

# Anwendungen der künstlichen Intelligenz zur Automatisierung in der Gynäkologie und Geburtshilfe – eine Standortbestimmung

## The Use of Artificial Intelligence in Automation in the Fields of Gynaecology and Obstetrics – an Assessment of the State of Play




### Autoren

Jan Weichert<sup>1,2</sup>, Amrei Welp<sup>1</sup>, Jann Lennard Scharf<sup>1</sup>, Christoph Dracopoulos<sup>1</sup>, Wolf-Henning Becker<sup>2</sup>, Michael Gembicki<sup>1</sup>

### Institute

- 1 Klinik für Frauenheilkunde und Geburtshilfe, Bereich Pränatalmedizin und Spezielle Geburtshilfe, Universitätsklinikum Schleswig-Holstein, Campus Lübeck, Lübeck
- 2 Zentrum für Pränatalmedizin an der Elbe, Hamburg

### Key words

pregnancy, sonography, gynecology, malformation

### Schlüsselwörter

Schwangerschaft, Sonografie, Gynäkologie, Fehlbildung

eingereicht 22.4.2021

angenommen 1.6.2021

### Bibliografie

Geburtsh Frauenheilk 2021; 81: 1203–1216

DOI 10.1055/a-1522-3029

ISSN 0016-5751

© 2021. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,  
70469 Stuttgart, Germany

### Korrespondenzadresse

Prof. Dr. Jan Weichert

Bereich Pränatalmedizin und Spezielle Geburtshilfe, Klinik für Frauenheilkunde und Geburtshilfe, Universitätsklinikum Schleswig-Holstein, Campus Lübeck  
Ratzeburger Allee 160, 23538 Lübeck, Deutschland  
[jan.weichert@uksh.de](mailto:jan.weichert@uksh.de)

### ZUSAMMENFASSUNG

Mit dem Fortschreiten der längst notwendigen Digitalisierung werden täglich riesige Mengen an medizinischen Daten generiert, deren manuelle Analyse und zielgerichtete patientenorientierte Auswertung zunehmend schwieriger bis unmöglich wird. Dieser Umstand und die damit verbundenen, immer

komplexer werdenden Anforderungen an eine individualisierte Präzisionsmedizin verdeutlichen die Notwendigkeit moderner Softwarelösungen und Algorithmen im gesamten Gesundheitswesen. Tatsächlich haben in den letzten Jahren in nahezu allen Bereichen der Medizin mit dem Einsatz modernster Apparaturen und Techniken nun auch automatisierte Prozessabläufe – zumindest teilweise – Einzug in die klinische Routine gehalten. Diese Systeme nutzen verschiedenste Anwendungen der künstlichen Intelligenz (KI) und sind mehrheitlich dazu konzipiert worden, um die medizinische Bildrekonstruktion, Rauschunterdrückung, Qualitätssicherung, Triage, Segmentierung, computergestützte Erkennung und Klassifizierung oder Radiogenomics als neuartiges Forschungsfeld zu optimieren. Die von der KI übernommenen Aufgaben werden dabei deutlich schneller und präziser erledigt, was sich spätestens mit den jährlichen Auswertungen der ImageNet Large-Scale Visual Recognition Challenge (ILSVCR) ab 2015 mit Fehlerraten deutlich unterhalb der von Menschen zeigt. In diesem Übersichtsartikel soll auf die potenziellen Möglichkeiten und die derzeit verfügbaren Anwendungen der KI in der gynäkologisch-geburtshilflichen Diagnostik eingegangen werden. Besonderes Augenmerk wird hierbei auf automatisierte Techniken in der pränatalen sonografischen Diagnostik gelegt werden.

### ABSTRACT

The long-awaited progress in digitalisation is generating huge amounts of medical data every day, and manual analysis and targeted, patient-oriented evaluation of this data is becoming increasingly difficult or even infeasible. This state of affairs and the associated, increasingly complex requirements for individualised precision medicine underline the need for modern software solutions and algorithms across the entire healthcare system. The utilisation of state-of-the-art equipment and techniques in almost all areas of medicine over the past few years has now indeed enabled automation processes to enter – at least in part – into routine clinical practice. Such systems utilise a wide variety of artificial intelligence (AI) techniques, the majority of which have been developed to optimise medical image reconstruction, noise reduction, quality

assurance, triage, segmentation, computer-aided detection and classification and, as an emerging field of research, radiogenomics. Tasks handled by AI are completed significantly faster and more precisely, clearly demonstrated by now in the annual findings of the ImageNet Large-Scale Visual Recognition Challenge (ILSVCR), first conducted in 2015, with error

rates well below those of humans. This review article will discuss the potential capabilities and currently available applications of AI in gynaecological-obstetric diagnostics. The article will focus, in particular, on automated techniques in prenatal sonographic diagnostics.

## Einleitung

Künstliche Intelligenz (KI) ist nicht erst nach dem Erfolg eines Computers (Deep Blue) über den damaligen amtierenden Schachweltmeister G. Kasparov im Jahr 1997 in das allgemeine Bewusstsein gerückt. Tatsächlich liegen die ersten erfolgreichen Anwendungen der KI deutlich weiter zurück und sind seitdem fest in unser tägliches Leben implementiert worden und als solche (häufig auch unerkannt) akzeptiert [4]. Nach Ansicht von KI-Experten in großen Unternehmen bezeichnen 79 Prozent der Befragten KI-Ansätze als strategisch sehr bedeutend oder gar erfolgskritisch im Hinblick auf einen nachhaltigen Geschäftserfolg. Anders ausgedrückt ist künstliche Intelligenz inzwischen weltweit zur Mainstream-Technologie egal welcher Branche geworden. Kompetenzen im Bereich der KI-Kerntechnologien (Machine Learning mit Deep Learning, Natural Language Processing und Computer Vision) sind für größere Unternehmen heutzutage unverzichtbar [5].

Wenngleich es keine einheitlich akzeptierte Definition der „künstlichen oder artifiziellen Intelligenz“ gibt, besteht wesentlicher Konsens darüber, dass von KI gesprochen wird, wenn eine Maschine oder ein System komplexe Aufgaben übernimmt, für deren Ausführung gemeinhin menschliche (oder andere biologische) Gehirnleistung erforderlich ist [1–3]. Die künstliche Intelligenz ist dabei nicht nur eine singuläre Technologie, vielmehr umfasst sie ein Konglomerat von Anwendungen in verschiedensten Bereichen (► **Abb. 1**). Das maschinelle Lernen (ML) ist eine Form der künstlichen Intelligenz, die sich darauf konzentriert, einem Computer das Ausführen von Aufgaben mit einem bestimmten Ziel beizubringen, ohne dabei die Regeln für die Ausführung dieser Aufgabe explizit zu programmieren. Sie kann als eine statistische Methode verstanden werden, die sich kontinuierlich verbessert, indem sie zunehmenden Mengen an Daten ausgesetzt wird, dadurch konsekutiv die Fähigkeit erwirbt, aus Bildern, Texten oder akustischen Daten Objekte richtig zu erkennen, und nach gemeinsamen Eigenschaften und Gesetzmäßigkeiten sucht, um letztlich ein Muster aus diesen zu extrahieren.

Beim Deep Learning, der Königsdisziplin des ML, muss der Mensch im Gegensatz zu anderen Methoden des maschinellen Lernens nicht mehr selbst eingreifen. Ermöglicht werden derartige maschinelle Lernprozesse durch sog. Artificial oder Convolutional Neural Networks (ANN/CNN), die sich vereinfacht aus mehreren Convolutional Layern („gefaltete“ Filter/Schichten) zusammensetzen, gefolgt von einem Pooling Layer, der die Daten der Filter aggregiert und überflüssige Informationen eliminiert (► **Abb. 2**). Mit jeder dieser Filterebenen erhöht sich so das Abstraktions-Level eines CNN. Die Entwicklungen auf dem Gebiet der computergestützten Signalverarbeitung und der Ausbau der Rechenleistung mit modernsten schnellen Grafikprozessoren er-

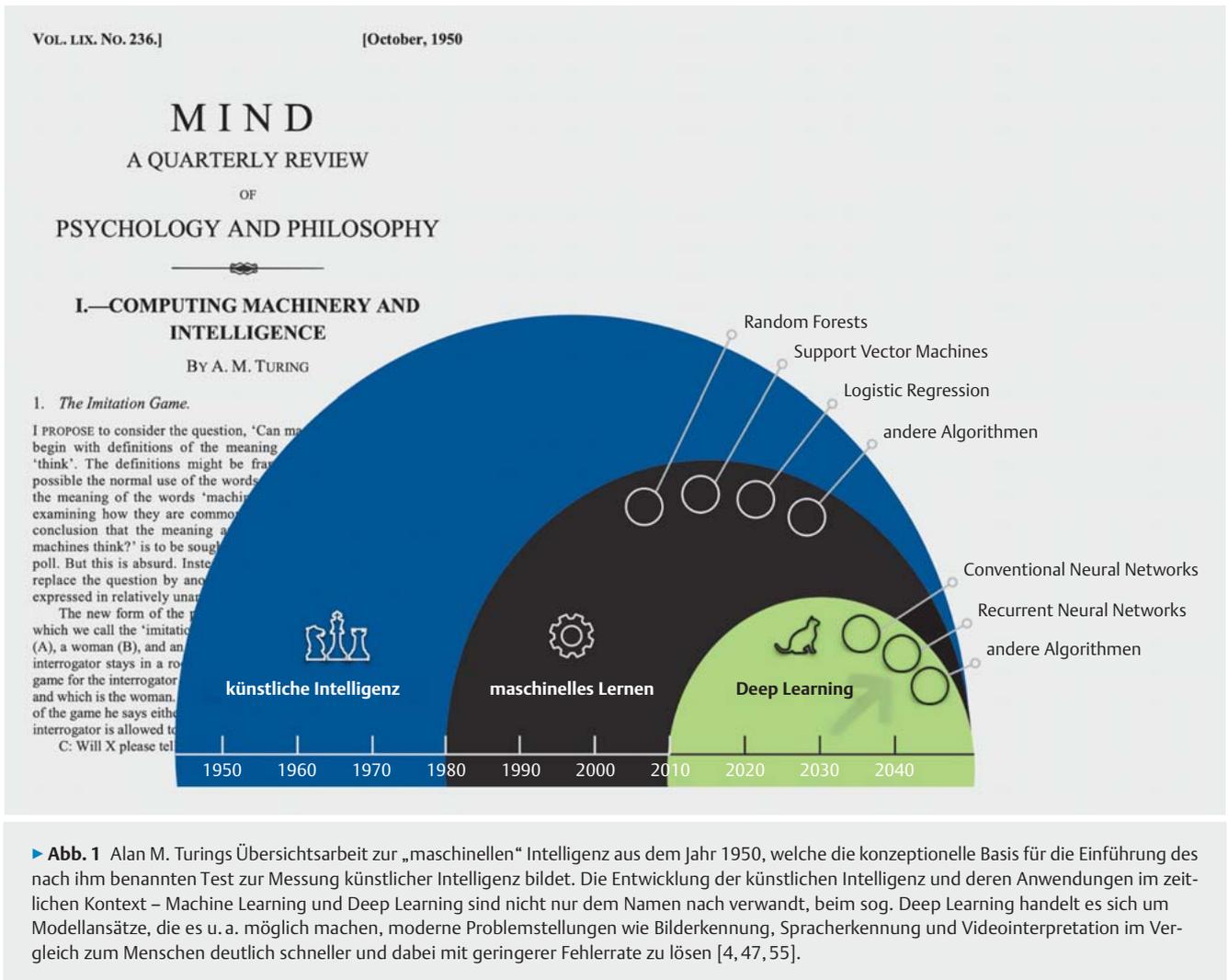
lauben mittlerweile eine unbegrenzte Zahl an Filterschichten innerhalb eines CNN, was letztlich mit dem Ausdruck „deep“ verdeutlicht wird (im Gegensatz zu sog. konventionellen Shallow Neural Networks, die in der Regel nur aus einem Filter bestehen). Das Lernen ist ein adaptiver Prozess, bei dem sich die Gewichtung aller miteinander verbundenen Neuronen ändert, um letztlich die bestmögliche Reaktion (Output) auf alle Eingangsvariablen zu erzielen. Neuronale Netze können dabei auf 2 Arten lernen: überwacht (supervised) oder unüberwacht (unsupervised). Bei ersterem verwenden ML-Algorithmen einen vorkodierten Datensatz, um das gewünschte Ergebnis vorherzusagen. Im Gegensatz dazu erhalten unüberwachte Ansätze nur nicht gelabelte (kodierte) Eingangsdaten, um in diesen versteckte Muster zu finden und entsprechend neuartige Vorhersagen treffen zu können.

