

# KÜNSTLICHE INTELLIGENZ IN DER OPHTHALMOLOGISCHEN DIAGNOSTIK

---

**Prof. Dr. med. Nicole Eter**  
Klinik für Augenheilkunde  
Universitätsklinikum Münster, Münster

**VNR: 2760909011803320015 | Gültigkeit: 30.03.2022 – 30.03.2023**

## 1 EINLEITUNG

Mit der zunehmenden Lebenserwartung der Bevölkerung steigt auch der Anteil älterer Menschen und damit verbunden die Häufigkeit von Alterserkrankungen, darunter Augenerkrankungen, wie z. B. die altersbedingte Makuladegeneration (AMD). Um einen Sehverlust bei den Betroffenen zu vermeiden und die Lebensqualität zu erhalten, sind die Früherkennung und eine zielgerichtete Behandlung essenziell [Lu et al. 2018]. Bislang stützte sich die Diagnostik nahezu ausschließlich auf die Expertise der Behandler. Durch immer leistungsfähigere neue technologische Entwicklungen werden jedoch in der klinischen Ophthalmologie immer mehr und größere bildbasierte Datenmengen generiert, deren systematische Auswertung vom medizinischen Personal allein nicht mehr bewältigt werden kann. Die Daten bergen somit ein großes, häufig ungenutztes Potenzial in Bezug auf diagnostische und prognostische Auswertungen. Um eine Automatisierung der Auswertungsprozesse zu ermöglichen, werden gegenwärtig zahlreiche auf künstlicher Intelligenz (KI) basierende Technologien entwickelt und zum Teil bereits erfolgreich eingesetzt. Es ist zu erwarten, dass entsprechende Methoden zukünftig immer mehr Einzug in die ophthalmologische Diagnostik halten.

Unter dem Begriff KI werden computerbasierte Technologien zusammengefasst, die in der Lage sind, menschliches Verhalten nachzuahmen [Moraru et al. 2020]. Im medizinischen Bereich wird das *Machine Learning*, eine Unterkategorie der KI, bereits seit den 70er-Jahren im Bereich der Diagnostik und des Screenings sowie zu prognostischen Zwecken eingesetzt [Patel et al. 2009]. Beim *Machine Learning* werden Computer-Algorithmen angewandt, die es dem System ermöglichen, auf Grundlage großer Datenmengen mit nur geringfügigem Eingreifen durch den Menschen zu lernen [Moraru et al. 2020]. Mittels *Machine Learning*, und insbesondere der Unterkategorie *Deep Learning*, lassen sich pathologische Merkmale von Augenerkrankungen automatisch erkennen und klassifizieren. Hier schreitet die Leistungsfähigkeit entsprechender auf KI basierender Systeme immer weiter voran. Die KI-basierten Anwendungen sollen Ophthalmologen zukünftig dabei unterstützen präzisere Diagnosen zu stellen und eine personalisierte Gesundheitsversorgung zu ermöglichen [Tong et al. 2020].

Ziel der vorliegenden Fortbildung ist es, einen Überblick über aktuelle und zukünftige Entwicklungen KI-basierter Systeme im Bereich der Ophthalmologie zu geben sowie die Möglichkeiten und Grenzen dieser Anwendungen aufzuzeigen.

## 2 TECHNISCHE GRUNDLAGEN

### 2.1 MACHINE LEARNING

Der Begriff „künstliche Intelligenz“ bezeichnet Maschinen, die menschliches, intelligentes Verhalten nachahmen. Das u. a. im medizinischen Bereich eingesetzte *Machine Learning* (ML) stellt dabei einen Teilbereich der KI dar. Anhand eines geeigneten Datensets wird bei dieser Methode das Computersystem dahingehend trainiert, dass es im Anschluss in der Lage ist korrekte Voraussagen hinsichtlich einer bestimmten Fragestellung zu treffen [Moraru et al. 2020, Tong et al. 2020]. Hierbei kann weiterhin das *Supervised Learning* (überwachtes Lernen) vom *Unsupervised Learning* (nicht überwachtes Lernen) unterschieden werden. Beim *Supervised Learning* wird die Maschine mithilfe von Daten trainiert, die von Menschen im Vorfeld bereits kategorisiert wurden. Dies erfordert einen erheblichen Arbeitsaufwand und ist daher zeitaufwendig [Tong et al. 2020]. Beim *Unsupervised Learning* wird die Maschine hingegen mit unbearbeiteten Daten trainiert, wobei die Maschine selbstständig Muster und Strukturen in den Daten identifiziert [Tong et al. 2020].

Algorithmen, die im Bereich des ML angewandt werden, sind Entscheidungsbäume (*Decision Trees*) [Rokach und Maimon 2014], der naive Bayes-Algorithmus [Lowd und Domingos 2005], das *Random-Forest*-Verfahren [Cutler et al. 2012], die *Support-Vector-Machine*-Methode [Cortes und Vapnik 1995] und der *K-Nearest-Neighbor*-(KNN)-Algorithmus [Cover und Hart 1967].

Die Entwicklung von ML-Algorithmen erfordert große Datensets. Diese umfassen eine Sammlung von **Instanzen**, die repräsentativ für einzelne Objekte stehen, z. B. Patienten oder Augen. Die Instanzen werden wiederum durch **Variablen** näher beschrieben, die klinische Einzelheiten wie Alter, Geschlecht oder Augennendruck bzw. bestimmte Bildelemente, bestehend aus einzelnen Pixeln, Pixelgruppen oder Voxeln (3-dimensionalen Pixelsets), repräsentieren. Die Instanzen werden in verwandten Gruppen zusammengefasst, die **Klassen** genannt werden (z. B. Diabetespatienten, Glaukomapatienten, gesunde Patienten) [Armstrong und Lorch 2020].

