versehene Rollenkontur, die durch den simulierten Verschleiß der Rolle entsteht, wohingegen die rote Linie die vom CNN prädizierte Rollengeometrie auf Basis der erwarteten Druckeinflüsse darstellt.

Es wird deutlich, dass das entwickelte datenbasierte Modell grundsätzlich sehr gute Prädiktionsergebnisse erzielt und damit zur Anwendung im Use Case geeignet ist.

Messrauschen wird tendenziell vom CNN rausgeglättet und das nichtlineare Verschleißverhalten kann über diese multiple Regression für jeden abgetasteten Punkt einer Ummantelungsrolle sehr gut approximiert werden.



**Abb. 42: Exemplarische Ergebnisse der ML-basierten Verschleißvorhersage für mehrere, aufeinanderfolgende, simulierte Verschleißiterationen einer Ummantelungsrolle von a) bis f)**

Ausgehend von dieser Architektur konnte im weiteren Projektverlauf sowie in der anschließenden Verwertung ein fortlaufendes Training auf sich entwickelnden Datenströmen erfolgen, beispielsweise durch ein Mini-Batch-Training (online), wie in Abschnitt 2.3.1 beschrieben. Zukünftig ist ein Zwischentraining auf den historischen, aber korrupten Daten denkbar, die vor der Anpassung der Datenakquise entstanden sind. Diese können beispielsweise unter Nutzung einer kleinen Lernrate für das Erlernen von Zusammenhängen, die in den synthetischen Daten nicht abgebildet werden, verwendet werden. Insgesamt kann so ausgehend von dem eintrainierten Vorwissen aus simulativen Daten die Performance auf den zunehmenden Realdaten sukzessive verbessert werden und eine langfristig robuste Verschleißprädiktion ermöglicht werden. Damit wird auch eine Abschätzung im Vorfeld möglich, ob die an der Anlage vorhandenen Rollen für die Erreichung einer den Anforderungen der Ummantelung genügenden Produktqualität geeignet sind, oder ob eine Nachbearbeitung von Rollen oder gar eine Anfertigung neuer Rollen von Nöten ist.

**Tabelle 12: Zuordnung der Arbeiten für den Use Case Predictive Quality für AP 4**

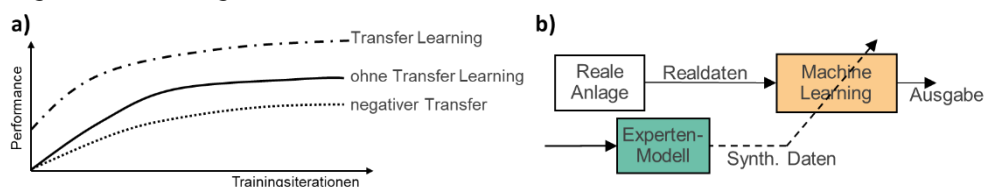
Training, Validierung und Bewertung von ML-Modellen					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl					X

## AP 5: Laufzeitoptimierung

Die Möglichkeiten zur Laufzeitoptimierung der Algorithmen wurden im Rahmen einer Studie untersucht (vgl. Abschnitt 2.3.1). Grundsätzlich stellen diese einen Trade-off zwischen der finalen Performance der maschinellen Lernverfahren und den zur Prädiktion benötigten Rechenressourcen dar. Basierend auf den Erkenntnissen des Forschungsvorhabens kann zukünftig anhand der Anforderungen an die Prädiktionsausgabe (Häufigkeit, Prädiktionsgüte, etc.) und den verfügbaren Rechenressourcen auf dem an der Anlage vorhandenen Edge Device eine Laufzeitoptimierung nach Bedarf erfolgen. Geeignete Möglichkeiten zur Laufzeitoptimierung wie Quantisierung und Pruning sowie benötigte Tools wurden im Rahmen von AP 5 erarbeitet.

Die Firma düspohl hat im Rahmen des Arbeitspaketes die Anforderungen an die Laufzeitfähigkeit überarbeitet und hinsichtlich der Ausführbarkeit validiert. Im Austausch mit dem IEM wurde die Datenstrategie optimiert, um die Prädiktion auf der bestehenden Hardware zu verschlanken und die Rechenaufwände zu reduzieren. Bei der Validierung an der Anlage durch düspohl, die durch das Fraunhofer IEM begleitet wurde, konnten Prädiktionen auf Basis der aktuellen Datenlage ohne Probleme ausgegeben werden. Durch die Parallelisierung der Berechnungen innerhalb des verwendeten CNNs sowie aufgrund der verfügbaren Rechenleistung waren keine stärkeren Reduktionen der Modellkomplexität notwendig. Daher wurde von der weiteren Laufzeitoptimierung mittels Quantisierung oder Pruning abgesehen, die mit einer Reduktion der Prädiktionsgüte bei der Verschleißvorhersage und der darauf aufbauenden Bewertung der Ummantelungsqualität einhergeht.

Wesentliche rechentechnische Aufwände bei der datenbasierten Prädiktion mittels ML-Verfahren entstehen nicht während der Nutzung, sondern während des Trainings der Modelle. Um Trainingsaufwände zu reduzieren, wurden Möglichkeiten zur Übertragung von Wissen mittels Daten sowie zwischen ML-Modellen untersucht. Das Thema Transfer Learning, bei dem durch einen positiven Transfer von Wissen – insbesondere durch die Nutzung eines vortrainierten Netzes beim Training – stand dabei in Untersuchungen des IEM im Fokus. Das Ziel ist das Realisieren eines Performance-Vorteils in der Nutzung des Netzes bei gleichzeitig verringertem Trainingsaufwand durch einen „Warm Start“, wie in Bild Abb. 43 a) in Anlehnung an Abb. 7 dargestellt. Wichtig ist dabei zu beachten, dass je nach Datenlage und Ähnlichkeit der Lernaufgabe auch ein negativer Transfer möglich ist. Diesen gilt es zu vermeiden.



