

Gynäkologe

<https://doi.org/10.1007/s00129-021-04808-2>

Angenommen: 14. April 2021

© Springer Medizin Verlag GmbH, ein Teil von Springer Nature 2021

Redaktion

K. Vetter, Berlin

W. Janni, Ulm

R. Kreienberg, Mainz



Daniel Sonntag

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Universität Oldenburg, Oldenburg, Deutschland

## Künstliche Intelligenz in der Medizin und Gynäkologie – Holzweg oder Heilversprechen?

**Die Bundesregierung ist entschlossen, sowohl Forschung und Entwicklung als auch Anwendung von Künstlicher Intelligenz (KI) in Deutschland und Europa auf ein weltweit führendes Niveau zu bringen. Ende Juli 2018 sind die Eckpunkte für eine Strategie KI veröffentlicht. Im Beitrag werden die wichtigsten aufgegriffen, um sie in den Anwendungsbereich der Medizin und der Gynäkologie zu übertragen. Beleuchtet werden Ausgangssituation, Handlungsfelder und Transfer in die medizinische Versorgung vor dem Hintergrund der KI-Forschung in Deutschland, Europa und den USA. Welche Bedeutung hat KI für das Gesundheitswesen? Welche Potenziale und Limitationen können aufgeführt werden? Welche positiven und welche negativen Beispiele gibt es?**

### Hintergrund

Entscheidend für die erfolgreiche Anwendung von KI sind der Zugang zu Daten und die Integration in komplexe medizinische Dienstleistungen im klinischen und nichtklinischen Umfeld. Darum muss die Menge an nutzbaren, qualitativ hochwertigen Daten deutlich erhöht werden.

Meiner Auffassung nach wird in Zukunft der Schlüssel zum Erfolg in der Wissensakquisitionsstrategie liegen.

Aktualisierte Version der Originalpublikation in *HNO* (2019) 67:343–349. <https://doi.org/10.1007/s00106-019-0665-z>.

Technologien können relativ leicht repliziert werden, qualitative Daten nicht, weil sie auf längere Sicht zur Verarbeitung mit KI-Technologien mühsam aufgebaut werden müssen. Darum kann Deutschland längerfristig damit punkten, die Erfassung patientenrelevanter Basisdaten und anderer digitaler Informationen zum Krankheitsstand zu motivieren und auszubauen.

### Entwicklung der KI in Deutschland

Nach der ersten Jahresversammlung der 1969 gegründeten Gesellschaft für Informatik (GI) lässt sich die KI in Deutschland auf das Jahr 1975 datieren, etwa 20 Jahre nach der *Dartmouth Conference* in den USA, auf welcher der Begriff *Artificial Intelligence* geprägt wurde. Der KI-Unterausschuss der GI war nach 1972 entscheidend für die Etablierung des Forschungsfeldes in der Bundesrepublik. Dieser Unterausschuss wurde sehr durch die Arbeitsgruppe Mustererkennung geprägt, das Feld der KI, das momentan die größte Aufmerksamkeit in der Medizin genießt. Der heutige Fachbereich der KI (FBKI, <https://fb-ki.gi.de>) zählt über tausend Mitglieder.

Die Teilgebiete der KI sind Datenbanken und Wissensrepräsentation, Deduktion und Verifikation, Sprachverarbeitung, Computersehen und Bildverarbeitung, Robotik, Multiagentensysteme, maschinelles Lernen (ML) und andere. Wie von Siekmann [33] angedeutet, gibt es keine offizielle Gliederung des Faches in Teildisziplinen, und die Veränderun-

gen sind selbst in kurzen Zeiträumen von 2–3 Jahren oft erheblich. Einen Anhaltspunkt bieten nach wie vor die großen internationalen Konferenzen (IJCAI [International Joint Conferences on Artificial Intelligence Organization], AAAI [Association for the Advancement of Artificial Intelligence], ECAI [European Conference on Artificial Intelligence]), neuerdings Konferenzen mit starker Betonung der Mensch-Maschine-Interaktion, wie die ACM (Association for Computing Machinery)-Konferenzen in diesem Bereich (IUI [Intelligent User Interfaces], ACM-CHI [ACM Conference on Human Factors in Computing Systems], UMAP [User Modeling, Adaptation and Personalization]), und Lehrbücher [38]. Die Gliederung auf der Webseite des Fachbereichs KI der GI ist auch ein guter Anhaltspunkt.

Die *Deutsche Zeitschrift für Künstliche Intelligenz* (*German Journal of Artificial Intelligence, KI-Journal*) ist das offizielle Sprachrohr der deutschen KI-Community, neben der deutschen KI-Konferenz. Beide entwickelten sich aus dem Rundbrief zur Vorbereitung der oben erwähnten GI-Fachgruppe KI im Jahre 1975. Die Hefte ab 1988 sind bei DBLP (Digital Bibliography and Library Project) indexiert und verfügbar und immer noch sehr lesenswert, beispielsweise die über Expertensysteme im medizinischen Bereich [46]. Seit 2010 wird das *KI-Journal* in Zusammenarbeit mit dem Springer-Verlag mit weiterer Internationalisierung herausgegeben [36]. Das letzte Heft mit Fokus auf KI in der Medizin wurde 2015 veröffentlicht [15], mit dem Ver-

such, klinische und nichtklinische Systeme in der KI-Community zu fördern. Die Publikationsorgane der KI-Community sind zum gegenwärtigen Zeitpunkt immer noch stark getrennt von den medizinischen Publikationsorganen. Ausnahmen sind Fachkonferenzen und Journals im bildgebenden Bereich, wie z. B. RSNA (Radiology and Machine Learning: from Dialogue to Clinical Practice), MICCAI (Medical Image Computing and Computer Assisted Intervention Society) oder CBMS (Computer-Based Medical Systems). Es bleibt zu hoffen, dass sich Journals, wie beispielsweise *Artificial Intelligence in Medicine* oder *JAMIA*, als Brücke in Zukunft weiter etablieren können, auch im Hinblick auf frei verfügbare KI-Software für den medizinischen Bereich, was auch am DFKI (Deutsches Forschungszentrum für Künstliche Intelligenz) stark gefördert wird [40], oder Analysen von Fehlfunktionen klinischer Entscheidungsunterstützungssysteme [48].