Am häufigsten werden in der Ophthalmologie neuronale Netzwerke eingesetzt, deren Algorithmus dem menschlichen Gehirn nachempfunden ist [Armstrong und Lorch 2020]. In den neuronalen Netzwerken werden die Neuronen erster Ordnung durch einzelne Variablen (z. B. Alter) für eine einzelne Instanz repräsentiert. Jedes Neuron im Algorithmus ist mit einer Reihe weiterer nachgeschalteter Neurone verbunden, wobei die Verbindungen unterschiedlich stark gewichtet sind, sodass jedes Neuron erster Ordnung den Wert des Neurons zweiter Ordnung entweder verkleinert oder vergrößert (Abbildung 1A). Der Wert des Neurons zweiter Ordnung wird somit durch die Summe der gewichteten Werte aller mit ihm in Verbindung stehenden Neurone erster Ordnung bestimmt (Abbildung 1B) [Armstrong und Lorch 2020].

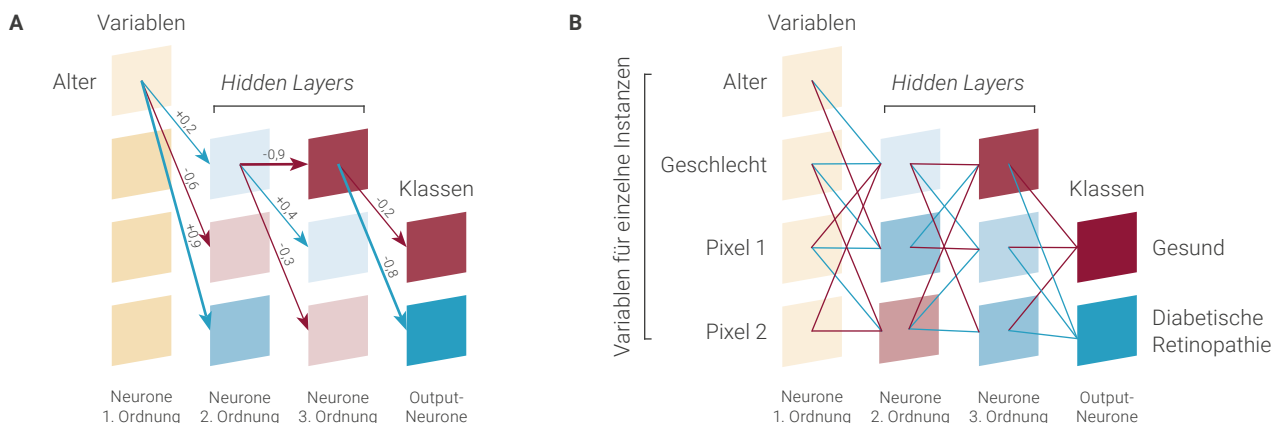
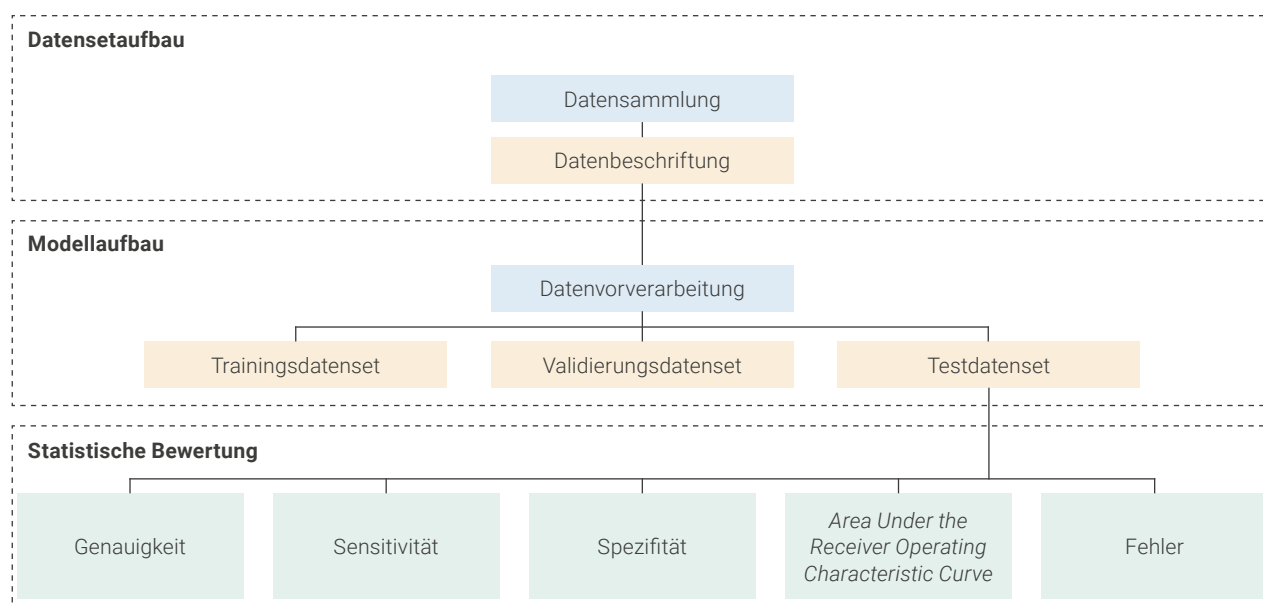


Abbildung 1: Prinzip eines neuronalen Netzwerks; modifiziert nach [Armstrong und Lorch 2020].