**Abb. 43: a) Vorteile bei positivem Transfer Learning [17], b) Transfer Learning zur Parametrierung mit synthetischen Daten**

Dazu entwickelte und untersuchte das IEM Strategien für den Use Case Predictive Quality mittels derer die Potenziale von Transfer Learning erschlossen werden können. So ist eine Nutzung individueller Modelle je Anwendung zu wählen, da die Abnutzung von Rollen sowie die Auswirkungen auf die Prozessqualität leicht veränderten Zusammenhängen unterliegen, je nachdem ob zum Beispiel Fußleisten aus MDF oder Fensterprofile aus Kunststoff bzw. Aluminium ummantelt werden. Grundsätzliche Zusammenhänge bleiben aber vorhanden, womit eine Übertragung von ML-Modellen im Betrieb mittels Transfer Learning zukünftig zwischen verschiedenen Anwendungen zu empfehlen ist. Von Vorteil wäre auch das Nutzen einer zentralen Datenbank zum Vortrainieren von Modellen auf einer großen Datenbasis; die möglichst viele Anwendungen umfasst. Initial soll zudem bereits das Wissen aus dem erstellten Prozessmodell (AP 3) initial mittels Transfer Learning in die Algorithmen eingespielt werden, wie in Abb. 43 b) skizziert. So soll eine bessere initiale Parametrierung vor dem Training auf Realdaten erreicht werden.

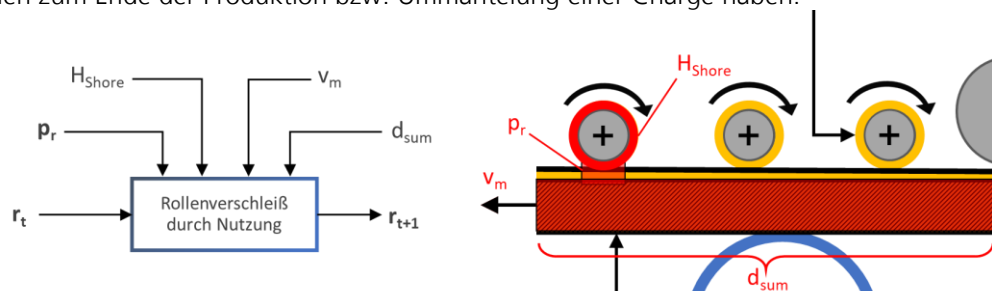
**Tabelle 13: Zuordnung der Arbeiten für den Use Case Predictive Quality für AP 5**

Optimierung Datenstrategie					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM			X	X	X
Düspohl	X	X	X	X	X
Transfer Learning für Warm Start bei Training					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	
Düspohl					X

**AP 6: Use Case Predictive Quality**

Im diesem Arbeitspaket sollten die entwickelten Methoden des maschinellen Lernens genutzt werden, um die Qualität des Ummantelungsprozesses kontinuierlich zu überwachen und dem Bediener eine Rückmeldung bezüglich kritischer Zustände zu geben. Der Fokus lag dabei insbesondere auf der Vorhersage von Verschleiß an Ummantelungsrollen. Diese werden in einer gewissen Qualität benötigt, um bei wiederkehrenden Produkten und den dafür bekannten Maschineneinstellungen reproduzierbar eine gute Ummantelungsqualität zu liefern. Aufgrund der mechanischen Belastungen sowie des Materials unterliegen diese Rollen aber einem Verschleiß durch mechanische Abnutzung. Dieser muss frühzeitig erkannt und prognostiziert werden, damit eine Wartung der Rollen im Magazin durch Nach- oder Neuschleifen bereits frühzeitig im Vorfeld zu einem kritischen Verschleißzustand vorgesehen werden kann, um Produktionsstillstand und Ausschuss zu vermeiden.

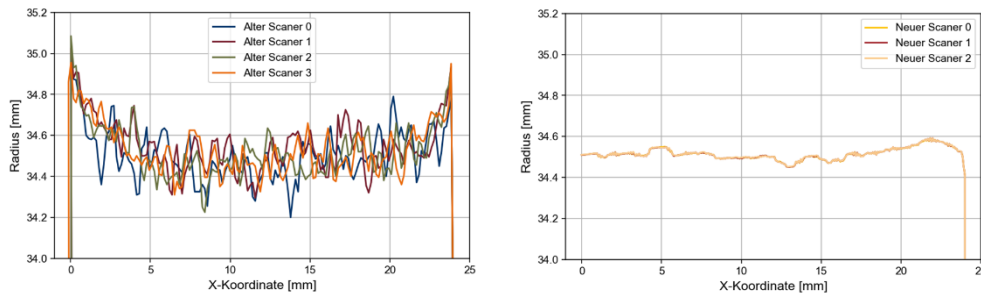
In diesem Zusammenhang hat die Firma düspohl die Integration benötigter Hard- und Software an der Anlage durchgeführt. Aufbauend auf den Ergebnissen aus AP 1 und AP 2 resultierte die Anforderung, zusätzliche Daten in der Anlage während der Produktion aufzuzeichnen. Diese wurden durch das IEM gemeinsam mit düspohl identifiziert. Anschließend nahm düspohl die Anpassung an einer Testanlage vor. Für die Prädiktion der Rollenabnutzung war dazu zusätzlich zu den Rüsteinstellungen und den Scans der Rollengeometrie vor und nach deren Nutzung die Erfassung weiterer Informationen erforderlich. Zu diesen zählen die Anzahl der produzierten Meter ummanteltes Profil sowie der Vorschub der Profile während der Produktion, wie in Abb. 44 ersichtlich. Basierend auf den vorangegangenen Analysen wurde abgeleitet, dass diese Größen einen massiven Einfluss auf die Abnutzung und damit den resultierenden Radius von Ummantelungsrollen zum Ende der Produktion bzw. Ummantelung einer Charge haben.



