

Bildererkennung zur Prozessautomatisierung am Beispiel eines medizinischen Labors

KI-basierte Bildanalyse zur Prozessautomatisierung

C. Hofmann, M. Benfer, G. Lanza, T. Rüdiger

ZUSAMMENFASSUNG In pathologischen Laboren werden Gewebeproben in einem Produktionsprozess in mehreren Schritten für die anschließende Diagnose aufbereitet. Während ein Großteil der Prozesse automatisiert stattfindet und Robotik in der Handhabung Einzug hält, ist der erste Arbeitsschritt des Zuschnitts noch durch manuelle Schritte geprägt. Dieser Beitrag stellt eine KI-basierte Automatisierung der Probenbeschreibung und -transkription dar. Damit beschreibt der Beitrag einen wichtigen Schritt zur Verbesserung heterogener Prozesse in der Pathologie.

STICHWÖRTER

Digitalisierung, Industrie 4.0, Künstliche Intelligenz

AI-based image analysis for process automation

ABSTRACT In pathological laboratories, tissue samples are prepared for subsequent diagnosis in a multi-step production process. While most of the processes are automated and use robotics for handling, the initial sample sectioning is still largely manual. This paper introduces an AI-based automatization of the sample description, including categorization and transcription. Thus, this work represents an important step towards improving heterogeneous processes in pathological laboratories.

1 Einleitung

1.1 Histologische Untersuchung im Pathologielabor als Produktionsprozess

Die Pathologie gewinnt aus Gewebeproben, die zu diagnostischen oder therapeutischen Zwecken entnommen wurden, Informationen. Häufig wird anhand einer kleinen Probe (Biopsie) die Art einer Erkrankung bestimmt, um Aussagen über die Notwendigkeit und mögliche Optionen einer Behandlung zu treffen. Die aus dem Gewebe gewonnenen Erkenntnisse werden als Befund dokumentiert und haben maßgeblichen Einfluss auf den weiteren Therapieverlauf. Gerade in der Onkologie (Tumormedizin) ist die Genauigkeit dieser Befunde entscheidend. Der Laborprozess gliedert sich in sieben aufeinanderfolgende Schritte (**Bild 1**).

Nach Eingang und Erfassung der Probe erfolgt zunächst die makroskopische Beschreibung und Vorbefundung. Ziel ist es, festzulegen, welche Teile des Materials für die feingewebliche (histologische, mikroskopische) Analyse ausgewählt werden. Dazu wird die Probe zugeschnitten: ein Arbeitsschritt, der das Präparat unwiderruflich verändert und bei dem die makroskopischen Merkmale verloren gehen, sofern sie nicht zuvor dokumentiert wurden. Anschließend wird das zugeschnittene Gewebestück oder die Biopsie in eine Probekassette eingelegt, entwässert, in Paraffin eingebettet, ausgehärtet und in etwa 4 µm dünne Schnitte zerteilt. Diese histologischen Schnitte werden eingefärbt, auf Glasobjektträger aufgebracht und unter dem Mikroskop ausgewertet. Abschließend wird der Befund schriftlich festgehalten und an die anfordernde Stelle übermittelt. [1]

Während die meisten Laborprozesse bereits stark automatisiert sind und Handhabungsschritte ebenfalls mithilfe kollaborativer Robotik eine Automatisierung erfahren [2] bleibt Schritt 2 des Prozesses, die makroskopische Beschreibung und der Zuschnitt, weitestgehend manuell und ohne technische Unterstützung. Dieser Umstand ist dem zugrunde liegenden Produktionsschritt geschuldet. Während in den späteren Schritten die Gewebeproben bereits eine einheitliche, gut automatisierbare Gestalt aufweisen, liegen die Gewebeproben in der makroskopischen Untersuchung noch in ihrer heterogenen Ausgangsform vor.

Im manuellen Prozess werden die eingesandten Gewebeproben von Hand aus dem Probenbehälter entnommen und es wird ein Foto aufgenommen, dann wird die Probe vermessen. Anschließend wird eine makroskopische Beschreibung diktiert, die später transkribiert und in den Befund aufgenommen wird. Nach der makroskopischen Beschreibung wird das Gewebe in Scheiben geschnitten und in Probekassetten gelegt. Je nach Gewebe werden mehrere Probekassetten benötigt.

Dieser Beitrag stellt Ansätze vor, die zeigen, wie diese heterogenen Prozesse mittels KI automatisiert werden können. Für die Untersuchung wird die häufig vorkommende Gewebeart der Hautproben genutzt. Die vorliegenden Arbeiten zeigen, wie:

- die makroskopische Beschreibung automatisiert erstellt werden kann,
- die Verwechslung von Gewebeproben durch die Qualitätssicherung ausgeschlossen werden kann.

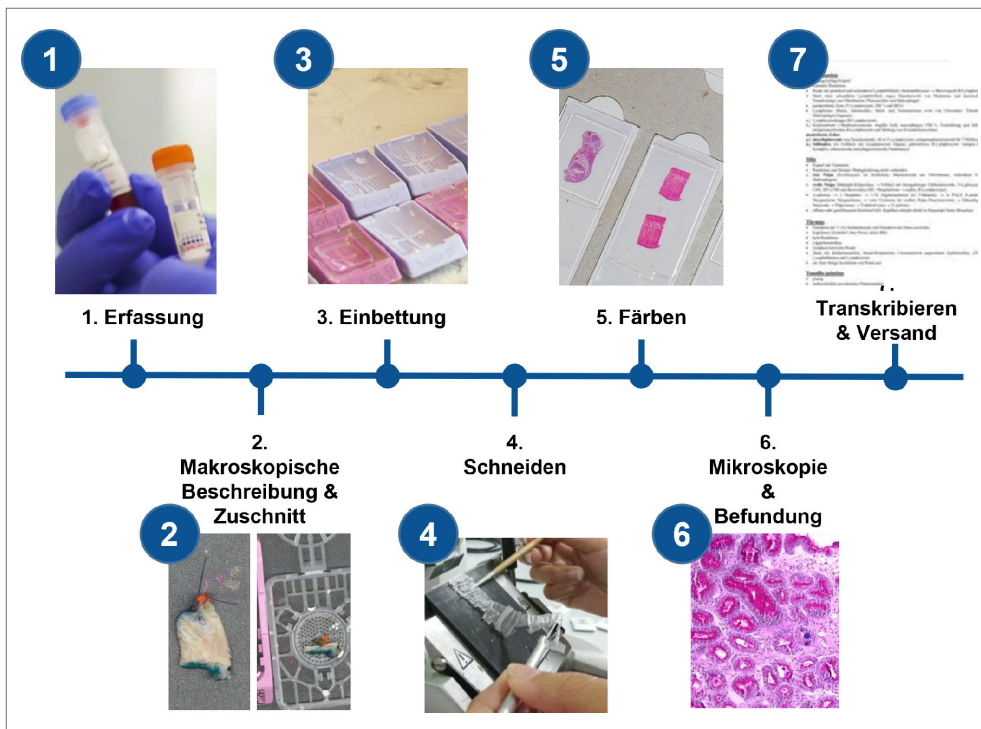


Bild 1 Der Produktionsprozess eines Pathologielabors.
Grafik: asmedos

1.2 KI-basierte Ansätze in der Pathologie

Während der betrachtete Arbeitsschritt heute manuell und mit geringer technischer Unterstützung durchgeführt wird, sind KI-Methoden im späteren Diagnostikprozess bereits verbreitet.