Der Output besteht aus den vordefinierten Klassen des Datensets (z. B. Gesund vs. diabetische Retinopathie). Das neuronale Netzwerk wird anhand von Datensets, mit Instanzen bekannter Klassen, trainiert. Dabei prüft das System selbstständig, ob es eine korrekte oder falsche Voraussage für eine Klasse gemacht hat, und verändert die Gewichtung der Verbindungen dahingehend, dass die Aussage korrekt wird. Nach der Anpassung erfolgt eine erneute Prüfung anhand eines Trainingsdatensets, solange, bis alle Voraussagen korrekt ausfallen (das System „lernt“). Im nächsten Schritt wird der Algorithmus an einem unbekanntem Validierungs-Datenset überprüft, bevor

es auf ein aktuelles Testdatenset angewandt wird (Abbildung 2) [Armstrong und Lorch 2020]. Anhand der Ergebnisse dieser Validierung lassen sich die Sensitivität (die wahre Positivrate), die Spezifität (die wahre Negativrate) und die Präzision (Verhältnis von wahren positiven und wahren negativen Ergebnissen) und eine *Receiver Operating Characteristic Curve* (ROC) berechnen [Armstrong und Lorch 2020]. Die *Area Under the Receiver Operating Characteristic Curve* (AUC) wird häufig verwendet, um die Qualität des Vorhersagemodells zu bewerten. Die AUC von effektiven Modellen bewegen sich dabei im Bereich zwischen 0,5 und 1,0 (Maximalwert) [Lu et al. 2018].



**Abbildung 2:** Arbeitsablauf KI-basierter Prozesse; modifiziert nach [Tong et al. 2020]. KI: künstliche Intelligenz

## 2.2 DEEP LEARNING

Eine wichtige Unterkategorie des ML ist das *Deep Learning* (DL). Bei dieser Methode werden mehrere Lagen künstlicher neuronaler Netzwerke hintereinandergeschaltet und der Lernprozess des Systems erfolgt unabhängig von menschlichen Eingriffen. Die Technik wird aktuell bei der Bilderkennung (z. B. Gesichtserkennung bei Facebook und Bildersuche bei Google), virtuellen Assistenten und diagnostischen Assistenzsystemen (z. B. IBM *Watson for Oncology*) eingesetzt. Im Vergleich zum ML verwendet die Software-Architektur des DL mehr Hidden Layers, d. h.

Neurone höherer Ordnung, die zwischen Input- und Output-Neuronen geschaltet sind (Abbildung 3, Seite 4). Die zusätzlichen Neurone ermöglichen die Dekodierung von Bildrohdaten, ohne dass eine Notwendigkeit zur manuellen Vorbereitung der Bilddaten besteht. Häufig verwendete Algorithmen des DL sind *Deep Belief Networks* (DBN), *Convolutional Neural Networks* (CNN) und *Recurrent Neural Networks* (RNN) (Tabelle 1, Seite 4) [Tong et al. 2020]. Am häufigsten eingesetzt wird bei der Bilderkennung das CNN, darunter LeNet, AlexNet, ResNet und GoogleNet [Tong et al. 2020]. DL ist mittlerweile auch die Standardmethode bei der medizinischen Bildauswertung [Tan et al. 2020].

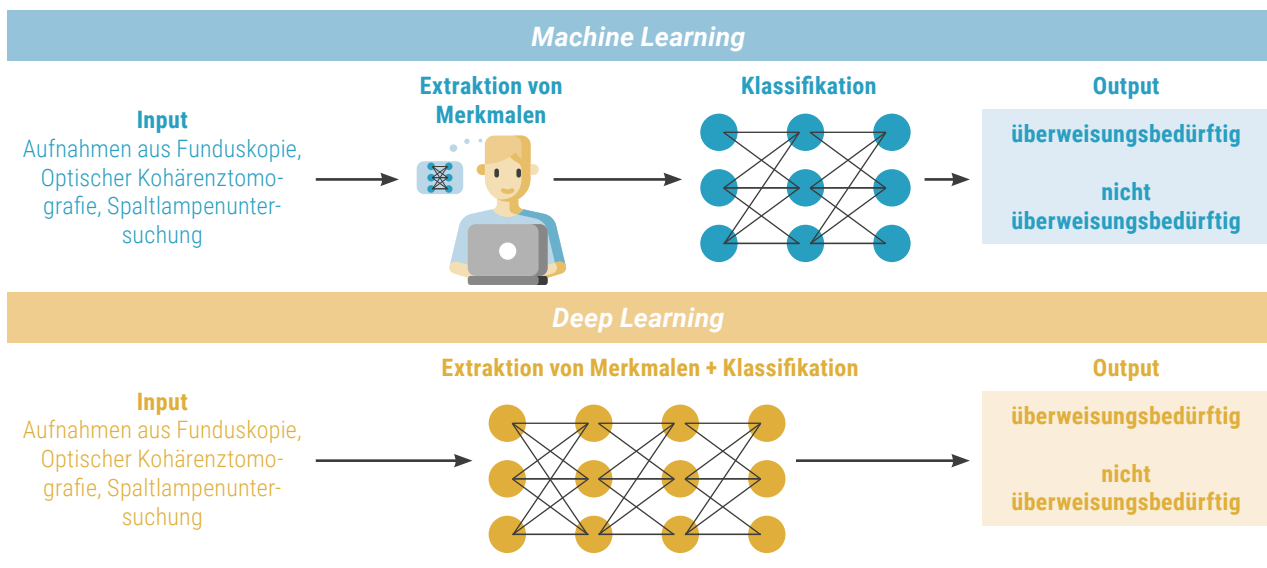


Abbildung 3: Unterschiede zwischen *Machine Learning* und *Deep Learning*; modifiziert nach [Schmidt-Erfurth et al. 2018a].

Tabelle 1: Wichtige Algorithmen des *Machine Learning* und *Deep Learning* [Tong et al. 2020].