Neben Entscheidungsunterstützungssystemen setzen sich Monitoringsysteme für Patienten und Roboter im Operationssaal durch; dies deckt sich mit internationalen Vorhersagen [45].

Ein Blick auf die KI-Strategie der Bundesregierung von 2018 identifiziert 12 Handlungsfelder, u. a. Daten verfügbar und nutzbar zu machen (Handlungsfeld 8) sowie nationale und internationale Vernetzung (Handlungsfeld 11).<sup>1</sup>

## Transferziele und Beispiele

Mit der Verfügbarkeit höherer Rechenleistung und der Fokussierung auf spezielle, realitätsnahe Aufgaben (z. B. Bilderkennung in der Dermatologie mit über 100.000 Bildern; [13]) oder auf Röntgenaufnahmen [47] mit über 80.000 Bildern für ML wird es der KI-Forschung wahrscheinlich in den nächsten 4–5 Jahren gelingen, sich erfolgreich zu etablieren. Hier ist die Qualität der Ergebnisse bei ganz speziellen Fragestellungen mit der von Fachärzten vergleichbar.

Ein anderer Anreiz für Ärzte ist eine schnelle und kostengünstige Zweitmeinung. Die Bedeutung für das Gesund-

heitswesen kann beispielhaft aufgezeigt werden: Die KI wird erstens dazu beitragen, dass Ärzte durch geschickte Vorverarbeitung von medizinischen Texten und medizinischen Bildern Dinge erkennen, die ihnen nicht ohne Weiteres aufgefallen wären.

Ein zweites Beispiel ist im nichtklinischen Bereich zu finden: Um den Angehörigen von Pflegebedürftigen mehr Alltagsfreiheit zu verschaffen, werden Roboter einfache Haushaltsaufgaben mithilfe ihrer sensomotorischen Intelligenz übernehmen und durch Sprach- und Videofunktionen eine Brücke zur Außenwelt schaffen.

Drittens wird durch KI-Systeme auch jeder Patient die Möglichkeit haben, von zuhause per Knopfdruck eine Meinung zur Diagnose und Therapie einzuholen, schon bevor er zum Arzt geht.

Es gibt prinzipiell 3 inhaltliche Problemstellungen: Erstens, die robuste Vorhersage, zweitens, die Abdeckung der Krankheitsfälle der Entscheidungsunterstützung und drittens die Nachvollziehbarkeit der Handlungsentscheidungen. Die ersten beiden Challenges beziehen sich auf die Qualität und Verfügbarkeit von Trainingsdaten. Eine Lösungsarchitektur wird mit inkrementeller Wissensakquisition bereitgestellt. Ein technisches Problem stellt die Forderung zur Kontrolle und Nachvollziehbarkeit KI-basierter Prognose- und Entscheidungssysteme dar. Daher muss von Fall zu Fall unterschieden werden, wann diese Entscheidungssysteme Anwendung finden sollen und wann nicht. In diesem Beitrag gehe ich explizit auf inkrementelle Wissensakquisition ein. Das Thema der Nachvollziehbarkeit ist Gegenstand von aktuellen Forschungsgruppen wie z. B. im Forschungsbereich Interaktives Maschinelles Lernen (IML) am DFKI, und wird im Beitrag ausgeklammert.<sup>2</sup> Die menschliche Intelligenz ist der Maschinenintelligenz in wesentlichen Aspekten, wie Nachvollziehbarkeit unter Berücksichtigung von Kontextfaktoren und Patientenmodellen, immer noch deutlich überlegen. Der Blick auf die Geschichte der Kognitions-

wissenschaft lohnt sich, um mögliche Zukunftsszenarien medizinischer Versorgung realistischer einzuschätzen [6], auch im Pflegebereich mit Robotern und einhergehenden ethischen Fragen [7]. Die Industrie beginnt auch, transparente Modelle wieder in den Vordergrund zu rücken [49].

---

## » Technologien können relativ leicht repliziert werden, qualitative Daten nicht

---

Oft scheidet es nicht an der KI-Methodik, sondern daran, dass viele Kliniken keine vernetzten digitalen Daten ihrer Patienten bereitstellen. Mit der Durchsetzung einer standardisierten elektronischen Patientenakte würde sich das schnell ändern, auch wenn alle technisch nachrüsten müssen. Dann können Diagnostik, Behandlungs- und Versorgungsprozesse und sogar Heilungschancen in verschiedenen Gebieten entscheidend verbessert werden. Onkologiezentren, die bereits KI einsetzen, oder dies in Zukunft machen wollen, gibt es beispielsweise am DKFZ (Deutsches Krebsforschungszentrum) und DKTK (Deutsches Konsortium für Translationale Krebsforschung, mit 8 Standorten).<sup>3</sup> Im BMBF (Bundesministerium für Bildung und Forschung)-Projekt *pAIent* wird in den nächsten 3 Jahren ein Innovationszentrum mit integrierter, rechtssicherer Umgebung zur Entwicklung, Testung, und klinischen Bewertung KI-basierter Anwendungen aufgebaut.<sup>4</sup> Bei der molekularen Diagnostik von Krebserkrankungen werden in Zukunft KI-Verfahren eingesetzt, zusätzlich zu der Früherkennung von Krebserkrankungen, die durch bildgebende Verfahren bereits heute mit KI-Methoden unterstützt wird.

