



Deepfake-Röntgenbilder – selbst Radiologen erkennen KI nur begrenzt zuverlässig

Katja Kupfer

Künstlich erzeugte Röntgenbilder sind inzwischen so realistisch, dass selbst erfahrene Radiologen Schwierigkeiten haben, sie von echten Aufnahmen zu unterscheiden. Das zeigt eine aktuelle Studie im Fachjournal *Radiology*.

Untersucht wurde, wie gut Ärzte und KI-Systeme sogenannte „Deepfake“-Röntgenbilder erkennen können. Hintergrund ist die rasante Entwicklung generativer KI, die inzwischen in der Lage ist, medizinische Bilder allein auf Basis von Texteingaben zu erzeugen. Das wirft zugleich die Frage auf, wie verlässlich Bilddaten künftig noch sind. Für die Studie wurden 17 Radiologen aus verschiedenen Ländern einbezogen. Sie beurteilten in mehreren Durchläufen unterschiedliche Bild Datensätze.

Zunächst analysierten sie 154 Röntgenbilder aus verschiedenen Körperregionen, jeweils zur Hälfte echte und KI-generierte Aufnahmen. Später folgte ein zweiter Datensatz mit 110 zusätzlichen Bildern, die mit einem spezialisierten KI-Modell erzeugt worden waren. Im ersten Schritt wussten die Radiologen nicht, dass sich künstliche Bilder darunter befanden. Nur ein Teil bemerkte überhaupt Auffälligkeiten. Erst nachdem sie informiert wurden, sollten sie gezielt unterscheiden, ob es sich um echte oder synthetische Aufnahmen

handelt. Die Trefferquote lag dabei im Schnitt bei rund 75 Prozent. Parallel wurden auch verschiedene KI-Systeme getestet. Sie erhielten dieselben Bilder und sollten ebenfalls entscheiden, ob diese echt oder künstlich erzeugt sind. Keines der Systeme konnte alle Bilder korrekt einordnen.

Die Radiologen konnten krankhafte Veränderungen auf den KI-generierten Bildern nahezu genauso zuverlässig erkennen wie auf echten Röntgenaufnahmen. Die Bilder wirken also nicht nur realistisch, sondern sind auch medizinisch plausibel. Die Autoren warnen deshalb vor möglichen Folgen. Künstlich erzeugte Bilder könnten in Forschungsdatensätze gelangen oder gezielt eingesetzt werden, etwa um Diagnosen zu beeinflussen oder falsche Befunde zu erzeugen. Gleichzeitig sehen sie auch Chancen, etwa für Ausbildung und Training mit seltenen Krankheitsbildern. Um Risiken jedoch zu begrenzen, fordern die Autoren unter anderem eine bessere Schulung im Umgang mit KI-Bildmaterial sowie technische Lösungen zur Kennzeichnung und Nachverfolgung künstlich erzeugter Daten.

Quellen: ZWP online; *Radiology*, Volume 318, Number 3, <https://doi.org/10.1148/radiol.252094>.