**Abb. 44: Relevante Einflussfaktoren zur Prädiktion von Rollenverschleiß beim Ummanteln**

Weiterhin konnte bei der Validierung der ML-Modelle auf den Produktionsdaten festgestellt werden, dass diese für die Prädiktion des Verschleißes qualitativ ungeeignet sind. Wie in Abb. 45 ersichtlich, liefern die alten Scanner beim Scannen der gleichen Rolle zwar eine Information über die grobe Rollengeometrie. Eine feine Erkennung der genauen Rollengeometrie ist aber nicht möglich. Diese Erkenntnis, die IEM und düspohl gemeinsam auf Basis der Datenanalysen des IEM erlangten, führten zu der Notwendigkeit, neue

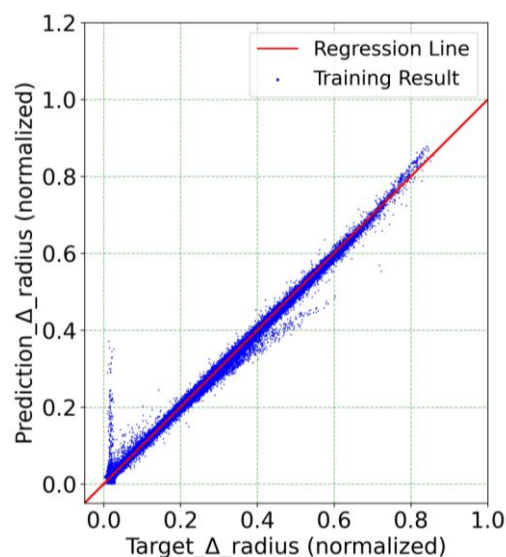
Profils Scanner in die Anlage zu integrieren. Diese Integration sowie die Einstellung der Scanner und deren Validierung war maßgeblicher Teil der Arbeiten von düspohl in diesem Zusammenhang, da der Wechsel der verwendeten Scanner auch Einfluss auf zahlreiche andere Funktionalitäten der RoboWrap nimmt.



**Abb. 45: Vergleich der Scan-Qualität alter und neuer Scanner für Rauschen und Reproduzierbarkeit**

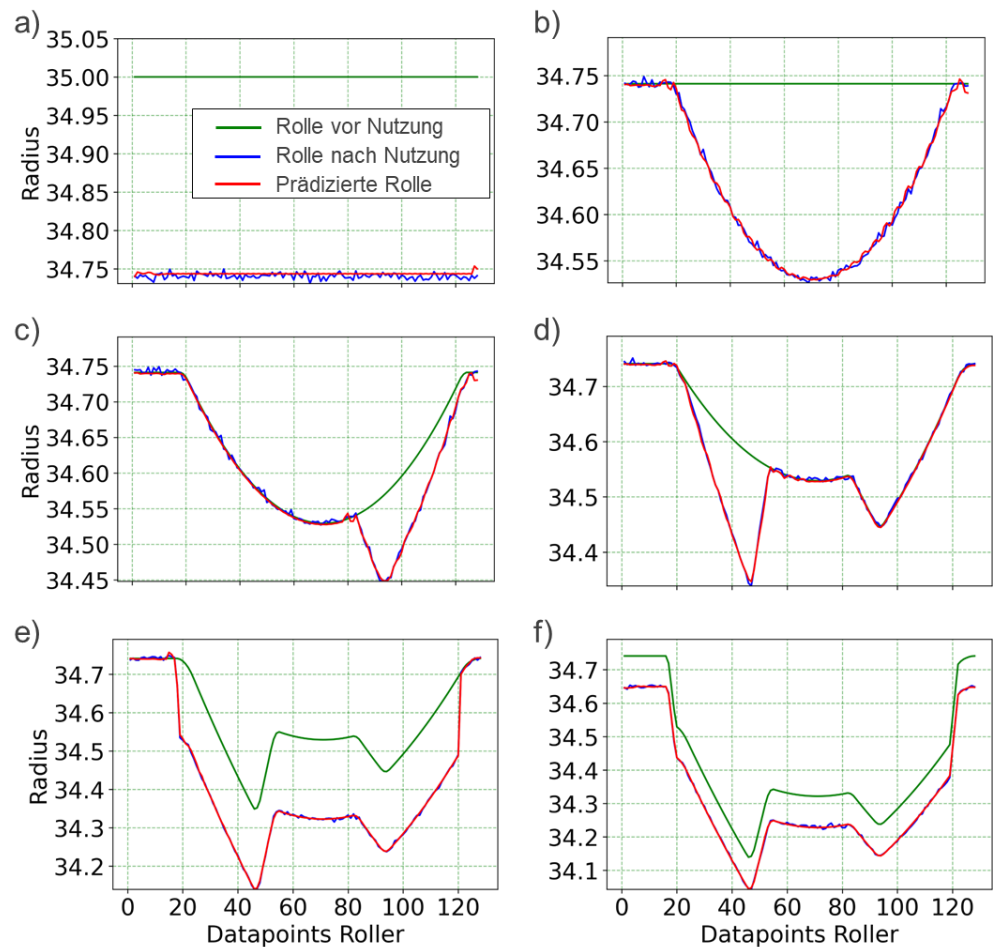
Mit den neuen Scannern ist, wie ebenfalls in Abb. 45 ersichtlich, eine Erfassung der Rollengeometrie mit deutlich höherer Auflösung und weniger Rauschen möglich und das vor allem reproduzierbar. Durch diese Anpassung startete aber auch die Aufnahme der Daten aus dem Feld neu, da die Daten der alten Scanner für das Machine Learning nicht geeignet sind. Im Laufe des Projektes konnten durch die kostenneutrale Laufzeitverlängerung um drei Monate zwar noch erste neue Scans von Rollen an der Anlage aufgezeichnet werden, die Validierung der Algorithmik erfolgte aber aufgrund der nicht ausreichend großen Realdatenbasis mit neuen Scannern basierend auf den synthetischen Daten aus dem Verschleißmodell.