Seit Anfang der frühen 2000er-Jahre werden Deep-Learning-Netze z. B. erfolgreich zur Erkennung und Segmentierung von Objekten und Bildinhalten genutzt. Ähnliches gilt auch für KI-gestützte Sprachsteuerung und Spracherkennung. So basieren die Sprachassistenten von Amazon Alexa, Google Home und Apple Siri auf Deep-Learning-Ansätzen. Die Anwendungen hierbei sind generell vielfältig – so ist es nach Einschätzung ihrer Entwickler beispielsweise aktuell möglich, durch eine App-basierte Analyse von Soundevents („Coughing“ App) aus der Stimme COVID-19-relevante Merkmale zu erkennen [6]. Durch automatisierte Auswertung von Stimm-Spektrogrammen ist es darüber hinaus auch möglich, Stimmen-Biomarker für eine Vielzahl von Erkrankungen, wie z. B. Depression, zu identifizieren [6, 7].

Bezogen auf das Gesundheitswesen ist einer der maßgeblichen Vorteile der KI zweifelsohne in der computergestützten Analyse von Bilddaten zu sehen. In den letzten Jahren ist hierzu ein neuartiges und sich rasant entwickelndes Forschungsfeld entstanden, das unter dem Überbegriff „Radiomics“ die systematische, KI-gestützte Analyse und Quantifizierung von patientenbezogenen Bilddaten entlang einer hohen Anzahl unterschiedlicher Bildmerkmale in Korrelation und klinischen Differenzierungsmerkmalen zum Ziel hat. Demgegenüber versteht man unter „Radiogenomics“ eine spezielle Anwendung, bei der radiomische oder sonstige Bildgebungsmerkmale mit genomischen Profilen verknüpft werden [8].

## KI und Vorteile für gynäkologisch-geburts-hilfliche Bildgebung und Diagnostik

Während die anfängliche Hysterie, dass KI-Technologien perspektivisch den klinisch tätigen Radiologen sukzessive ersetzen könnten, bereits überwunden wurde und sich stattdessen das Bewusstsein eingestellt hat, dass mit den Möglichkeiten des maschinellen

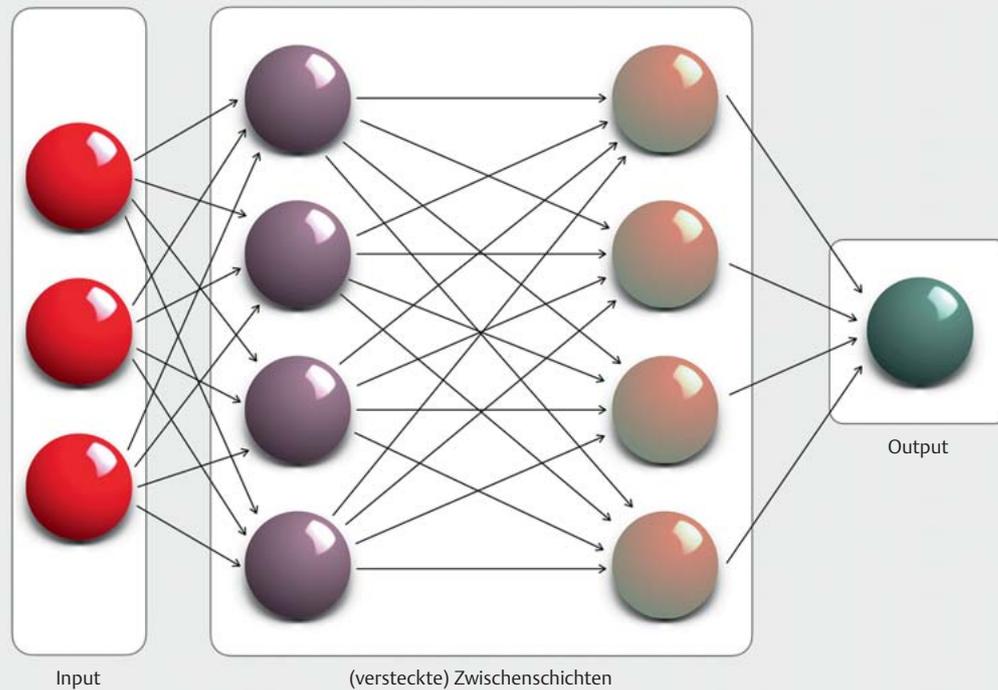


▶ **Abb. 1** Alan M. Turings Übersichtsarbeit zur „maschinellen“ Intelligenz aus dem Jahr 1950, welche die konzeptionelle Basis für die Einführung des nach ihm benannten Test zur Messung künstlicher Intelligenz bildet. Die Entwicklung der künstlichen Intelligenz und deren Anwendungen im zeitlichen Kontext – Machine Learning und Deep Learning sind nicht nur dem Namen nach verwandt, beim sog. Deep Learning handelt es sich um Modellansätze, die es u. a. möglich machen, moderne Problemstellungen wie Bilderkennung, Spracherkennung und Videointerpretation im Vergleich zum Menschen deutlich schneller und dabei mit geringerer Fehlerrate zu lösen [4, 47, 55].

Lernens benutzerdefinierte KI-basierte Softwarealgorithmen mit interaktiver Visualisierung und automatisierter Quantifizierung die klinische Entscheidungsfindung und das Zeitintervall für die Analyse beschleunigen werden, ist die Akzeptanz der KI in anderen klinischen Bereichen noch eher überschaubar bzw. zögerlich [9–11]. Dabei werden, insbesondere in der Mammadiagnostik, schon seit mehr als 25 Jahren computerassistierte Diagnosesysteme (CAD) genutzt [12, 13]. Mit neuartigen Deep-Learning-Algorithmen werden die diagnostischen Möglichkeiten sowohl im Rahmen der Mammografie als auch bei der KI-gestützten Befundung mammasonografischer Datensätze optimiert; damit wird auch den Einschränkungen im Umgang mit konventionellen CAD-Systemen (hohe Entwicklungskosten, generelle Kosten-/Workflow[ineffizienz, rel. hohe Falsch-positiv-Rate, Limitierung auf bestimmte Läsionen/Entitäten) begegnet [14]. So konnte in einer aktuellen US-amerikanisch-britischen Studie ein CNN anhand von 76 000 Mammografie-Scans so trainiert werden, dass eine signifikante Reduktion falsch positiver bzw. falsch negativer Befunde um 1,2 bzw. 2,7% (UK) bzw. 5,7 und 9,4% (USA) verglichen mit der initialen Expertenbefundung eindrucksvoll erreicht werden konnte [15]. Darüber hinaus kann durch konsequente KI-

Unterstützung auch die Arbeitsbelastung i. S. der konventionell zu beurteilenden Mammografie-Scans nachhaltig reduziert werden, indem eine automatisierte Vorauswahl unauffälliger Screeningbefunde getroffen wurde [16]. O’Connell et al. publizierten ähnlich erfolgversprechende Daten zur KI-unterstützten Auswertung von Mammasonografie-Befunden. Sie konnten anhand von 300 Patientinnen zeigen, dass mithilfe eines kommerziellen Diagnosetools (S-Detect) die automatisierte Detektion von Brustläsionen unter Anwendung einer Reihe von BI-RADS-Deskriptoren mit den Resultaten von 10 Radiologen mit entsprechender Expertise übereinstimmte (Sensitivität, Spezifität > 0,8) [17].

Auch in anderen Anwendungsbereichen unseres Fachgebietes konnte der Vorteil von Deep-Learning-Algorithmen explizit herausgearbeitet werden. So hatten Cho et al. Deep-Learning-Modelle entwickelt und validiert, um zervikale Neoplasien auf kolposkopischen Bildern automatisch zu klassifizieren. Vorab trainierte CNN wurden für 2 Bewertungssysteme optimiert: das CIN-System (Cervical Intraepithelial Neoplasia) und das LAST-System (Lower Anogenital Squamous Terminology). Biopsiewürdige Befunde konnten dabei effizient identifiziert werden (AUC 0,947) [18]. Bei der automatisierten Befundung mikroskopischer Zervixzellabstri-



► **Abb. 2** Schematischer Aufbau eines (Feed-Forward-)Konvolutionsnetzwerks mit 2 versteckten Schichten (hidden layers). Die Quellinformationen werden zur Mustererkennung in diesen Schichten segmentiert und abstrahiert und letztlich an die Ausgabeschicht weitergegeben. Die Kapazität derartiger neuronaler Netzwerke kann durch Variation ihrer Tiefe (Anzahl der Schichten) und Breite (Anzahl Neurone/Perzeptrone pro Schicht) gesteuert werden.

che konnten in der Arbeit von Shanthi et al. anhand verschiedener mit augmentierten Datensätze trainierter CNN (Originalkolposkopie, konturiert extrahierte und binäre Bilddaten) in 94,1%, 92,1% und 85,1% eine korrekte Klassifikation in normal bzw. milde, moderate, schwere oder karzinomatöse Zellveränderungen vorgenommen werden [19]. Eine der maßgeblichen Herausforderungen in der verstärkten Integration von KI-Algorithmen bei der pathologischen Beurteilung und Diagnostik von histomorphologischen Präparaten generell besteht nach Ansicht von Försch et al. darin, dass derzeit nur ein Bruchteil der histopathologischen Daten tatsächlich in digitaler Form vorliegt und somit überhaupt einer automatisierten Auswertung zugänglich ist [20]. Letzteres trifft im Augenblick noch auf die überwiegende Mehrzahl potenzieller klinischer Anwendungsgebiete der KI zu [1, 21].

Durchaus vergleichbare Ansätze wurden innerhalb der letzten 5 Jahre auch in der Reproduktionsmedizin verfolgt, wo u. a. erfolgreich versucht wurde, KI-Technologien zur Embryoselektion einzusetzen. Dabei wurden z. B. CNN derart trainiert, dass diese in der Lage waren, anhand von entsprechenden Bilddaten und/oder morphokinetischen Daten qualitative Aussagen zu treffen, die hinsichtlich der Prädiktion einer erfolgreichen Implantation ausgewertet werden konnten [22, 23]. Der Studie von Bori et al. an mehr als 600 Patientinnen zufolge wurden neben den erwähnten morphokinetischen Eigenschaften auch neuartige Parameter wie die Distanz und Geschwindigkeit der pronukleären Migration, innere Zellmassenfläche, erweiterter Durchmesser der Blastozyste

und Länge des Trophektoderm-Zellzyklus analysiert. Von den 4 Testalgorithmen zeigte sich der mit der Kombination konventioneller morphokinetischer und o. a. morphodynamischer Merkmale am effizientesten (AUC 0,77), wobei insbesondere die beiden letztgenannten Parameter signifikant häufiger Unterschiede bei implantierten und nicht implantierten Embryonen aufwiesen [24].

