

KI für die Produktion – Iterative und interaktive Verbesserung der KI

Interaktives ML zur Adaption von KI-Systemen

M. Köhler, C. Jauch, D. Karelina

ZUSAMMENFASSUNG Flexible Produktionsprozesse und das Beherrschen von Variantenvielfalt werden für Unternehmen immer wichtiger. Künstliche Intelligenz unterstützt Unternehmen, um trotz dieser Anforderungen effizient zu agieren. Allerdings sind der Datenbedarf und die Intransparenz eine Hemmschwelle für den Einsatz von KI. Hierfür kann interaktives Maschinelles Lernen eine Lösung sein. Die Einbindung von Domänenexperten in das Training ermöglicht es, die Akzeptanz des Systems zu verbessern, während der Einsatz von Active Learning den Datenbedarf senkt. Montage und Qualitäts sicherung sind nur zwei Einsatzbereiche.

STICHWÖRTER

Künstliche Intelligenz, Montage, Qualitätssicherung

Interactive ML for the adaptation of AI systems

ABSTRACT Flexible production processes and mastering a wide range of variants are becoming increasingly important for companies. Artificial intelligence helps companies to operate efficiently despite these requirements. However, data requirements and a lack of transparency are a barrier to the use of AI. Interactive machine learning can be a solution to this. Involving domain experts in training can improve acceptance of the system, while the use of active learning reduces data requirements. Assembly and quality assurance are just two areas of application.

1 Einleitung

Unternehmen in der Wirtschaftswelt stehen aktuell vor der Herausforderung, ihre Produktion immer flexibler gestalten zu müssen. Produktwechsel müssen schnell erfolgen und das Umsetzen individueller Kundenwünsche erfordert, die damit einhergehende steigende Variantenvielfalt zu beherrschen [1]. Vor diesem Hintergrund kann der Einsatz von Künstlicher Intelligenz (KI) Unternehmen unterstützen, effizient und agil zu handeln [2].

Generell lässt sich zwischen Expertensystemen und allgemeinen KI-Systemen unterscheiden. Während Expertensysteme spezialisiertes Wissen in einem eng umgrenzten Bereich anbieten, sind allgemeine KI-Systeme darauf ausgelegt, ein breiteres Spektrum an Problemen zu bewältigen. Zudem wird zwischen generativen und diskriminativen Ansätzen unterschieden: Generative KI-Modelle sind in der Lage, selbstständig neue Daten zu generieren, während diskriminative Modelle darauf spezialisiert sind, zwischen verschiedenen Daten zu unterscheiden. Ein aktuelles Beispiel für allgemeine und generative KI ist ChatGPT [3], ein Sprachmodell, das versucht, die verschiedensten Sprachenfragen möglichst natürlich zu beantworten. Beispiele für ein diskriminatives Expertensystem sind eine End-of-Line-Qualitätskontrolle oder eine Wareneingangskontrolle.

Um zuverlässig und performant zu sein, brauchen diskriminative Expertensysteme viele Daten für das Training. Da sich die Produktpalette oft ändert, tritt dieser Bedarf ebenso häufig und in wiederkehrenden Zyklen auf, sodass KI-Experten das bestehende KI-System regelmäßig anpassen müssen. Insbesondere kleinere

Unternehmen sind dabei auf externe Dienstleister angewiesen [4]. Denn für den Domänenexperten, zum Beispiel den Werker, der keinen KI-Hintergrund hat, ist die Funktionsweise des KI-Systems oft undurchsichtig. Die Interaktion zwischen Mensch und KI-System ist im Idealfall für den Regelbetrieb optimiert. Das Einbeziehen des Domänenexperten in den Trainingsvorgang des KI-Systems ist meist nicht vorgesehen. Dabei könnte diese Interaktion die Akzeptanz eines KI-Systems verbessern, indem die Bedürfnisse der Nutzenden berücksichtigt werden und ein besseres Verständnis des KI-Systems sowie seiner Funktionsweise erreicht wird. Dies ist besonders effektiv, wenn der Domänenexperte in den Trainingsprozess eingebunden wird und direkt mit der KI interagiert [5].