Abb. 46 zeigt die Ergebnisse der Validierung nach Training auf den synthetischen Daten. Dazu wurde mittels ML-Modell der Verschleiß von Rollen in Abhängigkeit deren Nutzung (Kontaktsituation Rolle und Profil), des Verschiebs und der Distanz an ummanteltem Profil für einen unabhängigen Testdatensatz, der nicht Teil des Trainings war, prädiziert. Der dargestellte Regressions-Plot zeigt deutlich, dass die Verschleißzusammenhänge auf Basis der Daten sehr exakt gelernt werden konnten. Es gibt einige leichte Abweichungen, insbesondere wenn im Ergebnis der Testdaten der Verschleiß nahe Null war. Dies liegt vor allem daran, dass im Simulationsmodell auch ein Messrauschen implementiert wurde, um das Rauschen in den echten Messungen im Feld mit bei der Validierung des entwickelten ML-Verfahrens abzubilden.



**Abb. 46: Darstellung der Regressionsgüte der Rollenverschleißprädiktion auf den Testdaten**

Die Unterteilung in Trainings-, Validierungs- und Testdatensatz wurde anhand von Rollen-IDs, wie es sie auch an der realen Anlage gibt (abrufbar durch automatisches Scannen von Barcodes an den Rollen), durchgeführt. Während das CNN auf den Validierungsdaten nicht trainiert wurde, diese aber zur Auswahl einer Architektur im Rahmen der Hyperparameteroptimierung genutzt wurden, dienten die Testdaten zur unabhängigen Bewertung der datenbasierten Verschleißprädiktion. Exemplarische Ergebnisse für die ersten 6 aufeinander folgenden Verschleißzyklen der gleichen Rolle aus dem Testdatensatz sind in Abb. 42 ersichtlich, um die Prädiktion des Verschleißes über mehrere Nutzungen der Rolle in die Zukunft zu bewerten. Die grüne Linie charakterisiert dabei die Rollenkontur vor der jeweiligen Nutzung im Ummantlungsprozess. Die blaue Linie hingegen zeigt die mit Messrauschen versehene Rollenkontur, die durch den simulierten Verschleiß der Rolle entsteht, wohingegen die rote Linie die vom CNN prädizierte Rollengeometrie auf Basis der erwarteten Druckeinflüsse darstellt.



**Abb. 47: Exemplarische Ergebnisse der ML-basierten Verschleißvorhersage für mehrere, aufeinanderfolgende, simulierte Verschleißiterationen einer Ummantlungsrolle von a) bis f)**

Es wird deutlich, dass das entwickelte Lernverfahren den Verschleiß grundsätzlich sehr gut vorhersagen kann. Messrauschen wurde tendenziell vom CNN rausgeglättet und das nichtlineare Verschleißverhalten konnte über diese multiple Regression für jeden abgetasteten Punkt einer Ummantlungsrolle sehr gut approximiert werden. Das CNN ist sehr gut für die Verschleißprädiktion in dieser Anwendung geeignet und kann die Abnutzung von Rollen in die Zukunft prädizieren. Damit gilt es im Anschluss an das Vorhaben, diese Architektur in der für die anschließende Verwertung aufzubereiten und fortlaufend auf den sich entwickelnden Realdatenströmen weiter zu trainieren, beispielsweise durch ein Transfer Learning. Auch kann zuvor ein Zwischentraining auf den historischen, aber

korrupten Daten erfolgen, die vor der Anpassung der Datenakquise durch neue Scanner aufgezeichnet worden sind. Diese können beispielsweise unter Nutzung einer kleinen Lernrate für das Erlernen von Zusammenhängen, die in den synthetischen Daten nicht abgebildet werden, verwendet werden.

Insgesamt kann so ausgehend von dem eintrainierten Vorwissen aus simulativen Daten die Performance auf den zunehmenden Realdaten sukzessive verbessert werden und eine langfristig robuste Verschleißprädiktion ermöglicht werden. Damit wird auch eine Abschätzung im Vorfeld möglich, ob die an der Anlage vorhandenen Rollen für die Erreichung einer den Anforderungen der Ummantelung genügenden Produktqualität geeignet sind, oder ob eine Nachbearbeitung von Rollen oder gar eine Anfertigung neuer Rollen von Nöten ist. Dazu werden verschiedene geometrische Eigenschaften von Rollen aus der Prädiktion ermittelt. Insbesondere der mittlere Radius für verschiedene Teilbereiche einer Rolle ist dabei von Bedeutung und wird mit den Sollwerten für diese Rolle verglichen. Bei zu starker Abweichung ist die Rolle zu ersetzen oder nachzubearbeiten.

-----  
Eingehende Darstellung  
-----

**Tabelle 14: Zuordnung der Arbeiten für den Use Case Predictive Quality für AP 6**

Erweiterung der Datenbasis und Sensorik um Produktions-Metainformationen					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM					X
düspohl	X	X	X	X	X
Analyse, Integration und Implementierung neuer Scanner					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM		X		X	X
düspohl	X	X	X	X	X
Validierung Verschleißprädiktion und Qualitätsbewertung					
	Koordinierend	Entwickelnd	Implementierend	Unterstützend	Bewertend
IEM	X	X	X	X	X
düspohl		X		X	X

### 2.3.4 Ergänzende Inhalte

Die im Projekt geleisteten Arbeiten zur Adaption der betrachteten technischen Systeme und der Dateninfrastruktur in den Use Cases waren erforderlich, um notwendige Daten verfügbar und Algorithmen des maschinellen Lernens auf Embedded- und Edge Devices in der unmittelbaren Nähe der technischen Anlagen zu erproben. Während des Projektes wurden zudem die wissenschaftlichen Entwicklungen verfolgt und bei im Vorgehen berücksichtigt. So wurden beispielsweise modernste MEMS als Sensorik verwendet und aktuelle Verfahren zur Modellkomprimierung verwendet und bewertet. Wissenschaftliche Fortschritte waren insbesondere auch im Bereich der Integration von Expertenwissen in ML-Verfahren zu verzeichnen. An diesen hat sich insbesondere das IEM als Kompetenzträger für die Konzipierung und Implementierung hybrider Lernverfahren orientiert. Zudem tragen die Projektergebnisse maßgeblich zu Fortschritten in zwei Promotionsvorhaben auf Seiten des Fraunhofer IEM bei. Zudem erfolgte die Veröffentlichung zweier Publikationen, in denen ergänzend zu den Ergebnissen aus anderen Projekten auch zentrale Erkenntnisse und Teilergebnisse aus dem Projekt EASY eingeflossen sind.