Unstrittig ist, dass bislang nur relativ wenige KI-basierte Ultraschallanwendungen den gesamten Weg vom akademischen Konzept zur klinischen Anwendung und Kommerzialisierung genommen haben. Neben der in den nachfolgenden Absätzen abgehandelten Bedeutung der KI in der pränatalen Diagnostik, zeigen sich die Vorteile von KI-gestützten automatisierten Algorithmen sehr eindrucksvoll in der Befundung gynäkologischer Auffälligkeiten, was insbesondere in Anbetracht der eingeschränkten Qualität der derzeitigen Ultraschallausbildung eine zusätzliche Wertschätzung erlangen wird [25–27]. Wenngleich die ersten Arbeiten hierzu schon mehr als 20 Jahre zurückliegen [28], konnte insbesondere in der letzten Dekade maßgebliche Pionierarbeit, nicht zuletzt durch die umfangreichen Studien der IOTA Working Group, geleistet werden. Modellanalysen zur Risikoquantifizierung von sonografisch detektierten Adnexitäten haben zeigen können, inwieweit ein standardisiertes Vorgehen zur qualifizierten Beurteilung zum einen und ein anhand von tausenden Patientinnenverläufen validiertes Mehrklassenrisikomodel (IOTA Adnex – Assessment of Different Neoplasias in the AdnEXa) zum anderen

die Qualität der sonografischen Befundung von Adnexprozessen exakt und reproduzierbar möglich gemacht und damit anderen Studienansätzen ( $\pm$  KI) auf diesem Gebiet maßgeblichen Vorschub gewährt haben [29–31]. Ein augenscheinlicher Beleg für Vorangestelltes ist die nun erfolgte Übernahme des ADNEX-Modells in eine Consensus-Guideline des American College of Radiologists (ACR) – bemerkenswert, da die US-amerikanischen Fachgesellschaften traditionell fachübergreifend ohnehin eher als Ultraschall-skeptisch gelten [32]. Christiansen et al. konnten in einer aktuellen Untersuchung zur Validität zweier KI-Modelle hinsichtlich der Dignität von Adnexläsionen (trainiert an B-Mode und Farb-Doppler-Bilddateien) eine Sensitivität von 96 bzw. 97,1% bei einer Spezifität von 86,7 bzw. 93,7% erreichen – ohne signifikante Unterschiede im Vergleich zur Experteneinschätzung [33]. Der Mehrwert verschiedener ML-Klassifikatoren allein oder in Kombination wurde in verschiedenen anderen Ansätzen untersucht, mit dem ähnlichen Ergebnis, dass KI-Ansätze zukünftig mehr und mehr Einsatz in der (Früh-)Entdeckung ovarieller Neoplasien finden werden [34–38]. Al-Karawi et al. untersuchten im Rahmen einer aktuellen Arbeit mittels ML-Algorithmen (Support Vector Machine Classifier) 7 unterschiedliche bekannte Bildtexturparameter in US-Standbildern, die nach Vorstellung der Autoren Auskunft über die veränderte zelluläre Zusammensetzung i. R. d. Karzinogenese geben können. Durch Kombination der Merkmale mit den besten Testergebnissen war eine Genauigkeit von 86–90% beschrieben worden [39].

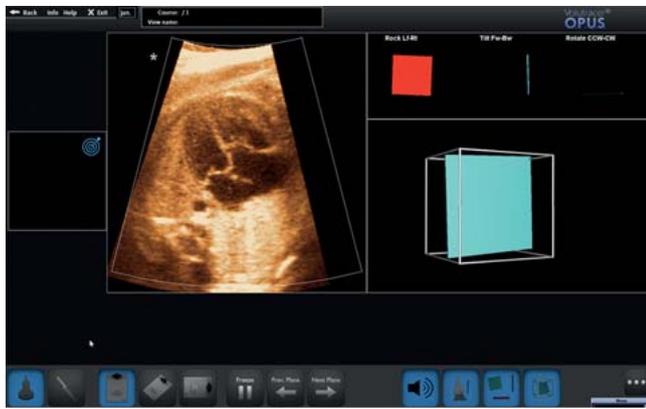
## KI in der fetalen Echokardiografie

Selbstverständlich stellt sich bei der Analyse zum Stellenwert der KI auch die Frage, wie automatisierte Ansätze im Rahmen der fetalen Herzuntersuchung, einem der wesentlichen, gleichsam aber auch komplexeren Bestandteile der vorgeburtlichen sonografischen Untersuchung, einen Mehrwert für die Diagnostik und den Untersucher erbringen können. Hierzu ist es wichtig zu wissen, dass obwohl die Detektionsraten von fetalen Herzfehlern (congenital heart defects, CHD) in nationalen oder regionalen Screeningprogrammen im letzten Jahrzehnt nachweislich verbessert werden konnten, die Sensitivität nach wie vor zwischen 22,5 und 52,8% liegt [40]. Die Ursachen hierfür sind vielschichtig – einer der wesentlichen Faktoren ist sicherlich, dass die überwiegende Mehrheit der CHD tatsächlich im Low-Risk-Kollektiv und nur in etwa 10% bei Schwangeren mit bekannten Risikofaktoren auftreten. Laut einer niederländischen Studie scheinen darüber hinaus neben der mangelnden Expertise im Routinesetting auch Faktoren wie eingeschränkte adaptive visuell-motorische Fähigkeiten in der Erlangung der korrekten Herzebenen sowie eine verminderte Vigilanz bei der Erkennung von Herzauffälligkeiten eine entscheidende Rolle zu spielen [41].

Die Erfahrungen aus der Erwachsenenkardiologie haben u. a. gezeigt, dass der Einsatz automatisierter Systeme (im Übrigen keineswegs ein neuartiger konzeptioneller Ansatz) nachweislich effizienter als ein konventionelles (manuelles) Vorgehen ist und perspektivisch die Lücke zwischen dem Experten und dem weniger erfahrenen Untersucher zu schließen vermag, während gleichzeitig die Inter- und Intraobservervarianz verringert wird. Pilotstudien zur automatisierten Analyse linksventrikulärer (Funktions-)Para-

meter wie Kammervolumen und Ejektionsfraktion anhand von 2-D-Bildern bzw. Untersuchungen zum KI-basierten Tracing von Endokardkonturen in apikalen Zwei- und Vierkammeransichten unter Verwendung transthorakal akquirierter 3-D-Datensätze konnten eine der manuellen Evaluation vergleichbare Genauigkeit belegen [42,43]. Kusunose definierte in diesem Zusammenhang entscheidende 4 Schritte für relevante KI-Modelle in der Echokardiografie (neben der Gewährleistung einer adäquaten Bildqualität sind nachfolgend Ebenenklassifikation, Messansätze und letztlich die Detektion von Anomalien zu nennen) [44]. Zhang et al. untersuchten die Validität eines komplett automatischen KI-Ansatzes zur echokardiografischen Befundung im klinischen Kontext, in dem sie anhand von > 14000 kompletten Echokardiogrammen tiefe Konvolutionsnetzwerke derart trainierten, dass diese in der Lage waren, 23 verschiedene Bildpunkte aus 5 unterschiedlichen Schnittbildreferenzen zu identifizieren. In bis zu 96% konnten die einzelnen kardialen diagnostischen Ebenen exakt erkannt und darüber hinaus 11 verschiedene Messparameter mit vergleichbarer oder sogar höherer Genauigkeit als manuelle Ansätze quantifiziert werden [45]. Werden KI-Algorithmen also zukünftig den Echokardiographen oder gar Kardiologen ersetzen? Sollten wir besorgt sein? Sind wir Teil der nutzlosen Klasse, wie von Harari provokativ formuliert [46]? Die Antwort auf diese Fragen ist dabei eindeutig wie insgesamt komplex und in gleichem Maße auch für die fetale Echokardiografie gültig. Wenngleich KI-Ansätze sehr bald ein integraler Bestandteil in der kardialen Routinediagnostik sein werden, obliegt es dem Untersucher weiterhin oder sogar verstärkt, mit klinischer Expertise die automatisierten Abläufe zu verstehen, zu überwachen und beurteilen und bei Fehlern entsprechend zu korrigieren [47]. Arnaout et al. haben anhand von 107823 Ultraschallbildern von > 1300 fetalen Echokardiogrammen erfolgreich ein Modell zur Erstellung diagnostischer Schnittebenen trainieren können [48]. In einem separaten Modellansatz konnten sie so zwischen strukturell normalen Herzen und solchen mit komplexen Vitien unterscheiden. Die Ergebnisse der KI waren dabei mit denen von Experten vergleichbar. Eine etwas geringere Sensitivität/Spezifität (0,93 bzw. 0,72, AUC 0,83) dokumentierten Le et al. 2020 in deren KI-Ansatz an knapp 4000 Feten [49]. Dong et al. demonstrierten, wie genau ein dreigestaffertes CNN anhand von 2-D-Bilddateien verschiedene Darstellungen des Vierkammerblickes zu detektieren in der Lage ist und gleichzeitig eine Aussage zur Vollständigkeit der abgebildeten kardialen Kernstrukturen rückmelden kann [50].

Zusammenfassend muss festgehalten werden, dass die essenzielle Voraussetzung für die bereits angeklungene Effizienz der kardialen Diagnostik in der Erstellung exakter Schnittbilder im Rahmen der Untersuchung zu sehen ist, was letztlich insbesondere in der bildgebenden Funktionsdiagnostik fachübergreifend generell zu fordern ist. Hinton formulierte hierzu treffend „To recognize shapes first learn to generate images“ [51]. In diesem Zusammenhang ist die kürzlich erfolgte Zulassung eines adaptiven US-Systems (Caption AI) zur Unterstützung und Optimierung der Schnittebenenerstellung (und Aufnahme von Videosequenzen) im Rahmen der adulten Echokardiografie durch die US-amerikanische Gesundheitsbehörde FDA zu erwähnen. Nach Ansicht der Entwickler zeigt sich hierbei, wie das enorme Potenzial künstlicher Intelligenz und der Technologien des maschinellen Lernens gezielt

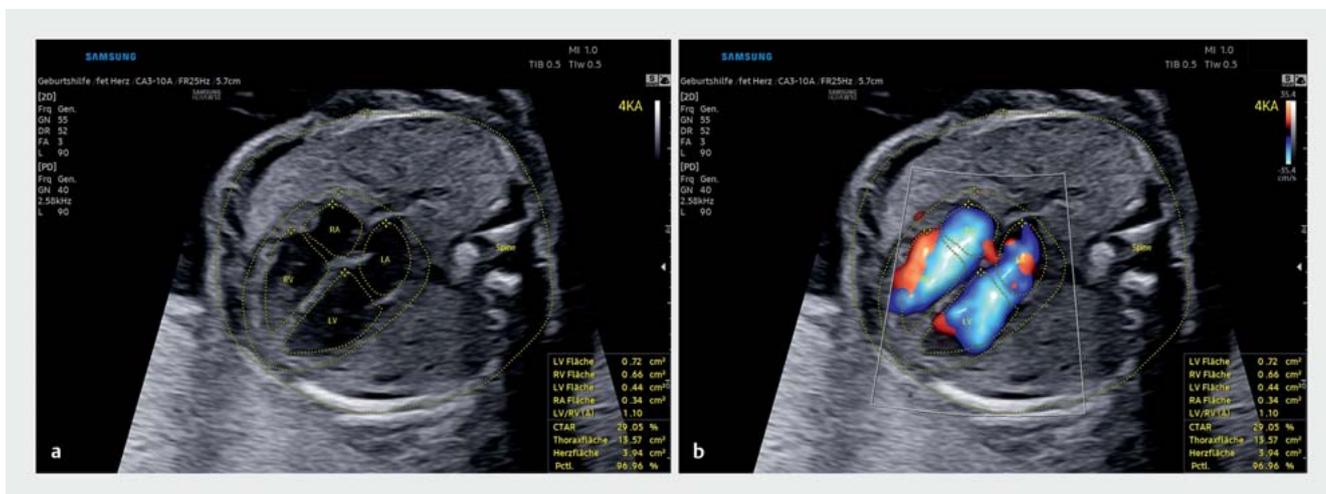


► **Abb. 3** Darstellung des optischen Ultraschallsimulators Volutracer O. P. U. S.. Ein beliebiger Volumendatensatz (s. a. ► **Abb. 5**) kann hochgeladen und z. B. für Lehrzwecke entsprechend ebenenadaptiert nachbearbeitet werden (sog. Freestyle-Modus – ohne Simulatorinstruktionen). Im oberen rechten Bildrand gibt das System ein grafisches Feedback an, um die Bewegungen für die Einstellung der korrekten Zielebene zu unterstützen. Die Simulationssoftware bietet darüber hinaus eine Vielzahl cloudbasierter Trainings-Datensätze an, die über ein GPS-Trackingsystem und akustische Simulatorinstruktionen mit eingblendeten Animationen die korrekten Einstellungen erlernen lassen. Das System misst u. a. die Position, Rotationswinkel und Zeit bis zur Einstellung der geforderten Zielebene und vergleicht diese mit einer ebenfalls abrufbaren Expertenreferenz.