2 KI und Mensch Hand in Hand

KI bietet neue Möglichkeiten, die Arbeitswelt nachhaltig zu verändern. Bei der Entscheidung, welche Aufgabenstellung mit KI gelöst werden soll, ist es wichtig, die Grenzen von KI zu betrachten. Diese Grenzen sind darin bedingt, dass KI auf statistischen Lernverfahren basiert. Das heißt, die KI hängt von großen Datensätzen ab und kann nur lernen, was in den Daten vorhanden ist. Trotz dieser Einschränkungen bietet KI enorme Chancen, besonders in der Ausführung spezifischer Expertenaufgaben, bei denen sie menschliche Fähigkeiten übertreffen kann. Beispielsweise kann KI Muster und Zusammenhänge in Daten erkennen, die Menschen kaum oder nur mit sehr viel Erfahrung wahrnehmen können. Diese Fähigkeit ist Chance und Risiko zugleich. Sie ermög-

licht zwar die Umsetzung neuer Anwendungen, gleichzeitig ist dem Nutzenden aber eventuell nicht klar, warum etwas nicht erkannt wurde, da die Muster für den Menschen nicht ersichtlich sind. Abhilfe kann hier die genannte Interaktion zwischen Mensch und KI schaffen: Die KI versucht zu erklären, welche Muster für die Entscheidung ausschlaggebend waren und der Mensch bewertet, ob die KI korrekt lernt.

Diese Symbiose zwischen Mensch und KI kann auf verschiedene Weise gestaltet werden. Zum einen können KI-Systeme Entscheidungen treffen, die der Mensch anschließend überprüft und kontrolliert, sei es durch eine vollständige Überprüfung, Stichproben oder die Prüfung des vom System Gelernten. Das Feedback kann iterativ für die Weiterentwicklung des KI-Systems verwendet werden. Zum anderen kann die KI unterschiedliche Vorschläge unterbreiten, aus denen der Mensch den passenden auswählen muss. Auch diese Rückmeldung kann in das KI-System zurückgegeben werden. Diese Zusammenarbeit ermöglicht es, bei sinnvoller Gestaltung des Gesamtsystems die Stärken sowohl der KI als auch des Menschen bestmöglich zu nutzen. Voraussetzung für sinnvolle Vorhersagen und Entscheidungen des KI-Systems ist eine breite Datenbasis, auf der die KI trainiert wurde.

Der immense Datenbedarf ist ein Problem für die industrielle Anwendung. Vor allem bei einem Produktwechsel stehen vom aktualisierten oder neuen Produkt noch sehr wenige Daten zur Verfügung. Zudem müssen die Mitarbeitenden individuell geschult werden. Dies ist eine Aufgabe, die im Idealfall ein KI-gestütztes digitales Assistenzsystem übernehmen kann. Vor allem in Zeiten des Arbeitskräftemangels können solche Assistenzsysteme dazu beitragen, Arbeitszeit zu sparen und schnell neues Wissen zu vermitteln. Oft sind Assistenzsysteme aufgrund der schlechten Datenlage genau dann, wenn der Schulungsbedarf bestünde, noch nicht einsatzbereit. Darin liegt die zu lösende Herausforderung: Die Bereitstellung ausreichender und relevanter Datenmengen in Hochlaufphasen, um eine KI effektiv zu trainieren. Interaktives maschinelles Lernen (iML) ist eine Technologie, die in solchen Szenarien besonders wertvoll sein kann, um Expertensysteme zielgerichtet zu verbessern.

3 Interaktives Maschinelles Lernen

3.1 Was ist interaktives Maschinelles Lernen?

Interaktives maschinelles Lernen (iML) gibt dem Menschen deutlich mehr Kontrolle über die KI als die typischen maschinellen Lernansätze. Der hohe Datenbedarf sowie die Intransparenz darüber, was genau das Modell gelernt hat (Stichwort „Black Box“), sind Gründe, warum man KI wenig vertraut und es nicht gerne verwendet wird oder oft nicht auf Akzeptanz stößt [6]. Beim iML interagiert ein Mensch mit dem ML-Modell und ist fest in der Lernschleife des Prozesses eingebunden. Ein solcher Lernprozess wird auch als „Human in the Loop“ – der Mensch in der Schleife – bezeichnet. Ziel ist es, ein KI-Modell schrittweise zu optimieren, indem das Modell nach dem Lernen Feedback über seinen Zustand gibt. Der Anwendende interpretiert dieses Feedback und kann der KI dann zielgerichtet Daten zurückgeben, bei denen sie noch nicht gut abgeschnitten hat. Somit entsteht eine Interaktion zwischen Mensch und KI, die dem Menschen die Kontrolle über die Black Box zurückgibt.

