

## KI-gestützte Bildanalyse: aussagekräftig nur mit passender Metrik

**Wie gut lösen die Algorithmen, die in der KI-gestützten Auswertung medizinischer Bilder genutzt werden, ihre jeweiligen Aufgaben? Das hängt im hohen Maße davon ab, anhand welcher Messgrößen – Metriken – ihre Leistung bewertet wird. Ein internationales Konsortium hat das weltweit verfügbare Wissen über die spezifischen Stärken, Schwächen und Limitationen der verschiedenen Validierungs-Metriken zusammengetragen.**

Mit „Metrics Reloaded“ stellen die Forschenden, unter der Federführung von Wissenschaftlerinnen vom Deutschen Krebsforschungszentrum (DKFZ) und dem Nationalen Centrum für Tumorerkrankungen (NCT) Heidelberg, nun ein breit verfügbares online-Tool zur Verfügung, das Nutzer bei der Auswahl des für ihre Aufgabenstellung geeigneten Algorithmus unterstützt.

Immer mehr Bereiche der Medizin setzen auf Unterstützung durch künstliche Intelligenz (KI). Das gilt besonders für das breite Spektrum an Fragestellungen, die auf der Auswertung von Bilddaten beruhen: So suchen Ärzte in Mammographien nach kleinsten Krebsherden oder berechnen das Volumen eines Hirntumors anhand der Schichtbilder aus dem MRT. Mit endoskopischen Aufnahmen des Darms spüren sie Polypen auf, bei der Auswertung von mikroskopischen Gewebeschnitten müssen subtile Änderungen einzelner Zellen erfasst werden.

Doch sind die Algorithmen, die für diese verschiedenartigen Bildanalysen eingesetzt werden, tatsächlich immer für die jeweilige Aufgabe geeignet? Das hängt in hohem Maße davon ab, welche Messgrößen, im Fachbegriff als „Metriken“ bezeichnet, sie erfassen – und ob diese tatsächlich zur jeweiligen Aufgabe passen.

„Wir bemerken oft, dass für bestimmte Aufgaben Validierungs-Metriken genutzt werden, die aus klinischer Perspektive für die Aufgabenstellung gar nicht relevant sind“, sagt Lena Maier-Hein vom DKFZ und nennt ein Beispiel dafür: „Bei der Suche nach Metastasen im Gehirn ist es zunächst wichtiger, dass der Algorithmus auch kleinste Läsionen erfasst, als dass er hochpräzise die Konturen jeder einzelnen Metastase definieren kann.“

Lena Maier-Hein und ihre Kolleginnen und Kollegen befürchten, dass die Verwendung ungeeigneter Validierungs-Metriken den wissenschaftlichen Fortschritt hemmen und die Einführung wichtiger Bildanalyseverfahren in die klinische Praxis verzögern kann.

Doch welche Metriken eignen sich unter Einbeziehung aller Stärken, Schwächen und Limitationen für eine gegebene klinische Fragestellung? Dazu befragten die Heidelberger Datenwissenschaftlerinnen in einem mehrstufigen, strukturierten Verfahren Meinungsführer aus der Wissenschaft und Industrie aus über 70 Forschungsinstitutionen weltweit. Mit der Befragung trugen sie Informationen zusammen, die bisher nur weltweit verstreut vorlagen.

„Mit dieser Arbeit machen wir der Fachwelt erstmals zuverlässige und umfassende Informationen über Probleme und Fallstricke im Zusammenhang mit Validierungs-Metriken in der Bildanalyse zugänglich“, sagt Annika Reinke, eine der federführenden Autorinnen. Als strukturiertes Informationswerk, auf das Forschende aller Fachrichtungen zugreifen können, soll die Arbeit das Verständnis eines Schlüsselproblems bei der KI-assistierten Bildauswertung steigern. Auch wenn der Schwerpunkt auf der Analyse medizinischer Bilder liegt, so lassen sich die Informationen auch auf andere Gebiete der Bildauswertung übertragen.

In einer zweiten Arbeit beschreibt das Expertenkonsortium unter der Federführung der Heidelberger Forscherinnen nun „Metrics Reloaded“: Ein umfassendes Rahmenwerk, das Ärzten und Wissenschaftlern bei der problemgerechten Auswahl von Metriken hilft. „Metrics Reloaded“ kann als online-Tool genutzt werden. „Die Anwender werden durch einen umfangreichen Fragenkatalog geführt, damit erstellen sie sozusagen einen präzisen Fingerabdruck ihres Bildanalyse-Problems. Das Tool macht außerdem auf spezifische Probleme aufmerksam, die bei bestimmten biomedizinischen Fragestellungen auftreten“, erklärt Paul Jäger, der ebenfalls federführend an der Metrics Reloaded Initiative beteiligt war.

Metrics Reloaded ist für alle verschiedenen Problemkategorien der Bildanalyse geeignet, das heißt, für die Klassifizierung der Bilder, die Objektdetektion oder die Zuordnung einzelner Pixel (semantische Segmentation). Das Tool funktioniert völlig unabhängig von der Bildquelle, lässt sich also für CT oder MRT-Aufnahmen ebenso gut einsetzen wie für mikroskopische Bilder. Auch für Bildanalysen jenseits biomedizinischer Fragestellungen können Forschende sich von Metrics Reloaded anleiten

lassen.

„Mit Metrics Reloaded liefern wir erstmals eine systematische Anleitung, die den Nutzern von KI-basierten Bildanalysen den Weg zum passenden Algorithmus weist. Wir hoffen, dass Metrics Reloaded möglichst rasch möglichst weite Verbreitung findet, denn dadurch könnte sich die Qualität und Verlässlichkeit der Ergebnisse KI-gestützter Bildanalysen erheblich verbessern lassen. Das würde auch das Vertrauen in KI-gestützte Bildanalysen in der klinischen Routine fördern“, sagt Minu Tizabi, eine der Erstautorinnen.

Die Projekte wurden maßgeblich gefördert von Helmholtz Imaging, einer von fünf Forschungsplattformen, die vom Helmholtz Information & Data Science Incubator initiiert wurden.

#### **Publikationen:**

Reinke, A./Tizabi, M., ... Jäger, P., Maier-Hein L.: Understanding Metric-Related Pitfalls in Image Analysis Validation. Nature Methods 2024, DOI: 10.1038/s41592-023-02150-0

Maier-Hein, L./Reinke, A., ... Jäger, P.: Metrics Reloaded: Recommendations for Image Analysis Validation. Nature Methods 2024, DOI 10.1038/s41592-023-02151-z

---

#### **Pressemitteilung**

12.02.2024

Quelle: Deutsches Krebsforschungszentrum

---

#### **Weitere Informationen**

► [Deutsches  
Krebsforschungszentrum](#)