

# Diagnostik der diabetischen Retinopathie: Googles Computer unterstützen Augenärzte

Corina Klaeger, Ulrike Novotny

**Augenärztinnen und Augenärzte werden dadurch nicht überflüssig, aber: Ein Computer-Algorithmus diagnostiziert eine diabetische Retinopathie (DR) ebenso zuverlässig wie ein Retinaspezialist. Im Vorteil ist der Algorithmus aber in Bezug auf Kosten und Schnelligkeit, und überraschend liefert er zudem Aussagen, die seinen Nutzen noch deutlich erweitern können (z. B. indem er kardiovaskuläre Risiken erkennt).**

## «Deep Learning»

Das Thema «Deep Learning» (DL) ist in der Retinadiagnostik ein Hot Topic, so dass ihm auch am SOG-Kongress in Davos eine Keynote Lecture gewidmet wurde. Die Referentin dieses Vortrags, Dr. Lily Peng, San Francisco, leitet bei Google das Deep Mind-Programm zur Retinadiagnostik. Es hat zum Ziel, eine diabetische Retinopathie in grossen Bevölkerungsgruppen möglichst früh zu entdecken.

Das Unternehmen Google Deep Mind befasst sich mit Machine Learning. «Lernen» in diesem Sinne bedeutet, dass ein Algorithmus selber herausfindet, wie ein gegebenes Problem zu lösen ist, ohne dass das konkrete Vorgehen dazu im Voraus bekannt ist. Je nach Art des Problems kommen unterschiedliche Techniken des Machine Learnings zur Anwendung: «supervised» oder «unsupervised» learning sowie «reinforcement learning».

**«supervised learning»:** Der Algorithmus wird vorab mit einer grossen Datenmenge von korrekt gelösten Beispielen des Problems trainiert. Dabei erhält er keine Anleitung, wie diese Daten zu verarbeiten sind, sondern der Algorithmus errechnet sich selbstständig ein geeignetes Vorgehen mit dem Ziel, dass seine eigene Lösung möglichst genau den korrekt gelösten Beispielen entspricht. Beispiele von Problemstellungen, bei denen sich dieser Ansatz eignet, sind: Erkennen von handschriftlichem Text, Erkennen von Bildinhalten und somit auch Erkennen einer diabetischen Retinopathie auf Fundusbildern.

**«unsupervised learning»** kommt dann zur Anwendung, wenn zur Problemstellung keine korrekt gelösten Trainingsdaten produziert werden können. Stattdessen kann der Algorithmus mit statistischen Methoden in grossen Datenmengen Gruppierungen und Muster erkennen und dabei Ausreisser aufspüren, was z. B. ermöglicht, Kreditkartenmissbrauch aufgrund von ungewöhnlichen Transaktionen festzustellen.

**«reinforcement learning»** ist, was wir alle aus der Psychologie kennen. Man trainiert gewünschtes Verhalten durch Belohnung von gutem und Bestrafen von schlechtem Verhalten. So kann z. B. ein Algorithmus ein Computerspiel erlernen, ohne dass man ihm die Regeln und ein optimales Vorgehen programmiert, sondern indem man lediglich Feedback über das erreichte Resultat gibt. Durch viele Wiederholungen findet der Computer selber die beste Strategie.

Deep Learning ist eine konkrete Art, wie man diese Lernansätze implementieren kann, nämlich mit sogenannten Artificial Neural Networks. Das sind Algorithmen, die in ihrer Funktionsweise der Biologie nachempfunden sind. Sie bestehen aus aktivierbaren «Neuronen», die miteinander verbunden sind und ihre Aktivierung anderen Neuronen weitergeben können. Die Möglichkeit zu lernen besteht darin, diese Verbindungen zwischen den Neuronen zu schwächen oder zu stärken, analog den synaptischen Verbindungen in der Neurophysiologie.

