

Lichtblick für die Diagnostik: Mit einem Infrarotlaser analysiert Christina Hofer, die am Max-Planck-Institut für Quantenoptik promovierte, Blutproben. In den Infrarotspektren lassen sich mit Methoden des maschinellen Lernens diverse Krankheiten erkennen.



EIN TROPFEN BLUT IM LASERLICHT

TEXT: THOMAS BRANDSTETTER

FOTO: THORSTEN NAESER / MPI FÜR QUANTENOPTIK

Die Biomoleküle in unserem Blut enthalten wichtige Informationen über unseren Gesundheitszustand. Forschende des Max-Planck-Instituts für Quantenoptik in Garching bei München setzen auf maschinelles Lernen, um die Analyse von Blutproben zu erleichtern. So wollen sie einfache und kostengünstige Gesundheitschecks zur Früherkennung von Krankheiten ermöglichen.

Es ist ein leiser, aber kraftvoller Wandel, der sich aus den medizinischen Laboren bis in die Arztpraxen und Krankenhäuser ausbreitet. Maschinelles Lernen erobert rasant auch die Welt der Gesundheitsversorgung. Denn überall, wo es Muster zu erkennen gilt, können die neuen, lernfähigen Algorithmen ihre Stärken voll ausspielen. Sie unterstützen Ärztinnen und Ärzte bei der Auswertung von Röntgenaufnahmen, analysieren Gewebeschnitte oder unterscheiden harmlose Muttermale von bösartigen Melanomen. Ihre Diagnosen sind dabei immer öfter auch genauer als die menschlicher Fachleute. Der größte Teil der neuen Methoden, die künstliche Intelligenz nutzen, ist allerdings noch gar nicht für den medizinischen Einsatz zugelassen. Was wir gerade erleben, ist erst der Anfang des medizinischen Fortschritts auf der Basis von Daten.

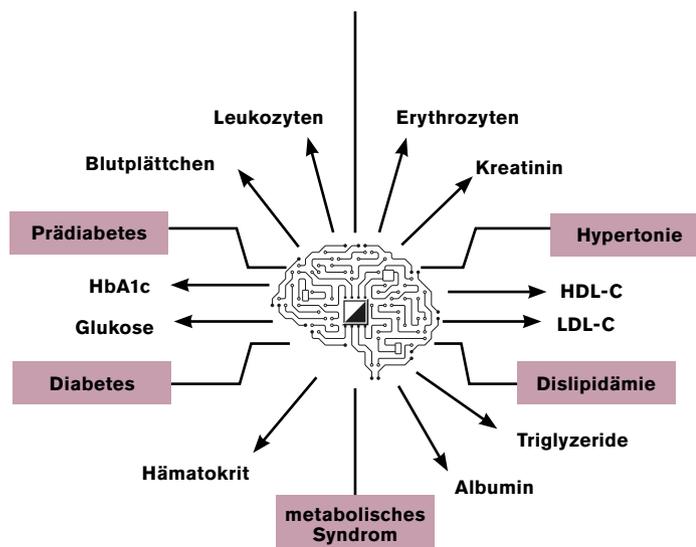
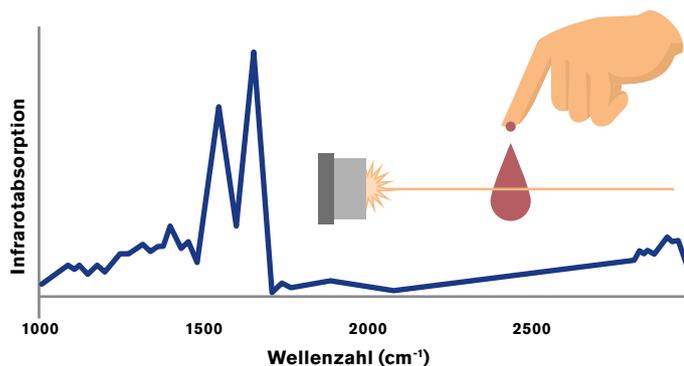
Am Max-Planck-Institut für Quantenoptik arbeitet auch Kosmas Kepesidis daran, maschinelles Lernen für die Medizin nutzbar zu machen, insbesondere für die Auswertung von Messungen an menschlichem Blutplasma. Gemeinsam mit seiner Kollegin Mihaela Zigman, die für die Messungen verantwortlich ist, hat der Datenwissenschaftler eine Methode entwickelt, die Gesundheits-screensings zur Früherkennung von Krankheiten wie Diabetes oder Lungenkrebs ermöglichen soll. Dafür reicht den beiden Forschenden eine einzige, wenige Minuten dauernde Messung an nur einem einzigen Tropfen Blut. Die Methode ist so schnell und kostengünstig, dass sie auch für einen massenhaften Einsatz geeignet wäre. Damit werden umfassende Gesundheits-screensings möglich, die manche Erkrankungen schon frühzeitig erkennen und so eine bessere Therapie ermöglichen. Denn dem RKI zufolge leiden in Deutschland gut eine Million Menschen unter einem unerkannten Diabetes, noch mehr wissen nichts von ihrem Bluthochdruck.

