

Optische Erfassung von Messwerten

Maschinelle Lernverfahren zur Digitalisierung manueller Ablese- und Messvorgänge

Matthias Mühlbauer, Hubert Würschinger, Nico Hanenkamp, Universität Erlangen-Nürnberg und Svyatoslav Funtikov, ROI-EFESO MANAGEMENT CONSULTING AG

Im Fabrikbetrieb kommen häufig Messmittel zum Einsatz, bei denen noch keine automatische Abspeicherung oder Weiterverarbeitung des gemessenen Werts erfolgt. Mitarbeiter müssen in diesem Fall den Messwert manuell erfassen und weiterverarbeiten. Folgend wird ein Ansatz zur optischen Erfassung und Digitalisierung von Messwerten mithilfe Maschinelles Lernverfahren vorgestellt. Dies soll zur Entlastung der Mitarbeiter beitragen, Ablesefehler reduzieren sowie eine automatisierte Dokumentation ermöglichen.

Trotz fortschreitender Digitalisierung und Automatisierung stellen manuelle Tätigkeiten einen wesentlichen Faktor im Fabrikbetrieb dar. Als ein Teilbereich können manuelle Ablese- und Messvorgänge betrachtet werden. Derartige Prozesse sind beispielweise Bestandteil von Qualitätsprüfungen, Einrichtungsvorgängen oder Inspektionen. Hierbei werden zum Teil Messmittel eingesetzt, die den Messwert analog oder digital darstellen, jedoch keine Konnektivität zu einem zentralen System aufweisen. Die Dokumentation des Messwerts respektive die Durchführung eines Soll-Ist-Abgleichs sowie die Einleitung der entsprechenden Folgemaßnahmen müssen somit durch den Werker durchgeführt werden. Aufgrund steigender Komplexität, Zeitdruck, Müdigkeit oder unterschiedlichen Qualifikationsniveaus können hierbei beispielweise Ablesefehler oder Schwankungen in der Dauer der Durchführung auftreten [1]. Insbesondere bei Ablesefehlern sind erhebliche Folgekosten durch Nacharbeit und Ausschuss sowie mögliche Ausfälle von Bauteilen mit Personenschäden nicht auszuschließen.

Im Rahmen dieses Beitrags wird ein kameragestützter Ansatz zur Schließung dieser Digitalisierungslücke vorgestellt. Das Messmittel bzw. der Ablese- oder Messvorgang wird durch eine vom Mitarbeiter geführte bzw. getragene Kamera aufgezeichnet und der Messwert durch Bildanalysetechniken extrahiert. Zentrales Element stellen hierbei Objektdetektoren dar, die

es erlauben, das Messmittel auch bei unterschiedlichen Rahmenbedingungen im Bild zu lokalisieren.

Folgend wird der Stand der Technik kurz dargestellt. Im zweiten Teil des Beitrags wird die angewandte generische Vorgehensweise zur Messwertextraktion beschrieben. Diese wird anhand von zwei repräsentativen Anwendungsfällen validiert. Abschließend erfolgt die Diskussion der Ergebnisse.

Stand der Technik – Bildanalyse mit maschinellen Lernverfahren

Eine Möglichkeit zur Erkennung und Lokalisierung von Objekten im Bild stellen überwachte maschinelle Lernverfahren dar. Bei überwachten Lernverfahren werden auf Basis eines Trainingsdatensatzes Muster im Datensatz erlernt. Jede Instanz des Trainingsdatensatzes ist hierbei bereits mit der Information über die Zielgröße (Klasse, Regressionswert, Objektposition etc.) gekennzeichnet. Nach dem erfolgreichen Training können auch neue, unbekannte Instanzen verarbeitet und somit deren Zielgröße prognostiziert werden. Eine Unterkategorie des Maschinellen Lernens stellen künstliche neuronale Netze dar. Diese finden, aufgrund der Möglichkeit der Verarbeitung von komplexen, mehrdimensionalen Daten, beim maschinellen Sehen häufig Anwendung.