Der Fokus der Forschung im Bereich der Bilderkennung für die Pathologie liegt auf der automatisierten Auswertung der "whole-slide images" zur Erkennung von Krebserkrankungen [3, 4]. Dabei werden die etwa 100-fach vergrößerten farbigen Bilder der Gewebeschnitte analysiert, um unterschiedliches Gewebe zu erkennen und Erkrankungen zuzuordnen. Dabei sind die zu verarbeitenden Bilder aufgrund der Färbung und Vergrößerung im Vergleich zur Radiologie um ein Vielfaches größer. Dies erfordert besonders effiziente Modelle und stellt hohe Anforderungen an die IT-Infrastruktur, um Dateien dieser Größe zu versenden und langfristig zu archivieren [4].

Im Bereich des Zuschnittsarbeitsplatzes gibt es zwei Entwicklungen: Erstens, Ansätze zur Automatisierung des gesamten Prozesses und, zweitens, Ansätze zur Unterstützung des bestehenden Prozesses. Technische Ansätze zur Prozessautomatisierung erfordern hohe Investitionen und lohnen sich erst ab entsprechend hohem Durchsatz. Ein Beispiel ist die Automatisierungslösung von Inveox, die eingesendete Probe aus dem Einsendebehälter auspackt und vermisst [5]. Zu den Ansätzen der Prozessunterstützung gehören Verfahren, die das Diktat der makroskopischen Befundung mit Speech-to-Text-Verfahren umgehend verschriftlichen [6]. Unterschiedliche Anbieter bieten dafür Softwarelösungen an. Der Zuschnittsarbeitsplatz ist typischerweise aber hoher Geräuschbelastung der vorgeschriebenen, leistungsstarken Formalinabzüge ausgesetzt. Diese Hintergrundgeräusche machen die Spracherkennung herausfordernd. Die Spracherkennung erfordert außerdem weiterhin manuellen Aufwand durch Ärztinnen und Ärzte, deren Kapazität sehr begrenzt ist.

Das vorgestellte Konzept soll durch KI-Einsatz Arbeitsabläufe automatisieren und Ärztinnen und Ärzte im Zuschnitt entlasten,

ohne große technische Eingriffe und große Folgekosten zu erfordern.

1.3 KI-Anwendung in der optischen Prüfung heterogener Produkte

Im Remanufacturing werden gebrauchte Produkte mit unbekanntem Zustand untersucht, müssen identifiziert und hinsichtlich ihres Zustands bewertet werden. Dieser Anwendungsfall zeigt einige Parallelen zum betrachteten Laborbetrieb. In beiden Fällen sind die zu befundenden Objekte sehr heterogen [7]. Entsprechend müssen die verwendeten Methoden gut generalisieren können. KI-basierte Verfahren mit guter Generalisierungsfähigkeit werden zum Beispiel erfolgreich in der Defekterkennung eingesetzt [8–10].

2 KI-basierte makroskopische Befundung

Bild 2 präsentiert schematisch das Vorgehen zur KI-basierten Befundung. Ausgangspunkt des Verfahrens bildet ein Foto des Arbeitsbereichs. Dieses Foto zeigt neben den relevanten Objekten, der Gewebeprobe und der Kassette weitere Gegenstände, die für die weitere Analyse nicht relevant sind. Dies sind beispielsweise Lineale, Skalpelle, Zellstofftücher und Probenhalter. Nach einer technischen Bildvorverarbeitung folgt eine Objekterkennung, um die Bildbereiche mit relevanten Objekten zu erkennen und zu lokalisieren. Es werden drei relevante Klassen unterschieden: (i) Hautexzizat, (ii) Kassette und (iii) Transportbehälter.

Nach erfolgter Objekterkennung werden die erkannten Objektklassen unterschiedlich weiterverarbeitet. Für die Gewebeproben folgt eine Größenermittlung, um die Länge und die Breite zu ermitteln. Dann werden die Gewebeproben nach der Form und der Position einer Markierung des Gewebes mittels Faden klassifiziert. Zusammen können die Ergebnisse genutzt werden, um eine makroskopische Beschreibung zu generieren.