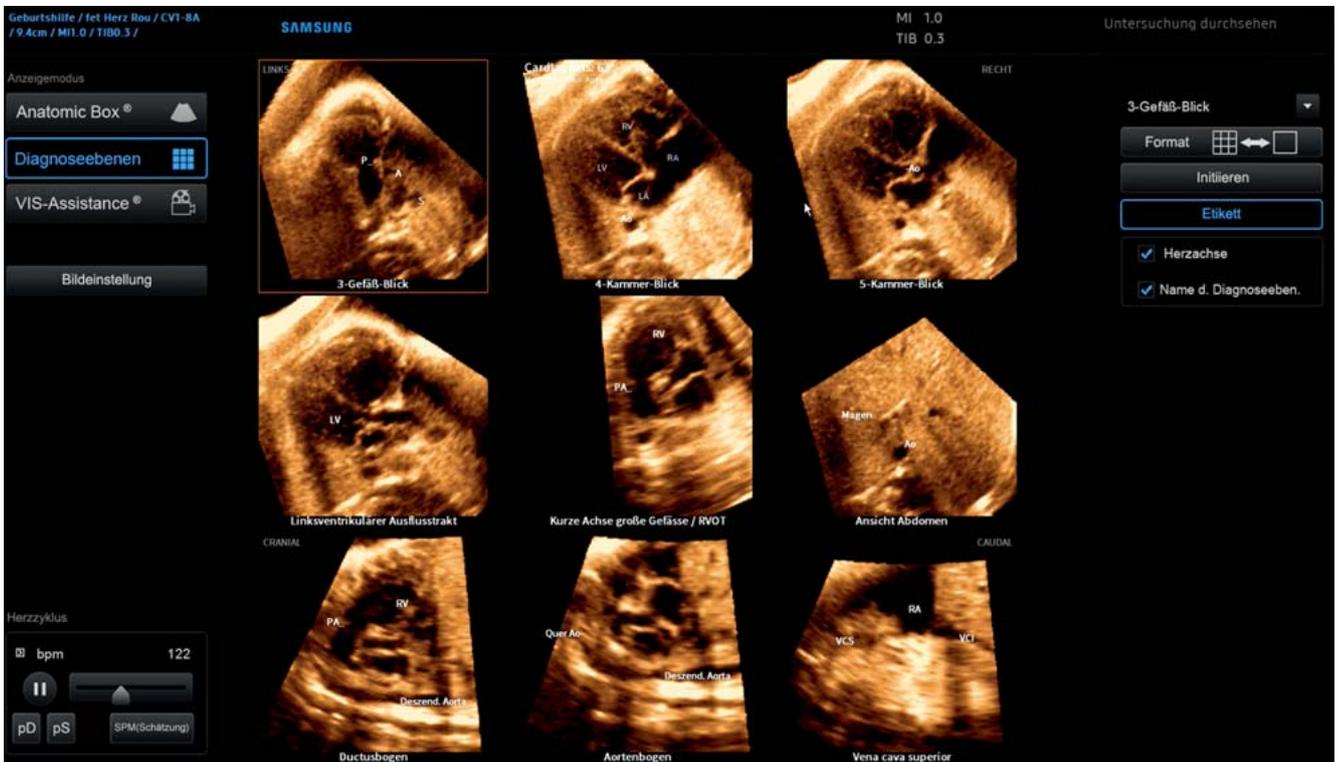
genutzt werden kann, um den Zugang zu sicherer und effektiver Herzdiagnostik zu verbessern [52]. Einen vergleichbaren KI-Workflow besitzt ein ebenfalls kommerziell erwerblicher hochsensitiver Ultraschallsimulator (Volutracer O. P. U. S.), der die händischen Einstellungen und Schallkopfbewegungen, um eine exakte Zielebene in beliebigen 2-D-Bildsequenzen (egal welcher anatomischen Struktur) zu erreichen, in Echtzeit steuert und adaptiv

korrigiert (► **Abb. 3**) [53]. Ein wesentlicher Vorteil dieser Systeme liegt dabei zweifelsohne in der gezielten Anwendung in Aus- und Weiterbildung, da u. a. der integrierte Selbstlernmodus die automatische Schulung, Bewertung und Zertifizierung des jeweiligen Anwenders ermöglicht, ohne dass dabei die vorgenommenen Einstellungen von Expertenseite physisch supervidiert werden müssen [54].

Aufgrund der vergleichsweise geringen Größe nimmt das fetale Herz in der Regel einen eher kleinen Teil des US-Bildes ein, was im Umkehrschluss bedeutet, dass jeder Algorithmus lernen muss, zumindest einen Teil der verfügbaren Bilddaten zu ignorieren. Ein weiterer Unterschied zur postnatalen Echokardiografie besteht darin, dass die relative Ausrichtung und Position des Herzens im Bild in Relation zur fetalen Position in utero sehr unterschiedlich sein kann, was in der Konsequenz zu einer weiteren Komplexität bei der Bildanalyse führt [55, 56]. Ein interessanter und kurz vor der Markteinführung stehender Ansatz zur automatisierten Erkennung, Annotation und Messung kardialer Strukturen unter Verwendung von Deep-Learning-Algorithmen ist HeartAssist. Mit diesem intelligenten Softwaretool können im Rahmen der fetalen Echokardiografie aus 2-D-Standbildern (direkt oder als Single Frame herausgelöst aus Videosequenzen) von kardialen diagnostischen Schnittebenen Targetstrukturen (axial, sagittal) erkannt und ausgewertet werden (► **Abb. 4**). Bemerkenswert ist dabei, dass zum einen selbst teilweise verdeckte Bildinformationen erfasst werden und in die Analyse mit einfließen und zum anderen, dass die Bildererkennung auch bei eingeschränktem sonografischen Fenster funktioniert. Dieser Ansatz beruht wie die meisten Algorithmen zur Automatisierung der (fetalen) Diagnostik (z. B. BiometryAssist, Smart OB oder SonoBiometry) auf der Segmentierung (Abstrahierung) fetaler Strukturen. Im Detail kommen dabei unterschiedlichste Verfahren der automatischen Segmentierung (pixel-, kanten-, regionenorientierte, sowie modell- und texturbasierte Modelle) zur Anwendung, die zur Erzielung besserer Ergebnisse in der Regel auch kombiniert werden [57, 58].



► **Abb. 4** Vierkammerblick (4KB) eines fetalen Herzens in der 23. SSW. Die fetale Wirbelsäule liegt bei 3 Uhr, der 4KB kommt in einer leicht schrägen Lage zur Darstellung. Zusätzlich zu Bauch- und Herzumfang wird im Standbild die Innenkontur der Vorhöfe und Ventrikel automatisch erkannt, umfahren und quantifiziert. In ähnlicher Weise können alle anderen kardialen diagnostischen Schnittebenen (axial/longitudinal) durch HeartAssist annotiert und ausgemessen werden.



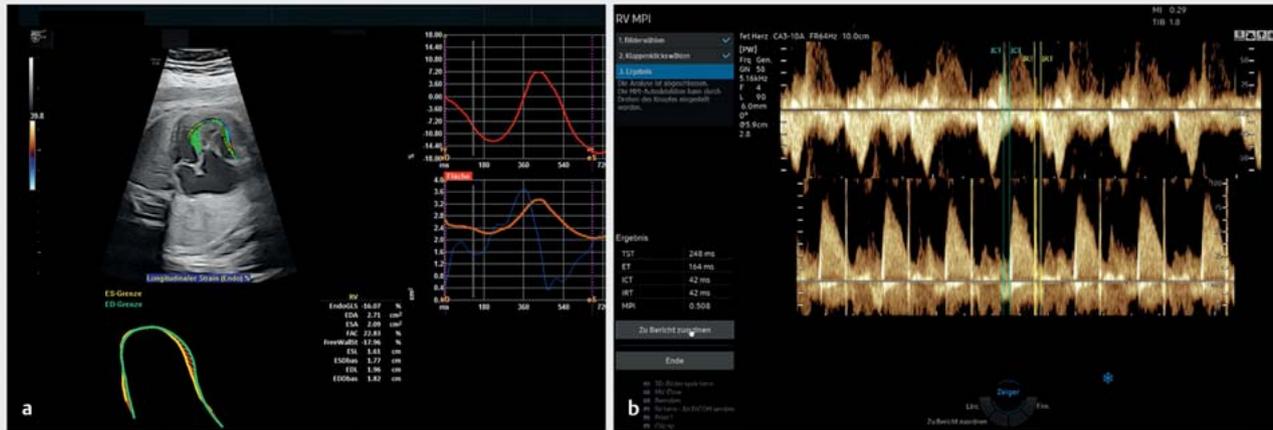
► **Abb. 5** Programmoberfläche von 5DHeart (Fetal Intelligent Navigation Echocardiography, FINE) mit automatisch rekonstruierten diagnostischen Ebenen einer Ebstein-Anomalie eines Feten in der 33. SSW (STIC-Volumen). Der atrialisierte rechte Ventrikel ist im linksrotierten Vierkammerblick (Herzachse > 63°) als Leitstruktur eindeutig darzustellen. Der fetale Rücken ist nach Applikation der automatisierten Software standardmäßig bei 6 Uhr lokalisiert (Volumenakquise erfolgte dagegen bei 7–8 Uhr, s. ► **Abb. 5**). Bei der Analyse der korrespondierenden Ebenen fällt zudem eine tubuläre Aortenstenose auf (visualisiert in den Ebenen: Dreigefäß-/Fünfkammerblick, LVOT, Aortenbogen).

## Besonderer Stellenwert der 3-D/4-D-Technologie

Spätestens mit der Einführung und mittlerweile auf den meisten US-Systemen ab Werk vorinstallierten 3-D/4-D-Technik sind nunmehr eine Reihe verschiedener Darstellungsmöglichkeiten verfügbar geworden, die auch im Rahmen der automatisierten Bildanalyse und Ebenenerstellung mehr und mehr Einsatz finden. So bieten einige Hersteller kommerzielle Softwaretools an, die einen volumenbasierten Ansatz zur fetalen Echokardiografie und dessen standardisierte Auswertung erleichtern sollen (Fetal Heart Navigator, SonoVCADheart, Smart Planes FH und 5D Heart). Letzterer Algorithmus ermöglicht durch Implementierung der Fetal Intelligent Navigation Echocardiography (FINE) eine standardisierte workflowbasierte 3-D/4-D-Evaluation der fetalen Herzanatomie (► **Abb. 5**). Analysiert werden dabei STIC-Volumina (Spatial Temporal Image Correlation) mit dem Vierkammerblick als Ausgangsebene der Volumenakquise. Im nächsten Schritt werden vordefinierte, anatomische Zielstrukturen markiert und die 9 diagnostischen, für eine komplette fetale echokardiografische Untersuchung notwendigen Ebenen automatisch rekonstruiert. Jede Ebene kann dabei im Nachtrag unabhängig von den anderen ausgewertet (z. B. quantitative Analyse der Ausflusstrakte etc.) und sofern notwendig noch manuell angepasst werden. Yeo et al.

konnten so zeigen, dass kardiale Vitien mit einer Sensitivität von 98% detektiert werden können [59]. Die Methode ist dabei nachweislich einfach zu erlernen und in der Lage Arbeitsabläufe bei der Evaluation des fetalen Herzens untersucherunabhängig zu vereinfachen, was insbesondere bei der detaillierten Erfassung kongenitaler Anomalien von Bedeutung ist [60, 61].

Die Erfassung und Quantifizierung objektiver fetaler kardialer Funktionsparameter ist ähnlich anspruchsvoll und damit untersucherabhängig wie die manuelle Ebenenrekonstruktion. Hier ist im Besonderen die Speckle-Tracking-Echokardiografie zu erwähnen, die anhand der sog. Speckles, die durch Interferenzen von zufällig auftretenden Streuechos im Ultraschallbild entstehen, quantitative Informationen über 2-dimensionale globale und segmentale myokardiale Wandbewegungen und Deformierungsparameter (Strain/Strain rate) liefert. Mit der Einführung einer semiautomatischen Software (fetalHQ), die anhand eines 2-D-Videoclips des Herzens und manueller Anwahl eines Herzzyklus und entsprechender Markierung von Klappenring und Apex ist es mittlerweile auch weniger geübten Untersuchern möglich, durch KI-unterstützte Analyse dieser Speckles quantitative Aussagen zur Größe, Form und Kontraktilität an 24 unterschiedlichen Segmenten des fetalen Herzens vorzunehmen [62–64] (► **Abb. 6**). Des Weiteren sind in den letzten Jahren KI-Methoden zur dopplerbasierten kardialen Funktionsanalyse (modifizierter myokardialer



► **Abb. 6** Softwaretools für die Funktionsanalyse des Fetalherzens. Semiautomatischer Ansatz zur Speckle-Tracking-Analyse mittels fetalHQ bei dem in ► **Abb. 3** und **5** untersuchten Feten mit Ebstein-Anomalie (**a**). Die Analyse eines manuell ausgewählten Herzzyklus erfolgt hierbei durch automatische Konturierung des Endokards für den linken u./o. rechten Ventrikel und subsequente Quantifizierung funktioneller Größen wie Kontraktilität und Verformung. Automatisierte Berechnung des (modifizierten) myokardialen Performance-Index (MPI; Tei-Index) durch Spektral-Doppler-Erfassung des Blutflusses über der Trikuspidal- und Pulmonalklappe durch Verwendung von MPI+ (**b**).