Viele Krankheiten spiegeln sich in der molekularen Zusammensetzung etwa des Blutplasmas wider. Moderne Labortechniken wie Massenspektromie oder Kernspinreso-

nanzspektroskopie haben zwar bereits Potenzial gezeigt, anhand der Messung einzelner Biomarker wie Proteinen oder Stoffwechselprodukten Krankheiten wie Krebs oder Diabetes zu erkennen. Doch diese Methoden sind aufwendig und teuer. „Sie messen üblicherweise nur wenige Moleküle gleichzeitig und benötigen meist eine komplizierte Aufbereitung der Proben“, erklärt Kepesidis.

Ein molekularer Fingerabdruck

Verdächtige Veränderungen in den Biomarkern lassen sich aber auch mit Infrarotlicht aufspüren. Diesen Teil des elektromagnetischen Spektrums nehmen wir als Wärmestrahlung wahr. Das Licht versetzt Moleküle in Schwingungen. Verschiedene Substanzen sprechen dabei auf charakteristische Frequenzen des Infrarotspektrums an. In komplizierten Biomolekülen wie etwa Proteinen sind es mehrere Hundert bis tausend Atome, die so in einem wilden Durcheinander zum Pulsieren gebracht werden.



Umfassender Gesundheitscheck: Mit maschinellem Lernen verknüpft, liefert die Infrarotspektroskopie einen molekularen Fingerabdruck des Blutplasmas und ermöglicht es, mit einer einzigen Messung Volkskrankheiten wie Diabetes, Bluthochdruck oder eine Fettstoffwechselstörung, die zusammen das sogenannte metabolische Syndrom bilden, zu erkennen.

Kepesidis und Zigman setzen dabei auf extrem kurze Laserpulse mit hoher Bandbreite, die alle in der Blutprobe möglichen Schwingungen gleichzeitig anregen. Durchleuchten sie damit eine Blutprobe, absorbieren die Moleküle darin genau die für sie typischen Frequenzen. Die Proteine, Fette oder Zucker im Blutplasma hinterlassen eine Art molekularen Fingerabdruck im Infrarotspektrum. Darin schlagen sich auch winzige, durch Krankheiten verursachte Veränderungen der Zusammensetzung nieder.

„Diese neue Messmethode hat uns auch in der Auswertung der Daten zu einem Paradigmenwechsel veranlasst“, sagt Kepesidis. Aufgrund der unzähligen molekularen Bindungen in den Proben ist das gemessene Spektrum nämlich sehr kompliziert, und die molekularen Ursachen für eine bestimmte Veränderung sind nur schwer zu erkennen. Anstatt also zu analysieren, warum ein bestimmter Peak in einem Spektrum etwas höher oder niedriger ausfällt als normalerweise, analysieren die maschinellen Lernmodelle den gesamten Fingerabdruck der Probe auf einmal. „Das ist der entscheidende Vorteil unseres Ansatzes“, sagt Kepesidis. Dafür trainieren die Forschenden ihr künstliches neuronales Netz kurzerhand mit den gesamten Spektren. Und die „Magie“ in den Tiefen ihrer Neuronenschichten ermöglicht es der KI, darin Muster zu erkennen, die auf bestimmte Krankheiten hindeuten.