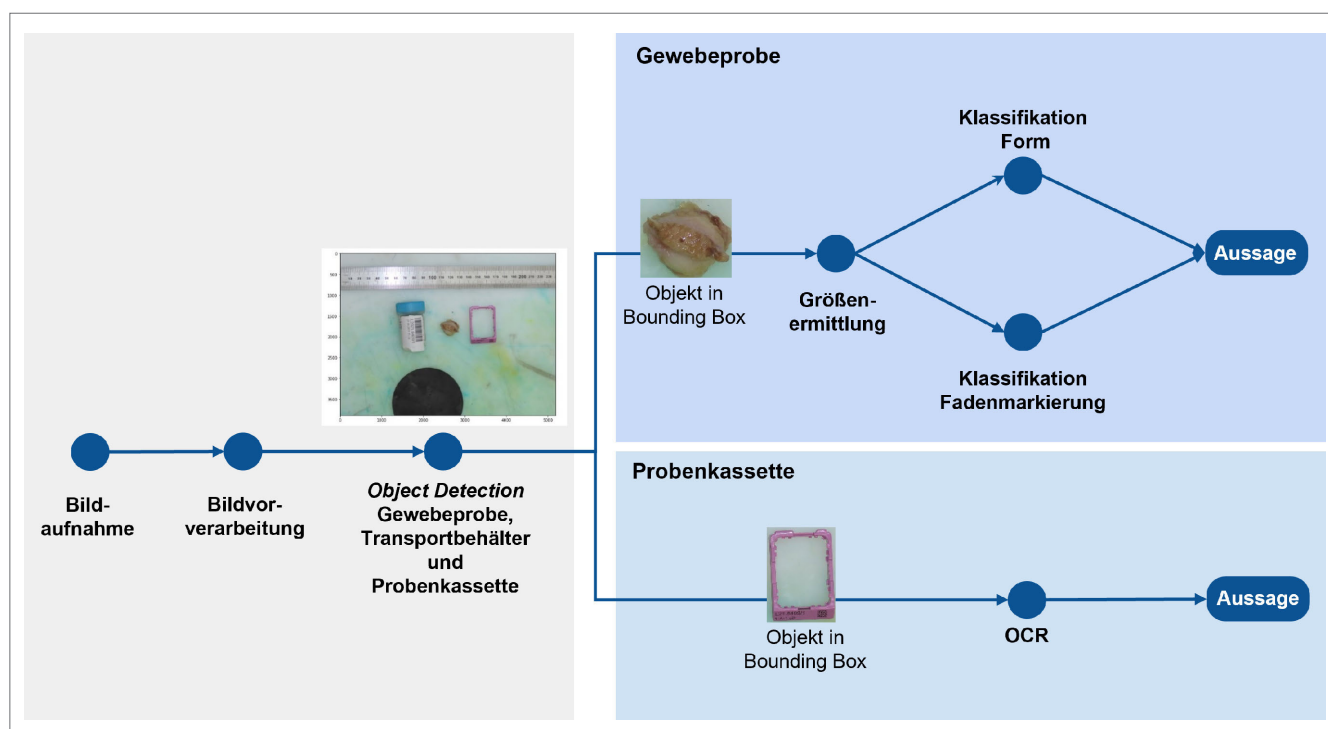


Bild 2 Schematisches Vorgehen zur Erstellung der makroskopischen Befunde. Grafik: asmedos

Nach dem Zuschneiden werden die Gewebeproben auf Kassetten verteilt. Die richtige Zuordnung der Gewebeschnitte auf die Kassetten ist für die korrekte Zuordnung von Diagnosen zu Fällen essenziell. Aus diesem Grund wird die Nummer auf der Kassette am Arbeitsplatz mit der erwarteten Nummer abgeglichen. So kann eine Vertauschung erkannt werden, die etwa aus dem Verbleib noch ungenutzter Kassetten des vorangegangenen Falls resultiert.

Dieser Beitrag behandelt vorrangig die Schritte der Objekterkennung und der Klassifikation, da diese beiden Schritte besonders herausfordernd sind.

2.1 Objekterkennung

Die Objekterkennung (Englisch: Object Detection) identifiziert und klassifiziert Gegenstände, Gewebeproben und Kassetten auf dem Foto. Die verwendeten Bilder wurden statisch von oben mit einem Abstand von circa 60 cm zum Arbeitsplatz aufgenommen, um eine hohe Flexibilität der Arbeiten zu gewährleisten. Die Gewebeproben waren zwischen 5 mm und 20 mm groß und machten somit nur einen geringen Anteil an der Bildfläche aus.

Für die Object Detection wurden Bilder eines Datensatzes aus insgesamt zwei Jahren annotiert. Der Trainingsdatensatz enthielt die vier Klassen Gewebeprobe, Probenbehälter, Kassette und Lineal. Der Datensatz bestand aus 1000 Bildern mit etwa 2200 Objekten. Es wurden ausschließlich Gegenstände gelabelt, welche vollständig im Bild waren, um Fehlklassifikationen zu vermeiden. Die Bounding Box, welche die Gegenstände umschließt, wurde möglichst eng um die Gegenstände gezogen. Waren Bilder nicht eindeutig, wurden sie nicht in den Datensatz aufgenommen.

Der Datensatz enthielt 40 % Gewebeprobe, 31,5 % Lineal, 17,6 % Kassette und 10,9 % Probenbehälter. Die Klasse Gewebeprobe wies die größte Variabilität auf, da die Gewebeproben individuell, die anderen Gegenstände hingegen immer exakt gleich

waren. Der Datensatz wurde gemäß Stand der Technik in ein Trainingsdatensatz, ein Validierungsdatensatz und ein Testdatensatz unterteilt. Dabei wurde sichergestellt, dass das Auftreten der vier Klassen in den drei Datensätzen dem Auftreten im Gesamtdatensatz entsprach. Das Trainingsdatensatz wurde genutzt, um im Transfer Learning an einem vortrainierten Netz die Klassen zu erlernen. Das Validierungsdatensatz diente der Hyperparameteroptimierung, während das Testdatensatz zur Ermittlung der Güte verwendet wurde.

Die Bewertung der Object Detection erfasst die korrekte Zuordnung des Objekts zu seiner Klasse und die korrekte Lokalisierung. Beide Aspekte lassen sich mithilfe der Maßzahl Intersection over Union (IoU) erfassen. Die IoU ist definiert als der Quotient aus der Schnittmenge der tatsächlichen Lokalisation mit der vorhergesagten Lokalisation, dividiert durch die Vereinigungsmenge der beiden Flächen. Ein Objekt gilt als korrekt lokalisiert, wenn der IoU über einem definierten Schwellenwert liegt. Als weitere relevante Maßzahlen werden Precision und Recall sowie Average Precision und Average Recall genutzt. Die Precision gibt an, wie viele der Vorhersagen aus allen Vorhersagen zutreffend sind. Der Recall hingegen gibt an, wie viele korrekte Vorhersagen im Vergleich zur zugrunde liegenden Wahrheit (Englisch: ground truth) korrekt erfolgen. Die Average Precision ist die Fläche unter der Kurve, wenn Precision als Funktion des Recalls ausgedrückt wird. Diese Maßzahl liegt zwischen 0 und 1.

Um das Lernverhalten des Verfahrens zu bewerten, wurde entsprechend dem Stand der Technik die Verlustfunktion betrachtet. Erprobt wurden unterschiedliche Varianten eines „ResNet-50 FPN“ [11] über die Bibliotheken „PyTorch“ beziehungsweise „Detecto“. Die Verlustfunktion zeigte bei einem Training über 14 Epochen und einer Batch-Size von 4 einen idealtypischen Verlauf. Die Verlustfunktion sank in den ersten Epochen exponentiell und verlief dann auf einem gleichbleibend niedrigen Niveau.

