

Zwischen Grenzen und Möglichkeiten

Wie Mensch und Maschine ***sicher*** gemeinsam entscheiden können

Asst. Prof. Dr. Christian F. Baumgartner – Universität Luzern

Das Versprechen der KI in der Medizin

- Frühere und genauere Diagnosen
- Weniger medizinische Fehler
- Effizienzgewinne in klinischen Abläufen
- **Übermenschliche Fähigkeiten**

Beispiele in denen KI bereits besser ist

Dermatologie

Hautkrebsklassifikation auf Facharzt-Niveau

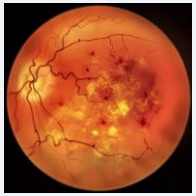
Ophthalmologie

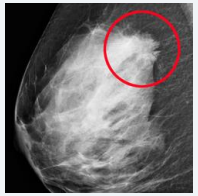
Diabetische Retinopathie-Erkennung auf Spezialisten-Niveau

Radiologie

Lungenknotenerkennung, Mammographie-Triage

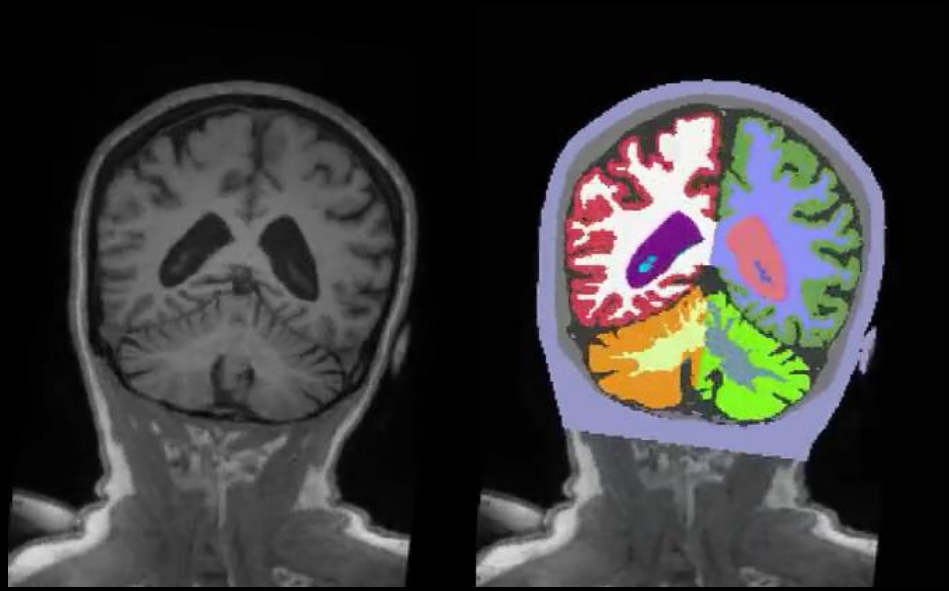






Jenseits menschlicher Fähigkeiten: zwei Beispiele





- KI kann das Risiko für schwere Herzereignisse aus Fundus Bildern prognostizieren
- KI kann inert Sekunden Gehirn-MRT Aufnahmen in anatomische Gehirnstruktur einteilen

Das Paradox

Wenn KI Dinge sieht, die wir nicht sehen können, wer kontrolliert sie?

Mensch → Mensch

Ein Radiologe kann einen anderen Radiologen überprüfen. Beide arbeiten im selben Wahrnehmungsraum. Fehler sind erkennbar.