## Anwendungen: Beispielsweise Imagenet und Deep Mind Health

Als ein Beispiel für ein Deep Learning-Programm nannte Peng IMAGE-NET. Hierfür wurde eine umfangreiche Bilddatenbank angelegt, und jedes Bild wurde mit insgesamt 1000 möglichen Bezeichnungen (Labels) versehen. Das Programm wurde daraufhin aufgefordert, neue Bilder selbst zu «labeln» (z. B. Milbe, Containerschiff oder Leopard). Im Jahr 2011 überholten Deep Learning-Algorithmen erstmals den Menschen bei der Aufgabe, die Bilder richtig einzuordnen. Google Deep Mind arbeitet mit dem Programm Tensor Flow, einer offenen Standardsoftware für generelles Maschinenlernen, die sich für das Deep Learning eignet. Das Projekt Deep Mind Health wurde im Februar 2016 lanciert. Etliche Studien hierzu wurden bereits publiziert, zwei fanden Eingang auf das Titelblatt von Nature.

## Diabetische Retinopathie

Zur diabetischen Retinopathie arbeiten Google Deep Mind und Dr. Pengs Forschergruppe mit dem Moorfields Eye Hospital in London zusammen.

Die diabetische Retinopathie ist deshalb besonders interessant, weil sie in den westlichen Ländern die häufigste Ursache für eine Blindheit im erwerbsfähigen Alter ist und weil ihre Inzidenz auch in den Schwellenländern sehr stark im Steigen begriffen ist. In diesen Ländern herrscht ein teilweise extrem gravierender Mangel an Augenärzten, in Indien beispielsweise fehlen



Dr. Lily Peng M. D., Product Manager with the Google Brain AI Research Group, Mountain View, CA. Sie hielt die spannende Keynote Lecture am SOG-Kongress 2017.

etwa 127 000 Augenärzte. Eine diabetische Retinopathie sollte entdeckt werden, bevor der Visus beeinträchtigt ist. Im Rahmen der Kooperation zwischen Google Deep Mind und dem National Health Service in Grossbritannien unter Leitung des Moorfields Eye Hospital (London) versah ein Team von 54 Ophthalmologen 130 000 Fundusfarbfotos mit einem von fünf Labels: Ohne diabetische Retinopathie (DR), milde, moderate, schwere und proliferative DR. Die Fundusfotos wurden durch Angabe des Geschlechts, des Alters, des Blutdrucks und HbA<sub>1c</sub>-Wertes ergänzt. Genauere Informationen, aufgrund welcher Kriterien das jeweilige Label vergeben worden war, wurden dem Programm nicht geliefert. Es verwendet zur Datenanalyse ein neuronales Netzwerk aus 26 Schichten, um letzten Endes nicht nur ein diagnostisches Label zu liefern, sondern auch die Bildqualität sowie die Grösse des untersuchten Bildausschnittes (Field of View) zu beurteilen. Von den 130 000 Bildern wurden 80 % für das Training des Programms und 20 % für das Feintuning eingesetzt. Klinisch validiert wurde das Programm in zwei Studien (Messidor-2 und EyePACS-1).<sup>1</sup> Im Ergebnis übertraf das Programm die diagnostische Sicherheit von Allgemeinophthalmologen. Der F-Score (ein Wert für die Treffsicherheit eines Tests, definiert als harmonisches Mittel von Genauigkeit und Trefferquote) erreichte für die Beurteilung der Fundusbilder bez. diabetischer Retinopathie bei Allgemeinophthalmologen den Wert von 0.91. Der

Algorithmus erzielte dagegen einen F-Score von 0.95. Die Zuverlässigkeit hängt von der Anzahl der beurteilten Bilder ab; bei grossen Datenmengen nähert sie sich 100 %.

Fehlende Übereinstimmung zwischen «maschinendiagnostizierter» und Ophthalmologen-diagnostizierter DR beruhte am häufigsten auf Nichterkennen von Mikroaneurysmen, Artefakten und fehlbeurteilten Hämorrhagien. Welche Zonen das Programm als relevant für seine jeweilige Beurteilung ansieht, kann es in Form sogenannter Heat maps anzeigen.