Diese Interaktion ist ressourcenschonender als ein herkömmliches ML-Training, da der Mensch stets eingebunden ist und bei einem zufriedenstellenden Ergebnis das Training jederzeit beendet

den kann. Da diese Interaktion iterativ stattfindet, kann der Nutzende ein eigenes Verständnis dafür entwickeln, wie das KI-Modell funktioniert. Während der einzelnen Schleifen wird nicht nur das KI-Modell verbessert, zudem nähert sich das Verständnis des Nutzenden immer weiter dem KI-Modell an. Dies wird auch als Bildung eines mentalen Modells beim Nutzenden bezeichnet. So entsteht Vertrauen in das Ergebnis und die Intransparenz wird aufgelöst.

Historisch entstand iML aus Reinforcement Learning (RL) und später aus Active Learning (AL) [7]. Reinforcement Learning ist ein Teilbereich des maschinellen Lernens, in dem es darum geht, einen Agenten durch die Interaktion mit seiner Umwelt zu belohnen oder zu bestrafen. Das Feedbacksignal führt zu einem Lerneffekt, ähnlich dem klassischen Pavlov'schen Behaviorismus bei Menschen oder Tieren. Im Gegensatz zu RL geht es bei AL nicht darum, wie etwas gelernt wird, sondern mit welchen Daten gelernt wird [8]. Bei einem AL-Ansatz werden relevante und wichtige Datenpunkte beispielsweise dem RL-Agenten oder dem ML-Modell gezeigt, sodass möglichst effizient und effektiv gelernt werden kann. Relevante Datenpunkte bilden dabei vorrangig Beispiele, die besonders informativ für das KI-Modell sind oder bei deren Entscheidung sich das KI-Modell besonders unsicher ist. Zeigt man dem KI-Modell etwa immer wieder ähnliche Datenpunkte, so kann das Modell später nur eine sehr schmale Bandbreite an Datenpunkten innerhalb des Anwendungsbereichs verstehen. Werden jedoch automatisch unähnliche und möglichst diverse Datenpunkte gezeigt, wird das KI-Modell auch später eine größere Bandbreite an unbekannten Datenpunkten vorhersagen können. Diese Vorhersagen basieren auf den erlernten Mustern und Eigenschaften des Modells. Active Learning ist die automatische Auswahl solcher relevanten Datenpunkte.

Durch den Einsatz von AL in einem iML-Tool werden dem Nutzenden Datenpunkte priorisiert angezeigt, welche für das Modell relevanter sein können als andere. Hierzu können in einer einfachen Ausbaustufe zum Beispiel die Vorhersagen des Modells auf den unbekannten Daten verwendet werden. Dies hat den Vorteil, dass der zu annotierende Anteil des Datensatzes für eine gleichbleibende Modellgenauigkeit sinkt. Gleichzeitig wird der Mensch selbst zum sogenannten „Active Learner“, weil er dem Modell relevante Punkte zeigt, von denen er denkt, dass sie der KI helfen, das Problem besser zu verstehen. Das prinzipielle Vorgehen der Interaktion bei der Verwendung von iML ist in **Bild 1** dargestellt.

Damit die Feedbackschleife zwischen Mensch und KI nicht unterbricht, muss die KI Feedback über ihren Zustand geben. Dies kann sowohl über Statistiken gemacht werden als auch über direktes Feedback zu bestimmten Datenpunkten. Wie gut lief beispielsweise die Objekterkennung auf einem Bild, oder wie gut konnte eine E-Mail als Spam erkannt werden? Dies gibt dem Menschen die Chance, entweder selbst oder über AL-Vorschläge zu erkennen, wo das Modell noch unsicher ist. Die Unsicherheit wird dabei über den vom Modell berechneten Konfidenzwert quantifiziert.