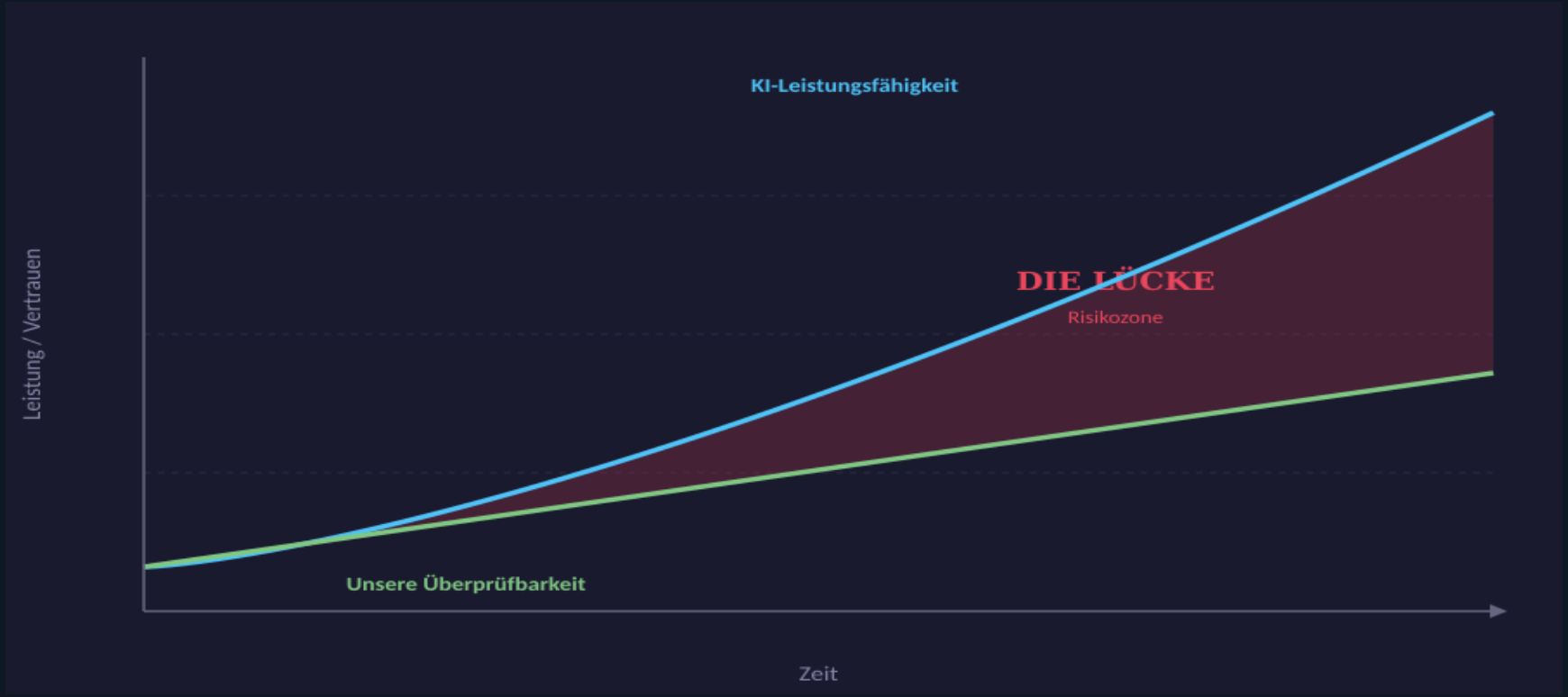
Mensch → KI

KI erkennt Muster, die für Menschen unsichtbar sind. Die Überprüfungsschleife ist unterbrochen. Wir können nicht prüfen, was wir nicht wahrnehmen.

*Übermenschliche Leistung klingt grossartig… bis sie die **Überprüfungsschleife** durchbricht.*