Beim maschinellen Sehen kann unter anderem zwischen Bildklassifizierung und Objekterkennung unterschieden

Optical Detection of Measured Values – Machine Learning Methods for Digitalizing Manual Reading and Measuring Processes

In factory operations, measuring equipment is often used without automatic storage or further processing possibilities of the measured value. In this case, employees must capture and process the measured values manually. In this article, an approach for the optical detection and digitization of measured values with the help of machine learning methods is presented. This aims to reduce the workload of the employees, avoid reading errors and enable automated documentation.

Keywords:

digitalization, image processing, machine learning, object detection

Matthias Mühlbauer, M. Sc. ist Doktorand am Lehrstuhl für Ressourcen- und Energieeffiziente Produktionsmaschinen an der Universität Erlangen-Nürnberg. Sein Forschungsfokus liegt auf der Analyse von manuellen Prozessen auf Basis von Bild- bzw. Videoaufzeichnungen.

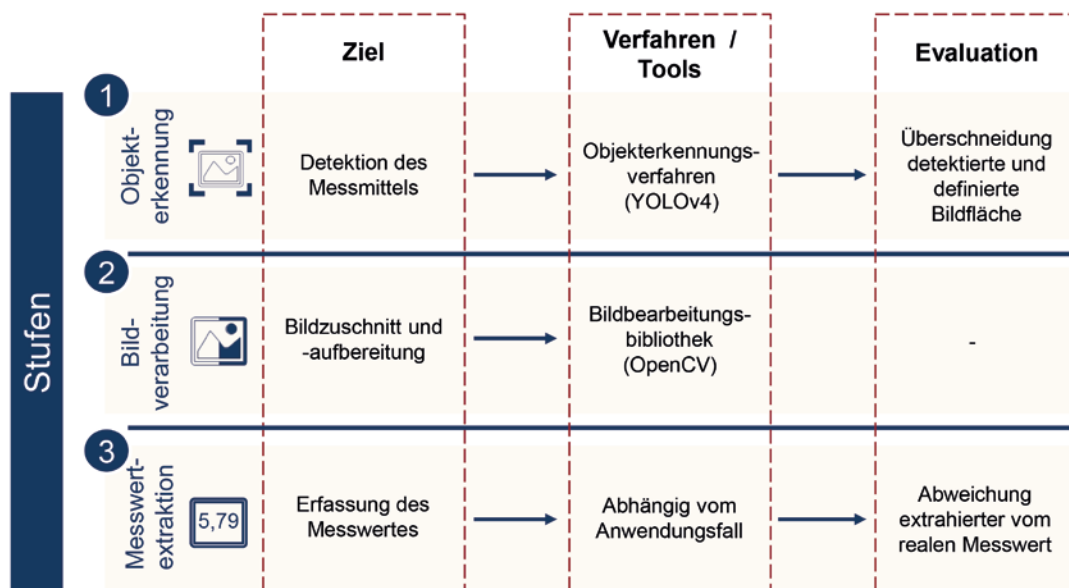
Hubert Würschinger, M. Sc. arbeitet als wissenschaftlicher Mitarbeiter am Lehrstuhl für Ressourcen- und Energieeffiziente Produktionsmaschinen an der Universität Erlangen-Nürnberg. Im Rahmen seiner Forschungstätigkeit untersucht er die Nutzung von druckluftinduzierten Strömungsgeräuschen für die Überwachung von Zerspanungswerkzeugen.

Prof. Dr.-Ing. Nico Hanenkamp leitet den Lehrstuhl für Ressourcen- und Energieeffiziente Produktionsmaschinen an der Universität Erlangen-Nürnberg. Schwerpunkte des Lehrstuhls stellen die Optimierung der zerspanenden Bearbeitung und die Gestaltung nachhaltiger Produktionssysteme dar.

Svyatoslav Funtikov, M. Sc. arbeitet als Consultant bei der ROI-EFESO MANAGEMENT CONSULTING AG und berät seine Klienten rund um Smart Factory Transformation, Digitalisierung & Industrie 4.0 und globale Footprint-Themen.