---

<sup>3</sup> <https://www.bmbf.de/de/deutsches-konsortium-fuer-translazionale-krebsforschung-395.html>

<sup>4</sup> <https://www.dfki.de/web/forschung/projekte-publikationen/projekte-uebersicht/projekt/paient/>

---

<sup>1</sup> <https://www.ki-strategie-deutschland.de/home.html>

---

<sup>2</sup> <https://dfki.de/iml>

## Inkrementelle Wissensakquisition

Wie Wissensakquisition durchzuführen und anwendbar zu machen ist, ist dargestellt in **Abb. 1** und im Folgenden:

① Strukturiertes Wissen in Form medizinischer Ontologien ist die Basis für die medizinische Wissensverarbeitung („structured/structural knowledge“). In radiologischen und anderen bildverarbeitenden Anwendungsszenarien werden oft medizinische Ontologien, wie FMA („foundational model of anatomy“; [23]), RadLex („radiology lexicon“; [20]) und die aufbereitete ICD-10 (International Classification of Diseases; [24]), einzeln oder zusammen [41] verwendet. Strukturiertes Wissen kann durch Sensorinterpretationen, wie Herzfrequenz oder Hautleitfähigkeit, ergänzt werden.

② Um Bildannotationen zu erhalten, werden klinische Desktop-Tools verwendet („desktop annotation“). Alternativ wird an strukturierter Befundung geforscht, die eine direkte Digitalisierung und Annotation erlaubt [43].<sup>5</sup>

③ Dann können annotierte Datensätze zur automatischen Bild- und Videoanalyse manuell erzeugt werden („automatic image recognition, spatial reasoning“; [39, 42]).

④ Anschließend stehen dem Arzt dialogbasierte multimodale Interaktionssysteme zur Verfügung, in neuester Form in virtueller Realität [21, 31] mit 3-D-Bild-Effekten [30].<sup>6</sup>

### » Für eine optimierte Wissensakquisition ist eine Erhöhung der nationalen Fördermittel sinnvoll

Konkrete Anwendungsfelder neben Tumorboards sind z. B. die Vermeidung schmerzhafter Biopsien oder anderer Eingriffe, wie in der Prostataagnostik oder bei Brustkrebsdiagnostik [2]. Ein weiteres Beispiel ist die Segmentierung

<sup>5</sup> <https://medicalcps.dfki.de/wp-content/uploads/2016/05/BIRADS-30-seconds.mp4>

<sup>6</sup> [https://medicalcps.dfki.de/wp-content/uploads/2017/08/KDI\\_V2\\_Pro\\_v04\\_2.mp4](https://medicalcps.dfki.de/wp-content/uploads/2017/08/KDI_V2_Pro_v04_2.mp4)

Gynäkologie <https://doi.org/10.1007/s00129-021-04808-2>

© Springer Medizin Verlag GmbH, ein Teil von Springer Nature 2021

D. Sonntag

## Künstliche Intelligenz in der Medizin und Gynäkologie – Holzweg oder Heilversprechen?

### Zusammenfassung

Künstliche Intelligenz (KI) hat in den letzten Jahren eine neue Reifephase erreicht und entwickelt sich zum Treiber der Digitalisierung in allen Lebensbereichen. Die KI ist eine Querschnittstechnologie, die für alle Bereiche der Medizin mit Bild-, Text- und Biodaten von großer Bedeutung ist. Es gibt keinen medizinischen Bereich, der nicht von KI beeinflusst werden wird. Dabei spielt die klinische Entscheidungsunterstützung eine wichtige Rolle. KI-Methoden etablieren sich gerade beim medizinischen Workflow-Management und bei der Vorhersage des Behandlungserfolgs bzw. des Behandlungsergebnisses. KI-Systeme können bereits in Bilddiagnose und im Patientenmanagement unterstützen, aber keine kritischen Entscheidungen

vorschlagen. Die jeweiligen Präventions- oder Therapiemaßnahmen können mit KI-Unterstützung sinnvoller bewertet werden, allerdings ist die Abdeckung der Krankheiten noch viel zu gering, um robuste Systeme für den klinischen Alltag zu erstellen. Der flächendeckende Einsatz setzt Fortbildungsmaßnahmen für Ärzte voraus, um die Entscheidung treffen zu können, wann auf automatische Entscheidungsunterstützung vertraut werden kann.

### Schlüsselwörter

Workflow · Maschinelles Lernen · Fortbildung · Technologie · Klinische Entscheidungsunterstützungssysteme

## Artificial intelligence in medicine and gynecology—the wrong track or promise of cure?

### Abstract

Artificial intelligence (AI) has attained a new level of maturity in recent years and is becoming the driver of digitalization in all areas of life. AI is a cross-sectional technology with great importance for all areas of medicine employing image data, text data and bio-data. There is no medical field that will remain unaffected by AI, with AI-assisted clinical decision-making assuming a particularly important role. AI methods are becoming established in medical workflow management and for prediction of treatment success or treatment outcome. AI systems are already able to lend support to imaging-based diagnosis and patient management,

but cannot suggest critical decisions. The corresponding preventive or therapeutic measures can be more rationally assessed with the help of AI, although the number of diseases covered is currently too low to create robust systems for routine clinical use. Prerequisite for the widespread use of AI systems is appropriate training to enable physicians to decide when computer-assisted decision-making can be relied upon.