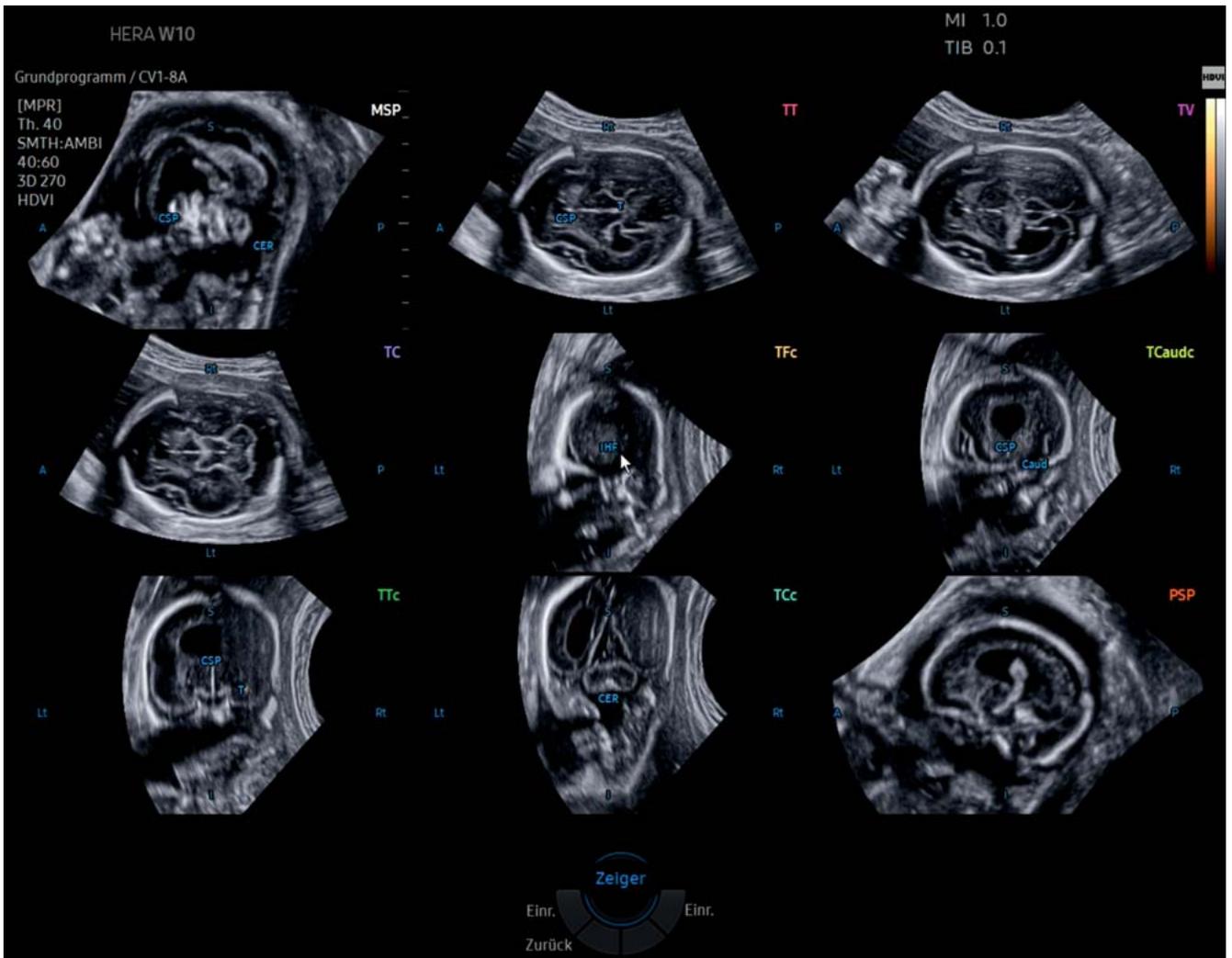
Performanceindex; Mod-MPI o. Tei-Index) entwickelt worden und mittlerweile auch kommerziell verfügbar [65, 66].

## KI in der standardisierten Diagnostik des fetalen ZNS

Wie im Vorangestellten ausgeführt, ist der entscheidende Vorteil automatisierter Techniken im Rahmen der vorgeburtlichen Diagnostik sicher darin zu sehen, dass es so auch weniger erfahrenen Untersuchern möglich sein wird, diagnostische Aussagen zu hochkomplexen anatomischen Strukturen wie dem Fetalherzen oder dem ZNS untersucherunabhängig und standardisiert vornehmen zu können. Die Basis für derartige Tools bilden transthalämisch (TT), (analog der für die Quantifizierung des biparietalen Durchmessers notwendige Schnittbildeinstellung) akquirierte 3-D-Volumendatensätze, die KI-unterstützt postprozessiert und ausgewertet eine Basisuntersuchung des fetalen ZNS mit Herauslösen auch der transversalen (TV) bzw. transzerebellären (TC) Ebene aus dem Volumenblock (SonoCNS, Smart Planes CNS) oder gar ein vollständiges Neurosonogramm (5DCNS+) erlauben (► **Abb. 7**). Bei letzterem Algorithmus wird nach axialer Ausrichtung der korrespondierenden B- und C-Ebenen und Markierung von Thalamuskernen bzw. Cavum septi pellucidi eine Rekonstruktion zusätzlich auch der koronaren und sagittalen Schnittebenen, die für ein komplettes Neurosonogramm gefordert sind, automatisch vorgenommen (► **Abb. 5**). Letztere Arbeitsgruppe konnte in einer prospektiven Folgestudie mit der Programmerweiterung 5DCNS+ erfolgreiche Visualisierungsraten von 97,7 bis 99,4% für axiale, 94,4 bis 97,7% für sagittale und 92,2 bis 97,2% für koronale Ebenen dokumentieren [67]. Zu ähnlichen Ergebnissen kam auch eine retrospektive klinische Validierungsstudie an mehr als 1100 Schwangeren [68]. Im Gegensatz zu den Daten von Pluym et al. konnten die Autoren in ihrer Studie u. a. zeigen, dass mit diesem

standardisierten Ansatz vergleichbar valide und reproduzierbare biometrische Parameter wie im Vergleich zu manuellem Vorgehen erhoben werden können [69]. Ähnlich deutlich äußerten sich Ambroise-Grandjean et al., die in einer Machbarkeitsstudie zeigten, dass die 3 Basisebenen inklusive biometrischer Messungen KI-basiert (Smart Planes CNS) mit geringer Intra- und Interobservervarianz (ICC > 0,98) regelhaft rekonstruiert und quantitativ ausgewertet werden konnten [70].

Sind für diese klinisch bereits angewandten Algorithmen in der Regel noch Zwischenschritte durch Eingreifen des Untersuchers notwendig, wird es perspektivisch möglich sein, mittels speziell trainierter CNN alle Schnittebenen vollständig automatisch aus Roh-Volumina zu extrahieren. Huang et al. haben zeigen können, dass durch sog. ansichtsbasierte VP-Nets (CNN) aus postprozessierten 3-D-Volumina (axiales Ausgangsvolumen und korrespondierende 90°-Rotationen sagittal/koronal) 5 vorgegebene anatomische ZNS-Strukturen parallel in 3 unterschiedlichen 3-D-Projektionen zuverlässig erkannt und dargestellt werden konnten, wobei die besten Detektionsraten auch hier für die axial dargestellten Zielstrukturen registriert wurden [71]. Letzteres ist u. a. der naturgemäß graduell verminderten Bildqualität in den orthogonalen B- und C-Ebenen zuzuschreiben. Die Autoren nutzten für ihre Analyse die Datensätze der INTERGROWTH-21-Studiengruppe. Eine exakte B-Bildqualität und Genauigkeit in der Schnittbildherstellung ist unabdingbare Voraussetzung für 2-D-basierte KI-Ansätze, insbesondere bei der automatisierten Detektion auffälliger ZNS-Befunde, wie kürzlich von Xie et al. publiziert [72]. In dieser Arbeit wurden CNN mithilfe von 2-D- und 3-D-Datensätzen von jeweils etwa 15000 normalen und auffälligen axialen Standardebenen trainiert und hinsichtlich der Effizienz der Segmentierung, der binären Klassifikation in normal und auffällig und der Lokalisation der ZNS-Läsion ausgewertet (Sensitivität/Spezifität 96,9 bzw. 95,9%; AUC 0,989). Um derartige KI-Ansätze dort zum Einsatz bringen zu können, wo der größte Nutzen zu erwarten wäre, näm-

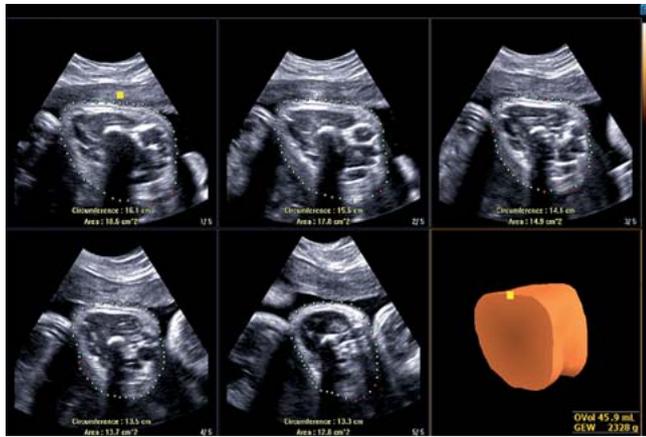


► **Abb. 7** (Semi-)automatische Rekonstruktion nach Applikation von 5DCNS+ eines axial akquirierten 3-D-Volumens des fetalen ZNS (biparietale Ebene) bei einem Feten mit einer semilobaren Holoprosenzephalie in der 23. SSW. Das aus dem Quellvolumen rekonstruierte komplette Neurosonogramm besteht aus den 9 geforderten diagnostischen Schnittebenen (3 axiale, 4 koronale und 2 sagittale Ebenen). In den axialen Ebenen werden automatische biometrische Messungen (nicht gezeigt) vorgenommen, die jederzeit noch manuell angepasst werden können.

lich in der Routinediagnostik, sind noch einige „Hürden“ zu nehmen, die vornehmlich an die initialen Schritte in der bildgebenden Diagnostik zu adressieren sind (entsprechend der von Hinton geforderten Qualität im Zuge der Bilderstellung), was letztlich auch auf andere fetale Zielstrukturen im Rahmen der pränatalen Diagnostik so zutrifft [51]. Interessant wäre in diesem Zusammenhang z.B. festzustellen, inwieweit diese automatisierten Ansätze zur standardisierten Ebenenrekonstruktion in Kombination mit DL-Algorithmen zur Klassifizierung und damit exakten und reproduzierbaren Detektion, Annotation und Quantifizierung von 2- und 3-dimensionalen Messparametern eine deutlich von der Expertise des Untersuchers unabhängige Diagnostik in Zukunft ermöglichen. Von besonderem Interesse sind ganz sicher auch solche Ansätze, in denen z. B. mittels spezieller neuronale Netze Bildaufnahmeprotokolle in der geburtshilflichen US-Diagnostik dahingehend optimiert werden, dass sie die Untersuchungszeit verkürzen und umfassende anatomische Informationen auch aus z.T.

verdeckten Bildbereichen bereitstellen können. So demonstrieren Cerrolaza et al. (analog zu Deep-Reinforcement-Learning-Modellen bei inkompletten CT-Scans), dass selbst wenn nur 60% des fetalen Schädels in einem Volumendatensatz erfasst wurden, eine KI-Rekonstruktion dennoch möglich war [73, 74].