Matthias.Muehlbauer@fau.de
www.rep.tf.fau.de

Bild 1: Dreistufige Vorgehensweise zur optischen Messwertextraktion.



werden. Die Bildklassifizierung ordnet ein Bild einer oder mehreren Klassen (z. B. Auto, Flugzeug, Mensch) zu. Die Objekterkennung kann hingegen neben der Bildklassifikation auch eine Lokalisierung eines Objekts im Bild vornehmen. Die Position wird durch einen Begrenzungsrahmen (Bounding Box) ausgegeben. Weiter wird ein Vertrauenswert (Confidence-Score) ermittelt. Dieser Wert gibt Aufschluss darüber, wie sicher das System bei der Erkennung des jeweiligen Objekts ist. Zu den bekanntesten Objektdetektoren gehören die Algorithmen der You Only Look Once (YOLO)-Familie [2].

Weiter werden klassische Verfahren zur Bildanalyse, wie beispielsweise Filter [3] und Kantendetektoren [4] angewandt, um den identifizierten Bildausschnitt weiter zu verarbeiten. Für eine Einführung in diese Verfahren wird auf die angegebenen Quellen verwiesen.

Ein industrieller Einsatz der Textextraktion wurde beispielsweise von Muresan u. a. untersucht [5]. Ziel war das Auslesen der Beschriftung von Trinkflaschen mithilfe eines neuronalen Netzwerks. Zum Einsatz kam der Objektdetektor „Mask R-CNN“. Als vereinfachender Umstand war hierbei die Position der Flaschen im Bild bereits bekannt.

Vorgehensweise zur Messwertextraktion

Die optische Erfassung von visualisierten Messwerten auf Messmitteln im Rahmen manueller Prozesse geht mit mehreren Herausforderungen einher. Zum einen existieren eine Vielzahl unterschiedlicher Messmittel und Möglichkeiten der Messwertdarstellung. Zum anderen unterscheidet sich deren Einsatz im Rahmen manueller Prozesse deutlich hinsichtlich Betrachtungsperspektive, Belichtungsverhältnissen oder Reflexionen.

Zur Bewältigung dieser Herausforderungen wird eine dreistufige Vorgehensweise angewandt. Als Voraussetzung wurde eine Aufnahme des Messmittels aus einer egozentrischen Perspektive festgelegt. Dies deckt sich mit einem üblichen Mess- oder Ablesevorgang, da hierbei ebenfalls eine Sichtlinie zwischen Messmittel und dem Mitarbeiter bestehen muss. Die drei Stufen zur Messwertextraktion sind in Bild 1 veranschaulicht.

Durch die drei Stufen erfolgt eine schrittweise Lösung der Problemstellung. In der ersten Stufe wird das Ziel verfolgt, das Messmittel bzw. den relevanten Bereich (Region of Interest, ROI) im Bild zu lokalisieren. Im Anschluss wird der identifizierte Bildausschnitt aufbereitet. Dies stellt die Basis für die Messwertextraktion in der dritten Stufe dar.

Als Grundlage für die Durchführung der ersten Stufe "Objekterkennung" wird der gewählte Objektdetektor YOLOv4 [2] modelliert und trainiert. Für das Training muss ein Bild-Datensatz, welcher das relevante Objekt beinhaltet, gelabelt werden. Nach dem Aufteilen der Daten in einen Trainings-, Validierungs- und Testdatensatz, kann das Modell trainiert werden. Anschließend können Objekte auch auf unbekanntem Bildaufnahmen lokalisiert werden. Das Ergebnis ist die Detektion des relevanten Objekts und dessen Kennzeichnung mit einem Begrenzungsrahmen.

In der zweiten Stufe wird der identifizierte Bildbereich als Vorbereitung für die Messwertextraktion aufbereitet. Hierfür wird zunächst das Bild entsprechend des detektierten Begrenzungsrahmens zugeschnitten. Dies dient zur Entfernung irrelevanter Bereiche und somit der Reduktion der Fehleranfälligkeit der Folgeschritte. Weiter wird der detektierte Bereich hinsicht-

lich der Reduktion von Verdrehungen, der Bildschärfe, der Extraktion vorhandener Konturen und der Entfernung irrelevanter Informationen aufbereitet. Zur Durchführung der Bildbearbeitung wird die Bildverarbeitungsbibliothek OpenCV verwendet.