### Keywords

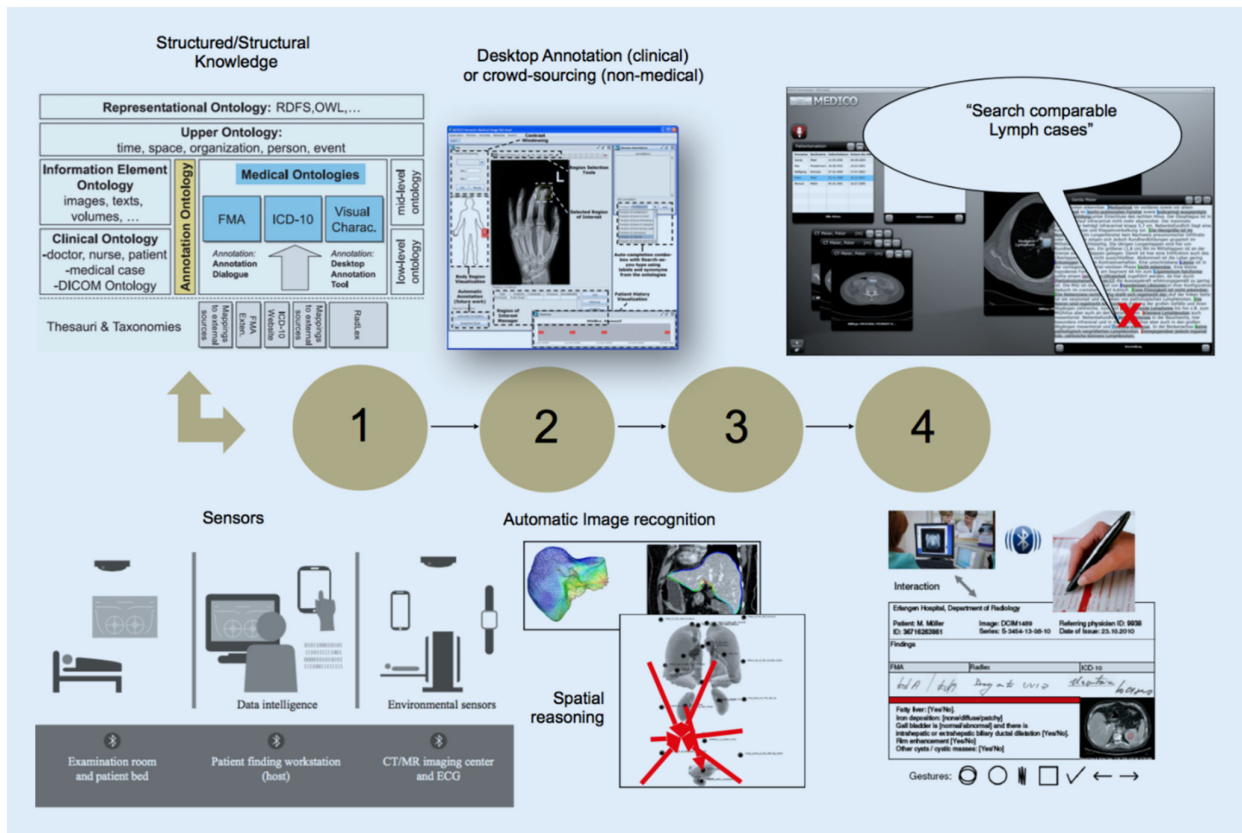
Workflow · Machine learning · Education, continuing · Technology · Decision support systems, clinical

und Klassifikation von Weichteilen bis hin zu automatisierten Brustultraschalluntersuchungen [9].

Inkrementelle Wissensakquisition dient dazu, für zukünftige KI-Systeme eine Brücke zwischen wissensfreien und wissensbasierten KI-Verfahren zu schlagen. Statistische und neuronale Lernverfahren (Deep Learning), die z. B. in der Bildverarbeitung eingesetzt werden, können nur dann robuster und plausibler werden, wenn sie mit Wissensgraphen und Ontologien kombiniert und damit

erklärbarer werden. Verfügbare Daten müssen auf ihren Aggregierungsstand im Sinne der Wissensakquisition und ontologischer Modellierung zur Steigerung der Datenqualität hin untersucht werden.

Es wäre sinnvoll, den nationalen Anteil der Fördermittel im Bereich Digitalisierung zu verdoppeln und im Bereich der Medizin zu verdreifachen, um die Wissensakquisition zu verbessern. Davon profitieren würden beispielsweise die schnelle Umsetzung von digitalen



**Abb. 1** ▲ Inkrementelle Datenakquisition. CT, „computed tomography“, DICOM Digital Imaging and Communications in Medicine, ECG „electrocardiogram“, ICD International Classification of Diseases, MR „magnetic resonance“, OWL Web Ontology Language, RDFS „Resource Description Framework Schema“ (Mit freundl. Genehmigung, ©D. Sonntag, alle Rechte vorbehalten)

Testverfahren und Entscheidungsunterstützungssystemen, computergestützten Navigations- und Assistenzsystemen und Lösungen einer digitalen elektronischen Patientenakte mit webbasiertem Lese- und Schreibzugriff, basierend auf den Entwicklungen abgeschlossener Forschungsprojekte [16, 35, 44].

### Verfügbare Daten

Das 2016 initiierte Programm Medizin-informatik<sup>7</sup> hat zum Ziel, in Deutschland Datenzentren bereitzustellen. Für die Anwendung der KI, multimodale Daten zu verarbeiten [3, 29], um dann beispielsweise gekoppeltes Deep Learning [5, 19] und Wissensmanagement zu betreiben [1], müssen umfangrei-

che Aggregationen der Daten in den nächsten Jahren vorgenommen werden (vgl. inkrementelle Wissensakquisition). Es ist sinnvoll, extern verfügbare strukturierte Datenbanken anzuzapfen, beispielsweise Linked Open Drug Data (LODD; [32]); diese bieten allerdings keine Bilddaten und Patientendaten, wie sie in einer elektronischen Patientenakte und technischem Patientenmodell vorhanden wären [44]. Auch europäische Projekte zur Datenmodellierung und Standardisierung existieren.<sup>8</sup>

### Diagnostik und Entscheidungsunterstützung

Neben dem ersten Beispiel zur klinischen Entscheidungsunterstützung des Arztes gibt es beispielsweise die Notwendigkeit

einer Entlastung von Patienten, Ärzten und Gesundheitssystem in Form computergestützter Hilfssysteme, die den Arzt oder Patienten bei der Vor-Interpretation von Krankheiten unterstützt. Spezielle KI-Anwendungen für spezielle onkologische Krankheitsfälle sind deswegen vielversprechend. Ein Beispiel ist die Dermatologie mit speziellen Untererkrankungen.