Die wachsende Lücke

KI entwickelt sich schneller, als wir sie sicher überprüfen können



KI versagt anders als Menschen

Menschliche Fehler

Zufällig
Müdigkeit, Ablenkung, individuelle Variabilität

Einzelfälle
Ein Patient zur Zeit

Erkennbar
Andere Ärzte können sie erkennen

KI-Fehler

Systematisch
Derselbe Fehler, jedes Mal, mit Überzeugung

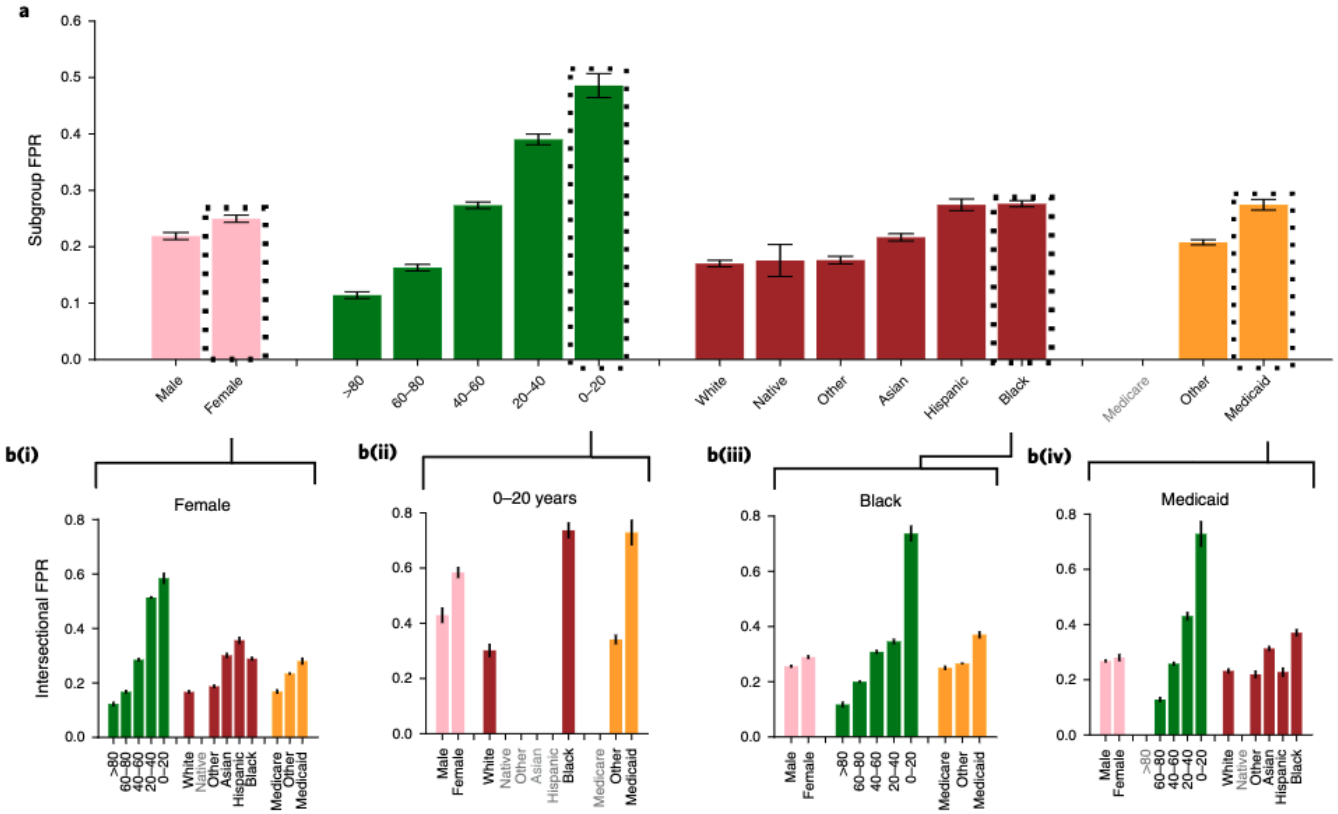
Im grossen Massstab
Tausende Patienten unbemerkt betroffen

Unsichtbar
Scheinkorrelationen, Verteilungsverschiebung

Ein einzelner systematischer Fehler kann Tausende von Patienten betreffen, bevor es jemand bemerkt.



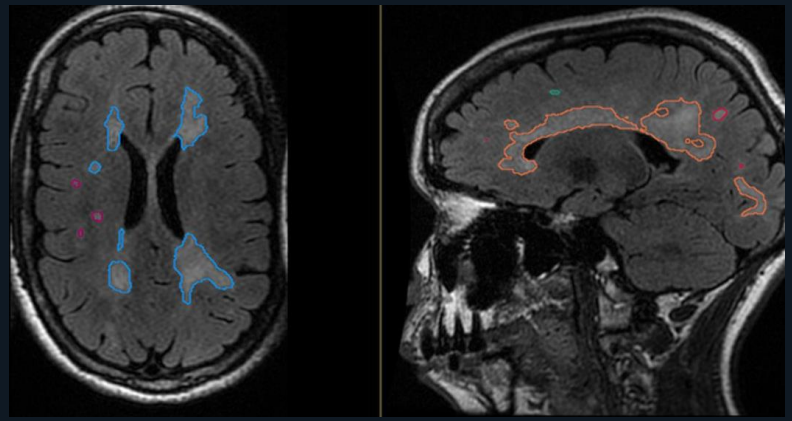




KI kann Röntgenbilder gewisser Gruppen deutlich Schlechter auswerten!

Seyyed-Kalantari et al., Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations, Nature Medicine (2021)

Unzureichende Überprüfung: Ein Beispiel



Ein von der FDA bewilligtes KI-Tool zur Erkennung von Störungen der weissen Gehirnsubstanz

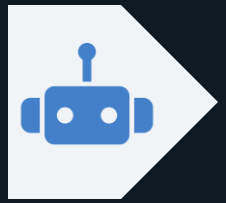
- Überprüft auf einem Datensatz mit **90 Patienten** (Alter: 22-99). **Keine** externe Validierung.
- **Intended use:** Patienten ab 22 Jahre

Unsere Überprüfungsinfrastruktur hält nicht Schritt mit der Technologie.