- Lenz, C., Henke, C., & Trächtler, A. (2023, July). A Methodical Approach to Hybrid Modelling for Contextual Anomaly Detection on Time-Series Data. In 2023 IEEE 21st International Conference on Industrial Informatics (INDIN) (pp. 1-5). IEEE.
- Koppert, S., Bause, M., Henke, C., & Trächtler, A. (2023, July). Learning the Automated Setup of Profile Wrapping Lines for New Products from Few Past Setups. In 2023 IEEE 21st International Conference on Industrial Informatics (INDIN) (pp. 1-6). IEEE.

Im Rahmen des Projektes entstanden zudem eine Bachelorarbeit, zwei Studienarbeiten sowie drei Masterarbeiten in den Bereichen Maschinenbau und Informatik an der Universität Paderborn. Eine weitere Publikation zum Thema hybride maschinelle Lernverfahren zur Zustandsüberwachung von Intralogistiksystemen befindet sich in Planung. Weiterhin sind im Rahmen des Projektes Dienstreisen aus verschiedenen Anlässen angefallen, u. a. für Workshops und Validierungen von Software und Infrastruktur vor Ort an den technischen Systemen im Rahmen der Use Cases. Zudem gab es einen projektübergreifenden Austausch mit dem Projekt ARISE sowie dem Projekt it's 3DL, die ebenfalls in deutsch-kanadischer Partnerschaft durchgeführt wurden. So wurde unter anderem das Abschluss-Meilensteintreffen der Projekte EASY und it's 3DL gemeinsam vor Ort in Lemgo durchgeführt, um den Austausch über Projektergebnisse zwischen den beteiligten Unternehmen und wissenschaftliche Einrichtungen zu stützen. Hervorzuheben ist auch die deutsch-kanadische Kooperation, die im Rahmen eines Symposiums gegen Projektende im Mai 2023 durch einen tiefen fachlichen Austausch vor Ort in Kanada gefestigt werden konnte. Dazu besuchte das IEM unter anderem den kanadischen Projektpartner NTWIST in Edmonton, bevor die Teilnahme am dreitägigen Symposium an der University of Victoria vom 17. Bis zum 19. Mai 2023 erfolgte. Bei diesem Anlass wurden auch die zentralen Ergebnisse aus dem Projekt EASY präsentiert und diskutiert.

## 2.4 Abgeschlossene Meilensteine

-----  
Eingehende Darstellung  
-----

### Meilenstein 1 wurde erreicht, die wesentlichen Ergebnisse bestehen in

- einer **Dokumentation** der **Anforderungen**
- einem **Machine Learning Konzept**
- der **Projektfeinplanung**

### Meilenstein 2 wurde erreicht, die wesentlichen Ergebnisse bestehen in

- evaluierten **Embedded KI-Frameworks**
- einer **Auswahl** an **ML-Verfahren**, die für beide Use Cases auf **Edge Devices** umsetzbar sind

### Meilenstein 3 wurde erreicht, die wesentlichen Ergebnisse bestehen in

- verschiedenen, **validierten Ansätzen zur Wissensrepräsentation**, individuell angepasst an die Use Cases
- einem allgemeinen **Konzept zur Integration** von verschiedenen formalisiertem Wissen

### Meilenstein 4 wurde erreicht, die wesentlichen Ergebnisse bestehen in

- Verfahren des **maschinellen Lernens**, die **auf sich dynamisch entwickelnden Datenströmen** lernen
- einem Konzept zum langfristigen **Online-Lernen** auf **Edge Devices**

### Meilenstein 5 wurde erreicht, die wesentlichen Ergebnisse bestehen in

- angepassten, **laufzeitoptimierten ML-Algorithmen**
- einer Auswahl an **Tools** und **Möglichkeiten** zur bedarfsgerechten Laufzeitoptimierung

### Meilenstein 6 wurde erreicht, die wesentlichen Ergebnisse bestehen in

- erfolgreich validierten Algorithmen in den Use Cases
- Demonstration intelligenter Diagnoseverfahren zur Fehlerdiagnose in der Infrastruktur eines Hochregallagers für den Use Case Predictive Quality
- einem validierten ML-Lösungskonzept zur Verschleißprädiktion, integriert in ein Framework zur Qualitätsprognose für die Profillummantelung

## 2.5 Zusammenfassung

Im Rahmen des Projektes „EASY – Embedded Artificial Intelligence for Production Systems“ wurden Algorithmen aus dem Bereich der künstlichen Intelligenz (KI) für Embedded und Edge Devices entwickelt. Die in diesem Vorhaben betrachteten Use Cases waren die intelligente Qualitätsprädiktion von Profilmantelungsmaschinen und die vorausschauende Wartung von ausfallkritischen Komponenten in den fördertechnischen Anwendungen von Hochregallagern. Zudem wurde auf Seiten der kanadischen Partner die ganzheitliche Fertigungsplanung in Form eines intelligenten Manufacturing Execution Systems unter Nutzung maschineller Lernverfahren untersucht.

Ein wesentliches Ziel dieses Vorhabens bestand dabei in der Entwicklung und Umsetzung von Lernverfahren in verteilten, eingebetteten Systemen unter Berücksichtigung von Ressourcenbeschränkungen sowie die Integration von Expertenwissen in die Lernverfahren. Ausgangspunkt hierfür waren Daten, die aus multimodalen, aktiven Sensorsystemen gewonnen werden und mit eingebetteten, realzeitfähigen Datenanalysemethoden sowie mit ML-Methoden ausgewertet werden konnten.