KI ist da gut, wo eine ganz spezifische Klassifikation gemacht werden soll, nicht aber bei zu vielen Entscheidungsmöglichkeiten.<sup>9</sup> Darum muss beispielsweise die Anwendung in der Dermatologie mit über 2000–3000 definierbaren Krankheitsbildern stark eingeschränkt werden. Dies gilt insbesondere für statistische und neuronale

<sup>7</sup> [http://www.gesundheitsforschung-bmbf.de/mediaMedizininformatik\\_englisch\\_barrierefrei.pdf](http://www.gesundheitsforschung-bmbf.de/mediaMedizininformatik_englisch_barrierefrei.pdf)

<sup>8</sup> <https://www.gesundheitsforschung-bmbf.de/de/medizininformatik.php>

<sup>9</sup> Es gibt aber auch Ansätze, die Diagnostik vorwiegend auf das Finden ähnlicher Patientenfälle zu beschränken [4].

Tab. 1 Dermatologiedatensätze für maschinelles Lernen			
Name	#Bilder	Klassen	Bemerkung
ISIC	~23.000	mal/ben and 18 Subklassen	Unbalanciert: über 19.000 ben
HAM 10000	10.000	mal/ben	Enthalten in ISIC. Unbalanciert: 6702 ben
MedNode	170	melanoma/naevus	Verteilung: 70/100
PH2	200	common nevus, atypical nevus, melanoma	Balanciert

*mal* maligne, *ben* benigne, *ISIC* International Skin Imaging Collaboration, *HAM10000* Human Against Machine with 10000 training images, *MedNode/MED-NODE* A Computer-Assisted Melanoma Diagnosis System using Non-Dermoscopic Images, *PH2* A dermoscopic image database for research and benchmarking

Lernverfahren, wie sie im Moment stark eingesetzt werden, da die reine Bildinterpretationsleistung basierend auf den Bildpixeln nur dann sehr hoch ist, wenn geschätzt mindestens 1000 Bilder pro Krankheitsbild vorliegen.

Die momentan für die Forschung öffentlich verfügbaren Bilddatensätze, die auch für ML-Klassifikation verwendet werden können, zeigt [Tab. 1](#).<sup>10</sup> Seit 2016 wird eine ML-Challenge organisiert, um die besten ML-Verfahren bei Krankheitsklassifikation, Feature-Klassifikation oder Segmentierung zu vergleichen [10, 22].

Potenzielle Vorreiter für klinische Anwendungen sind beispielsweise Segmentierungsmasken [8] und Tumordiagnosen [14] sowie Klassifizierungen anderer Hautkrankheiten [50]. Eigene DFKI-Arbeiten in diesem Bereich drehen sich darum, die Ergebnisse des ML nachvollziehbarer zu machen, beispielsweise durch Maskierung von Teilbereichen der Bilder [28], visuelle Erklärungen mit Heatmaps [25, 34] und Fusion der Bilddaten mit Patientenmetadaten [27].

Bildererkennung im COVID-19-Bereich ist auch Gegenstand aktueller Forschungsarbeit [26]. In der Pathologie gibt es erste anonyme Datensätze zur Metastasenerkennung, um international KI-Forschungssysteme zu vergleichen (Camelyon).<sup>11</sup>

<sup>10</sup> International Skin Imaging Collaboration, <https://isic-archive.com/> letzter Download am 5. Februar 2019. Dank an Fabrizio Nunnari.

<sup>11</sup> <https://camelyon17.grand-challenge.org>

## Gynäkologie-Transfer

Wir gehen aus vom HNO-Transfer, der in [37] beschrieben wurde: Analysen von Fehlfunktionen klinischer Entscheidungsunterstützungssysteme legen offen, dass oft keine Transfermöglichkeit der KI-Methodik, insbesondere des gelernten Datenmodells, von einem Anwendungsfall zum nächsten besteht [48]. Dennoch lässt sich der Ansatz der inkrementellen Datenakquisition für KI in die HNO (und Gynäkologie) übertragen. Glücklicherweise sind vor allem für Bildererkennung mittlerweile Standard-Pipelines mit Bildvorverarbeitung vorhanden, die bei entsprechender Datenlage im HNO-Bereich oder Gynäkologiebereich ML-Experimente prinzipiell zulassen (z. B. Niftynet; [17]). Dies betrifft vor allem komplexe Vorverarbeitungsschritte, beispielsweise aus den Bereichen der medizinischen Bildregistrierung. Es sei an dieser Stelle an die zentralen Aufgaben der Bildverarbeitung, Abgrenzung, Analyse, Identifizierung und Visualisierung medizinischer Bildobjekte (Gewebe, Tumoren, Läsionen, Gefäßsysteme), erinnert [18].

Viele HNO-Daten und Daten in der Gynäkologie sind Bild- und Ultraschall-daten. Mit einem ontologischen Ansatz könnten diese nicht nur digital vorliegen, sondern auch schon mit wichtigen Meta-informationen angereichert werden [41], bevor die entsprechenden Pipelines des ML und des Klassifizierens angewendet werden. Diagnostische Verfahren werden zunehmend automatisiert, d. h. es besteht ein Bedarf, Datenmaterial für ML zu erzeugen. Digitale ML-basierte Volumetomographie erlaubt auch in der Gynäkologie die Darstellung von

Weichteilen und potenziell die automatische Beurteilung von Strukturen. Computergestützten Navigations- und Assistenzsystemen für die komplexe dreidimensionale Anatomie wird ebenfalls großes Potenzial zugemessen. Dementsprechend sind 3-D-Bildverfahren mit Darstellung automatisch gefundener anatomischer Strukturen und Auffälligkeiten in Virtual Reality ein potenzielles Anwendungsfeld (vgl. virtuelle Endoskopie).