Für das Training der Netze konnten Kepesidis und sein Team auf Ergebnisse von Fall-Kontroll-Studien zurückgreifen, also klinischen Studien, mit denen herausgefunden werden soll, ob bestimmte Faktoren wie eben die Zusammensetzung des Blutplasmas mit einer bestimmten Krankheit zusammenhängen. Dafür bilden Mediziner

AUF DEN PUNKT GEBRACHT

Methoden des maschinellen Lernens identifizieren in Infrarotspektren Signaturen verschiedener Erkrankungen wie Diabetes, Bluthochdruck, oder metabolisches Syndrom und Krebserkrankungen wie Lungenkrebs.

Die künstliche Intelligenz wird dabei zum einen darauf trainiert, die gesamten Spektren zu analysieren, zum anderen kann sie lernen, diese auf die für die Diagnostik wesentliche Information zu reduzieren.

Da die Infrarot-Fingerabdrücke die gesundheitlichen Veränderungen frühzeitig erkennen, eignen sie sich für präventive bevölkerungsweite Gesundheitsscreenings.

Sowohl beim Umgang mit den Trainingsdaten als auch beim Einsatz datenbasierter Methoden müssen Datensicherheit und -schutz gewährleistet sein.

üblicherweise zunächst zwei Gruppen: Personen, die die Krankheit bereits haben (Fälle), und gesunde Personen (Kontrollen). Dann untersuchen sie, ob sich die Gruppen in bestimmten Merkmalen zum Beispiel im Blutplasma unterscheiden. Während ein Mensch aber lediglich einzelne Auffälligkeiten in einem Infrarotspektrum mit einer bestimmten Erkrankung in Verbindung bringen kann, entdeckt eine künstliche Intelligenz auch komplexere Zusammenhänge und Muster in den Spektren. So ist es Kepesidis und Zigman gelungen, mit einer einzigen Messung des molekularen Fingerabdrucks eine Reihe chronischer Erkrankungen zu erkennen. Erste Ergebnisse dazu haben sie bereits vergangenes Jahr im Fachjournal *Cell Reports Medicine* veröffentlicht.

Für ihre Studie haben sie gemeinsam mit Kolleginnen und Kollegen der Ludwig-Maximilians-Universität München und der Technischen Universität München über 5000 Blutproben per Infrarotspektroskopie untersucht und dabei eine Reihe gesundheitlicher Auffälligkeiten von Fettstoffwechselstörungen über Bluthochdruck bis zu Diabetes mit bis zu 95-prozentiger

Treffsicherheit erkannt. Dabei kann das System im Gegensatz zu herkömmlichen Diagnostiken gleichzeitig die Wahrscheinlichkeiten für unterschiedliche und voneinander unabhängige Krankheiten berechnen. „Das Ergebnis könnte also etwa lauten, dass ein Patient mit einer Wahrscheinlichkeit von 70 Prozent an Bluthochdruck leidet und gleichzeitig mit 40-prozentiger Wahrscheinlichkeit Diabetes hat“, erklärt Kepesidis. „Das sind wichtige Informationen, die Ärzten dabei helfen können, ihre Diagnosen zu stellen.“ Außerdem konnten die Forscher vorhersagen, ob jemand in Zukunft ein sogenanntes metabolisches Syndrom entwickeln wird – und das schon Jahre, bevor sich die Erkrankung tatsächlich bemerkbar macht.

