

Original-Titel

Comprehensive review of publicly available colonoscopic imaging databases for artificial intelligence research: availability, accessibility, and usability

Autoren

Britt B. S. L. Houwen, MD, Karlijn J. Nass, MD, Jasper L. A. Vleugels, MD, PhD, Paul Fockens, MD, PhD, Yark Hazewinkel, MD, PhD, Evelien Dekker, MD, PhD; *Gastrointest Endosc* 2023;97:184-99

Kommentar

PD Dr. Alexander Hann, Yun Chiang, Uniklinikum Würzburg, 28.03.2023

Die Verwendung eines umfangreichen und vielfältiger Datensatzes ist für die Entwicklung und Validierung von künstlicher Intelligenz (KI) unerlässlich. Dies führte zur Generierung mehrerer koloskopischen Bilddatenbanken. Medizinische Daten unterliegen jedoch einem besonderen Schutz. Daher ist ein Großteil dieser Bilddatenbanken mono-zentrisch und nicht frei verfügbar. Dies hat direkte Auswirkungen auf die Qualität der KIs, welche für den Einsatz am Patienten entwickelt werden. Des Weiteren hat es direkte Auswirkungen wie akkurat die Validierung der KIs ist bevor der Einsatz am Patienten erfolgt. Frei verfügbare Datenbanken können daher nur einen kleinen Teil unserer Patientenpopulationen abbilden und sind darüber hinaus nur mit wenig Daten über die Patienten und Patientinnen ausgestattet, um deren Anonymität zu wahren. Ebenfalls fehlt eine Liste an verfügbaren Datenbanken, was die Suche nach geeigneten Trainingsdaten zeitaufwändig und umständlich macht.

In der vorliegenden Arbeit wurde durch Britt B. S. L. Houwen et al. eine umfassende und systematische Literaturrecherche mit dem Ziel der Identifikation von koloskopischen Bilddatenbanken durchgeführt. Dabei ging es vorrangig um die Erkennung, Lokalisation, Segmentierung und Charakterisierung von Polypen. Datenbanken mit koloskopischen Bildern wurden durch eine Suche von relevanten Publikationen in MEDLINE und Embase identifiziert. Anschließend wurde die Zugängigkeit der Daten bestimmt. Es folgte eine ausführliche Beschreibung der Datenbanken einschließlich Veröffentlichungsdatum, Anzahl der an der Datenerhebung beteiligten Zentren und Anzahl der Bilder. Zum Schluss wurde die Vollständigkeit der dazugehörigen Daten über die Patienten und Patientinnen bewertet.

Diese Studie bietet einen klaren und umfassenden Überblick über derzeit frei verfügbare Datenbanken zur koloskopischen Bildgebung. Es vereinfacht den Prozess der Ermittlung solcher Quellen und damit das Training und die Validierung eigener KIs. Es offenbart jedoch auch, dass viel zu wenige Daten frei zur Verfügung stehen, um der Vielfalt unserer Patienten, Patientinnen und verschiedenen Endoskoptypen gerecht zu werden. Hier schlagen die Autoren zu Recht vor, dass eine frei verfügbare Datenbank installiert werden soll, in der Bilddaten mit Zusatzinformationen zentral gelagert werden sollen.