Vorhersage im Vergleich mit den tatsächlichen Werten

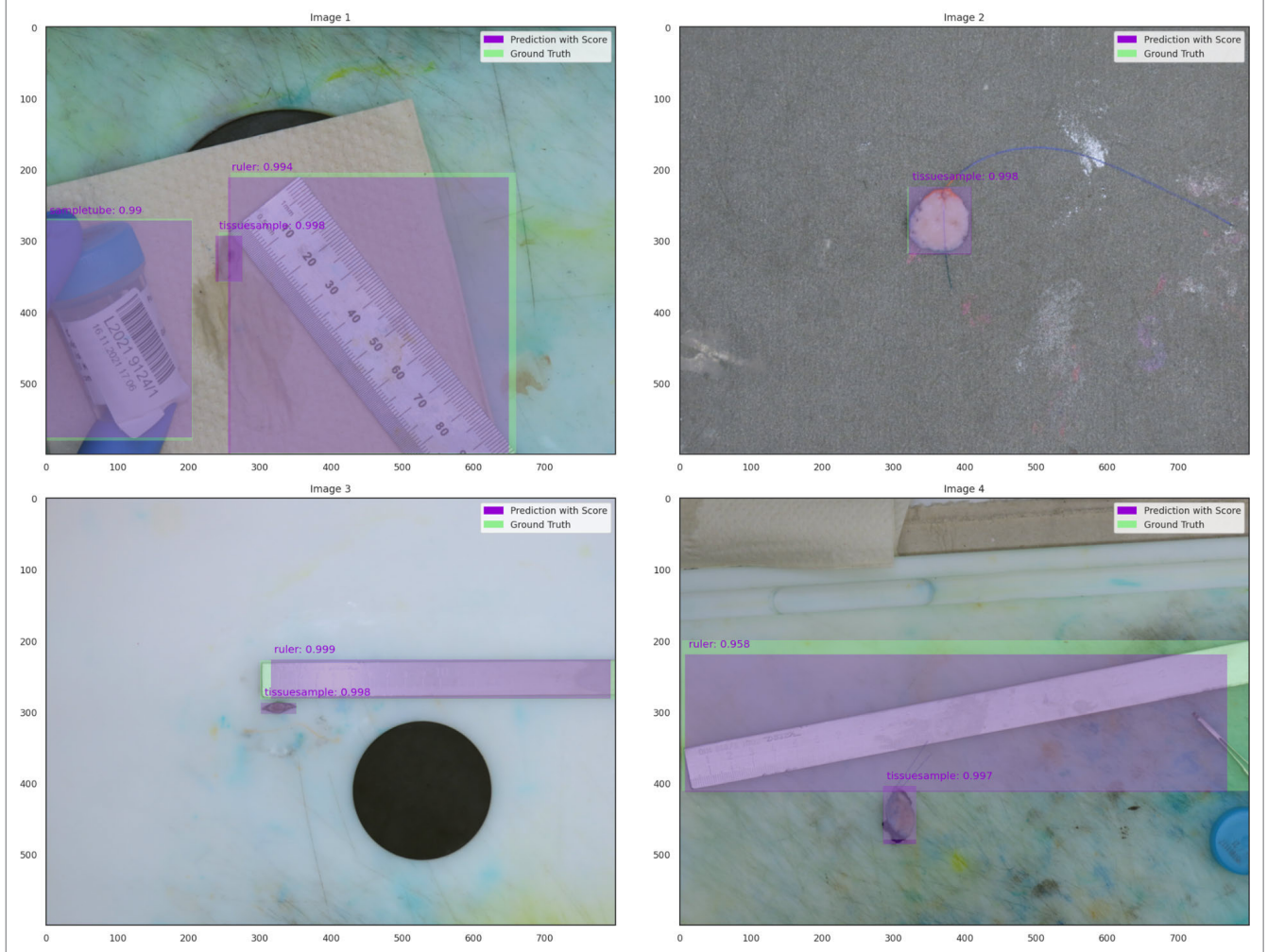


Bild 3 Klassifikation und Lokalisierung der erkannten Gegenstände anhand eines Beispielbildes. Grafik: asmedos

Die Güte wurde für unterschiedliche Anforderungsniveaus hinsichtlich der Überdeckung, gemessen als IoU, ermittelt. Es zeigte sich der erwartbare Zusammenhang, dass bei höherer IoU-Anforderung die Güte sinkt. Die Werte für die Average Precision waren folgendermaßen:

- Kasette [95,7%–98%]
- Lineal [93,4%–97,9%]
- Probenbehälter [92,9%–99,9%]
- Gewebeprobe [80,8%–98,6%]

Die Gewebeprobe, welche naturgemäß die höchste Variabilität aufweist, erzielte die geringste Average Precision, während die standardisierten Gegenstände besser erkannt wurden. Insgesamt erzielte die Objekterkennung eine ausreichend hohe Güte, sodass in den weiteren Schritten auf diesen Arbeiten aufgebaut werden kann. **Bild 3** zeigt die Ergebnisse anhand eines Beispielbildes.

2.2 Größenermittlung

Für die Größenermittlung wurden die Ergebnisse aus der Objekterkennung der Klasse Gewebeprobe genutzt. Die Bestimmung der Länge und Breite der Gewebeprobe erfolgte entweder anhand eines bekannten Referenzobjekts im Bildbereich oder mithilfe der

Kameraparameter. Da eine Prämisse der Arbeiten ein möglichst geringer Eingriff in das Arbeitsumfeld war, kam der zweite Ansatz zum Einsatz, bei dem der statische und bekannte Abstand der Kamera zum Arbeitsplatz genutzt wird.

Der Quotient aus der Breite des Kamerasensors und dem Abstand zwischen Kamera und Arbeitsplatz entspricht dem Quotienten aus der Brennweite und dem Bildbereich der Aufnahme. Aufgrund der bekannten Werte lässt sich der Bildbereich der Aufnahme als geometrische Größe bestimmen.

2.3 Klassifikation nach Grundform und Position der Fadenmarkierung

Für die makroskopische Befundung sind zwei Klassifikationsschritte nötig. Erstens die Klassifikation nach der Grundform und zweitens die Klassifikation nach der Position der Fadenmarkierung.

2.3.1 Klassifikation der Grundform

Für die Klassifikation nach Form wurden drei grundlegende Klassen definiert:

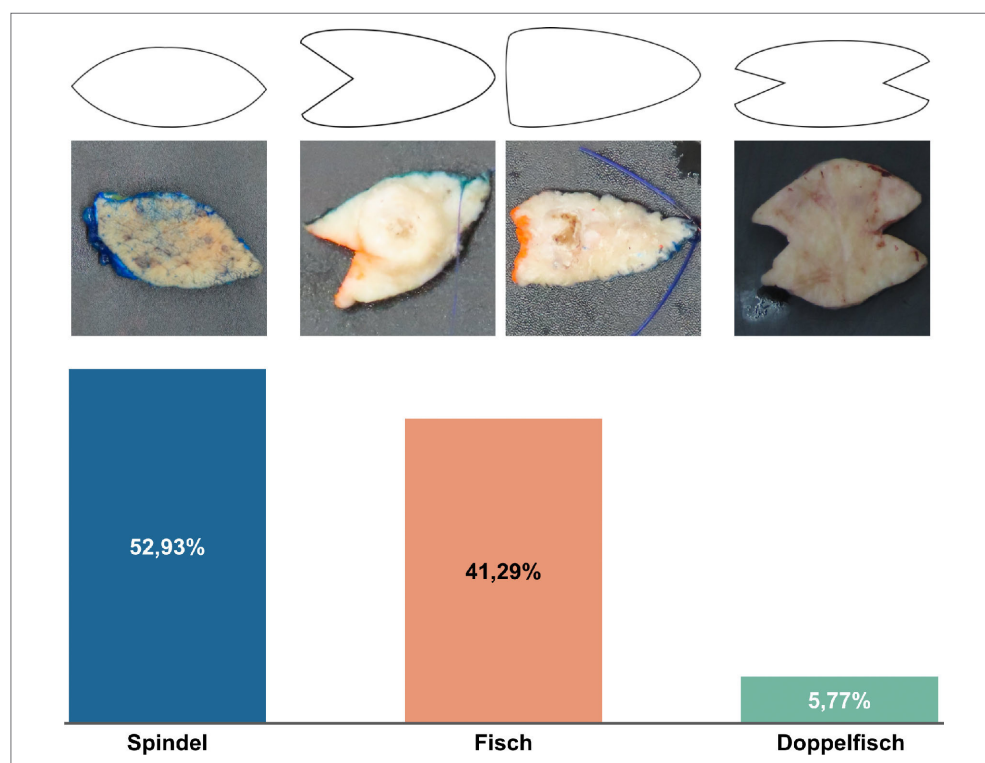


Bild 4 Auftreten der Formklassen im Datensatz. Grafik: asmedos

1. Spindel: Diese Form ist punktsymmetrisch und elliptisch, wobei die langen Seiten spitz zulaufen.
2. Fisch: Diese Form ähnelt der „Spindel“, weist jedoch zwei Varianten auf:
 - a. Die Spindel ist einseitig eingeschnitten, sodass eine Öffnung entsteht, die an ein Fischmaul erinnert.
 - b. Statt eines Einschnitts ist eine der langen Seiten gerade abgeschnitten.
3. Doppelfisch: Form mit zwei Einschnitten an den kurzen Seiten.

Als erster Schritt für die Klassifikation wurde ein annotierter Datensatz erstellt. Dazu wurden etwa 14 000 Bilder analysiert und ein Datensatz aus etwas mehr als 1000 annotierten Bildern erstellt. Der Datensatz wies eine deutliche Unausgewogenheit auf. Die Klassen Spindel (53 %) und Fisch (41 %) waren annähernd gleich stark vertreten, während die Klasse Doppelfisch mit lediglich 6 % deutlich unterrepräsentiert war (**Bild 4**).

Es wurden alle vorhandenen Bilder in den Grunddaten (14 000 Bilder) genutzt, dennoch war der Anteil der Klasse Doppelfisch im resultierenden Datensatz im einstelligen Prozentbereich. Dieser niedrige Anteil war für das Training des KI-Modells problematisch, da das Modell eine hohe Vorhersagegüte erreichen kann, indem die häufig vorkommenden Klassen prognostiziert werden.

Um diesem Problem zu begegnen, kam das Verfahren des „Oversamplings“ zum Einsatz. Da nicht mehr Bilder der Klasse Doppelfisch vorhanden waren, mussten zusätzliche Bilder generiert werden. Dafür wurde Data Augmentation eingesetzt. Dabei wurde jedes vorhandene Bild der Klasse Doppelfisch auf neun verschiedene, zufällige Weisen modifiziert. Diese Modifikationen umfassten unter anderem Farbveränderungen und Verzerrungen, sodass eine größere Vielfalt an Trainingsdaten entstand. Durch diesen Ansatz wurde sichergestellt, dass keine identischen Bilder mehrfach in den Trainings-, Validierungs- oder Testdatensatz

gelangten, um eine ungewollte Verzerrung der Ergebnisse zu vermeiden.

Für die Klassifikation wurde ein ResNet-50 genutzt [11]. Aus der Literatur ist bekannt, dass bereits vortrainierte Netze ohne Anpassungen bei hochspezialisierten Bildern wie den vorliegenden Gewebeproben keine guten Ergebnisse liefern, da das Training mit Gegenständen des alltäglichen Lebens vorgenommen wurde. Im Rahmen des Trainings wurden die Gewichte des Netzes neu bestimmt, sodass sich das künstliche neuronale Netz an die Vorhersage der drei Klassen anpasste (Transfer Learning). Im Training wurde fünffache Kreuzvalidierung genutzt, um die Abhängigkeit von der zufälligen Aufteilung in Trainings-, Validierungs- und Testdatensätze zu verringern. Hierbei wurden fünf unterschiedliche, zufällige Datensätze jeweils bestehend aus Trainings-, Validierungs- und Testdatensätzen, erstellt. Für jeden dieser Datensätze wurde ein Modell trainiert. Zur Bewertung wurde das arithmetische Mittel der jeweils betrachteten Maßzahl aus den fünf Modellen gebildet.

Die Ergebnisse zeigen, dass die zusätzlichen Maßnahmen zur Erhöhung des Anteils der Klasse Doppelfisch in den Daten erforderlich waren. Wenn alle Klassen gemäß ihrem Auftreten im Datensatz vertreten sind, werden nur 9 % der Doppelfische als solche erkannt. Durch den Einsatz von Oversampling lassen sich die Ergebnisse deutlich optimieren. Dabei wurden 80 % der Doppelfische, 94 % der Fische und 99 % der Spindeln korrekt identifiziert. Verwechslungen traten fast ausschließlich zwischen den Kategorien Doppelfisch und Fisch sowie zwischen Fisch und Spindel auf (**Bild 5**).

2.3.2 Klassifikation der Positionierung der Fadenmarkierung am Hautexzizat

Die Bestimmung der Position der Fadenmarkierung spielt eine entscheidende Rolle, um eine präzise Positionsangabe für die