## » Pipelines mit Bildvorverarbeitung lassen ML-Experimente im Bereich der Gynäkologie prinzipiell zu

Meine Recherchen haben ergeben, dass es heute bereits zukunftssträchtige Forschungsarbeiten für spezielle Anwendungen in der Gynäkologie gibt. KI in der Geburtshilfe (z. B. zur Verbesserung der Sensitivität der Kardiotokographie [CTG]), in der onkologischen Gynäkologie (wie oben genannt, aber auch zur automatischen Krebsprognose anhand molekularer Biomarker und FIGO[Fédération Internationale de Gynécologie et d'Obstétrique]-Klassifizierung), generell das Feld der personalisierten Medizin (wo die ontologische Modellierung gut zu gebrauchen ist) und die Embryoauswahl bei IVF (In-vitro-Fertilisation; [12]). Weiter zeigt die Forschungsarbeit [11] bei der Interpretation von Ultraschallbildern und Videos nachvollziehbare KI-Anwendungsfelder auf. Diese Anwendungsfelder sind Sondenführung und Navigation, fetale biometrische Ebenenfinder, Anomalie-Scans, Highlighting und Zystenklassifikation. Die gynäkologische Diagnostik wird wahrscheinlich auch indirekt durch KI beim Brustkrebscreening beeinflusst. Hier lassen sich direkte Erfolge in der automatischen Klassifikation der Mammographien beobachten sowie in der Dokumentation auf einem digitalem BI-RADS(Breast Imaging – Reporting and Data System)-Bogen [43], um die Ergebnisse besser fusionieren zu können, z. B. beim zukünftigen digitalen Tumorboard.

## Ausblick

Entscheidend für die erfolgreiche Anwendung von KI sind der Zugang zu qualitativ hochwertigen Daten und die Integration in komplexe medizinische Dienstleistungen im klinischen und nichtklinischen Umfeld. Darum muss die Menge an nutzbaren, qualitativ hochwertigen Daten deutlich erhöht werden. Die Mensch-Maschine-Schnittstelle gewinnt an Bedeutung, weil der Arzt in einer Behandlungssituation, auch während einer Operation, schnell entscheiden muss. Ein technisches Problem stellt die Forderung zur Kontrolle und Nachvollziehbarkeit KI-basierter Prognose- und Entscheidungssysteme dar. Es ist daher von Fall zu Fall zu entscheiden werden, wann diese Anwendung finden und kommuniziert werden soll. Die Zukunft wird so aussehen, dass KI-Computer mit Klinikern zusammenarbeiten werden.

### » In Zukunft werden Kliniker und KI-Computer zusammenarbeiten

Man muss sich auch für einen flächen-deckenden Einsatz folgende Frage stellen: „Welche KI-Methoden brauchen tatsächlich vollständige Transparenz bei welchen Diagnoseentscheidungen?“ und „Wo können statistische Ansätze direkt verwendet werden?“ Dies bedarf nicht nur der Ausbildung neuer KI-Experten für die Medizin, sondern auch einer Schulung der Ärzte. Bei lebensbedrohlichen Entscheidungen fehlt oft die nötige Robustheit der KI-Systeme. Viele medizinische Fragestellungen müssen kausale Zusammenhänge aufdecken oder nutzen, das wird relevant, wenn die Bildverarbeitung weiter integriert wird.

Große Herausforderungen bestehen in der Standardisierung medizinischer Begriffswelten in Form von technischen Ontologien und in der Informationsextraktion aus Texten (hier nicht behandelt) und vor allem aus Bildern.

### Fazit für die Praxis

- Einige Datentöpfe stehen bereits für KI(Künstliche Intelligenz)-Anwen-

dungen bereit, vor allem im Bereich der Bildverarbeitung.

- Es gibt bereits Prototypenwendungen, die mithilfe hochqualitativer Daten einer klinischen Studie im BMWi(Bundesministeriums für Wirtschaft und Energie)-Projekt KDI (Klinische Datenintelligenz) realisiert wurden [44].
- Viele ähnliche Entscheidungsunterstützungssysteme wären heute schon möglich, wenn Datenpartnerschaften zwischen Kliniken und Forschungsinstituten existierten. Dann würden die Potenziale des Einsatzes von KI in der Medizin besser ausgeschöpft werden können. Analysen von Fehlfunktionen klinischer Entscheidungsunterstützungssysteme sind aber unbedingt notwendig.

### Korrespondenzadresse



**Prof. Dr. Daniel Sonntag**  
Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Universität Oldenburg  
Marie-Curie-Str. 1,  
26129 Oldenburg,  
Deutschland  
sonntag@dfki.de

### Einhaltung ethischer Richtlinien

**Interessenkonflikt.** D. Sonntag gibt an, dass kein Interessenkonflikt besteht.

Für diesen Beitrag wurden vom Autor keine Studien an Menschen oder Tieren durchgeführt. Für die aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