Zusätzlich gibt es die Möglichkeit, dass der Mensch über die (eigene) Interpretation der Ergebnisse der erklärbaren KI (xAI) weiter aufschlüsselt, warum eine bestimmte Entscheidung getroffen wurde und warum bestimmte Datenpunkte noch nicht gut genug erkannt werden können. Genau diese Datenpunkte werden dann für das Modell klar beschrieben, indem man zum Beispiel auf dem Bild bestimmte Stellen markiert, die Aufschluss geben,

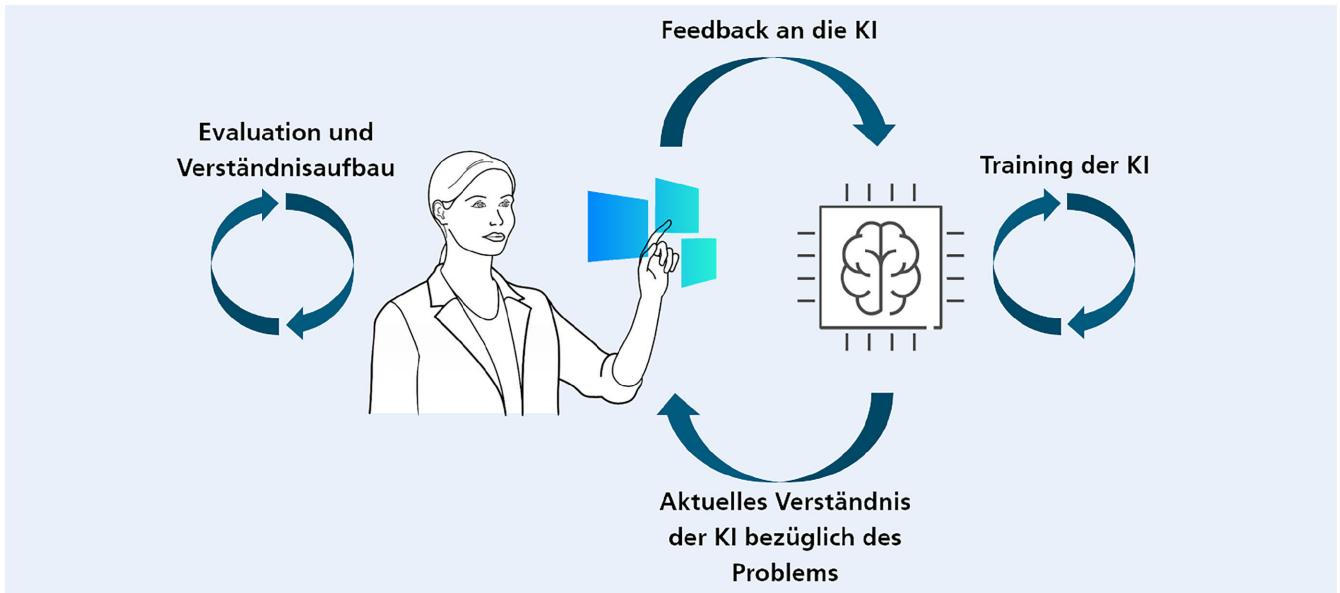


Bild 1. Prinzipielles Vorgehen für interaktives maschinelles Lernen (iML). *Grafik: Fraunhofer IPA*

warum dieses Objekt auf dem Bild zu der Kategorie gehört. Wenn die KI Kratzer auf einer Oberfläche erkennen soll, sollte der Mensch diesen Kratzer markieren können und diese Markierung muss wiederum die KI verstehen. Auf diesen neuen Informationen kann das ML-Modell erneut trainiert werden, was es schrittweise verbessert.