Das Potenzial neuronaler Netzwerke zeigen auch aktuelle Arbeiten von Cai et al., die ein Multi-Task-CNN entwickelten, das durch Erfassen von Augenbewegungen des Untersuchers bei der Betrachtung von Videosequenzen die Detektion von axialen Standardebenen wie dem fetalen Bauch- und Kopfumfang (transventrikuläre Schnittebene) erlernt [75]. Baumgartner et al. konnten zeigen, dass mittels eines speziell trainierten Konvolutionsnetzwerks (SonoNet) 13 verschiedene fetale Standardebenen in Echtzeit detektiert und Targetstrukturen korrekt erfasst wurden [76]. Einen ähnlichen Ansatz wählten Yacub et al., die ein neuronales Netzwerk einsetzten, was zum einen der Feststellung der Vollständigkeit einer sonografischen Fehlbildungsdiagnostik und zum an-



► **Abb. 8** Automatisierte Schnittebenen-Rekonstruktion eines fetalen Oberschenkels in der 35. SSW zur Schätzung des Fetalgewichts (Oberschenkel-Weichteilmantel nach 5DLimb-Programmapplikation). Nach 3-D-Volumenakquise des transversal ausgerichteten Oberschenkels kann das auf diese Weise errechnete Weichteilvolumen für eine exaktere Schätzung des fetalen Gewichts herangezogen werden.

deren der Qualitätskontrolle der gewonnenen Bilddaten (entsprechend den Vorgaben internationaler Richtlinien) diene. Hier zeigten sich im Vergleich zur manuellen Expertenbegutachtung keine Unterschiede [77, 78]. Die gleichen Modellansätze bilden mittlerweile auch die (intelligente) Basis für das weltweit erste vollständig integrierte KI-Tool (SonoLyst) zur automatisierten biometrischen Erfassung fetaler Zielstrukturen und ebenfalls KI-gestützter Qualitätskontrolle [5]. Welches Potenzial neuronale Netze besitzen, zeigen auch aktuelle Daten einer britischen Arbeitsgruppe zur KI-basierten 2-D-Videoanalyse des Workflows erfahrener Untersucher, die Vorhersagen für die wahrscheinlichsten Schallkopfbewegungen zur Erstellung exakter Zielebenen im Rahmen der Fehlbildungsdiagnostik erlauben [79]. Initiale KI-Modelle der gleichen Arbeitsgruppe konnten zum einen zeigen, dass sie in der Lage waren, Videoinhalte (Schnittebenen) automatisiert zu erkennen und mit entsprechenden Untertiteln zu versehen, und zum anderen konnten speziell trainierte CNN kombinierte Daten eines Bewegungssensors und einer US-Sonde auswerten und in Signale zur augmentierten korrekten Schallkopfführung umwandeln [80, 81].

## KI und weitere klinische Anwendungen in der geburtshilflichen Überwachung

Ein anderes Einsatzgebiet der KI mit direktem klinischen Bezug ist in der Optimierung der biometrischen Genauigkeit zu sehen, die unabhängig von den bereits erwähnten Assistenzsystemen (s. o.) und ungeachtet der in den letzten Jahren deutlich verbesserten apparativen US-Diagnostik nach wie vor eine Herausforderung bleibt. Die Mehrzahl der Schätzmodelle zum fetalen Gewicht basieren auf den im Rahmen der konventionellen 2-D-Ultraschalluntersuchung erhobenen Messparametern (Kopfumfang, biparietaler Durchmesser, Abdomenumfang, Femurdiaphysenlänge). Die

nicht direkt biometrisch quantifizierbare Entwicklung des Weichteilgewebes der oberen und unteren Extremität konnte in der Vergangenheit als guter Surrogatparameter für den fetalen Ernährungszustand etabliert werden [82]. Mit der 3-dimensionalen Erfassung des fraktionierten Extremitätenvolumens (z. B. fractional limb volume, FLV) von Oberarm und/oder -schenkel konnte eine nachweislich höhere Präzision in der Schätzung des fetalen Gewichts selbst bei Mehrlingsschwangerschaften erreicht werden [83]. Mittels automatisierter Techniken, die eine wesentlich schnellere und vor allem untersucherunabhängigere Prozessierung der 3-D-Volumina (effiziente Erkennung und Tracing der Weichteilgrenzen) erlauben, konnte der klinische Nutzen der volumetrischen Erfassung des FLV klar herausgearbeitet werden (5DLimbVol). Hierbei werden workflowbasiert entsprechende axial akquirierte 3-D-Datensätze des Oberarms oder -schenkels analysiert und in die konventionelle Gewichtsschätzung implementiert (► **Abb. 8**) [84, 85].

Auch sub partu können KI-unterstützt sonografische Parameter wie Angle of Progression (AoP) und Head Direction (HD) zum Geburtsfortschritt automatisiert erhoben werden. Erste Ergebnisse hierzu publizierten Youssef et al. 2017 und konstatierten, dass ein automatisierter Ansatz möglich ist und reproduzierbar eingesetzt werden kann [86]. Inwieweit kommerziell verfügbare Softwarelösungen wie LaborAssist einen klinischen Mehrwert bieten, muss allerdings noch abgewartet werden.

Exemplarisch für den mitunter dennoch schwierigen unmittelbaren klinischen Umgang mit automatisierten Techniken zeigt die kontrovers geführte Diskussion über mögliche Vorteile einer computerunterstützten Auswertung der peripartalen fetalen Herzaktion (Electronic Fetal Heart Rate Monitoring), die aufgrund der evidenten Interobservervariabilität und Subjektivität in der Beurteilung von CTG-Auffälligkeiten zumindest theoretisch von einer objektivierbaren automatisierten Analyse profitieren könnte. Die prospektiv-randomisiert erhobenen Daten der INFANT-Studien-Gruppe konnten tatsächlich keinen Vorteil gegenüber der konventionellen visuellen Beurteilung durch das Kreißsaalteam zeigen, weder im neonatalen Kurzzeitoutcome noch bei der Betrachtung des Outcomes nach 2 Jahren [87]. Die Frage, inwieweit methodologische Schwächen im Studiendesign diese nicht signifikanten Unterschiede zwischen den Studienarmen mit verursacht haben (Hawthorne-Effekt), bleibt offen [88, 89], zumal andere computerbasierte Ansätze deutlich erfolversprechende Daten liefern konnten [90].

Fung et al. nutzten hierzu Daten zweier großer populationsbasierter Kohortenstudien (INTERGROWTH 21st und deren Phase-II-Studie – INTERBIO 21st) um zu zeigen, dass mittels maschinellen Lernens die biometrischen Daten eines Ultraschalls zwischen der 20. und 30. SSW und einer erneuten Messung innerhalb der folgenden 10 Schwangerschaftswochen das Gestationsalter mit einer Genauigkeit von 3 Tagen bestimmt und eine Vorhersage der Wachstumskurve der nächsten 6 Wochen individualisiert für jeden Feten vorgenommen werden können [91]. Zweifelsohne wird den Möglichkeiten der KI z. B. bei der Einschätzung und Prädiktion von fetomaternalen Risikokonstellationen wie Frühgeburtlichkeit, Gestationsdiabetes oder hypertensiven Schwangerschaftserkrankungen perspektivisch zunehmend Bedeutung zukommen [92].

## Fazit

Die Autoren einer aktuellen webbasierten Umfrage an 8 Universitätskliniken gaben u. a. an, dass die Mehrzahl der Befragten eher eine positive Einstellung zur KI haben und letztlich davon ausgehen, dass die Zukunft der klinischen Medizin von einer Kombination aus menschlicher und künstlicher Intelligenz geprägt sein und die Patientenversorgung durch sinnvollen Einsatz von KI-Technologien erheblich verbessert werden wird. Das größte Potenzial räumten die Studienteilnehmer dabei der Analyse von sensorgestützten, kontinuierlich erhobenen Daten im Rahmen von Elektrokardio-/Elektroenzephalografie, der intensivmedizinischen Patientenüberwachung oder Bildgebungsverfahren in der gezielten Diagnostik und Workflow-Unterstützung ein [93]. Ganz konkret bleibt bezogen auf unser Fachgebiet festzuhalten, dass die stetige Weiterentwicklung der US-Systeme und Ausstattung dieser mit z. B. hochauflösenden US-Sonden/Matrixsonden für die gynäkologisch-geburtshilfliche Diagnostik sowie der unaufhaltsame Einzug effizienter automatisierter Segmentierungstechniken 2- und insbesondere 3-dimensionaler Bildinformationen die gesamte Prozesskette von der Erstellung, Analyse, Verarbeitung bis zur Verwaltung von Bilddaten in Zukunft zunehmend beeinflussen und optimieren werden.

In einem aktuellen systematischen Review von mehr als 80 Studien zur automatisierten Bildanalyse konnte eine äquivalente Präzision der KI im Vergleich zu der von Experten auf dem jeweiligen Gebiet festgestellt werden. Die Autoren stellten allerdings auch fest, dass in vielen Publikationen eine externe Validierung der jeweiligen KI-Algorithmen nicht oder nur unzureichend stattgefunden hatte. Ein Umstand, der zusammen mit der bereits in vielen Bereichen gut umgesetzten, aber nach wie vor zwingend zu intensivierenden Kollaboration zwischen KI-Entwicklern und Klinikern eine weitere Implementierung in relevante klinische Prozesse derzeit noch erschwert [94]. Allegorisch lässt sich die derzeitige Situation der KI und ihrer Anwendungen im Gesundheitswesen treffend mit einem brandneuen Auto vergleichen, das, um adäquat genutzt werden zu können, Treibstoff und Straßen benötigt. In die Welt des maschinellen Lernens transferiert, bedeutet dies, dass die jeweiligen Algorithmen z. B. mit (annotierten) Bilddaten „betankt“ werden müssen, diese aber auch nur bei entsprechender Infrastruktur i. S. von effizienten und skalierbaren Prozessen mit KI-bereitem Workflow ihr Potenzial ausschöpfen können [21].

## Ausblick

Die zunehmende Entwicklung von KI-Systemen und deren Integration in klinische Prozesse geht mit enormen Erwartungen an die Weiterentwicklung der medizinischen Versorgung einher. Fest steht, dass die Integration dieser Tools Arbeits- und Trainingsmethoden perspektivisch grundlegend verändern wird, indem das gesamte medizinische Personal durch die Bereitstellung von schnell und verlässlich erhobenen Daten und Sachverhalten in Befundinterpretation und Konsultationen unterstützt werden wird, wobei sich dieses idealerweise vermehrt auf die einzigartig menschlichen Elemente seines Berufsstandes konzentrieren kann. Diejenigen Aufgaben, die von einer Maschine nicht ausgeführt werden können, weil sie emotionale Intelligenz erfordern, wie z. B.

die gezielte Patienteninteraktion, um differenziertere Symptome zu eruieren und durch menschliche Intuition Vertrauen aufzubauen, unterstreichen die einzigartige Qualifikation und den hervorzuhebenden Stellenwert des menschlichen Faktors im Umgang mit klinischen KI-Anwendungen der Zukunft [95]. Nicht zuletzt deswegen sind wir noch weit davon entfernt, den Menschen effektiv durch maschinelle Lösungen zu ersetzen. An dieser Stelle sei auf die visionären Ausführungen von Fritz Kahn verwiesen („Der Arzt der Zukunft“), die im Licht der aktuellen und zukünftigen KI-Technologien in der Medizin eine derart plastische Form von Konstruktivismus beschreiben und so bereits vor fast 100 Jahren erahnen lassen, wie die technologische Zivilisation und experimentelle Wissenschaft in die biologische Natur des menschlichen Körpers synergistisch eingreifen können [96,97]. Was sich daraus u. a. auch erschließt, ist, dass sich der Mensch allem technischen Fortschritt zum Trotz bislang nicht überflüssig gemacht hat bzw. machen wird. Prognosen, nach denen bis 47% aller Jobs einer Automatisierung zum Opfer fallen werden, scheinen sich so nicht halten zu lassen, da insbesondere im Gesundheitswesen die Arbeitsplatzaufbaurate die -verlustrate nachweislich übersteigt [46,98,99].

Um das Potenzial von KI-Algorithmen optimal ausschöpfen zu können, ist die interdisziplinäre Kommunikation und ständige Einbeziehung von Ärztinnen und Ärzten als Hauptanwender dieser Tools in die Entwicklungsprozesse und Funktionsweisen von KI-Lösungen unerlässlich, da sonst die Medizin von morgen ausschließlich von der Vision der Ingenieure geprägt sein und weniger den tatsächlichen Ansprüchen im Hinblick auf eine personalisierte (Präzisions-)Medizin gerecht wird [47,100]. Auf einem Konsensus-Workshop radiologischer Fachgesellschaften aus 2018 formulierten die Teilnehmer perspektivisch die dringlichsten Forschungsschwerpunkte zur KI wie in ► **Tab. 1** aufgeführt [101,102]. Dem wäre aus gynäkologisch-geburtshilflicher Perspektive noch hinzuzufügen, dass z. B. hinsichtlich der KI-unterstützten sonografischen Diagnostik weitere Anstrengungen zur Bildoptimierung (Pre-/Postprocessing) sowohl in der konventionellen 2-D-Darstellung, als auch der 3-D/4-D-Volumensonografie unternommen werden und dass ähnlich der bestehenden Algorithmen mit automatisiertem Workflow weitere KI-Technologien mit intuitiver Bedienführung, einfacher Anwendbarkeit und genereller (geräteübergreifender) Verfügbarkeit zur effizienten Analyse der Bild- und Volumendaten entwickelt werden müssen. Darüber hinaus sollte im Rahmen der Routinediagnostik die Integration von Assistenzsystemen zur Ebeneneinstellung und Zielstruktur-Quantifizierung in Echtzeit weiterverfolgt werden. Interessant ist in diesem Zusammenhang zu erwähnen, dass das Einbinden bereits vortrainierter Algorithmen für die Analyse eigener populationsbasierter Daten (transfer learning) eine attraktive und vor allem verlässliche Methode ist, da das Training eines neuen neuronalen Netzes mit einer großen Anzahl von Daten rechen- und zeitintensiv ist [103]. Dabei werden die bestehenden, vortrainierten Schichten eines CNN übernommen und nur die Ausgabeschicht auf die Anzahl der zu erkennenden Objektklassen des neuen Netzes angepasst und nachtrainiert.