Nachdem der Bildausschnitt aufbereitet wurde, wird in der letzten Stufe der Messwert extrahiert. Die dritte Stufe variiert aufgrund der unterschiedlichen Darstellungsarten des Messwerts deutlich. Hierfür wurden dementsprechend anwendungsspezifische Konzepte erarbeitet. Diese werden im folgenden Abschnitt im Rahmen der beiden Anwendungsfälle erörtert.

Validierung anhand von zwei Anwendungsfällen

Für die Validierung der Vorgehensweise wurden zwei repräsentative Anwendungsfälle ausgewählt, welche zwei verbreitete Darstellungsarten von Messwerten abdecken:

- Analoges Manometer – Visualisierung des Messwerts anhand eines analogen rotatorischen Zeigers
- Digitaler Messschieber – Visualisierung des Messwerts anhand einer digitalen numerischen Anzeige

Anwendungsfall analoger Manometer

Häufig werden Messwerte über einen rotatorischen Zeiger visualisiert. Diese Messwertdarstellung wird im Anwendungsfall durch ein analoges Manometer adressiert. Folgende Abbildung visualisiert diesen Anwendungsfall und stellt die Messwertextraktion auf Basis der beschriebenen Vorgehensweise dar (Bild 2).

Wie bereits dargelegt, erfolgt in der ersten Stufe die Lokalisierung des Messmittels im Bild. Als relevanter Bereich wurde die Ablesefläche des Manometers definiert (Bild 2, links). In der zweiten Stufe wird der Bildbereich zugeschnitten und weiter aufbereitet. Die Messwertextraktion findet in der dritten Stufe statt. Für den Anwendungsfall „analoges Manometer“ wird eine geometrische Berechnung für die Erfassung des Messwerts verwendet.

Hierfür wird im ersten Schritt die Lage des Zeigers ermittelt. Dafür werden die vorhandenen Konturen in dem Bildbereich hervorgehoben. Folgend können der Mittel- bzw. Drehpunkt des Zeigers sowie die vorhandenen geraden Konturen, die den Zeiger selbst repräsentieren, ermittelt werden. Durch definierte Schwellenwerte,

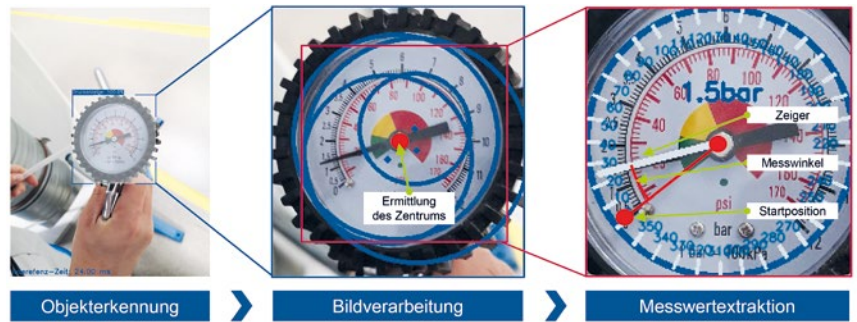


Bild 2: Visualisierung der Vorgehensweise für den Anwendungsfall analoger Manometer.

die den Zeigerabstand zum Zentrum als auch die minimale und maximale Länge des Zeigers umfassen, wird aus der Vielzahl der detektierten Konturen der tatsächliche Zeiger herausgefiltert. Auf Basis des detektierten Zeigers sowie durch den definierten Skalenbereich der Anzeige kann der Winkelversatz zur Startposition ermittelt und der Messwert berechnet werden.

Anwendungsfall digitaler Messschieber

Die Darstellung eines Messwerts durch eine digitale numerische Anzeige steht im Fokus des Anwendungsfalls „digitaler Messschieber“. Bild 3 visualisiert den Anwendungsfall sowie die dreistufige Vorgehensweise.