Frühzeitige Krebsdiagnose

In einer weiteren, ähnlichen Studie ist es Kepesidis und Zigman gelungen, die molekularen Fingerabdrücke von Krebserkrankungen zu identifizieren. Dafür haben sie das Blutplasma von über 2500 Personen untersucht, darunter Menschen mit Lungen-, Prostata-, Brust- und Blasenkrebs sowie gesunde Kontrollpersonen. Auch hier haben die KI-Algorithmen spezielle Infrarotmuster im Blut gefunden, die auf Krebs hindeuten – und das schon in einem sehr frühen Stadium. Mit einer Trefferquote von 88 Prozent funktionierte das bei

→

„Die neue Messmethode hat uns in der Auswertung der Daten zu einem Paradigmenwechsel veranlasst.“

KOSMAS KEPESIDIS

Lungenkrebs besonders gut. Bei den anderen Krebsarten lag die Genauigkeit zwischen 60 und 70 Prozent. Damit erwies sich die neue Technik bereits als vielversprechend, um auch Krebs frühzeitig und zuverlässig zu erkennen. Kepesidis und sein Team arbeiten nun daran, die Technik noch genauer zu machen. Für die Klassifikationsaufgaben bekommen die künstlichen neuronalen Netze für jedes Spektrum, das sie während ihres Trainings sehen, auch die Information, welche Krankheit sich hinter den Daten verbirgt. Im obigen Beispiel lernen sie so, verdächtige Muster in den Spektren mit Krebs in Zusammenhang zu bringen. Alternativ dazu beschäftigen sich Kepesidis und sein Team aber auch mit sogenannten Autoencodern. Das sind spezielle neuronale Netze, die das Rauschen aus den Spektren filtern und daraus die relevanten Informationen isolieren. „Das ist vergleichbar mit der Dateikomprimierung“, erklärt Kepesidis. „Man zippt die Daten in eine kompakte Form, kann sie dann aber auch wieder entpacken, um die volle Information wiederherzustellen.“ Anders als bei den Klassifikationsaufgaben erhalten die Autoencoder für ihr Training jedoch keinen Hinweis darauf, ob und welche Krankheit in einem Spektrum steckt. Sie werden lediglich darauf trainiert, die Spektren auf einige wenige sogenannte latente Variablen zu reduzieren, die aber immer noch genug Information enthalten, um daraus das ursprüngliche Spektrum zu rekonstruieren.

34

Man kann sich das tiefe neuronale Netz eines Autoencoders also wie einen Trichter vorstellen, der an seinem Eingang eine große Menge an Informationen aufnimmt und sie zur Verengung an seinem Ende weiterleitet. In diesem engen Bereich sind nur noch wenige Neuronen aktiv und bilden die stark reduzierte Darstellung der ursprünglichen Daten. Diese latenten Variablen enthalten gewissermaßen die Essenz der Eingangsdaten, unwichtige Informationen hat der Autoencoder dagegen entfernt. Das neuronale Netz des Decoders dagegen sieht wie ein umgedrehter Trichter aus, der diese latenten Variablen in seinem engen Eingang aufnimmt und versucht, daraus so gut wie möglich die ursprünglichen Daten wiederherzustellen. So lernt das Netzwerk, wie es Daten komprimieren kann, ohne dabei wichtige Informationen zu verlieren.

„So haben wir latente Variablen identifiziert, die mit verschiedenen Krankheiten korrelieren“, sagt Kepesidis. „Diese Werte haben also eine diagnostische Aussagekraft und könnten in Zukunft als Biomarker für bestimmte Krankheiten dienen.“ Wie die erst kürzlich veröffentlichten Ergebnisse zeigen, erkennt die neue Methode Lungenkrebs mit einer um 2,6 Prozentpunkte höheren Genauigkeit als die Interpretation der Infrarotspektren ohne das Autoencoding. „Das liegt daran, dass die komprimierten Spektren kein Rauschen und keine redundanten Informationen mehr enthalten, sondern sich hauptsächlich auf wenige, biologisch relevante Signale konzentrieren“, erklärt Kepesidis. Der Vorteil des

Autoencoding liegt also darin, dass es mehr Klarheit bei der Interpretation der Spektren schafft.