Im Bereich maschinelles Sehen gibt es wenige Vorreiter, die sich der iML-Umsetzung gewidmet haben. So wurde „Jaaba“ [9] entwickelt, um die Annotation von Tierbewegungen wie etwa bei Insekten für Biologen zu vereinfachen und schnell mithilfe von KI klassifizieren zu können. „Ilastik“ [10] ist ein Beispiel aus der Medizin, das Ärzte dabei unterstützt, medizinische Bilder teilweise automatisch zu annotieren und typische ML-Algorithmen ohne Programmierkenntnisse auszuführen. Diese Anwendungen haben gemeinsam, dass sie Domänenexperten unterstützen, ihr Fachwissen in ML-Modelle zu integrieren, ohne dass sie selbst dafür ML-Experten werden müssen. Die Domänenexperten lernen Schritt für Schritt, worauf es ankommt, um die KI-Modelle selbst zu verbessern. Diese Ko-Evolution zwischen Mensch und KI ist der große Vorteil, wenn man Mensch und Maschine zusammenbringt.

3.2 Anwendungsfälle von IML in der Produktion

In der Produktion wird eine datenarme und durch den Domänenexperten rekonfigurierbare KI-Lösung benötigt. Mit dieser können auch in KMU viele Anwendungsfälle mit KI-basierter Komponente so umgesetzt werden, dass sie noch nach mehreren Produktzyklen rentabel sind. Mit dem Wissen über iML und den bestehenden Lösungen zeigt sich neben dem Bedarf und dem Lösungsansatz auch eine Lücke. Zwar existieren bereits Lösungen für iML, diese sind jedoch nicht einfach auf die Anwendungsfälle in der Produktion übertragbar. Zudem wird teils ausführliches Vorwissen in Statistik vorausgesetzt (etwa bei Jaaba) oder die Lösungen sind für komplexere Aufgaben als notwendig ausgelegt (Bildsegmentierung bei Ilastik).

Daher wurde am Fraunhofer IPA ein iML-Tool für Anwendungsfälle in der Produktion entwickelt. Mit dem Tool kann ein

Domänenexperte direkt eine Objekterkennung trainieren. **Bild 2** zeigt die Oberfläche des iML-Tools, mit welcher der Anwendende das KI-Modell verbessern kann.

Der Anwendende bekommt oben links Informationen über die KI-Vorschläge der erkannten Objekte und darunter vom aktuellem Modell vor-annotierte Bilddaten angezeigt. Diese sind durch AL vorsortiert, um schnell signifikante Verbesserungen des KI-Modells zu erreichen und datenarm zu einem finalen KI-Modell zu gelangen. Wählt der Nutzende hier ein Bild aus, annotiert er dieses manuell in der Mitte des Bildschirms. Dabei werden ihm oben rechts die im aktuellen Bild annotierten Objekte sowie ihre Klasse angezeigt. Die Erklärbarkeit kann mit einer Schaltfläche unter dem Bild angezeigt werden. Der Nutzende gibt durch diese Anpassung der Annotation in der Mitte der KI Feedback und kann dann mit einem Klick das nächste Training anstoßen, in dem dieses Feedback verarbeitet wird. Ist dieses abgeschlossen, kann er das neue Modell auswählen, bekommt die Information zur Performance des neuen Modells präsentiert und kann dann entweder den Prozess abbrechen oder weiter annotieren und auf den vorangegangenen Ergebnissen aufbauen.

Dieses Tool kann bestehende Systeme zur optischen Qualitätskontrolle ergänzen. Für das KI-Modell sind unter anderem einzelne Bauteile beziehungsweise Bauteilgruppen sowie Kratzer oder andere Qualitätsmakel erlernbar. Dies kann zum Beispiel in einer Wareneingangskontrolle eingesetzt werden. Dort kann mithilfe des iML-Tools schnell auf Änderungen durch die Zulieferer reagiert werden. Das bestehende System dient als Basis und wird um weitere Daten durch die neuen Zulieferteile ergänzt. Dadurch kann das bestehende KI-System die umfassende Wareneingangsprüfung der neuen Teile ebenfalls durchführen. Statt der Wareneingangsprüfung ist so auch eine Endkontrolle umsetzbar und damit die KI-gestützte Qualitätsprüfung schnell an neue Produkte adaptierbar.