KI-Techniken	Klassifikation	Algorithmen
<b>Conventional Machine Learning</b>	<ul style="list-style-type: none"> <li>&gt; <i>Supervised Learning</i></li> <li>&gt; <i>Unsupervised Learning</i></li> <li>&gt; <i>Reinforcement Learning</i></li> </ul>	<ul style="list-style-type: none"> <li>&gt; <i>SVM, Lineare Regression, Logistische Regression, RF, KNN, Naive Bayes, Decision Tree, AdaBoost, Neuronale Netzwerke</i></li> <li>&gt; <i>Hauptkomponentenanalyse, K-Means, Expectation-Maximization, Mean Shift, Hierarchisches Clustern, Affinitätsausbreitung, Iterative Self-Organizing Data, Fuzzy C-Means-Systeme</i></li> <li>&gt; <i>Q-Learning, Temporal Difference Learning, State-Action-Reward-State-Action, Teaching-Box-Systeme, Maja-Systeme</i></li> </ul>
<b>Deep Learning</b>	<ul style="list-style-type: none"> <li>&gt; <i>DBN</i></li> <li>&gt; <i>CNN</i></li> <li>&gt; <i>RNN</i></li> </ul>	<ul style="list-style-type: none"> <li>&gt; <i>Convolutional Deep Belief Network, bedingt eingeschränkte Boltzmann-Maschine</i></li> <li>&gt; <i>AlexNet, GoogleNet, Visual Geometry Group Network (VGG), Deep Residual Learning, Inception v4 (v2, v3), Resnet-152 (34,50,101), LeNet</i></li> <li>&gt; <i>Bidirectional RNN, Long Short-Term Memory</i></li> </ul>

DBN: *Deep Belief Network*; CNN: *Convolution Neural Network*; RNN: *Recurrent Neural Network*; SVM: *Support Vector Machine*; RF: *Random Forest*; KNN: *K-Nearest Neighbor*

## 2.3 DATENSETS

Die medizinische Bildgebung spielt eine wichtige Rolle bei der Diagnose und Prognose von Augenerkrankungen. Die verfügbaren Techniken (u. a. optische Kohärenztomografie [OCT], Fundus-Fotografie und Spaltlampenuntersuchung) stellen hochaufgelöste Informationen hinsichtlich anatomischer und funktioneller Veränderungen bereit. Entscheidend für die Implementierung von ML/DL-Algorithmen für die automatische Auswertung der Bilder

ist die Verfügbarkeit großer Datenmengen. Eine wichtige Rolle spielen daher offen zugängliche Datenbanken, die beispielsweise für Fundus-Fotografien zu Trainingszwecken zur Verfügung stehen. Eine der größten derartigen Datenbanken ist Kaggle, die über 50.000 Bilder der Netzhaut umfasst. Weitere große Datenbanken sind z. B. EyePACS und MESSIDOR (Tabelle 2, Seite 5) [Tong et al. 2020]. Der Aufbau derartiger Datenbanken ist häufig aufgrund von Datenschutzbestimmungen, die sich von Land zu Land unterscheiden, limitiert.

**Tabelle 2:** Verbreitete öffentlich zugängliche Datenbanken für die Ophthalmologie; modifiziert nach [Tong et al. 2020].

Datensätze	Bildgebendes Verfahren	Population	Anzahl	Anmerkungen
<i>Kaggle</i>	FP	USA	53.576	DR
<i>EyePACS</i>	FP	USA	35.126	DR
<i>MESSIDOR</i>	FP	Frankreich	1200	DR, Makulaödem
<i>E-OPHTHA</i>	FP	Frankreich	463	DR
<i>HRF</i>	FP	Deutschland	45	DR, Glaukom, Sehnervenkopf, Gefäß
<i>DRIVE</i>	FP	Niederlande	40	DR, Gefäß
<i>RIGA</i>	FP	Frankreich; Saudi-Arabien	760	Glaukom
<i>ORIGA-650</i>	FP	Singapur	650	Glaukom
<i>DRISHTI-GS</i>	FP	Indien	101	Glaukom
<i>INSPIRE-AVR</i>	FP	USA	40	Glaukom
<i>REVIEW</i>	FP	Großbritannien	16	Gefäßkrankung

FP: Fundus-Fotografie; DR: diabetische Retinopathie

In der Ophthalmologie werden KI-basierte Systeme bislang in erster Linie zur Auswertung von Fundus-Fotografie, OCT und bei der Spaltlampen-Bildgebung angewandt. Hintergrund ist hier vor allem, dass in

diesen Bereichen größere Datensets zur Verfügung stehen, die es erlauben geeignete Algorithmen für die automatisierte Bildauswertung zu entwickeln.

## 3 KI IN DER OPHTHALMOLOGIE

Der Einsatz von KI bei medizinischen Fragestellungen, aber auch bei der Organisation von Patientendaten, besitzt ein hohes Potenzial. Hierdurch könnten zukünftig die Patientenversorgung und klinische Entscheidungsprozesse verbessert sowie die Ausführung von Routineaufgaben beschleunigt werden. In der Ophthalmologie werden entsprechende Anwendungen zum Teil schon erfolgreich eingesetzt. Schwerpunkt ist in erster Linie die automatische Erkennung klinischer Marker, welche der Diagnosestellung dienen und prognostische Aussagen erlauben [Tong et al. 2020]. Durch systematische Analysen großer Datenmengen lassen sich neben dem diagnostischen Nutzen auch neue wissenschaftliche Erkenntnisse gewinnen [Eter 2018]. Der Einsatz entsprechender Technologien konzentriert sich bislang auf häufige Erkrankungen wie die diabetische Retinopathie (DR), AMD, Glaukom, Frühgeborenen-Retinopathie, altersbedingte oder angeborene Katarakt sowie retinale

Verschlusskrankheiten [Du et al. 2018]. Diverse Studien zeigen eine vergleichbare Leistung oder zum Teil auch Überlegenheit von KI-basierten Systemen gegenüber klinischen Experten [Du et al. 2018, Lu et al. 2018].