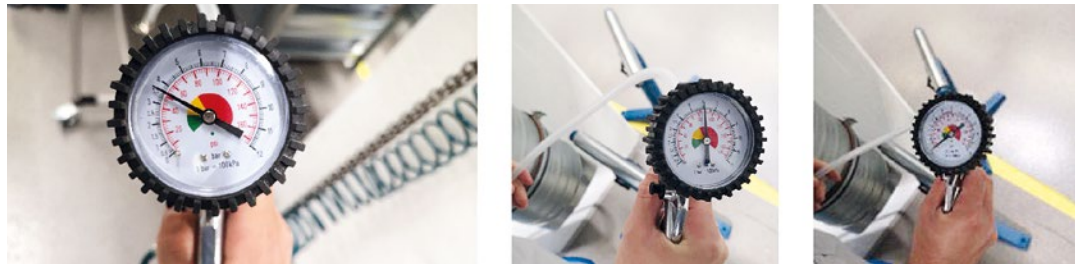
Die ersten beiden Stufen decken sich mit der bereits beschriebenen Vorgehensweise. Zur Reduktion der Fehleranfälligkeit werden in diesem Anwendungsfall zwei Bereiche des Messmittels durch den Objekterkennungsalgorithmus detektiert. Dies ist zum einen der vollständige Messschieber sowie zum anderen der Bildschirm für die Messwertdarstellung, der sich innerhalb des Begrenzungsrahmens des Messschiebers befindet bzw. befinden muss. Fehldetektionen von weiteren bildschirmähnlichen Objekten im Bild werden somit reduziert.

Für die Messwertextraktion wurden drei unterschiedliche Konzepte (Bild 3, rechts) entwickelt. Da die digitale numerische Anzeige der jeweiligen Ziffern auf sieben Segmenten beruht, wurde zunächst eine Analyse mithilfe eines Encoders realisiert. Hierfür werden die einzelnen Zif-

Bild 3: Visualisierung der Vorgehensweise für den Anwendungsfall digitaler Messschieber.



Bild 4: Beispielaufnahmen zum Anwendungsfall analoger Manometer.



fern-Segmente vorverarbeitet und für jede Ziffer (0 bis 9) das entsprechende Muster für die Decodierung hinterlegt.

Das zweite Konzept wird mit der Texterkennungs-Software „Tesseract“ umgesetzt. Die Herausforderungen stellen hierbei der Abstand der einzelnen Ziffern zueinander sowie der Abstand zwischen einzelnen Ziffer-Segmenten bei der vorliegenden Schriftart dar. Um eine Textextraktion zu ermöglichen, wird der Abstand zwischen einzelnen Ziffer-Segmenten mithilfe einer Bildbearbeitungsoperation in OpenCV „gefüllt“. Für die Messwertextraktion wurde ein neuronales Netz in Tesseract für die vorliegende Schriftart trainiert.

Im dritten Konzept „Doppel-YOLO“ wird das neuronale Netz YOLOv4 nicht nur zur Objekterkennung in der ersten Stufe, sondern auch für die Messwertextraktion verwendet. Zu diesem Zweck wurde für die Erkennung der Ziffern ein weiteres Modell erstellt. Der Vorteil dieses Konzepts liegt in der Wiederverwendbarkeit der Entwicklungen aus der ersten Stufe.

Für die Realisierung wurde das Edge-Gerät „OAK-1-POE“ von Luxonis gewählt [6]. Das Gerät ist trotz seiner kompakten Größe in der Lage komplexe neuronale Netze sowie weitere Vor- und Nachbearbeitungsschritte in Echtzeit auszuführen. Mittelfristig wird jedoch der Einsatz in Kombination mit einer Daten-Brille mit einer Vernetzung zu einem zentralen EDV-System angestrebt.

Diskussion der Ergebnisse

Folgend werden die erzielten Ergebnisse für die beiden Anwendungsfälle dargestellt und disku-

tiert. Eine kritische Analyse der entwickelten Vorgehensweise erfolgt anschließend.