Die sehr wahrscheinlich größte Herausforderung im Hinblick auf den gezielten Einsatz der KI im Gesundheitswesen generell besteht allerdings nicht darin, ob die automatisierten Technologien

► **Tab. 1** Empfehlungen des 2018 in Bethesda/USA abgehaltenen Konsensus-Workshops zur translationellen Forschung für die Weiterentwicklung und Integration von Anwendungen der künstlichen Intelligenz in klinischen Prozessen (nach Allen et al. 2019, Langlotz et al. 2019 [101, 102]).

#### Forschungsthemen zur künstlichen Intelligenz in der medizinischen Bildung

Erstellen strukturierter KI-Anwendungsbeispiele und Definition klinischer Problemstellungen, bei denen KI-Algorithmen tatsächliche Lösungsansätze bieten können

Entwicklung neuartiger Bildrekonstruktionsmethoden, mit denen aus Quelldaten effizient Bilder generiert werden können

Etablierung automatisierter Bildkennzeichnungs- und Annotationsmethoden, die effizient Trainingsdaten für die Erforschung moderner ML-Modelle liefern und deren intensivierten klinischen Einsatz ermöglichen

Forschungsbedarf zu Methoden des maschinellen Lernens, die den Anwenden KI-basierte Entscheidungshilfen konkreter vermitteln und veranschaulichen können

Etablierung von Methoden zur Validierung und objektiven Leistungsüberwachung von KI-Algorithmen, um behördliche Genehmigungsprozesse zu erleichtern

Entwicklung von Standards und gemeinsamer Datenplattformen für die problemlose Integration von KI-Tools in bestehende klinische Arbeitsabläufe

die an sie gestellten Anforderungen vollumfänglich erfüllen, sondern vielmehr darin, ihre Implementierung in die tägliche klinische Praxis im Detail sicherzustellen. Hierzu müssen u. a. entsprechende Genehmigungsverfahren eingeleitet, die (klinische) Infrastruktur hergestellt, eine Standardisierung erreicht und vor allem das klinische Personal ausreichend geschult werden. Dass diese Hürden perspektivisch überwunden werden, ist sicher, nur kann es durchaus länger dauern, bis die Technologien selbst ausgereift sind. Es wird daher erwartet, dass innerhalb der kommenden 5 Jahre ein noch begrenzter Einsatz von KI in der klinischen Praxis zu sehen sein wird (mit einem umfassenderen Einsatz innerhalb von 10 Jahren) [104].

#### Interessenkonflikt

Die Autoren (JW, MG) erklären, dass sie innerhalb der vergangenen 3 Jahre Vortragshonorare von Samsung HME und GE Healthcare erhalten haben.

#### Literatur

- [1] Drukker L, Droste R, Chatelain P et al. Expected-value bias in routine third-trimester growth scans. *Ultrasound Obstet Gynecol* 2020; 55: 375–382
- [2] Deng J, Dong W, Socher R et al. ImageNet: A large-scale hierarchical image database. Paper presented at: 2009 IEEE Conference on Computer Vision and Pattern Recognition; 20–25 June 2009, 2009
- [3] Russakovsky O, Deng J, Su H et al. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis* 2015; 115: 211–252
- [4] Turing AM. I – Computing Machinery and Intelligence. *Mind* 1950; LIX: 433–460

- [5] Deloitte. State of AI in the enterprise – 3rd ed. Deloitte, 2020. Online (Stand: 30.9.2021): [www2.deloitte.com/content/dam/Deloitte/de/Documents/technology-media-telecommunications/DELO-6418\\_State%20of%20AI%202020\\_KS4.pdf](http://www2.deloitte.com/content/dam/Deloitte/de/Documents/technology-media-telecommunications/DELO-6418_State%20of%20AI%202020_KS4.pdf)
- [6] Anthes E. Alexa, do I have COVID-19? *Nature* 2020; 586: 22–25
- [7] Huang Z, Epps J, Joachim D. Investigation of Speech Landmark Patterns for Depression Detection. *IEEE Transactions on Affective Computing* 2019. doi:10.1109/TAFFC.2019.2944380
- [8] Bodalal Z, Trebeschi S, Nguyen-Kim TDL et al. Radiogenomics: bridging imaging and genomics. *Abdom Radiol (NY)* 2019; 44: 1960–1984
- [9] Allen B, Dreyer K, McGinty GB. Integrating Artificial Intelligence Into Radiologic Practice: A Look to the Future. *J Am Coll Radiol* 2020; 17: 280–283
- [10] Purohit K. Growing Interest in Radiology Despite AI Fears. *Acad Radiol* 2019; 26: e75
- [11] Richardson ML, Garwood ER, Lee Y et al. Noninterpretive Uses of Artificial Intelligence in Radiology. *Acad Radiol* 2021; 28: 1225–1235
- [12] Bennani-Baiti B, Baltzer PAT. Künstliche Intelligenz in der Mammadiagnostik. *Radiologe* 2020; 60: 56–63
- [13] Chan H-P, Samala RK, Hadjiiski LM. CAD and AI for breast cancer—recent development and challenges. *Br J Radiol* 2019; 93: 20190580
- [14] Fujita H. AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. *Radiol Phys Technol* 2020; 13: 6–19
- [15] McKinney SM, Sieniek M, Godbole V et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020; 577: 89–94
- [16] Rodríguez-Ruiz A, Lång K, Gubern-Merida A et al. Can we reduce the workload of mammographic screening by automatic identification of normal exams with artificial intelligence? A feasibility study. *Eur Radiol* 2019; 29: 4825–4832
- [17] O’Connell AM, Bartolotta TV, Orlando A et al. Diagnostic Performance of An Artificial Intelligence System in Breast Ultrasound. *J Ultrasound Med* 2021. doi:10.1002/jum.15684
- [18] Cho BJ, Choi YJ, Lee MJ et al. Classification of cervical neoplasms on colposcopic photography using deep learning. *Sci Rep* 2020; 10: 13652
- [19] Shanthi PB, Faruqi F, Hareesha KS et al. Deep Convolution Neural Network for Malignancy Detection and Classification in Microscopic Uterine Cervix Cell Images. *Asian Pac J Cancer Prev* 2019; 20: 3447–3456
- [20] Försch S, Klauschen F, Hufnagl P et al. Künstliche Intelligenz in der Pathologie. *Dtsch Arztebl* 2021; 118: 199–204
- [21] Chang PJ. Moving Artificial Intelligence from Feasible to Real: Time to Drill for Gas and Build Roads. *Radiology* 2020; 294: 432–433
- [22] Tran D, Cooke S, Illingworth PJ et al. Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. *Hum Reprod* 2019; 34: 1011–1018
- [23] Zaninovic N, Rosenwaks Z. Artificial intelligence in human in vitro fertilization and embryology. *Fertil Steril* 2020; 114: 914–920
- [24] Bori L, Paya E, Alegre L et al. Novel and conventional embryo parameters as input data for artificial neural networks: an artificial intelligence model applied for prediction of the implantation potential. *Fertil Steril* 2020; 114: 1232–1241
- [25] DEGUM. Pressemitteilungen. DEGUM, 2017. Updated 29.11.2017. Online (Stand: 30.09.2021): [www.degum.de/aktuelles/presse-medien/pressemitteilungen/im-detail/news/zu-viele-kindliche-fehlbildungen-bleiben-unentdeckt.html](http://www.degum.de/aktuelles/presse-medien/pressemitteilungen/im-detail/news/zu-viele-kindliche-fehlbildungen-bleiben-unentdeckt.html)
- [26] Murugesu S, Galazis N, Jones BP et al. Evaluating the use of telemedicine in gynaecological practice: a systematic review. *BMJ Open* 2020; 10: e039457
- [27] Benacerraf BR, Minton KK, Benson CB et al. Proceedings: Beyond Ultrasound First Forum on Improving the Quality of Ultrasound Imaging in Obstetrics and Gynecology. *J Ultrasound Med* 2018; 37: 7–18