Der Vorsprung des Programms in der diagnostischen Zuverlässigkeit betraf besonders Frühstadien der DR. Bei schweren Fällen war die diagnostische Zuverlässigkeit bei grundversorgenden Augenärzten, Retinaspezialisten und dem Algorithmus mehr oder weniger gleich gut.

Das Deep Learning ermöglicht mithin, bekannte Signale zuverlässiger und vor allem schneller zu quantifizieren oder zu erkennen als ärztliche Beurteiler.

Mit verblüffender Genauigkeit konnte das Programm aber anhand der Fundusfotos auch das Geschlecht, das Alter, den systolischen und diastolischen Blutdruckwert sowie das HbA<sub>1c</sub> prognostizieren. Mit solchen Befunden erscheint es möglich, bisher noch unbekannt Signale zu entschlüsseln und so anhand der Retinadiagnostik auch das kardiovaskuläre Risiko bis hin zum Mortalitätsrisiko zu beurteilen, wenn zusätzliche longitudinale Daten eingegeben werden.

### Weitere Einsatzgebiete

Auch auf anderen medizinischen Gebieten lässt sich das Deep Learning nutzen, beispielsweise für die Analyse von Biopsaten mit Krebsverdacht. Das DL könnte das Risiko einer Unter- und einer Überdiagnose senken. In einem Vergleich zwischen Pathologen und DL-Programm erreichte das Programm einen Score von 0.89, die Pathologen (denen unbegrenzte Zeit gelassen wurde) einen Score von 0.73. Der Algorithmus hatte eine 92 %-Sensitivität mit 8 falsch Positiven, der Pathologe eine 73 %-Sensitivität mit 0 falsch Positiven.

### Bessere Ressourcen-Allokation

Die Vorteile der Diagnose auf Basis künstlicher Intelligenz liegen auf der Hand: Der Computer beurteilt in einem Bruchteil der Zeit grosse Datenmengen mit hoher Zuverlässigkeit, ohne Ermüdung und 24h am

Tag. Der Algorithmus kann weltweit zur Verfügung gestellt werden. Mit dieser Methode ist die Diagnostik von Personen mit geringem Schulungsaufwand schnell und bei einer hohen Zahl von Probanden durchführbar, eine ideales Szenario für das Screening grosser Populationen. In einem Land wie Indien, in dem 41 % der Betroffenen einen Visusverlust erleiden, bevor die Diagnose DR gestellt wird, kann dies entscheidend sein.

Selbst in der Schweiz, vor allem aber in der zweiten und dritten Welt stehen über kurz oder lang zu wenige Ärzte zur Verfügung, und dies bei steigender Inzidenz alters- und diabeteskorrelierter Netzhauterkrankungen. Programme wie das von Google Deep Mind können entlasten und Ressourcen besser auslasten. Sie werden kaum die Augenärzte verdrängen. Die Behandlung wird vermutlich die verfügbaren Ärzte mehr als genug beschäftigen. Therapien sind deutlich effektiver, wenn sie früh einsetzen und nicht erst bei unumkehrbarem Visusverlust.

Am Moorfields Eye Hospital geht die Forschung zum Deep Learning weiter. Integriert werden auch OCT-Daten. Die diesjährige Stipendiatin der Alfred-Vogt-Stiftung, Dr. Katrin Fasler, erhält diese Unterstützung für ihr Projekt «*Prospective clinical validation of automated diagnostic algorithms for imaging using machine learning techniques in retinal disease.*» •

### Referenz

1. Gulshan V, Peng L, Coram M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 2016;316(22):2402-2410.

### Danksagung

Wir danken Adrian Klaeger für die Durchsicht und Beratung zu den mathematischen und Informatik-relevanten Fragen.

### Skeptisch?

Zu möglichen Schattenseiten der Machine Learning Programme erschien in der NZZ der Beitrag: Betschon S. Wie sich künstliche Intelligenz narren lässt. 25.10.2017 S. 12.



### Korrespondenz

Dr. med. Corina Klaeger  
Augenärztin FMH  
Gurtengasse 2, CH-3011 Bern  
gurtengasse@gmail.com