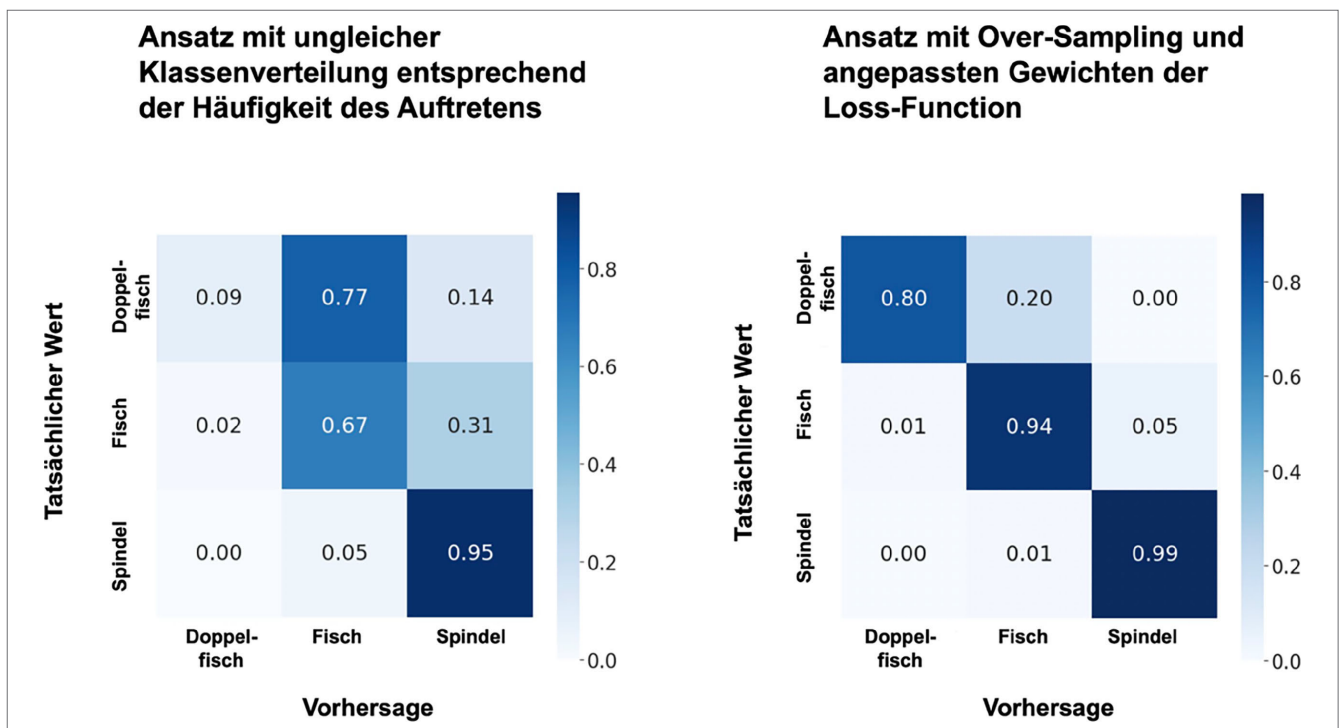


Bild 5 Klassifikation der Grundform. Grafik: asmedos

Diagnose zu ermöglichen. Diese Information hilft den Klinikern, die Lokalisation, die in den Diagnosen verwendet wird, der genauen Entnahmestelle des Gewebes zuzuordnen. Dazu werden die Gewebeprobe mit einem verknöteten Faden markiert. Die Rotationsposition des Fadens wird in 30°-Schritten angegeben.

Die Positionsangabe erfolgt relativ zur Entnahmestelle: Die Angabe „Faden auf 16:00 Uhr“ bedeutet etwa, dass der Faden bei 120° in Bezug auf den höchsten Punkt der Entnahmestelle gesetzt wurde. In der Praxis wird oft pauschal eine 12:00-Uhr-Position angenommen. Erst während der Befunderstellung wird diese angenommene Position auf die tatsächliche Position umgerechnet.

Zur Klassifikation der Fadenposition wurden zwei Ansätze untersucht:

1. Grobe Klassifikation in vier Kategorien – die Fadenmarkierung wird einer von vier Klassen zugeordnet:
 - unmarkiert
 - Seite
 - Spitze
 - Fischmaul

Dieser Ansatz setzt voraus, dass die Position „12:00 Uhr“ als Standard angenommen wird.

2. Feinere Klassifikation in 12 Winkelklassen – die Fadenmarkierung wird direkt einer von zwölf Rotationsklassen zugewiesen: ([0–30°), [30–60°), [60–90°), ..., [330–360°)).

Um eine korrekte Einordnung zu gewährleisten, muss die Gewebeprobe entsprechend der auf dem Einsendebogen angegebenen Position ausgerichtet werden. Ist dort beispielsweise „16:00 Uhr“ vermerkt, muss der Faden für das Foto zwischen 120° und 150° positioniert sein.

Für die Klassifikation der Fadenmarkierung nach dem ersten Ansatz wurde erneut ein Datensatz erstellt und annotiert, der in Bild 6 dargestellt ist. Hierbei wurden die vier genannten Klassen unterschieden, die wiederum in unterschiedlicher Häufigkeit auftraten.

Zur Modellierung kam erneut ein ResNet-50 zum Einsatz, wobei Transfer Learning angewendet wurde. Die erste Testreihe ohne Anpassungen zur Bewältigung der unausgeglichene Klassen zeigte Optimierungspotenzial: Zwar wurden insgesamt 83 % der Bilder korrekt klassifiziert, doch die seltenen Klassen Fischmaul und Seite wurden nur unzureichend erkannt. Die Klasse Seite erreichte eine Erkennungsrate von lediglich 56 %, wobei die häufigste Verwechslung mit der Klasse Spitze auftrat. Noch problematischer war die Klassifikation der Klasse Fischmaul, die nur in 33 % der Fälle richtig zugeordnet wurde – auch hier kam es häufig zu Verwechslungen mit der Klasse Spitze.

Durch den Einsatz von Random Oversampling sowie einer angepassten Loss Function zur besseren Berücksichtigung der seltenen Klassen konnten die Ergebnisse erheblich verbessert werden. Die Average Precision stieg auf 90,7 %. Vor allem die Erkennung der Klasse Seite verbesserte sich deutlich auf 90 % (vorher 56 %), während die Klasse Fischmaul nun in 75 % der Fälle korrekt zugeordnet wird (vorher 33 %).

Beim zweiten Klassifikationsansatz wurden die Gewebeprobe entsprechend der im Einsendebogen angegebenen Orientierung fotografiert. Da für diesen Ansatz kein bestehender Datensatz verfügbar war, musste ein künstlicher Datensatz generiert werden. Dazu wurden vorhandene Bilder zufällig ausgewählt und einer der zwölf Rotationsklassen ([0–30°), [30–60°), ..., [330–360°)) zugewiesen – indem das Bild gedreht wurde – sodass die Fadenmarkierung der jeweiligen Position entsprach. Der Hintergrund wurde entsprechend angepasst. Jede Rotationsklasse enthielt Bilder mit Fadenmarkierungen an der Seite, der Spitze und im „Fischmaul“. Zusätzlich wurde eine 13. Klasse für unmarkierte Gewebeprobe eingeführt.