### 3.1 FUNDUS-FOTOGRAFIE

Die Funduskopie ist eine in der Ophthalmologie häufig angewandte Methode zur Untersuchung der hinteren Augenabschnitte. Mit der Fundus-Fotografie werden hierbei digitale Bilder der Netzhaut aufgenommen, die der Diagnose, Therapieplanung und Verlaufskontrolle dienen. In den vergangenen Jahren wurden diverse auf DL-Algorithmen basierende Systeme für die automatische Analyse von Fundus-Fotografien entwickelt [Ahmad et al. 2020]. Im Fokus vieler Studien zur Anwendung dieser Systeme steht die DR, die eine häufige

Komplikation bei Diabetes darstellt und zu einem irreversiblen Visusverlust führen kann [Armstrong und Lorch 2020, Bellemo et al. 2019, Heydon et al. 2020, Olvera-Barrios et al. 2021]. Zentral für die Früherkennung und zeitnahe Behandlung einer DR ist ein regelmäßiges Screening der Patienten [Kashim et al. 2018, Vujosevic et al. 2020]. Spezifische Merkmale einer frühen DR wie Exudate, *Cotton-Wool*-Herde, Makulaödeme und Mikroaneurysmen können mit KI-basierten Systemen auf Grundlage von Fundus-Fotografien mittlerweile zuverlässig erkannt werden [Lu et al. 2018, Tong et al. 2020, Treder et al. 2020]. Hierbei bewegen sich die Werte für die erreichte Sensitivität und Spezifität im Bereich von etwa 75–100 % [Du et al. 2018]. Neben der DR-Erkennung lassen sich mit neueren Systemen auch verschiedene Krankheitsstadien der DR auf Basis von Fundus-Fotografien unterscheiden [Abbas et al. 2017, Takahashi et al. 2017].

Gulshan et al. stellten 2016 erstmals eine Applikation zur automatischen Erkennung einer DR vor. In Zusammenarbeit mit Google entwickelten sie einen DL-basierten Algorithmus zur DR-Diagnose auf Basis von Fundus-Fotografien, mit dem eine Sensitivität von bis zu 97,3 % und eine Spezifität von bis zu 93,9 % erreicht wurden (Tabelle 3) [Gulshan et al. 2016]. In einer weiteren größer angelegten multiethnischen Studie wurde eine Sensitivität von 90,5 % und eine Spezifität von 91,6 % für die Erkennung einer DR erreicht [Ting et al. 2017]. Auch in einer Analyse, welche die Eignung der beiden KI *EyeArt* und *RetMarker* für die *Real-World*-Anwendung untersuchte, wurden ähnlich hohe Werte erreicht [Tufail et al. 2017]. Der entscheidende Durchbruch hinsichtlich der *Real-World*-Anwendung von KI in der Ophthalmologie gelang 2018 mit der Zulassung des ersten autonomen Diagnosesystems für die DR – genannt *IDx-DR* – von der *U.S. Food and Drug Administration* (FDA) für das Screening von Diabetes-Patienten [Abràmoff et al. 2018]. Dies stellte einen wichtigen Schritt in Richtung Etablierung von KI in diesem Bereich dar.

Fundus-Fotografien eignen sich neben der Diagnose und Klassifikation einer DR auch für die Erkennung und Klassifikation weiterer häufiger Netzhauterkrankungen, darunter Glaukom und AMD [Tong et al. 2020]. Das Glaukom ist eine der häufigsten schweren Augenerkrankungen, die zu einem Sehverlust führen können [Höhn et al. 2018]. Kennzeichen sind ein erhöhter

Augeninnendruck, die Schädigung des Sehnervkopfs und Veränderungen der retinalen Nervenschicht. Ein wichtiges Erkennungsmerkmal bei der Diagnose ist das Verhältnis von Exkavation und Papille (*Cup-to-Disc Ratio*, CDR), das sich gut für die Identifizierung von Glaukom-Patienten eignet [Tangelder et al. 2006] und durch Segmentierung aus Fundus-Bildern extrahiert werden kann [Haleem et al. 2013]. Mit geeigneten KI-Modellen lässt sich daraus das CDR automatisch berechnen [Haleem et al. 2016, Liu et al. 2019, Singh et al. 2016]. Auf DL-Algorithmen basierende Methoden zur Glaukom-Diagnose sind bislang noch selten [Diener et al. 2021, Lu et al. 2018], doch auch hier werden zunehmend Anwendungen entwickelt [Chen et al. 2015, Li et al. 2016, Liu et al. 2019].

Die AMD ist eine sehr häufige Erkrankung bei älteren Menschen und kann im fortgeschrittenen Stadium zu einem irreversiblen Visusverlust führen. Auf KI-basierende Applikationen eignen sich, um das Screening auf eine AMD zu verbessern. Neben OCT-Aufnahmen (siehe Absatz 3.2) finden auch Fundus-Fotografien Verwendung. Hierbei können die aktuellen Systeme zwischen gesunden Augen und Augen mit AMD unterscheiden und auch eine Bestimmung des Schweregrads vornehmen. So erreichte ein von Grassmann et al. getesteter Algorithmus für die automatische Klassifikation des AMD-Schweregrads in 13 zuvor festgelegte Klassen eine (quadratisch gewichtete) Übereinstimmung  $\kappa$  von 92 % mit einer vom Menschen vorgenommenen Klassifizierung. Die Gesamtgenauigkeit lag bei 63,3 % [Grassmann et al. 2018]. Ein von Treder et al. entwickeltes System zur Klassifizierung einer geographischen Atrophie auf Basis von Fundus-Fotografien besaß eine Genauigkeit von 91 – 96 % [Tredet et al. 2018b]. Matsuba et al. erreichten mit ihrem Algorithmus für die Unterscheidung zwischen normaler und feuchter AMD einen Sensitivitätswert von 100 % sowie eine Spezifität von 97,31 % [Matsuba et al. 2019].

Neben der Erkennung von krankheitstypischen pathologischen Merkmalen sind KI-basierte Anwendungen auch in der Lage subtile, für das menschliche Auge nicht erkennbare Veränderungen zu detektieren. So kann anhand von Fundus-Bildern z. B. auf kardiovaskuläre Risikofaktoren wie Alter, systolischer Blutdruck, Hämoglobin A<sub>1c</sub> sowie das Geschlecht rückgeschlossen werden [Poplin et al. 2018].