### 2.5.1 Erreichte Resultate

Im Projekt konnten zentrale wissenschaftliche Ergebnisse in der Konzipierung, Implementierung und Validierung laufzeitoptimierter ML-Verfahren für Embedded und Edge Devices sowie bezüglich der Integration von Expertenwissen gewonnen und in praktischen Use Cases validiert werden. Insbesondere die Analysen der Daten aus den Use Cases haben zudem die Bedeutung der Datenbasis für den Erfolg von maschinellen Lernverfahren aufgezeigt. Hier konnte das IEM nicht nur Erkenntnisse gewinnen, sondern auch eine Methodik zur Analyse vorhandener Unternehmensdaten entwickeln, die Aussagen über die Realisierbarkeit von Zustandsüberwachungssystemen ermöglicht und als Grundlage bei der Erhebung notwendiger Anforderung für solche System von großem Nutzen ist. Weiterhin konnte das Fraunhofer IEM Pipelines zu Laufzeitoptimierung sowie Vorgehensweisen zum fortlaufenden Training auf Datenströmen konzipieren und in Zusammenarbeit mit den Kooperationspartnern validieren und damit seine Kompetenz bei der Entwicklung von Embedded- und Edge-AI-Lösungen ausbauen. Zudem wurden neuartige Konzepte zur Integration von Expertenwissen in Lernverfahren erarbeitet und in die Prozesse implementiert. Im Rahmen des EASY-Projekts wurden folgende Hauptergebnisse erzielt:

- Auswahl und Anwendung geeigneter Lernverfahren (ML und neuronale Netze) zur Erkennung mechanischer Fehler sowie zur Vorhersage einer schlechten Produktqualität.
- Eine umfangreiche Anforderungserhebung sowie die Integration von Expertenwissen führten zu spezifischen Verfahren zur Ermittlung mechanischer Schadensfrequenzen sowie zu einer hybriden Architektur für die Verschleißprädiktion von Rollen.
- Demonstration einer sicheren, laufzeitoptimierten und skalierbaren IoT-Architektur, die als Grundlage für zukünftige Projekte dient.
- Laufzeitoptimierte ML-Algorithmen und Möglichkeiten für deren effiziente und performante Übertragung auf Embedded- oder Edge-Devices.
- ML-Verfahren wurden hinsichtlich benötigter Laufzeit und Speicherkapazität optimiert und auf Embedded und Edge-Geräten getestet (z. B. auf einem selbst entwickelten Sensorboard mit zusätzlicher Vibrationssensorik).
- Hybride Architekturen wurden konzipiert, bewertet und ausgewählte Architekturen getestet sowie zum fortlaufenden Training auf Datenströmen verwendet.

Neben diesen Hauptergebnissen wurde durch intensiven Austausch ein Netzwerk mit Domänenexpertinnen und -experten in Kanada und weiteren it's OWL-Projekten etabliert. Während der Lösungsentwicklung wurden ein Anforderungserhebungsprozess für

Predictive Maintenance-Lösungen definiert und zur Entwicklung und Validierung eines Demonstrators von mechanischem Verschleiß an Regalbediengeräten im produktiven Einsatz genutzt. Zudem entstand ein CNN-basiertes, validiertes ML-Konzept zur datenbasierten Verschleißprädiktion von Ummantelungsrollen, das sich auch auf die Abnutzung anderer Rollen in gänzlich anderen technischen Systemen übertragen lässt. Die gesammelten Erfahrungen und erworbenen Kompetenzen führten zu einer Professionalisierung der lösungsorientierten Anwendung von Verfahren und Vorgehensweisen im Umgang mit dem Training und der Optimierung von Lernalgorithmen und ML-Verfahren. Im Use Case Predictive Maintenance wurden ML-Algorithmen und Embedded KI-Systeme entwickelt und genutzt, um Anomalien, Verschleißzustände und Ausfälle technischer Systeme in Hochregallagern basierend auf verfügbaren, multivariaten Sensordaten zu diagnostizieren. Ausgehend von einer Anforderungsanalyse in Hochregallagern wurden relevante Schadenbilder identifiziert und die notwendige Infrastruktur sowie eine Datengrundlage für eine zustandsorientierte Instandhaltung unter Nutzung von ML-Verfahren zur Zustandsprädiktion geschaffen. Ausgehend von identifizierten Lernverfahren wurde Verfahren zur Integration von Expertenwissen bei der Nutzung von ML-Verfahren entwickelt und validiert. Zudem wurden sowohl die Algorithmen als auch die Daten hinsichtlich der Ausführung auf Embedded oder Edge Devices optimiert und die laufzeitoptimierten Algorithmen auf der Zielhardware erprobt.

Im Use Case Predictive Quality wurde ein datenbasiertes Assistenzsystem entwickelt, das komplexe technische Zusammenhänge in der Profillummantelung bewertet und darauf aufbauend kritische Produktionsqualität anhand des Rüstzustandes und des Verschleißzustandes für die Ummantelungsanlage „Robowrap“ prädiziert. Dazu wurde ein ML-Verfahren zur Prädiktion von Rollenverschleiß unter Nutzung eines CNNs sowie unter Integration von Expertenwissen konzipiert und entwickelt. Ausgehend von diesem kann ein kritischer Rollenverschleiß erkannt werden und die Produktion von Ausschuss im Vorfeld prädiziert und durch gezielte Wartung der Ummantelungsrollen vermieden werden. Gleichzeitig wurde die Hardware der „RoboWrap“ prototypisch derart angepasst, dass eine Weiterentwicklung und Integration des Systems bei den Kunden der Firma düspohl in Nachlauf zu dem Projekt erfolgen kann. Für das entwickelte ML-Verfahren ist grundsätzlich eine Übertragung auf andere Systeme mit verschleißenden Rollen denkbar.

## 2.5.2 Zusammenarbeit mit den kanadischen Partnern

Mit den beiden kanadischen Partnern (University of Victoria (UVIC), NTWIST Inc.) wurden monatliche Webmeetings abgehalten. Aufgrund der COVID-19 Pandemie konnte ein persönlicher Austausch erst im Mai 2023 im Rahmen eines Symposiums in Kanada erfolgen. Dieses wurde durch den kanadischen Partner UVIC organisiert. Kanadische und deutsche Teilnehmer aus Wirtschaft und Wissenschaft kamen dabei zusammen um über Projekte, Innovationen und wissenschaftliche Erkenntnisse zu referieren und zu diskutieren.