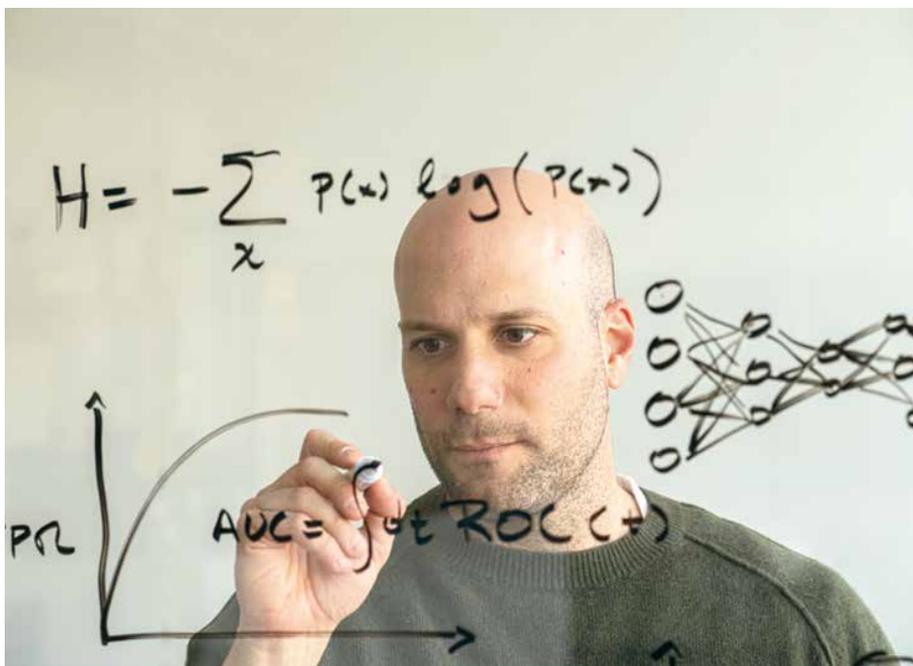
Kombinierte Techniken

Die Autoencoder sind auch Teil einer aktuellen Kooperation zwischen Kepesidis und Ferenc Krausz, der als Direktor am Max-Planck-Institut für Quantenoptik die Abteilung für Attosekundenphysik leitet. Er hat für die Erzeugung extrem kurzer Laserpulse und die Forschung damit 2023 den Physiknobelpreis erhalten. Im Unterschied zu den Fall-Kontroll-Studien mit Zigman geht es in diesem Projekt um Längsschnittstudien, die die Entwicklung der Teilnehmenden über mehrere Jahre verfolgen. „Unser Ziel ist es, Variablen zu identifizieren, die Abweichungen von einem gesunden Zustand signalisieren, und eine proaktive Gesundheitsfürsorge zu ermöglichen“, sagt Kepesidis. Die Aussagekraft solcher Studien ist dann am höchsten, wenn sie mehrere, sich ergänzende Techniken verwenden. So ist etwa die Massenspektrometrie besonders gut dafür geeignet, Proteine zu analysieren, während die Kernspinresonanzspektroskopie eher für Metaboliten, also die Zwischenprodukte in einem biochemischen Stoffwechselweg geeignet ist. Beide Verfahren verwenden die gleichen Blutproben wie die Infrarottechnik, liefern aber andere Erkenntnisse über die molekulare Zusammensetzung der Proben. Im Gegensatz zur Infrarotspektroskopie sind Massenspektrometrie und Kernspinresonanzspektroskopie allerdings teuer. Um möglichst gute Informationen zu gewinnen und gleichzeitig die Kosten zu minimieren, setzen die Forschenden auch hier zunächst auf günstige Methoden wie die Infrarotspektroskopie und maschinelles Lernen. „Wenn bei der Infrarotspektroskopie erste Anzeichen für eine Erkrankung zu erkennen sind, würde die KI gezielte Folgetests wie zum Beispiel eine Massenspektrometrie empfehlen“, erklärt Kepesidis.