**Tabelle 3:** Studien zur Anwendung von Künstliche-Intelligenz-Systemen zur Diagnose einer diabetischen Retinopathie basierend auf Fundus-Fotografien; modifiziert nach [Ahmad et al. 2020].

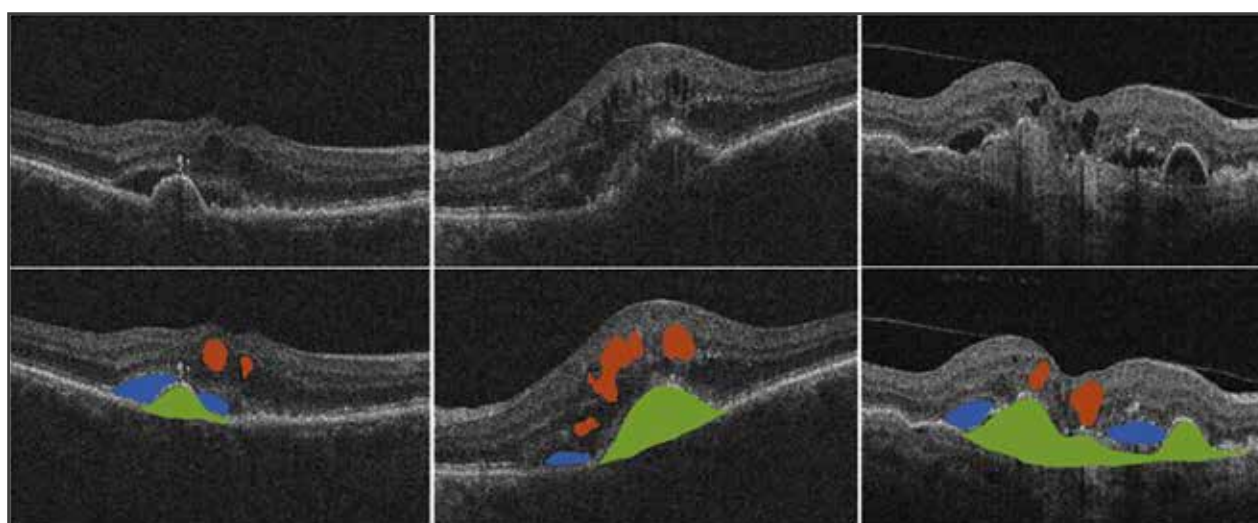
Publikation	Modell/Ziel	Für Training und Test/ Validierung verwendete Datensets	Ergebnis
Abràmoff et al. 2016	<i>Deep Learning, Supervised CNN</i> für überweisungsbedürftige DR	Fundus-Fotografien Training: n = 25.000 Validierung: n = 874	Sensitivität: 96,8 % Spezifität: 87,0 %
Abràmoff et al. 2018	<i>Deep Learning, Supervised CNN</i> für überweisungsbedürftige DR	Fundus-Fotografien Validierung: n = 900 Patientenbilder	Sensitivität: 87,2 % Spezifität: 90,7 %
Gulshan et al. 2016	<i>Deep Learning, Unsupervised CNN</i> für überweisungsbedürftige DR	Fundus-Fotografien Training: n = 128.175 Validierung: n = 9.963 (EyePACS)/n = 1.748 (MESSIDOR)	Sensitivität: 96,1 %/97,5 % Spezifität: 93,4 %/93,9 %
Gargeya und Leng 2017	<i>Deep Learning, Unsupervised CNN</i> für jegliche Retinopathie	Fundus-Fotografien Training: n = 75.137 Validierung: n = 15.000 (gemischte Quellen)	Sensitivität: 94 % Spezifität: 98 % AUC: 0,94 – 0,97
Ting et al. 2017	<i>Deep Learning, Unsupervised CNN</i> für Evaluation von überweisungsbedürftiger DR, möglichem Glaukom, AMD	Fundus-Fotografien Training: n = 76.370 (DR), n = 124.189 (mögliches Glaukom), n = 72.610 (AMD) Validierung: n = 71.896 Bilder von 14.880 Patienten	Sensitivität: 90,5 % (DR) 96,4 % (Glaukom) 93,2 % (AMD) Spezifität: 91,6 % (DR) 87,2 % (Glaukom) 88,7 % (AMD) AUC: 0,936 (DR) 0,942 (Glaukom) 0,931 (AMD)
Ramachandran et al. 2018	<i>Deep Learning</i> , Erkennung von überweisungsbedürftiger DR	Fundus-Fotografien Training: n > 100.000 Validierung: n = 485 (Otago)/n = 1.200 (MESSIDOR)	Sensitivität: 84,6 %/96 % Spezifität: 79,7 %/90,0 % AUC: 0,901/0,080

AUC: Area Under the Receiver Operating Characteristic Curve; AMD: altersbedingte Makuladegeneration;  
CNN: Convolutional Neural Networks; DR: diabetische Retinopathie.

**Tabelle 4:** Studien zur Anwendung von Künstliche-Intelligenz-Systemen zur Diagnose und Überwachung der altersbedingten Makuladegeneration basierend auf Fundus-Fotografien; modifiziert nach [Ahmad et al. 2020].