Evaluation Anwendungsfall analoger Manometer

Zur Evaluierung des Anwendungsfalls analoger Manometer wurden 360 unbekannte Bildaufnahmen herangezogen. Die Bildaufnahmen stellen typische Halteposition des Messmittels, aufgenommen aus einer egozentrischen Perspektive, dar (Bild 4).

Die Analyse der Ergebnisse für den Anwendungsfall „analoger Manometer“ zeigt gute Resultate über alle Stufen. Das Messmittel wird in allen Aufnahmen richtig detektiert. Für die Auswertung der Messwertextraktion in Stufe 3 wurde eine Toleranz von +/- 5 % festgelegt. Es kann festgestellt werden, dass der detektierte Messwert bei 96 % der analysierten Aufnahmen innerhalb der definierten Toleranz liegt. Als Unterstützung bei manuellen Ablesevorgängen wird die erreichte Genauigkeit als ausreichend eingestuft. Wird eine Messwernerfassung ohne weitere Kontrollinstanz angestrebt, muss eine weitere Verbesserung der Messwertextraktion erreicht werden.

Als Ursachen für die Fehldetektionen bzw. negative Einflussfaktoren auf die Genauigkeit sind eine starke Neigung, Drehung oder Rotation des Manometers im Raum zu nennen. Durch die Software werden aktuell noch keine derartig abweichenden Aufnahmeperspektiven des Haltewinkels des Messgeräts ausgeglichen. Um diese Fehlerquelle zu reduzieren, können weitere Kalibrierungsmaßnahmen integriert werden. Die vorgestellte Vorgehensweise kann echtzeit-

Bild 5: Beispielaufnahmen zum Anwendungsfall digitaler Messschieber.



1. Stufe - Objekterkennung			3. Stufe - Messwertextraktion		
Konzept	Richtig erkannte Aufnahmen		Konzept	Richtig erkannte Aufnahmen	
	Absolut	Relativ		Absolut	Relativ
YOLOv4	1044	100 %	Encoder	715	68 %
			Tesseract	684	66 %
			YOLOv4	960	92 %

fähig mit ca. 10 FPS (Bilder pro Sekunde) auf dem Edge-Gerät ausgeführt werden.

Evaluation Anwendungsfall digitaler Messschieber

1044 Bildaufnahmen dienten zur Evaluation des Anwendungsfalls digitaler Messschieber. In Bild 5 sind exemplarische Aufnahmen, die für die Bewertung herangezogen wurden, visualisiert. Zur Erstellung eines repräsentativen Test-Datensatzes wurden Aufnahmen von einem ruhenden wie auch von einem bewegten Messmittel mit unterschiedlichen Abständen, Winkeln und Belichtungsverhältnissen erstellt.

Die erzielten Ergebnisse sind in Bild 6 dargestellt. Auch bei diesem Anwendungsfall wird das Messmittel in der ersten Stufe in allen Aufnahmen richtig erkannt. Bei der Messwertextraktion in Stufe 3 werden ermittelte Messwerte als richtig klassifiziert, wenn sie exakt dem dargestellten Wert entsprechen. Somit müssen alle dargestellten Ziffern richtig erfasst werden. Es ist zu erkennen, dass durch die drei alternativen Konzepte deutlich unterschiedliche Ergebnisse erzielt werden.

Die Konzepte „Encoder“ und „Tesseract“ konnten bei 68 % bzw. 66 % der Aufnahmen die Messwerte richtig detektieren. Wesentliche Herausforderungen bei diesen Konzepten sind die hohe Varianz der Lichtverhältnisse sowie ungünstige Blickwinkel auf das Messmittel, welche durch die Bildverarbeitung in Stufe 2 nicht vollständig kompensiert werden können.