- [28] Timmerman D, Verrelst H, Bourne TH et al. Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses. *Ultrasound Obstet Gynecol* 1999; 13: 17–25
- [29] Froyman W, Timmerman D. Methods of Assessing Ovarian Masses: International Ovarian Tumor Analysis Approach. *Obstet Gynecol Clin North Am* 2019; 46: 625–641
- [30] Van Calster B, Van Hoorde K, Valentin L et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. *BMJ* 2014; 349: g5920
- [31] Vázquez-Manjarrez SE, Rico-Rodriguez OC, Guzman-Martinez N et al. Imaging and diagnostic approach of the adnexal mass: what the oncologist should know. *Chin Clin Oncol* 2020; 9: 69
- [32] Andreotti RF, Timmerman D, Strachowski LM et al. O-RADS US Risk Stratification and Management System: A Consensus Guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee. *Radiology* 2020; 294: 168–185
- [33] Christiansen F, Epstein EL, Smedberg E et al. Ultrasound image analysis using deep neural networks for discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment. *Ultrasound Obstet Gynecol* 2021; 57: 155–163
- [34] Acharya UR, Mookiah MR, Vinitha Sree S et al. Evolutionary algorithm-based classifier parameter tuning for automatic ovarian cancer tissue characterization and classification. *Ultraschall Med* 2014; 35: 237–245
- [35] Akazawa M, Hashimoto K. Artificial Intelligence in Ovarian Cancer Diagnosis. *Anticancer Res* 2020; 40: 4795–4800
- [36] Aramendia-Vidaurreta V, Cabeza R, Villanueva A et al. Ultrasound Image Discrimination between Benign and Malignant Adnexal Masses Based on a Neural Network Approach. *Ultrasound Med Biol* 2016; 42: 742–752
- [37] Khazendar S, Sayasneh A, Al-Assam H et al. Automated characterisation of ultrasound images of ovarian tumours: the diagnostic accuracy of a support vector machine and image processing with a local binary pattern operator. *Facts Views Vis Obgyn* 2015; 7: 7–15
- [38] Zhou J, Zeng ZY, Li L. Progress of Artificial Intelligence in Gynecological Malignant Tumors. *Cancer Manag Res* 2020; 12: 12823–12840
- [39] Al-Karawi D, Al-Assam H, Du H et al. An Evaluation of the Effectiveness of Image-based Texture Features Extracted from Static B-mode Ultrasound Images in Distinguishing between Benign and Malignant Ovarian Masses. *Ultrason Imaging* 2021; 43: 124–138
- [40] Bakker MK, Bergman JEH, Krikov S et al. Prenatal diagnosis and prevalence of critical congenital heart defects: an international retrospective cohort study. *BMJ Open* 2019; 9: e028139
- [41] van Nesselrooij AEL, Teunissen AKK, Clur SA et al. Why are congenital heart defects being missed? *Ultrasound Obstet Gynecol* 2020; 55: 747–757
- [42] Knackstedt C, Bekkers SC, Schummers G et al. Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: The FAST-EFs Multicenter Study. *J Am Coll Cardiol* 2015; 66: 1456–1466
- [43] Tsang W, Salgo IS, Medvedofsky D et al. Transthoracic 3D Echocardiographic Left Heart Chamber Quantification Using an Automated Adaptive Analytics Algorithm. *JACC Cardiovasc Imaging* 2016; 9: 769–782
- [44] Kusunose K. Steps to use artificial intelligence in echocardiography. *J Echocardiogr* 2021; 19: 21–27
- [45] Zhang J, Gajjala S, Agrawal P et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. *Circulation* 2018; 138: 1623–1635
- [46] Harari YN. Homo sapiens verliert die Kontrolle. Die Große Entkopplung. In: *Homo Deus – Eine Geschichte von Morgen*. 16. Aufl. München: C. H. Beck; 2020
- [47] Gandhi S, Mosleh W, Shen J et al. Automation, machine learning, and artificial intelligence in echocardiography: A brave new world. *Echocardiography* 2018; 35: 1402–1418
- [48] Arnaout R, Curran L, Zhao Y et al. Expert-level prenatal detection of complex congenital heart disease from screening ultrasound using deep learning. *medRxiv* 2020. doi:10.1101/2020.06.22.20137786
- [49] Le TK, Truong V, Nguyen-Vo TH et al. Application of machine learning in screening of congenital heart diseases using fetal echocardiography. *J Am Coll Cardiol* 2020; 75: 648
- [50] Dong J, Liu S, Liao Y et al. A Generic Quality Control Framework for Fetal Ultrasound Cardiac Four-Chamber Planes. *IEEE J Biomed Health Inform* 2020; 24: 931–942
- [51] Hinton GE. To recognize shapes, first learn to generate images. *Prog Brain Res* 2007; 165: 535–547
- [52] Voelker R. Cardiac Ultrasound Uses Artificial Intelligence to Produce Images. *JAMA* 2020; 323: 1034
- [53] Yeo L, Romero R. Optical ultrasound simulation-based training in obstetric sonography. *J Matern Fetal Neonatal Med* 2020. doi:10.1080/14767058.2020.1786519
- [54] Steinhard J, Dammeme Debbeh A, Laser KT et al. Randomised controlled study on the use of systematic simulator-based training (OPUS Fetal Heart Trainer) for learning the standard heart planes in fetal echocardiography. *Ultrasound Obstet Gynecol* 2019; 54 (S1): 28–29
- [55] Day TG, Kainz B, Hajnal J et al. Artificial intelligence, fetal echocardiography, and congenital heart disease. *Prenat Diagn* 2021; 41: 733–742. doi:10.1002/pd.5892
- [56] Garcia-Canadilla P, Sanchez-Martinez S, Crispi F et al. Machine Learning in Fetal Cardiology: What to Expect. *Fetal Diagn Ther* 2020; 47: 363–372
- [57] Meiburger KM, Acharya UR, Molinari F. Automated localization and segmentation techniques for B-mode ultrasound images: A review. *Comput Biol Med* 2018; 92: 210–235
- [58] Rawat V, Jain A, Shrimali V. Automated Techniques for the Interpretation of Fetal Abnormalities: A Review. *Appl Bionics Biomech* 2018; 2018: 6452050
- [59] Yeo L, Luewan S, Romero R. Fetal Intelligent Navigation Echocardiography (FINE) Detects 98% of Congenital Heart Disease. *J Ultrasound Med* 2018; 37: 2577–2593
- [60] Gembicki M, Hartge DR, Dracopoulos C et al. Semiautomatic Fetal Intelligent Navigation Echocardiography Has the Potential to Aid Cardiac Evaluations Even in Less Experienced Hands. *J Ultrasound Med* 2020; 39: 301–309
- [61] Weichert J, Weichert A. A “holistic” sonographic view on congenital heart disease: How automatic reconstruction using fetal intelligent navigation echocardiography eases unveiling of abnormal cardiac anatomy part II-Left heart anomalies. *Echocardiography* 2021; 38: 777–789
- [62] DeVore GR, Klas B, Satou G et al. Longitudinal Annular Systolic Displacement Compared to Global Strain in Normal Fetal Hearts and Those With Cardiac Abnormalities. *J Ultrasound Med* 2018; 37: 1159–1171
- [63] DeVore GR, Klas B, Satou G et al. 24-segment sphericity index: a new technique to evaluate fetal cardiac diastolic shape. *Ultrasound Obstet Gynecol* 2018; 51: 650–658
- [64] DeVore GR, Polanco B, Satou G et al. Two-Dimensional Speckle Tracking of the Fetal Heart: A Practical Step-by-Step Approach for the Fetal Sonologist. *J Ultrasound Med* 2016; 35: 1765–1781
- [65] Lee M, Won H. Novel technique for measurement of fetal right myocardial performance index using synchronised images of right ventricular inflow and outflow. *Ultrasound Obstet Gynecol* 2019; 54 (S1): 178–179
- [66] Leung V, Avnet H, Henry A et al. Automation of the Fetal Right Myocardial Performance Index to Optimise Repeatability. *Fetal Diagn Ther* 2018; 44: 28–35
- [67] Rizzo G, Aiello E, Pietrolucci ME et al. The feasibility of using 5D CNS software in obtaining standard fetal head measurements from volumes acquired by three-dimensional ultrasonography: comparison with two-dimensional ultrasound. *J Matern Fetal Neonatal Med* 2016; 29: 2217–2222

- [68] Welp A, Gembicki M, Rody A et al. Validation of a semiautomated volumetric approach for fetal neurosonography using 5DCNS+ in clinical data from > 1100 consecutive pregnancies. *Childs Nerv Syst* 2020; 36: 2989–2995
- [69] Pluym ID, Afshar Y, Holliman K et al. Accuracy of three-dimensional automated ultrasound imaging of biometric measurements of the fetal brain. *Ultrasound Obstet Gynecol* 2021; 57: 798–803
- [70] Ambrose Grandjean G, Hossu G, Bertholdt C et al. Artificial intelligence assistance for fetal head biometry: Assessment of automated measurement software. *Diagn Interv Imaging* 2018; 99: 709–716
- [71] Huang R, Xie W, Alison Noble J. VP-Nets: Efficient automatic localization of key brain structures in 3D fetal neurosonography. *Med Image Anal* 2018; 47: 127–139
- [72] Xie HN, Wang N, He M et al. Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal. *Ultrasound Obstet Gynecol* 2020; 56: 579–587
- [73] Cerralza JJ, Li Y, Biffi C et al. Fetal Skull Reconstruction via Deep Convolutional Autoencoders. *Annu Int Conf IEEE Eng Med Biol Soc* 2018; 2018: 887–890
- [74] Ghesu FC, Georgescu B, Grbic S et al. Towards intelligent robust detection of anatomical structures in incomplete volumetric data. *Med Image Anal* 2018; 48: 203–213
- [75] Cai Y, Droste R, Sharma H et al. Spatio-temporal visual attention modeling of standard biometry plane-finding navigation. *Medical Image Analysis* 2020; 65: 101762
- [76] Baumgartner CF, Kamnitsas K, Matthew J et al. SonoNet: Real-Time Detection and Localisation of Fetal Standard Scan Planes in Freehand Ultrasound. *IEEE Trans Med Imaging* 2017; 36: 2204–2215
- [77] Yaqub M, Kelly B, Noble JA et al. An AI system to support sonologists during fetal ultrasound anomaly screening. *Ultrasound Obstet Gynecol* 2018; 52 (S1): 9–10
- [78] Yaqub M, Sleep N, Syme S et al. ScanNav® audit: an AI-powered screening assistant for fetal anatomical ultrasound. *Am J Obstet Gynecol* 2021; 224 (Suppl.): S312. doi:10.1016/j.ajog.2020.12.512
- [79] Sharma H, Drukker L, Chatelain P et al. Knowledge representation and learning of operator clinical workflow from full-length routine fetal ultrasound scan videos. *Med Image Anal* 2021; 69: 101973
- [80] Droste R, Drukker L, Papageorgiou AT et al. Automatic Probe Movement Guidance for Freehand Obstetric Ultrasound. *Med Image Comput Assist Interv* 2020; 12263: 583–592
- [81] Alsharid M, Sharma H, Drukker L et al. Captioning Ultrasound Images Automatically. In: Shen D et al., eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. MICCAI 2019. Lecture Notes in Computer Science, vol 11767. Cham: Springer; 2019. doi:10.1007/978-3-030-32251-9\_37
- [82] Lee W, Deter RL, Ebersole JD et al. Birth weight prediction by three-dimensional ultrasonography: fractional limb volume. *J Ultrasound Med* 2001; 20: 1283–1292
- [83] Corrêa VM, Araujo Júnior E, Braga A et al. Prediction of birth weight in twin pregnancies using fractional limb volumes by three-dimensional ultrasonography. *J Matern Fetal Neonatal Med* 2020; 33: 3652–3657
- [84] Gembicki M, Offerman DR, Weichert J. Semiautomatic Assessment of Fetal Fractional Limb Volume for Weight Prediction in Clinical Praxis: How Does It Perform in Routine Use? *J Ultrasound Med* 2021. doi:10.1002/jum.15712
- [85] Mack LM, Kim SY, Lee S et al. Automated Fractional Limb Volume Measurements Improve the Precision of Birth Weight Predictions in Late Third-Trimester Fetuses. *J Ultrasound Med* 2017; 36: 1649–1655
- [86] Youssef A, Salsi G, Montaguti E et al. Automated Measurement of the Angle of Progression in Labor: A Feasibility and Reliability Study. *Fetal Diagn Ther* 2017; 41: 293–299
- [87] Brocklehurst P, Field D, Greene K et al. Computerised interpretation of fetal heart rate during labour (INFANT): a randomised controlled trial. *Lancet* 2017; 389: 1719–1729
- [88] Keith R. The INFANT study—a flawed design foreseen. *Lancet* 2017; 389: 1697–1698
- [89] Silver RM. Computerising the intrapartum continuous cardiotocography does not add to its predictive value: FOR: Computer analysis does not add to intrapartum continuous cardiotocography predictive value. *BJOG* 2019; 126: 1363
- [90] Gyllencreutz E, Lu K, Lindecrantz K et al. Validation of a computerized algorithm to quantify fetal heart rate deceleration area. *Acta Obstet Gynecol Scand* 2018; 97: 1137–1147
- [91] Fung R, Villar J, Dashti A et al.; International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st). Achieving accurate estimates of fetal gestational age and personalised predictions of fetal growth based on data from an international prospective cohort study: a population-based machine learning study. *Lancet Digit Health* 2020; 2: e368–e375
- [92] Lee KS, Ahn KH. Application of Artificial Intelligence in Early Diagnosis of Spontaneous Preterm Labor and Birth. *Diagnostics (Basel)* 2020; 10: 733
- [93] Maassen O, Fritsch S, Palm J et al. Future Medical Artificial Intelligence Application Requirements and Expectations of Physicians in German University Hospitals: Web-Based Survey. *J Med Internet Res* 2021; 23: e26646
- [94] Littmann M, Selig K, Cohen-Lavi L et al. Validity of machine learning in biology and medicine increased through collaborations across fields of expertise. *Nature Machine Intelligence* 2020; 2: 18–24
- [95] Norgeot B, Glicksberg BS, Butte AJ. A call for deep-learning healthcare. *Nat Med* 2019; 25: 14–15
- [96] Borck C. Communicating the Modern Body: Fritz Kahn's Popular Images of Human Physiology as an Industrialized World. *Canadian Journal of Communication* 2007; 32: 495–520
- [97] Jachertz N. Populärmedizin: Der Mensch ist eine Maschine, die vom Menschen bedient wird. *Dtsch Arztebl* 2010; 107: A-391–393
- [98] Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change* 2017; 114: 254–280
- [99] Gartner H, Stüber H. Strukturwandel am Arbeitsmarkt seit den 70er Jahren: Arbeitsplatzverluste werden durch neue Arbeitsplätze immer wieder ausgeglichen. 16.7.2019. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung; 2019
- [100] Bartoli A, Quarello E, Voznyuk I et al. Intelligence artificielle et imagerie en médecine fœtale: de quoi parle-t-on? [Artificial intelligence and fetal imaging: What are we talking about?]. *Gynecol Obstet Fertil Senol* 2019; 47: 765–768
- [101] Allen B jr., Seltzer SE, Langlotz CP et al. A Road Map for Translational Research on Artificial Intelligence in Medical Imaging: From the 2018 National Institutes of Health/RSNA/ACR/The Academy Workshop. *J Am Coll Radiol* 2019; 16 (9 Pt A): 1179–1189
- [102] Langlotz CP, Allen B, Erickson BJ et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019; 291: 781–791
- [103] Tolsgaard MG, Svendsen MBS, Thybo JK et al. Does artificial intelligence for classifying ultrasound imaging generalize between different populations and contexts? *Ultrasound Obstet Gynecol* 2021; 57: 342–343
- [104] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J* 2019; 6: 94–98