



Zweite Meinung vom KI-Assistenten

Der Einsatz künstlicher Intelligenz in der Medizin soll vertrauenswürdiger und nachvollziehbarer werden. Wiener Forschende bringen den Systemen bei, ihre Ergebnisse besser „erklärbar“ zu machen.

Alois Pumhösel



KI-Systeme durchleuchten Röntgen-, CT- und MRT-Bilder auf Krankheiten. Ihre Entscheidungen sollen auch transparenter werden.

Systeme der künstlichen Intelligenz (KI) sind gerade dabei, in die medizinische Praxis hineinzuwachsen. Eine der wichtigsten Anwendungen liegt in einer schnellen Analyse von Aufnahmen bildgebender Verfahren, um die Ärzte in ihrem Alltag zu unterstützen. Mit Deep-Learning-Ansätzen wird den Systemen beigebracht, etwa Röntgen-, CT- oder MRT-Bilder zu klassifizieren. Sie unterscheiden etwa Patienten, die an einer bestimmten Krebsform erkrankt sind, von jenen, die es nicht sind. Gleichzeitig können sie ihren Nutzern aber nicht darüber Aufschluss geben, an welchem Bildmerkmal sie ihre Entscheidung genau festmachen.

Künftige Systeme sollen in dieser Hinsicht nachvollziehbarer werden. Ihre Ergebnisse sollen in intuitiven Interaktionen und leicht erfassbaren Visualisierungen präsentiert werden. Das Wiener Forschungsinstitut VRVis, das sich mit Technologien in diesem Bereich beschäftigt, möchte in der bevorstehenden Förderperiode (siehe Wissen) seinen Schwerpunkt im Bereich dieser hu-

manzentrierten KI noch ausbauen. Forschungsfelder wie Interpretable und Explainable AI in Kombination mit Visual Computing stehen dabei im Vordergrund.

Während bei klassischen algorithmischen Methoden die Interpretierbarkeit eine Rolle spielt – etwa bei einem Modell mit einem Entscheidungsbaum, in den man „hineinzoomen“ kann, um den Weg zum Ergebnis nachvollziehbar zu machen –, sind Deep-Learning-Systeme zu komplex, um sie auf diese Art analysieren zu können, erklärt Katja Bühler, Leiterin der Biomedical Image Informatics Group des VRVis. „Die Vorgehensweise dieser neuronalen Netzwerke mit ihrer großen Zahl an Datenrepräsentationen, die aufgrund gelernter Gewichten entstehen, ist für den Menschen nicht mehr lesbar.“

Man muss sich also andere Strategien einfallen lassen, um etwas Licht in die oft als „Blackbox“ bezeichneten Abläufe zu bringen und sie „erklärbar“ zu machen. Bühler und ihr Team entwickeln beispielsweise Ansätze, die erkennbar ma-

chen, in welchem Bildbereich eine KI den entscheidenden Anhaltspunkt für eine Krankheit findet.

Ein gängiger Ansatz in diesem Bereich lässt sich mit folgender Frage umreißen: Wenn ich das Bild verändere, wie verändert sich dann die Entscheidung der KI? „Methoden dieser Art gibt es seit einigen Jahren. Doch wir beobachten, dass die Strategie oft zu ungenauen Ergebnissen führt“, sagt Bühler. „Wichtig ist, dass man die Bilder nicht beliebig verändern darf, sondern nur eine sorgfältige Vorgangsweise zum gewünschten Ergebnis führt.“

Bildertausch

Bühler und ihr Team haben die Frage anders gestellt. Sie lautet nun: Was passiert, wenn ich einzelne Bildbereiche durch solche ersetze, von denen ich weiß, dass sie einen gesunden Organismus abbilden? Der Ansatz sieht vor, dass das System eine Aufnahme automatisch in einzelne Abschnitte unterteilt, die den gesuchten Krankheitsmustern entsprechen. Die Bereiche werden dann jeweils mit künstlich generier-

ten Bildausschnitten, die keinen Krankheitshinweis enthalten, nahtlos überblendet. Ändert nun die KI ihr Urteil und klassifiziert die Gesamtaufnahme als gesund – dann hat man jene Bereiche identifiziert, denen das System den Hinweis auf die Krankheit entnommen hat. Die Entscheidung ist somit um ein gutes Stück nachvollziehbarer.

Ein weiterer Ansatz, der sich für bestimmte Krankheitsbilder eignet, ist dagegen, mehrere KI-Modelle, die jeweils auf Detailfragen spezialisiert sind, ein und dasselbe Bild untersuchen zu lassen. „Eine KI erkennt Tumoren, eine andere sieht eingelagerte Flüssigkeit, eine dritte bewertet Verkalkung“, veranschaulicht Bühler. „Diese Einzelbeobachtungen kann ich dann beispielsweise wiederum mit klassischem maschinellem Lernen kombinieren, das im Sinne der Interpretable AI besser nachvollziehbar ist.“

„Die Methoden vermitteln ein Gefühl dazu, wie sicher eine KI-Entscheidung ist“, sagt Bühler. „Ärzten kann damit eine transparente zweite Meinung gegeben werden.“

WISSEN

Neuzugänge im Comet-Programm

Das VRVis (Zentrum für Virtual Reality und Visualisierung) ist eine jener Forschungseinrichtungen, die im Rahmen des Comet-Programms mit Mitteln des Wirtschafts- und des Klimaschutzministeriums unterstützt werden. Die Abwicklung erfolgt über die Förderagentur FFG. Der Fokus des Programms mit seinen hohen Förderanteilen liegt darauf, Spitzenforschung und Wirtschaft für herausfordernde und risikoreichere Projekte zusammenzubringen.

Die Landkarte des Comet-Programms umfasst aktuell fünf K2-Comet-Zentren, 20 K1-Comet-Zentren – eines davon ist das VRVis –, sechs Comet-Module und 24 Comet-Projekte. Die einzelnen Einrichtungen unterscheiden sich durch Größe, Anzahl der Wirtschaftspartner und Fördervolumina.

Sieben der 24 Comet-Projekte – sie liegen in den Bereichen intelligente Produktion, Mobilität sowie Energie und Umwelt – sind Neulinge. Wirtschafts- und Klimaschutzministerium gaben im November mehr als zehn Millionen Euro für sie frei. Zusammen mit den Beiträgen von Ländern und beteiligten Partnern haben die Projekte ein Gesamtvolumen von gut 37,5 Millionen Euro. (pum)