In der (End-)Montage kann solch ein System auch zur Anpassung eines KI-basierten Assistenzsystems eingesetzt werden, das direkt eine Fehlerprüfung während der Montage durchführt. Der Domänenexperte kann bei Teileänderungen im Montageprozess mithilfe des iML-Systems direkt das Assistenzsystem anpassen,

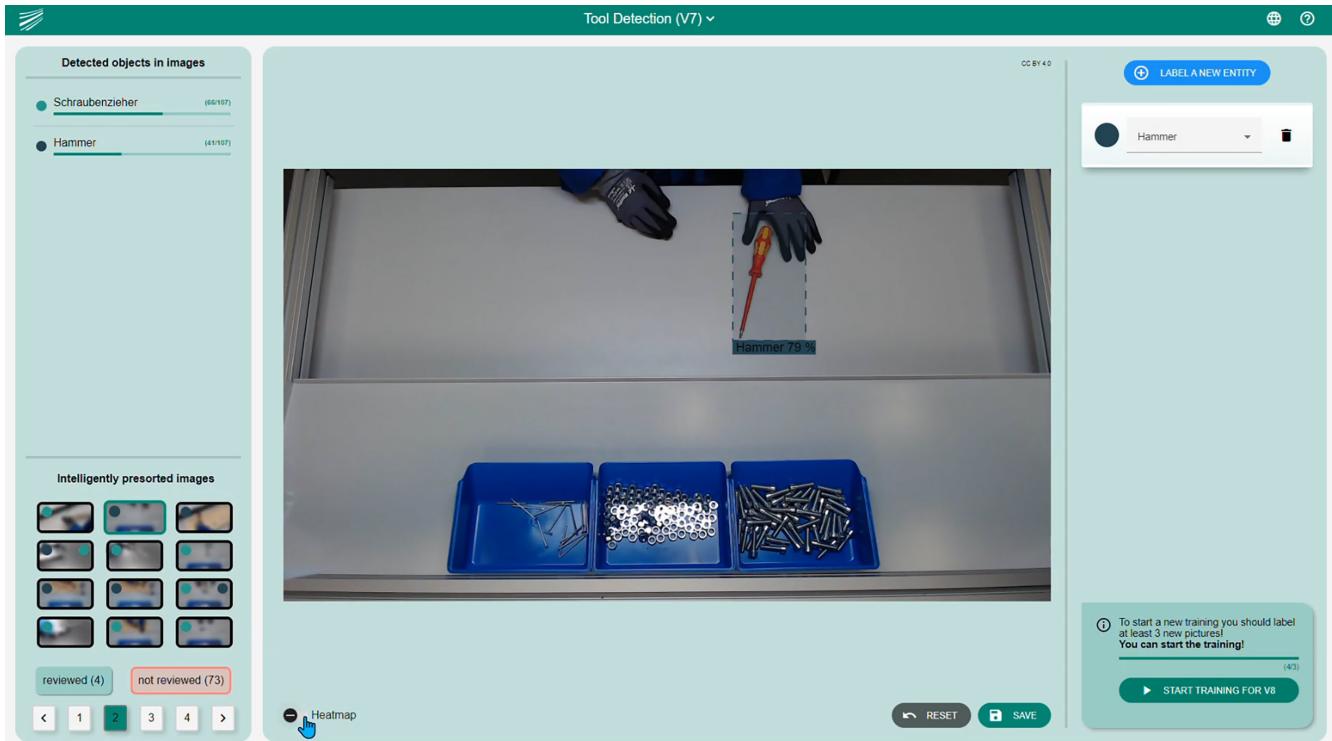


Bild 2. Oberfläche des am Fraunhofer IPA entwickelten iML-Tools. Foto: Fraunhofer IPA

ohne dass er einen KI-Experten hinzuziehen muss. Durch den Einsatz einer KI-basierten Lösung können im Assistenzsystem deutlich größere Freiheitsgrade in der Montage entstehen, sei es in der Belichtung oder Positionierung des Bauteils.

4 Fazit

In der Produktion sind KI-Systeme bisher nur selten vertreten. In Unternehmen mit häufig wechselnden Produkten sind Hemmnisse für den Einsatz von KI-basierten Assistenzsystemen der hohe Datenbedarf und die Intransparenz von KI-Systemen. Interaktives maschinelles Lernen zeigt, wie eine sinnvolle Einbindung des Menschen in den Trainingsprozess diese Nachteile stark reduzieren kann. Dabei wird ein Domänenexperte ohne KI-Kenntnisse dazu befähigt, eigenständig das KI-System auf neue Gegebenheiten anzupassen und schrittweise zu verbessern. Das parallel sich bildende Verständnis des Menschen, der iterativ mit dem iML-Tool interagiert, schafft Vertrauen und Transparenz. Durch die AL-Methoden wird zudem der Datenbedarf reduziert.