Auch für diesen Ansatz wurde ein ResNet-50-Modell verwendet. Die Optimierung der Hyperparameter, zum Beispiel der Lernrate, wurde anhand einer Design-of-Experiment-Studie

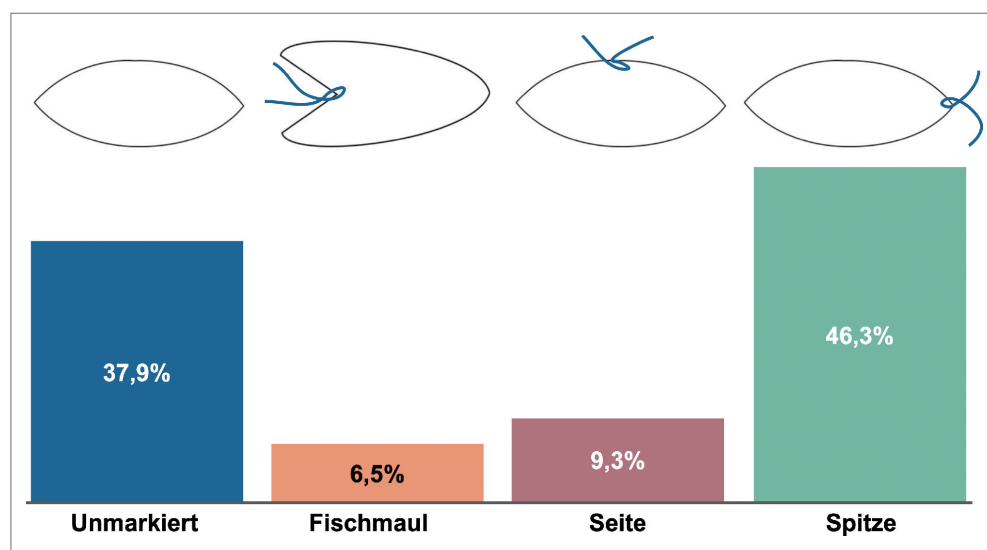


Bild 6 Datensatz zur Klassifikation der Position der Fadenmarkierung.
Grafik: asmedos

durchgeführt (DoE), wobei ein vordefinierter Parameterraum in festen Schritten getestet wurde.

Die Ergebnisse zeigten eine deutliche Verbesserung gegenüber dem ersten Ansatz: Die Average Precision lag bei 97 %. Zudem gab es keine erheblichen Unterschiede in der Erkennungsrate zwischen den einzelnen Klassen.

Abschließend lässt sich die Empfehlung aussprechen, den Arbeitsablauf zu adaptieren und die Gewebeprobe entsprechend den Angaben auf dem Einsendeschein auszurichten, um so eine noch höhere Genauigkeit bei geringem zusätzlichem Aufwand zu erreichen.

Durch die Zusammenführung der Ergebnisse der Object Detection und der Klassifikationen lassen sich makroskopische Beschreibungen der Art "Eingesendet wurde eine spindelförmige Gewebeprobe, mit einer Fadenmarkierung auf 16:00 Uhr und den Abmaßen 14 x 21 mm." generieren.

2.4 Texterkennung der Probenkassette

Ein potenziell schwerwiegender Fehler kann entstehen, wenn Gewebe in eine vorbedruckte Kassette einer anderen Patientin oder eines anderen Patienten gelangt, beispielsweise aus einem vorangegangenen Fall. Um dies zu verhindern, müssen die Identifikationsnummern der tatsächlich am Arbeitsplatz befindlichen Kassetten erfasst und mit den korrekten Nummern abgeglichen werden. Hierbei handelt es sich um eine Aufgabe der optischen Zeichenerkennung (OCR).

Die Erkennung der Kassetten erfolgte mithilfe des beschriebenen Object-Detection-Ansatzes. Dabei zeigte sich jedoch, dass die Kassetten im Bild nicht parallel zu den Bildrändern ausgerichtet waren, sondern in beliebigen Winkeln erscheinen können. Dadurch stand auch der aufgedruckte Text in variabler Orientierung. Da Texterkennung zuverlässiger funktioniert, wenn die Schrift horizontal oder in Vielfachen von 90° ausgerichtet ist, ist eine Korrektur der Orientierung von Vorteil.

Dazu wurde ein Verfahren mit orientierten Bounding Boxes entwickelt und mit dem bisherigen Ansatz verglichen. Ausgangspunkt waren die Bilder aus der initialen Objekterkennung, in denen die Kassetten noch in beliebiger Ausrichtung vorlagen. In einem zweiten Erkennungsschritt wurden Bounding Boxes

erzeugt, die tangential und parallel zu den Kanten der Kassette verlaufen, sodass eine korrekte Ausrichtung gewährleistet ist.

Für die Evaluation wurde ein Datensatz mit rund 1100 Bildern erstellt und mit den korrekten Zeichenfolgen annotiert. Die Auswertung erfolgte mittels fünffacher Kreuzvalidierung. Die Erkennungsleistung der orientierten Bounding Boxes wurde anhand der Kennzahl Intersection over Union (IoU) bewertet und erreichte Werte von über 90%.

Im Anschluss wurde eine Texterkennung auf die neu ausgerichteten Kassettenbilder angewendet. Zum Vergleich erfolgte die Texterkennung auch auf den ursprünglichen Bildern ohne Rotationskorrektur. Zum Einsatz kam das OCR-Modell "paddleOCR", das bereits in Voruntersuchungen gute Ergebnisse erzielt hatte. Die Resultate zeigten, dass die Korrektur der Orientierung die Erkennungsgenauigkeit der einzelnen Zeichen von etwa 85 % auf 93 % steigern konnte.

Im Unterschied zu typischen OCR-Aufgaben zur Erkennung von Wörtern einer bekannten Sprache, handelt es sich hier um alphanumerische Sequenzen. Die Rücksprache mit Fachpersonal der Pathologie des Städtischen Klinikums Karlsruhe ergab, dass die zugrunde liegenden Kennzeichnungsschemata stets formal beschreibbar sind.

Auf dieser Basis wurde untersucht, ob sich die Genauigkeit durch eine schemabasierte Nachverarbeitung weiter verbessern lässt. Dabei wurde das Format der Zeichenfolgen formal definiert: Für jede Position wurde festgelegt, ob dort eine Zahl oder ein Buchstabe erwartet wird. Teilweise konnten auch Abhängigkeiten zwischen Positionen beschrieben werden, etwa bei Jahreszahlen. Zudem wurden feste Trennzeichen zwischen Teilsegmenten der Zeichenfolgen berücksichtigt.

In der Nachverarbeitung wurden anschließend alle erkannten Zeichenfolgen ausgeschlossen, die nicht den definierten Kriterien entsprachen. Das vollständige Verfahren zur Texterkennung mit orientierten Bounding Boxes und schemabasierter Nachverarbeitung ist in **Bild 7** dargestellt.

Durch die Nachverarbeitung liess sich die Güte der Texterkennung mit und ohne orientierte Bounding Boxes steigern. Für den Fall der orientierten Bounding Boxes ergab sich eine Güte von mehr als 99 %. Bei Kassetten, die nicht achsparallel ausgerichtet waren, wurden nur fast 93 % der Zeichenfolgen korrekt erkannt.