Bei allen Vorteilen werfen die neuen, datengetriebenen Methoden aber auch ethische Fragen auf. „Ein wichtiger

„Unser Ziel ist es, eine proaktive Gesundheitsfürsorge zu ermöglichen.“

KOSMAS KEPESIDIS



Aspekt ist zum Beispiel der Umgang mit Patientendaten“, sagt Kepesidis. Das sieht auch Eva Winkler so. Die Krebsforscherin und geschäftsführende Direktorin am Nationalen Centrum für Tumorerkrankungen Heidelberg engagiert sich im Deutschen Ethikrat im Diskurs zu Innovationen in den Lebenswissenschaften und einer ethisch verantwortungsvollen Gesundheitsversorgung. „KI-Trainingsdatensätze sind riesig und nachträglich nicht lösbar“, gibt sie zu bedenken. „Zwar sind sie in den meisten Fällen anonymisiert, doch absolute Anonymität gibt es nicht.“ Was es gebe, sei lediglich eine gewisse Unwahrscheinlichkeit der Rückverfolgbarkeit. Es könne also nie ganz ausgeschlossen werden, dass die Daten, auf denen die Systeme trainiert werden, wieder mit den Patienten in Verbindung gebracht werden und ihnen daraus Nachteile entstehen – etwa durch Diskriminierung aufgrund höherer Krankheitsrisiken. Das könnte etwa dazu führen, dass Versicherungen Menschen ablehnen oder es etwa bei Verbeamten wegen psychiatrischer Vorgeschichten zu Benachteiligungen kommt. „Datensicherheit und Datenschutz sind Grundbedingungen für datenintensive Medizin“, sagt Winkler. Die klinischen Studien am Max-Planck-Institut für Quantenoptik folgen deshalb einem strengen Protokoll, das von einer Ethikkommission der jeweiligen medizinischen Kooperationspartner genehmigt wurde. Außerdem ist die Mustererkennung von KI-Systemen immer nur so gut wie die Daten, mit denen sie trainiert wurden. „Hier müssen wir darauf achten, dass die Menschen, bei denen solche Methoden angewendet werden, auch in den Trainingsproben repräsentiert sind, auf denen sie entwickelt wurden.“ Ein Test sollte also nicht nur Diabetes oder Bluthochdruck bei einer bestimmten Bevölkerungsgruppe erkennen, sondern bei allen Menschen – ein Problem, dessen sich auch

Kosmas Kepesidis und seine Gruppe entwickeln Algorithmen, um die medizinische Diagnostik zu verbessern und zu vereinfachen. Dabei spielen auch datenethische Aspekte eine Rolle.

Kepesidis bewusst ist. „Ein an Patienten in München trainierter Test muss möglicherweise für andere Bevölkerungsgruppen angepasst werden“, bestätigt der Datenwissenschaftler.

Dazu kommt ein Problem, das zwar alle KI-Anwendungen eint, in der Medizin jedoch besonders heikel ist:

die mangelnde Erklärbarkeit der Ergebnisse. Wenn KI-Modelle Spektroskopiedaten auswerten und dabei komplexe Muster erkennen, die für das menschliche Auge nicht nachvollziehbar sind, entsteht nämlich das berühmte Problem der „Blackbox“. „Als Arzt oder Labormediziner kann man nicht überprüfen, warum die KI eine Probe als auffällig einstuft“, sagt Winkler. „Trotzdem trägt grundsätzlich der Arzt die Letztverantwortung, wenn etwas schiefgeht.“ Besser wäre es aus Sicht der Medizinerin, wenn es ein klares Regelwerk gäbe, das die Verantwortung je nach Problemfall zwischen KI-Entwicklern, den Institutionen, die die Technik einsetzen, und den behandelnden Ärzten aufteilt.

Dennoch bleibt es ein zentraler Punkt beim Einsatz von KI, dass die Systeme nie selbst Entscheidungen treffen dürfen, sondern immer auch ein menschlicher Arzt eingebunden sein muss. „Das ist auch im europäischen KI-Gesetz so vorgeschrieben“, sagt Winkler. Zwar gebe es Studien, in denen die KI Fachärztinnen oder Fachärzte teilweise übertrifft, ein medizinisches Bild müsse aber immer im Kontext der Patientengeschichte betrachtet werden. Die meisten ihrer bekannten Studien würden jedenfalls eines klar zeigen: Die besten Ergebnisse für Patientinnen und Patienten ergeben sich aus der Zusammenarbeit zwischen Mensch und Maschine.



[www.mpg.de/podcasts/
ki-und-biomedizin](http://www.mpg.de/podcasts/ki-und-biomedizin)