Publikation	Modell/Ziel	Für Training und Test/ Validierung verwendete Datensets	Ergebnis
Burlina et al. 2017	<i>Deep Learning, Unsupervised</i> CNN für Unterscheidung zwischen keiner/früher AMD oder mittelschwerer/fortgeschrittener AMD	Fundus-Fotografie Training und Validierung: n > 130.000 Bilder von 4.613 Patienten (AREDS-Studie)	AUC: 0,94 – 0,96
Grassmann et al. 2018	<i>Deep Learning</i> , Hybrid aus CNN und <i>Random Forest</i> , Übereinstimmung mit menschlicher Beurteilung der AMD-Klassifikation	Fundus-Fotografien (AREDS-Studie) Training: n = 86.770 Validierung: n = 21.867	Gewichtete Übereinstimmung ( $\kappa$ ): 92 %
Treder et al. 2018b	<i>Deep Learning, Supervised</i> CNN zur Detektion und Klassifikation einer geographischen Atrophie in Autofluoreszenzbildern	Autofluoreszenz-Fundus-Fotografien Training: n = 600 Validierung: n = 60	Genauigkeit: 91 – 96 %
Matsuba et al. 2019	<i>Deep Learning</i> , nichtspezifiziertes CNN zur Klassifizierung in trockene und feuchte AMD	Optos®-Fundus-Fotografien Training: n = 253 (trockene AMD 158, feuchte AMD 95) Validierung: n = 111 (trockene AMD 69, feuchte AMD 42)	Sensitivität: 100 % Spezifität: 97,31 % AUC: 99,76 %
Keel et al. 2019	<i>Deep Learning, Unsupervised</i> CNN zur Detektion einer neovaskulären AMD	Fundus-Fotografie Training: 27.397 Validierung: 3.850	Sensitivität: 96,7 % Spezifität: 96,4 % AUC: 0,995

AMD: altersbedingte Makuladegeneration; AUC: Area Under the Receiver Operating Characteristic Curve;  
CNN: Convolutional Neural Networks.



**Abbildung 4:** Repräsentatives Beispiel einer automatischen Bild-Segmentierung von OCT-Scans: Dargestellt sind die intraretinale Flüssigkeit (rot), die subretinale Flüssigkeit (blau) und die Ablösung des Pigmentepithels (grün); [Schmidt-Erfurth et al. 2020b].  
OCT: Optische Kohärenztomographie



### 3.2 OPTISCHE KOHÄRENZ-TOMOGRAFIE

Die OCT ist eine häufig angewandte Technik im Rahmen der ophthalmologischen Diagnostik und Verlaufskontrolle. Mittels OCT lassen sich umfangreiche Informationen zur Netzhautmorphologie generieren [Huang et al. 1991]. Für die automatisierte Auswertung von OCT-Aufnahmen existieren inzwischen zahlreiche Algorithmen, deren Aufgaben sich in die Bereiche Klassifizierung und Segmentierung unterteilen lassen [Tong et al. 2020]. Während die Klassifizierung die Zuordnung der Bilder in bestimmte Gruppen erlaubt, können bei der Segmentierung Strukturen und Läsionen aus den OCT-Scans extrahiert und unterschieden sowie abweichende Bereiche oder Volumina identifiziert werden [Tong et al. 2020].

Die OCT gewinnt bei der Krankheitserkennung, Prognose und Überwachung der AMD zunehmend an Bedeutung, insbesondere bei der feuchten AMD und deren Behandlung mit VEGF-(*Vascular-Endothelial-Growth-Factor*-)Inhibitoren [Moraes et al. 2020, Tong et al. 2020, Treder und Eter 2019]. Grund ist, dass sich mit der OCT wichtige Krankheitsmerkmale der AMD, wie z. B. intraretinale Flüssigkeit (IRF) und subretinale Flüssigkeit (SRF), Drusen, Vernarbungen oder Atrophien, abbilden lassen, die über das Krankheitsstadium der AMD Auskunft geben [Kanagasingam et al. 2014]. Anhand von *Spectral-Domain*-OCT-Bildern lässt sich eine AMD bereits mit hoher Sicherheit mittels KI erkennen, wobei die Treffsicherheit 98 % beträgt [Treder et al. 2018a].

DL-Algorithmen ermöglichen außerdem eine automatisierte Flüssigkeitsanalyse in der Retina. So gelang es Schmidt-Erfurth et al., Volumina von IRF und SRF sowie eine Ablösung des Pigmentepithels (PED) anhand von 3D-OCT-Scans zu bestimmen (Abbildung 4) [Schmidt-Erfurth et al. 2020b]. Die OCT-Scans stammten hierbei von Patienten der Phase-III-Studie HARBOR, die eine Behandlung mit dem VEGF-Inhibitor Ranibizumab erhielten. Mithilfe der Methode konnte die Veränderung der einzelnen Parameter unter Anti-VEGF-Behandlung quantifiziert und außerdem eine negative Korrelation zwischen dem Verbleib von IRF und dem Visus sowie ein schwach prognostischer Effekt der SRF festgestellt werden. Ziel der KI-Anwendung ist die unabhängige

und reproduzierbare quantitative Bewertung der Flüssigkeitsantwort auf eine Anti-VEGF-Therapie, die eine Überwachung des Therapieerfolgs erlaubt [Schmidt-Erfurth et al. 2022]. Die Eignung der Methode konnte mit Patientendaten aus den Studien HAWK und HARRIER bestätigt werden. In beiden Studien wurde die Wirksamkeit der beiden VEGF-Inhibitoren Aflibercept und Brolucizumab verglichen. Bei der Analyse der OCT-Scans zeigte sich, dass die IRF-Volumina bei beiden Wirkstoffen infolge der ersten intravitrealen Injektion rasch um > 93 % abnahmen, bei den SRF-Volumina waren es > 75 %. Während der *Follow-Up*-Periode war die Reduktion der SRF unter Brolucizumab signifikant höher als unter Aflibercept [Schmidt-Erfurth et al. 2020a]. Beide Studien zeigen, dass mittels der Anwendung von KI wichtige Erkenntnisse aus den vorhandenen Bilddaten extrahiert werden können, die die Möglichkeiten einer manuellen Auswertung durch einen Facharzt übersteigen. Die erfolgreiche, für die klinische Anwendung wesentliche Analyse wurde inzwischen ebenfalls gezeigt [Schmidt-Erfurth et al. 2020b].