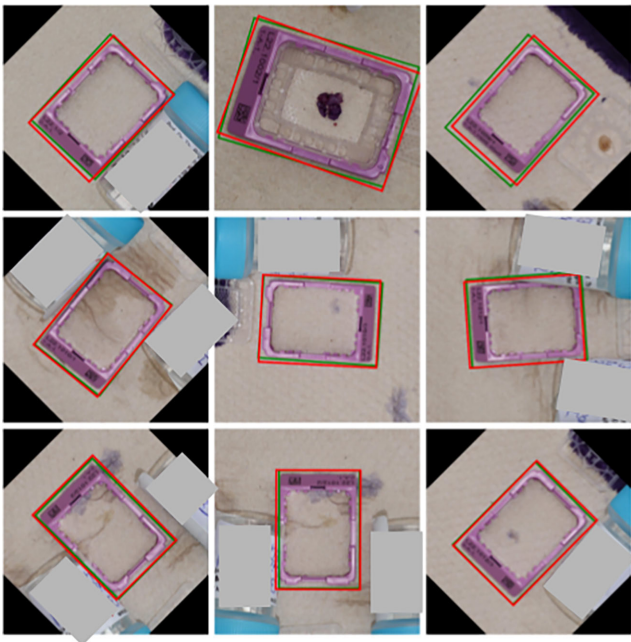


Bild 7 Vorgehen zur Texterkennung mit orientierten Bounding Boxes.
Grafik: asmedos

Das hier vorgestellte Verfahren kann dazu genutzt werden, die Identifikationsnummern auf den Probenkassetten zuverlässig auszulesen und mit den erwarteten Identifikationsnummern abzugleichen. Prozessfehler, die zu Vertauschung führen, können mit hoher Wahrscheinlichkeit erkannt werden.

4 Ausblick

Die vorgestellten Arbeiten zeigen, wie durch den Einsatz von künstlicher Intelligenz auch Arbeitsplätze mit sehr heterogenen Abläufen und Werkstücken automatisiert werden können. Der vorgestellte Ansatz begegnet der bestehenden Komplexität mit Digitalisierung und Methoden der künstlichen Intelligenz. Während durch dieses Vorgehen keine Handhabungsprozesse automatisiert werden können, lassen sich auf diese Art jedoch nicht-physische Prozesse automatisieren und somit Ärztinnen und Ärzte entlasten.

Das Beispiel illustriert auch den Wert durchgängiger Digitalisierung. Die Erkennung der Position der Fadenmarkierung wird durch die einsendende Praxis angegeben, liegt jedoch nicht digital vor und muss auf Seiten der Pathologie wieder erfasst werden. Diese Digitalisierungsbrüche sind insbesondere dann schwierig zu beheben, wenn sie Unternehmensgrenzen überbrücken.

Der vorgestellte Beitrag behandelt nur Hautproben. Ein weiterer wichtiger Schritt wäre eine Übertragung auf andere Gewebetypen.

LITERATUR

- [1] Mulisch, M.; Welsch, U. (Hrsg.): Romeis – Mikroskopische Technik. Heidelberg: Springer 2015
- [2] Munari, E.; Scarpa, A.; Cima, L.; et al.: Cutting-edge technology and automation in the pathology laboratory. *Virchows Archiv* 484 (2024) 4, pp. 555–566, <https://doi.org/10.1007/s00428-023-03637-z>
- [3] Försch, S.; Klauschen, F.; Hufnagel, P. et al.: Künstliche Intelligenz in der Pathologie. *Deutsches Ärzteblatt* 118 (2021) 12, S. 199–204
- [4] Parwani, A. V. (ed.): *Whole Slide Imaging: Current Applications and Future Directions*. Cham: Springer International Publishing 2021
- [5] Ciancia, S.; Vannozi, L.; Poliziani, A. et al.: Technologies for the Automation of Anatomic Pathology Processes: A Review. *IEEE Transactions on Medical Robotics and Bionics* 6 (2024) 3, pp. 888–902, <https://doi.org/10.1109/TMRB.2024.3421611>
- [6] Lopez-Ubeda, P.; Martin-Noguerol, T.; Aneiros-Fernandez, J. et al.: Natural Language Processing in Pathology. *American Journal of Pathology* 192 (2022) 11, pp. 1486–1495, doi.org/10.1016/j.ajpath.2022.07.012
- [7] Koch, D.; Kaiser, J.-P.; Stamer, F. et al.: Enhancing Visual Inspection in Remanufacturing: A Reinforcement Learning Approach with Integrated Robot Simulation. *Procedia CIRP* 134 (2025), pp. 939–944, <https://doi.org/10.1016/j.procir.2025.02.228>
- [8] Schlagenhaut, T.; Feuring, C.-P.; Hillenbrand, J. et al.: Camera Based Ball Screw Spindle Defect Classification System. In: Wulfsberg, J.P., Hintze, W., Behrens, B.A. (eds): *Production at the Leading Edge of Technology*. Heidelberg: Springer Vieweg 2019, pp. 503–512, https://doi.org/10.1007/978-3-662-60417-5_50
- [9] Kaiser, J. P., Enslin, C. L., Barczak, E. T., et al.: A framework for anomaly classification and segmentation in remanufacturing using autoencoders and simulated data. *Procedia CIRP*, 126 (2024), pp. 1005–1010, doi.org/10.1016/j.procir.2024.08.383
- [10] Wu, C.; Fu, H.; Kaiser, J.-P. et al.: 6D Pose Estimation on Point Cloud Data through Prior Knowledge Integration: A Case Study in Autonomous Disassembly. *Procedia CIRP* 122 (2024), pp. 192–198, <https://doi.org/10.1016/j.procir.2024.01.028>
- [11] Ren, S.; He, K.; Girshick, R. et al.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv* 2016, <https://doi.org/10.48550/arXiv.1506.01497>

Dr.-Ing. Constantin Hofmann

constantin.hofmann@asmedos.com

asmedos GmbH & Co. KG

Hirschstr. 29, 76133 Karlsruhe

Dr.-Ing. Martin Benfer

martin.benfer@kit.edu

Prof. Dr.-Ing. Gisela Lanza

gisela.lanza@kit.edu

wbk Institut für Produktionstechnik

Karlsruher Institut für Technologie (KIT)

Kaiserstr. 12, 76131 Karlsruhe

Prof. Dr. med. Thomas Rüdiger

thomas.ruediger@klinikum-karlsruhe.de

Städtisches Klinikum Karlsruhe

Moltkestr. 90; 76133 Karlsruhe

FÖRDERHINWEIS

Die vorgestellten Arbeiten wurden durch das Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg unter dem Förderkennzeichen BW1_3066 gefördert.

LIZENZ



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