Bei den regelmäßigen Projekttreffen wurden die technischen und wissenschaftlichen Erkenntnisse in Form von Präsentationen geteilt und im Nachgang entsprechend besprochen. Der wissenschaftliche Diskurs und die Sicht eines Partners, der die Ergebnisse von außen bewertet, ermöglichte die Entwicklung und Optimierung der ML-Verfahren und Expertenverfahren, aber auch die Optimierung der Datenbasis aus einem anderen Blickwinkel zu betrachten und damit bessere Lösungskonzepte zu erarbeiten. Zudem wurden inhaltliche Workshops durchgeführt. Durch den engen Austausch konnten beide Seiten andere Sichtweisen auf technische Fragestellungen gewinnen und neue Impulse in das eigene Teilprojekt einbinden. Auch wurden innerhalb des Projektes Daten und algorithmische Bausteine ausgetauscht. Dies ermöglichte den Partnern, ihre entwickelten Lösungen umfangreicher zu validieren und basierend auf dem Feedback der Partner zu verbessern.

Die kanadischen Partner haben sich mit der Entwicklung eines Agenten-basierten IMES beschäftigt. Am Beispiel eines Kraftwerkes wurden die Anwendungen einer verfahrenstechnischen Anlage analysiert und unterschiedliche Ansätze von neuronalen Netzen

analysiert. Weiterhin wurde eine umfassende Übersicht erarbeitet, welche Ansätze in der Forschung im Bereich von KI-basierten MES-Systemen existieren. Diese wurden systematisch aufgearbeitet und in Form von wissenschaftlichen Publikationen veröffentlicht.

### 2.5.3 Verwertung der Ergebnisse und Anschlussfähigkeit

Das **Fraunhofer IEM** konnte im Rahmen des Projektes wesentliche Erkenntnisse im Bereich der datenbasierten Überwachung von technischen Prozessen gewinnen. Die gewonnene Expertise bietet die Möglichkeit, Unternehmen bei der Einführung von datenbasierten Überwachungssystemen im Bereich Predictive Quality und Predictive Maintenance zu unterstützen und die erfolgreiche Nutzung maschineller Lernverfahren auf begrenzten Ressourcen in und an Maschinen technischer Prozesse zu ermöglichen. Von Bedeutung sind dabei auch die Erkenntnisse über die notwendige Datenbasis sowie Datenqualität und Datenquantität, die für eine erfolgreiche Umsetzung datenbasierter Lernverfahren gewährleistet sein müssen. Auch konnten Möglichkeiten zur Integration von Expertenwissen in ML-Verfahren identifiziert und in den Use Cases validiert werden. Hier ergeben sich große Potenziale für weitere Untersuchungen, insbesondere vor dem Hintergrund des EU-AI-Acts und den damit einhergehend steigenden Anforderungen an die funktionale Sicherheit von KI- und ML-Verfahren. Hybride Strukturen und die Einbindung von Expertenwissen können hier eine zentrale Rolle spielen. Zudem soll das generierte Wissen genutzt werden, um das gewonnene Technologiewissen aus den Projektergebnissen in weitere Industrieunternehmen zu transferieren. Zudem nutzt das IEM Webinare, Innovationsplattformen, Netzwerke und Vorträge, um das aufbereitete Wissen für Industrie und Wissenschaft aber auch die breite Öffentlichkeit zugänglich zu machen.

Weiterhin plant das **Fraunhofer IEM** das gewonnene Wissen zur Laufzeitoptimierung von ML-Modellen zu nutzen, um deren Leistungsfähigkeit auf Embedded Devices zu verbessern und Unternehmen bei der Entwicklung dieser Systeme als kostengünstige, aber sehr leistungsfähige dezentrale Lösungen für Produktion und Logistik zu unterstützen. Anschließende Forschung zu wissenschaftlichen Fragestellungen zum Beispiel zur Integration von Expertenwissen sowie zum Thema Transfer Learning sollen zudem in weiteren Projekten untersucht werden. Dazu wurde bereits mit Auslaufen des Projektes eine Vorhabensbeschreibung ausgearbeitet, die sich unter anderem mit dem intelligenten Datendesign für robustes Transferlernen beschäftigt.

Das Unternehmen **encoway** konnte seine Kenntnisse im Bereich der zustandsbasierten Instandhaltung vertiefen und neuartige Verfahren zur Detektion und Prädiktion von Defekten und schlechten Systemzuständen für die Instandhaltung von Intralogistiksystemen in die Anwendung bringen. Zudem konnten die Potenziale von Motorstromdaten für die Fehlererkennung identifiziert werden und damit ein Prototyp zum intelligenten Monitoring mittels ML-Verfahren auf Embedded und Edge Devices geschaffen werden, der ausgehend von der Applikation in eigenen Anwendungen zukünftig auch für die Ausbreitung bei Kunden weiterentwickelt werden soll. Die gesammelten Erfahrungen und erworbenen Kompetenzen führten zudem zu einer Professionalisierung der lösungsorientierten Anwendung von Verfahren und Vorgehensweisen im Umgang mit dem Training und der Optimierung von Lernalgorithmen.