### Literatur

- Alpaydin E (2018) Classifying multimodal data. In: Signal processing, architectures, and detection of emotion and cognition. The handbook of multimodal-multisensor interfaces, Bd. 2. Morgan & Claypool, San Rafael (chapter 2)
- Bahl M, Barzilay R, Yehia A, Locascio N, Yu L, Lehman C (2018) High-risk breast lesions: a machine learning model to predict pathologic up-grade and reduce unnecessary surgical excision. *Radiology* 286:810–818
- Baltrusaitis T, Ahuja C, Morency L-P (2018) Multimodal machine learning. In: Signal processing, architectures, and detection of emotion and cognition. The handbook of multimodal-multisensor interfaces, Bd. 2. Morgan & Claypool, San Rafael (chapter 1)
- Bates DW, Saria S, Ohno-Machado L, Shah A, Escobar G (2014) Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Aff* 33(7):1123–1131
- Bengio S, Deng L, Morency L-P, Schuller B (2018) Multidisciplinary challenge topic: perspectives on predictive power of multimodal deep learning: surprises and future directions. In: Signal processing, architectures, and detection of emotion and cognition. The handbook of multimodal-multisensor interfaces, Bd. 2. Morgan & Claypool, San Rafael (chapter 14)
- Boden MA (2008) Mind as machine: a history of cognitive science. Clarendon Press, Oxford
- Boden MA, Bryson J, Caldwell DG, Dautenhahn K, Edwards L, Kember S, Newman P, Parry V, Pegman G, Rodden T, Sorrell T, Wallis M, Whitby B, Winfield AFT (2017) Principles of robotics: regulating robots in the real world. *Connect Sci* 29(2):124–129. <https://doi.org/10.1080/09540091.2016.1271400>
- Burdick J, Marques O, Weinthal J, Furt B (2018) Rethinking skin lesion segmentation in a convolutional classifier. *J Digit Imaging* 31(4):435–440. <https://doi.org/10.1007/s10278-017-0026-y>
- Choi J-H, Kang BJ, Baek JE, Lee HS, Kim SH (2018) Application of computer-aided diagnosis in breast ultrasound interpretation: improvements in diagnostic performance according to reader experience. *Ultrasonography* 37(3):217–225. <https://doi.org/10.14366/ulg.17046>
- Codella NCF, Gutman D, Celebi ME, Helba B, Marchetti MA, Dusza SW, Kalloo A, Liopyris K, Mishra N, Kittler H, Halpern A (2018) Skin lesion analysis toward melanoma detection: a challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC). In: 2018 IEEE 15th International Symposium on Biomedical Imaging ISBI 2018. IEEE, Washington, DC, 5168–172. <https://doi.org/10.1109/ISBI.2018.8363547>
- Drukker L, Noble J, Papageorgiou A (2020) Introduction to artificial intelligence in ultrasound imaging in obstetrics and gynecology. *Ultrasound Obstet Gynecol* 56(4):498–505
- Emin EI, Emin E, Papalois A, Willmott F, Clarke S, Sideris M (2019) Artificial intelligence in obstetrics and gynaecology: is this the way forward? *In Vivo* 33(5):1547–1551
- Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542:115. <https://doi.org/10.1038/nature21056>
- Fujisawa Y, Otomo Y, Ogata Y, Nakamura Y, Fujita R, Ishitsuka Y, Watanabe R, Okiyama N, Ohara K, Fujimoto M (2018) Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *Br J Dermatol*. <https://doi.org/10.1111/bjd.16924>
- Gelissen J, Sonntag D (2015a) Special issue on health and wellbeing. *KI Künstliche Intell* 29(2):111–113. <https://doi.org/10.1007/s13218-015-0360-5>
- Gelissen J, Sonntag D (2015b) Special issue on health and wellbeing. *KI Künstliche Intell* 29(2):111–113. <https://doi.org/10.1007/s13218-015-0360-5>
- Gibson E, Li W, Sudre C, Fidon L, Shakeri DI, Wang G, Eaton-Rosen Z, Gray R, Doel T, Hu Y, Whyntie T, Nachev P, Modat M, Barratt DC, Ourselin S, Cardoso MJ, Vercauteren T (2018) Niftnet: a deep-