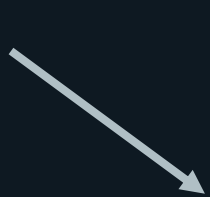
FDA Unterlagen verfügbar unter: https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices

Der fehlende Kommunikationskanal





Lungenödem









Wir brauchen einen Kommunikationskanal zwischen Mensch und KI

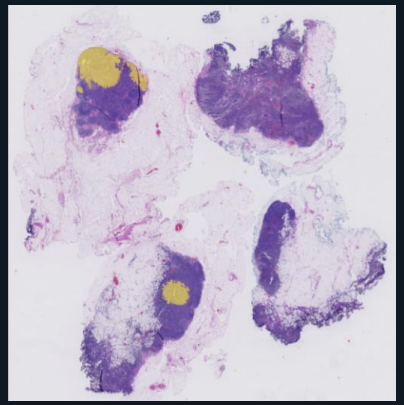
Interpretierbares ML: Ansätze aus unserer Forschung

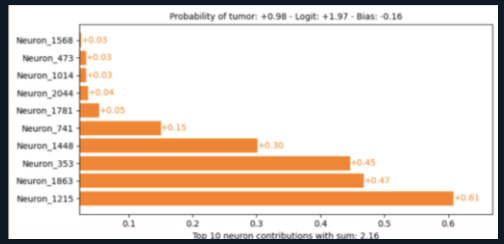
Konzeptbasierte Erklärungen

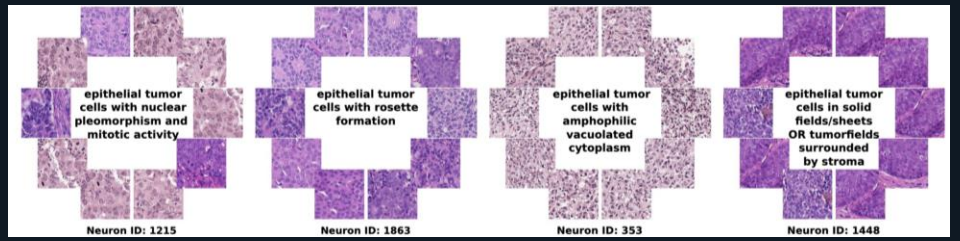
Pathologie-Diagnose

Das Modell denkt in menschlich verständlichen Konzepten, z.B. Gewebetyp, Zellmorphologie, strukturelle Muster.

Kliniker können nachvollziehen, worauf die KI ihre Entscheidung stützt.







Susu Sun, Dominique van Midden, Geert Litjens, and Christian F. Baumgartner. "Prototype-Based Multiple Instance Learning for Gigapixel Whole Slide Image Classification." *Proc. MICCAI* (2025).

Interpretierbares ML: Ansätze aus unserer Forschung

Kontrafaktische Erklärungen

Thorax-Röntgen-Diagnose

«Das Modell sagt Kardiomegalie. So müsste das Röntgenbild aussehen, damit es gesund sagt.»

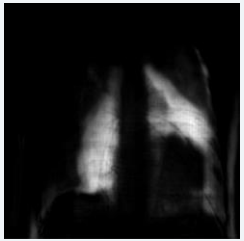
Gibt Klinikern etwas, das sie überprüfen und nachvollziehen können.



Vorhersage:
Kardiomegalie



”Counterfactual”:
Keine Kardiomegalie



Differenzbild

Susu Sun, Stefano Woerner, Andreas Maier, Lisa M. Koch, and Christian F. Baumgartner. ”AttriNet: A globally and locally interpretable model for multi-label classification using class-specific counterfactuals”. *MELBA* (2024).

Die Lücke schliessen

Wie können wir in Zukunft eine sichere Zusammenarbeit von Mensch und KI gewährleiten?

- Mit besserer KI muss auch deren Überprüfbarkeit wachsen
- KI entwickeln, die ihre Entscheidungen erklären kann
- KI-Transparenz gesetzlich verankern
- Zertifizierungsstandards von KI-tools verbessern
- Weg von reiner Tool-performance, hin zu “Kollaborations-performance”



Vielen Dank!

Asst. Prof. Dr. Christian F. Baumgartner

medical ai lab

UNIVERSITÄT LUZERN

Weitere Ressourcen:

- Unsere Webseite: https://med-ai-unilu.ch
- Symposium on AI in Medical Imaging – 18. Juni 2026 – Bern: https://www.saimi.ch
- Fakultät für Gesundheitswissenschaften und Medizin Uni Luzern: https://www.unilu.ch/fakultaeten/gmf/uebersicht/
- Competence Center for Health Data Science: https://www.unilu.ch/fakultaeten/gmf/forschung-und-zentren/zentren/competence-center-for-health-data-science/uebersicht/