Beim Konzept „Doppel-YOLOv4“ wird bei 92 % der Aufnahmen der Messwert richtig erfasst. Die deutlich besseren Ergebnisse im Vergleich zu den beiden alternativen Konzepten können insbesondere auf folgende Aspekte zurückgeführt werden: Die Besonderheit des Konzepts von zwei neuronalen Netzen liegt in der Verkürzung der zweiten Stufe. Hier findet lediglich ein Bildzuschnitt statt. Es entfällt somit eine mögliche Fehlerquelle, was zu einer höheren Robustheit des Systems sowie zu einem verringerten Berechnungsaufwand beiträgt. Darüber hinaus konnten bei den Trainingsdatensätzen viele Extremfälle bereits berücksichtigt sowie durch

das Augmentieren der Aufnahmen eine künstliche Erweiterung des Datensatzes durchgeführt werden. Dies hat weiter zur Steigerung der Robustheit beigetragen.

Die erzielten Ergebnisse bei den Konzepten „Tesseract“ und „Encoder“ werden als noch nicht ausreichend zur Unterstützung für Mess- und Ablesevorgänge eingestuft. Mit einer richtigen Detektion des Messwerts bei 92 % der Bildaufnahmen kann die Güte des Konzepts „Doppel-YOLOv4“ jedoch als ausreichend für die Unterstützung bei manuellen Ablesevorgängen bewertet werden. Ähnlich wie beim ersten Anwendungsfall wird eine weitere Optimierung der Messwertextraktion als erforderlich bewertet, sollte keine weitere Kontrollinstanz zur Verfügung stehen. Das Konzept „Doppel-YOLOv4“ kann auf dem Edge-Gerät ebenfalls mit einer Geschwindigkeit von ca. 10 FPS echtzeitfähig angewandt werden.

Zusammenfassung

Im Rahmen des Beitrags wurde eine generische Vorgehensweise zur Extraktion von bereits visualisierten Messwerten aus Bildaufzeichnungen dargestellt. Hierdurch sollen die Digitalisierungslücke, welche durch noch nicht vernetzte Messmittel entsteht, geschlossen und Mitarbeiter bei der Durchführung von Mess- und Ablesevorgängen unterstützt werden. Die Vorgehensweise wurde anhand von zwei repräsentativen Anwendungsfällen validiert. Dies ist zum einen die Erfassung eines digital dargestellten Messwerts bei einem Messschieber und zum anderen die Digitalisierung eines analog durch einen Zeiger visualisierten Messwerts bei einem Manometer. Die Ergebnisse zeigen eine zufriedenstellende und echtzeitfähige Erkennung der jeweiligen Messwerte. Wird jedoch eine Messwertextraktion ohne weitere Kontrollinstanz angestrebt, ist eine weitere Optimierung der Erkennung zu empfehlen. Weiter weist die Vorgehensweise hohes Skalierungspotenzial auf. Der Änderungsaufwand für eine Übertragung auf ähnlich gelagerte Anwendungsfälle wird als vertretbar bewertet.

Schlüsselwörter:

Digitalisierung, Bildverarbeitung, Maschinelles Lernen, Objekterkennung

Bild 6: Ergebnisse Anwendungsfall digitaler Messschieber.

Literatur

- [1] Schenk, M.: Produktion und Logistik mit Zukunft. Heidelberg 2015.
- [2] Wang, C.-Y.; Bochkovskij, A.; Liao, H.-Y. M.: Scaled-YOLOv4: Optimal Speed and Accuracy of Object Detection. 2020.
- [3] Caldeira, T.; Ciarelli, P. M.; Neto, G. A.: Industrial Optical Character Recognition System in Printing Quality Control of Hot-Rolled Coils Identification. In: Journal of Control, Automation and Electrical Systems 31 (2020) 1, S. 108-118.
- [4] Wang, H.; Tian, Q.; Hu, Z.; Guo, J.: Image Feature Detection Based on OpenCV. In: Journal of Research in Science and Engineering (JRSE) ISSN: 1656-1996 Volume-2, Issue-7, July 2020.
- [5] Muresan, M. P.; Szabo, P. A.; Nedevschi, S.: Dot Matrix OCR for Bottle Validity Inspection. In: 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP): IEEE, 2019 - 2019, S. 395-401.
- [6] Luxonis: Dokumentation OAK. URL: <https://docs.luxonis.com/projects/hardware/en/latest/>, Abrufdatum 15.08.2022.