- learning platform for medical imaging. *Comput Methods Programs Biomed* 158:113–122. <https://doi.org/10.1016/j.cmpb.2018.01.025>
18. Handels H (2015) *Medizinische Bildverarbeitung*. Springer, Heidelberg, Berlin
  19. Keren G, Mousa AE-D, Pietquin O, Zafeiriou S, Schuller B (2018) Deep learning for multisensorial and multimodal interaction. In: *Signal processing, architectures, and detection of emotion and cognition. The handbook of multimodal-multisensor interfaces*, Bd. 2. Morgan & Claypool, San Rafael (chapter 4)
  20. Langlotz CP (2006) Radlex: a new method for indexing online educational materials. *Radiographics* 26:1595–1597. <https://doi.org/10.1148/rg.266065168>
  21. Luxenburger A, Prange A, Moniri MM, Sonntag D (2016) Medicalvr: Towards medical remote collaboration using virtual reality. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct UbiComp '16*. ACM, New York, S 321–324. <https://doi.org/10.1145/2968219.2971392>
  22. Marchetti MA, Codella NC, Dusza SW, Gutman DA, Helba B, Kalloo A, Mishra N, Carrera C, Celebi ME, DeFazio JL, Jaimes N, Marghoob AA, Quigley E, Scope A, Yélamos O, Halpern AC (2018) Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images. *J Am Acad Dermatol* 78(2):270–277.e1. <https://doi.org/10.1016/j.jaad.2017.08.016>
  23. Mejino JL, Rubin DL, Brinkley JF (2008) FMA-RadLex: an application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. In: *Proc. of AMIA Symposium*, S 465–469
  24. Möller M, Sintek M, Biedert R, Ernst P, Dengel A, Sonntag D (2010) Representing the international classification of diseases version 10 in OWL. In: *Filipe J, Dietz JLG (Hrsg) KEOD 2010—Proceedings of the International Conference on Knowledge Engineering and Ontology Development Valencia*, October 25–28, 2010 SciTePress, Setúbal, S 50–59. ISBN 9789898425294
  25. Nguyen DMH, Ezema AO, Nunnari F, Sonntag D (2020) A visually explainable learning system for skin lesion detection using multiscale input with attention U-Net. In: *Schmid U, Klügl F, Wolter D (Hrsg) KI 2020: Advances in Artificial Intelligence—43rd German Conference on AI. Proceedings Bamberg, Germany, September 21–25, 2020. Lecture notes in computer science*, Bd. 12325. Springer, Cham, S 313–319. [https://doi.org/10.1007/978-3-030-58285-2\\_28](https://doi.org/10.1007/978-3-030-58285-2_28)
  26. Nguyen HMD, Nguyen DM, Vu H, Nguyen BT, Nunnari F, Sonntag D (2021) An attention mechanism using multiple knowledge sources for COVID-19 detection from CT images. In: *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21), Workshop: Trustworthy AI for Healthcare Vancouver, BC, Canada, February 2–9 (AAAI Virtual Conference)*
  27. Nunnari F, Bhuvaneshwara C, Ezema AO, Sonntag D (2020a) A study on the fusion of pixels and patient metadata in CNN-based classification of skin lesion images. In: *Holzinger A, Kieseberg P, Tjoa AM, Weippl ER (Hrsg) Machine Learning and Knowledge Extraction—4th IFIP TC 5, TC 12, WG 8.4, WG 8.9, WG 12.9 International Cross-Domain Conference CD-MAKE 2020, Dublin, Ireland, August 25–28, 2020. Lecture notes in computer science*, Bd. 12279. Springer, Cham, S 191–208. <https://doi.org/10.1007/978-3-030-57321-8>
  28. Nunnari F, Ezema AO, Sonntag D (2020b) The effects of masking in melanoma image classification with CNNs towards international standards for image preprocessing. In: *Ye J, O'Grady MJ, Civitarese G, Yordanova K (Hrsg) Wireless Mobile Communication and Healthcare—9th EAI International Conference MobiHealth 2020, 19. Nov. 2020. Lecture notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Bd. 362. Springer, Cham, S 257–273. <https://doi.org/10.1007/978-3-030-70569-5> (Virtual Event)
  29. Panagakis Y, Rudovic O, Pantic M (2018) Learning for multi-modal and context-sensitive interfaces. In: *Signal processing, architectures, and detection of emotion and cognition. The handbook of multimodal-multisensor interfaces*, Bd. 2. Morgan & Claypool, San Rafael (chapter 3)
  30. Prange A, Barz M, Sonntag D (2018) Medical 3d images in multimodal virtual reality. In: *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion IUI'18*. ACM, New York, S 19:1–19:2. <https://doi.org/10.1145/3180308.3180327>
  31. Rizzo A, Talbot T (2016) Virtual reality standardized patients for clinical training. In: *The digital patient*. John Wiley & Sons, Hoboken, S 255–272. <https://doi.org/10.1002/9781118952788.ch18>
  32. Samwald M, Jentzsch A, Bouton C, Kallesøe C, Willighagen EL, Hajagos J, Marshall MS, Prud'hommeaux E, Hassanzadeh O, Pichler E, Stephens S (2011) Linked open drug data for pharmaceutical research and development. *J Cheminform* 3:19–19. <https://doi.org/10.1186/1758-2946-3-19>
  33. Siekmann JH (2009) Die entwicklung der disziplin in deutschland. *KI* 23(1):47–52
  34. Somfai E, Baffy B, Fenech K, Guo C, Hosszú R, Korózs D, Nunnari F, Pólik M, Sonntag D, Ulbert A, Lorincz A (2021) Minimizing false negative rate in melanoma detection and providing insight into the causes of classification. *CoRR* (<https://arxiv.org/abs/2102.09199>)
  35. Sonntag D (2016) Medical cyber-physical systems. In: *Cyber-Physical system design with sensor networking technologies. Control, robotics and sensors*. Institution of Engineering and Technology, S 311–333
  36. Sonntag D (2018) AI in Germany: well-prepared and eager to do something. *KI Künstliche Intell* 32(2/3):97–99. <https://doi.org/10.1007/s13218-018-0555-7>
  37. Sonntag D (2019a) Künstliche Intelligenz in der Medizin – Holzweg oder Heilversprechen? *HNO* 67(5):343–349. <https://doi.org/10.1007/s00106-019-0665-z>
  38. Sonntag D (2019b) Medical and health systems. In: *The handbook of multimodal-multisensor interfaces: language processing, software, commercialization, and emerging directions. Association for Computing Machinery and Morgan & Claypool, USA*, S 423–476. <https://doi.org/10.1145/3233795.3233808>
  39. Sonntag D, Möller M (2010) A multimodal dialogue mashup for medical image semantics. In: *Proceedings of the 15th International Conference on Intelligent User Interfaces IUI '10*. ACM, New York, S 381–384. <https://doi.org/10.1145/1719970.1720036>
  40. Sonntag D, Profitlich H (2019) An architecture of open-source tools to combine textual information extraction, faceted search and information visualisation. *ArtifIntell Med* 93:13–28. <https://doi.org/10.1016/j.artmed.2018.08.003>
  41. Sonntag D, Wennerberg P, Buitelaar P, Zillner S (2009) Pillars of ontology treatment in the medical domain. *J Cases Inf Technol* 11(4):47–73
  42. Sonntag D, Schulz C, Reuschling C, Galarraga L (2012) Radspeech's mobile dialogue system for radiologists. In: *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces IUI '12*. ACM, New York, S 317–318. <https://doi.org/10.1145/2166966.2167031>
  43. Sonntag D, Weber M, Cavallaro A, Hammon M (2014) Integrating digital pens in breast imaging for instant knowledge acquisition. *AI Mag* 35(1):26–37
  44. Sonntag D, Tresp V, Zillner S, Cavallaro A, Hammon M, Reis A, Fasching PA, Sedlmayr M, Ganslandt T, Prokosch H, Budde K, Schmidt D, Hinrichs C, Wittenberg T, Daumke P, Oppelt PG (2016) The clinical data intelligence project—a smart data initiative. *Inform Spektrum* 39(4):290–300. <https://doi.org/10.1007/s00287-015-0913-x>
  45. Stone P, Brooks R, Brynjolfsson E, Calo R, Etzioni O, Hager G, Hirschberg J, Kalyanakrishnan S, Kamar E, Kraus S, Leyton-Brown K, Parkes D, Press W, Saxenian A, Shah J, Tambe M, Teller AS (2016) Artificial intelligence and life in 2030. Technical report, one hundred year study on artificial intelligence: report of the 2015–2016 study panel. Stanford University, Stanford
  46. Strecker H, Pfitzner K (1988) XRAY – ein prototypisches Konfigurierungs-Expertensystem für die automatische Röntgenprüfung. *KI Künstliche Intell* 2(2):4–8
  47. Wang I, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *CoRR* (<http://arxiv.org/abs/1705.02315>)
  48. Wright A, Hickman T, McEvoy D et al (2016) Analysis of clinical decision support system malfunctions: a case series and survey. *J Am Med Inform Assoc* 23:1068–1076
  49. Yang Y, Tresp V, Wunderle M, Fasching PA (2018) Explaining therapy predictions with layer-wise relevance propagation in neural networks. In: *2018 IEEE International Conference on Healthcare Informatics (ICHI)*, S 152–162. <https://doi.org/10.1109/ICHI.2018.00025>
  50. Zhang X, Wang S, Liu J, Tao C (2017) Computer-aided diagnosis of four common cutaneous diseases using deep learning algorithm. In: *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, Kansas City, S 1304–1306. <https://doi.org/10.1109/BIBM.2017.8217850>