Bestehende iML-Lösungen sind meist noch nicht für den Einsatz in der Produktion vorgesehen. Dabei könnte der Einsatz dieses in der Wissenschaft bewährten Verfahrens in der Produktion helfen, die Kosten bei Prozessänderungen, zum Beispiel durch die hohe Produktvielfalt, gering zu halten. Das am Fraunhofer IPA entwickelte iML-Tool versucht, diese Lücke zu schließen.

Aktuell beschränkt sich das iML-Tool auf die Objekterkennung und kann damit bereits viele Anwendungsfälle unterstützen. Im nächsten Schritt wird das iML-Tool um eine menschliche Aktivitätserkennung erweitert, um weitere Anwendungsfälle in der manuellen Montage zu ermöglichen. Dies soll die Domänenexperten unterstützen, ihre eigenen Arbeitsprozesse in ein Assistenzsystem mithilfe von iML eintrainieren zu können.

L iteratur

- [1] Reinhart, G.: Handbuch Industrie 4.0. Geschäftsmodelle, Prozesse, Technik. München: Carl Hanser Verlag 2017
- [2] Bitkom Research: Stimmen Sie den folgenden Aussagen zur Künstlichen Intelligenz in der Arbeitswelt zu? Internet: de.statista.com/statistik/daten/studie/1176941/umfrage/vorstellungen-von-kuenstlicher-intelligenz-am-arbeitsplatz-in-deutschland/. Zugriff am 30.07.2024
- [3] OpenAI; Achiam, J.; Adler, S. et al.: GPT-4 Technical Report. Proceedings, arXiv 2023, doi.org/10.48550/arXiv.2303.08774
- [4] Lundborg, M.; Gull, I.: Künstliche Intelligenz im Mittelstand. So wird KI für kleine und mittlere Unternehmen zum Game Changer. Stand: 2021. Internet: www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/ki-Studie-2021.pdf?__blob=publicationFile&v=1. Zugriff am 30.07.2024
- [5] Guo, L.; Daly, E. M.; Alkan, O. et al.: Building Trust in Interactive Machine Learning via User Contributed Interpretable Rules. Proceedings of the 27th International Conference on Intelligent User Interfaces (IUI '22), New York/USA: ACM 2022, pp. 537–548, doi.org/10.1145/3490099.3511111
- [6] Wondimu, N. A.; Buche, C.; Visser, U.: Interactive Machine Learning: A State of the Art Review. arXiv 2022, doi.org/10.48550/arXiv.2207.06196
- [7] Holzinger, A.: Interactive machine learning for health informatics: when do we need the human-in-the-loop? Brain Informatics 3 (2016) 2, pp. 119–131
- [8] Zhan, X.; Wang, Q.; Huang, K. et al.: A Comparative Survey of Deep Active Learning. arXiv 2022, doi.org/10.48550/arXiv.2203.13450
- [9] Kabra, M.; Robie, A. A.; Rivera-Alba, M. et al.: JAABA: interactive machine learning for automatic annotation of animal behavior. Nature Methods 10 (2013) 1, pp. 64–67
- [10] Sommer, C.; Straehle, C.; Kothe, U. et al.: Ilastik: Interactive learning and segmentation toolkit. IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Chicago, IL/USA 2011, pp. 230–233, doi.org/10.1109/ISBI.2011.5872394



Martina Köhler, M.Sc.

martina.köhler@ipa.fraunhofer.de

Tel. +49 711 / 970-1832

Foto: Autorin

Christian Jauch, M.Sc.

Dascha Karelina, M.Sc.

Fraunhofer-Institut für Produktionstechnik

und Automatisierung IPA

Nobelstr. 12, 70569 Stuttgart

www.ipa.fraunhofer.de

LIZENZ



Dieser Fachaufsatz steht unter der Lizenz Creative Commons
Namensnennung 4.0 International (CC BY 4.0)