Im Use Case Predictive Quality hat die Firma **düspohl** wertvolle Erkenntnisse über die Potenziale von ML-Verfahren zur Verschleiß- und Qualitätsvorhersage in der Produktion erlangt. Zudem konnten Schwachstellen in der bestehenden Datenbasis identifiziert und ausgebessert werden. Durch die Analyse und Adaption der Anlage hinsichtlich der identifizierten Anforderungen an die aufgezeichneten Daten konnte der Grundstein für die zukünftige Markteinführung von KI-Features bei düspohl gelegt werden. Das resultierende und virtuell simulierte Konzept zur Qualitätsprädiktion fügt sich nahtlos in das datenbasierte Ökosystem von düspohl ein. Diese Integration soll im Anschluss an das Projekt erfolgen, auch wenn die Qualitätsprädiktion im Rahmen des Projektes an der Anlage „RoboWrap“ selbst aufgrund der umfangreich notwendigen Adaption und der zu Projektlaufzeit zu kleinen Datengrundlage nicht vollständig umgesetzt werden konnte. Auch wenn bis zum Ausrollen der Qualitäts- und Verschleißvorhersage am Markt noch einige

Zwischenschritte notwendig sind, ist davon auszugehen, dass zukünftig ein Mehrwert für die Kunden durch KI-Assistenzsysteme geschaffen werden kann, indem die Produktion optimiert und der Ausschuss minimiert wird. Mit dem Projekt konnte düspohl seine Kompetenzen im Bereich KI weiter ausbauen und den Technologievorsprung im Bereich der Profilmantelung im internationalen Vergleich ausbauen.

-----  
Eingehende Darstellung  
-----

## 2.6 Literaturverzeichnis

- [1] THEISLER, A.; PÉREZ-VELÁZQUEZ, J.; KETTELGERDES, M.; ELGER, G.: Predictive Maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability engineering & system safety* 215, 2021
- [2] JIMENEZ, J. J. M.; SCHWARTZ, S.; VINGERHOEDS, R. A.; GRABOT, B.; SALAÜN, M.: Towards multi-model approaches to predictive maintenance: A systematic literature survey on diagnostics and prognostics. *Journal of manufacturing systems*, 56, 2020, S. 539-557
- [3] BEYERER, J.; RICHTER, M.; NAGEL, M.: *Pattern Recognition – Introduction, Features, Classifiers and Principles*. De Gruyter Oldenbourg, Berlin, Boston, 2018
- [4] ANDRE, A. B.; BELTRAME, E.; WAINER, J.: A combination of support vector machine and k-nearest neighbors for machine fault detection. *Applied Artificial Intelligence*, 27(1), 2013, S. 2013
- [5] GUPTA, P.; SEHGAL, N. K.: *Introduction to machine learning in the cloud with python: Concepts and practices*. Springer Nature, 2021
- [6] ZHU, T.; RAN, Y.; ZHOU, X.; WEN, Y.: A Survey of Predictive Maintenance: Systems, Purposes and Approaches. *arXiv preprint, arXiv:1912.07383*, 2019
- [7] O'MAHONY, N.; CAMPBELL, S.; CARVALHO, A.; HARAPANAHALLI, S.; HERNANDEZ, G. V.; KRPALKOVA, L.; RIORDAN, D.; WALSH, J.: Deep learning vs. traditional computer vision. *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1* 1, Springer International Publishing, 2020, S. 128-144
- [8] SERRADILLA, O.; ZUGASTI, E.; RODRIGUEZ, J.; ZURUTUZA, U.: Deep learning models for predictive maintenance: a survey, comparison, challenges and prospects. *Applied Intelligence*, 51(10), 2022, S. 10934-10964
- [9] VON RUEDEN, L.; MAYER, S.; BECKH, K.; GEORGIEV, B.; GIESSELBACH, S.; HEESE, R.; KIRSCH, B.; PFROMMER, J.; PICK, A.; RAMAMURTHY, R.; WALCZAK, M.; GARCKE, J.; BAUCKHAGE, C.; SCHUECKER, J.: *Informed Machine Learning – A taxonomy and survey of integrating prior knowledge into learning systems*. *IEEE Transactions of Knowledge and Data Engineering* 35.1, 2021, S. 614-633
- [10] KARNIADAKIS, G. E.; KEVREKIDIS, I. G.; LU, L.; PERDIKARIS, P.; WANG, S.; YANG, L.: Physics-informed machine learning. *Nature Reviews Physics*, 3 (6), 2021, S. 422-440
- [11] MENG, C.; SEO, S.; CAO, D.; GRIESEMER, S.; LIU, J.: When physics meets machine learning: A survey of physics-informed machine learning. *arXiv preprint, arXiv:2203.16797*, 2022
- [12] TERCAN, H.; GUAJARDO, A.; HEINISCH, J.; THIELE, T.; HOPMANN, C.; MEISEN, T.: Transfer-learning: Bridging the gap between real and simulation data for machine learning in injection molding. *Procedia Cirp*, 72. 2018, S. 185-190

- [13] KAZMI, H. S.; SUYKENS, J.; DRIESEN, J.; CORRADO, V.; FABRIZIO, E.; GASPARELLA, A.; PATUZZI, F.: Large-scale transfer learning for data-driven modelling of hot water systems. Proceedings of the International Building Performance Association, 2019
- [14] YOUNESIAN, D.; HOSSEINKHANI, A.; ASKARI, H.; ESMAILZADEH, E.: Elastic and viscoelastic foundation: a review on linear and nonlinear vibration modeling and applications. Nonlinear Dyn 97, 2019, S. 853-895
- [15] TEODORU, I. B.: EBBEF2p-A Computer Code for Analysing Beams on Elastic Foundations. IntersctII/Intersections, 6 (1), 2009
- [16] ARGATOV, I. I.; CHAI, Y. S.: Artificial neural network modeling of sliding wear. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 235 (4), 2021, S. 748-757
- TOMMASI, T.; ORABONA, F.; CAPUTO, B.: Learning categories from few examples with multi model knowledge transfer. IEEE transactions on pattern analysis and machine intelligence, 36(5), S. 928.942, 2013

## 2.7 Abkürzungsverzeichnis

AP	Arbeitspaket
CAD	Computer Aided Design
CNN	Convolutional Neural Network
IMES	Intelligentes Manufacturing Execution System
KI	Künstliche Intelligenz
MCC	Mechatronics Competence Campus
MEMS	Micro Electro Mechanical System
MES	Manufacturing Execution System
MQTT	Message Queuing Telemetry Transport
ML	Machine Learning
RBG	Regalbediengerät
TL	Transfer Learning