Eine weitere zentrale Fragestellung in Zusammenhang mit einer Anti-VEGF-Therapie ist die optimale Frequenz der Injektionen für die Patienten, die sich individuell unterscheiden kann. Bogunovic et al. adressierten diese Frage in einer Post-hoc-Analyse der HARBOR-Studie [Bogunovic et al. 2017b]. Eingesetzt wurde hierbei eine auf ML basierende Methode (*Random Forest*) unter Einbeziehung von quantitativen spatio-temporalen Merkmalen. Hierbei wurde eine Segmentierung in innere Retina (IR), äußere Körnerschicht (ONL), Außensegmente der Photorezeptoren plus retinales Pigmentepithelium (OR) und Gesamtreinadicke (TRT) vorgenommen sowie Bereiche mit Flüssigkeit berücksichtigt. Die Analyse der OCT-Bilder erfolgte zu Studienbeginn, nach einem und nach zwei Monaten, mit dem Ziel den Injektions-Bedarf in den folgenden Monaten vorauszusagen. In diesem Zeitraum hatten alle Studienteilnehmer Injektionen mit Ranibizumab (0,5 mg bzw. 2,0 mg) im monatlichen Abstand erhalten. Verglichen wurden die Ergebnisse des Modells dabei mit der tatsächlichen Häufigkeit der Injektionen nach den ersten drei Behandlungsmonaten, die anhand von den Studienleitern festgelegten Kriterien bestimmt wurden. Die Vorhersage des Injektionsbedarfs während der VEGF-Inhibitor-Therapie war auch das Ziel einer weiteren Studie, die hierfür ein *Deep Convolutional*

Neural Network (DCNN) verwendete [Prahs et al. 2018]. Die Genauigkeit des anhand von 153.912 OCT-Scans trainierten Systems für die Vorhersage des Injektionsbedarfs betrug im Validierungs-Datenset 95,5 %. Für einzelne retinale B-Scans im Validierungs-Datenset konnte eine Sensitivität von 90,1 % und eine Spezifität von 96,2 % sowie eine AUC von 0,968 erreicht werden.

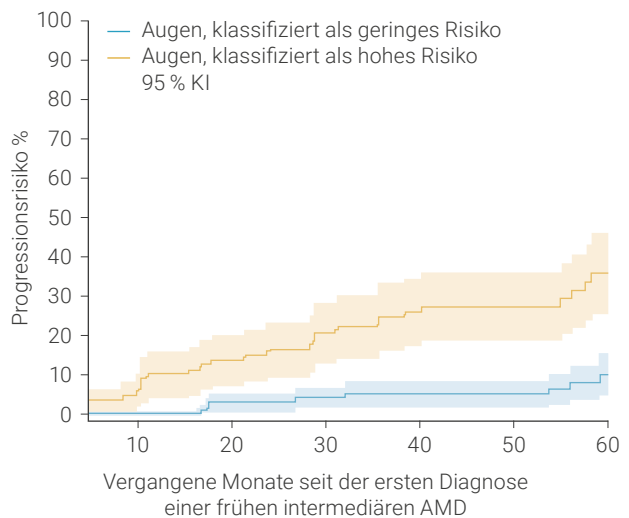
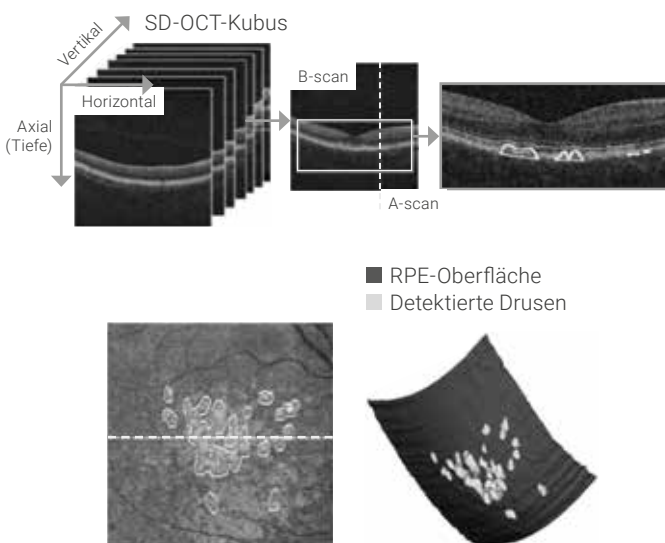
Auf Basis von OCT-Scans können noch weitere Aussagen getroffen werden. Dazu zählen die Auskunft über den Ist-Zustand und die wahrscheinliche zukünftige Entwicklung der Krankheitsaktivität bei AMD. So ließ sich auf Basis von 1.210 OCT-Scans die Sehschärfe von AMD-Patienten anhand diverser OCT-Merkmale wie SRF und SHRM (*Subretinal Hyperreflective Material*) abschätzen [Aslam et al. 2018]. Darüber hinaus können aus OCT-Aufnahmen potenzielle Biomarker extrahiert werden. So ließ sich die Progression einer AMD anhand der in der Retina vorhandenen Drusen mit recht hoher Genauigkeit (AUC 0,75) abschätzen [Bogunovic et al. 2017a]. Gleiches gilt für die Wahrscheinlichkeit eines Übergangs in eine fortgeschrittene AMD (Abbildung 5) [Schmidt-Erfurth et al. 2018b]. Entscheidende quantitative Merkmale in diesem Zusammenhang sind die äußere Netzhautdicke, hyperreflektive Bereiche und Bereiche, die Drusen aufweisen.

Neben der Bedeutung im Bereich der AMD-Diagnostik und -Therapie besitzt die KI-basierte OCT-Analyse auch

großes Potential bei weiteren Augenerkrankungen. So können beispielsweise die automatisierte Auswertung quantitativer Messungen der retinalen Nervenfaserdicke und die Bestimmung des CDR für die Früherkennung eines Glaukoms genutzt werden [Barella et al. 2013, Bizios et al. 2010, Wu et al. 2015]. Weiterhin findet sie Anwendung bei der Verlaufskontrolle nach Hornhauttransplantationen (*Descemet Membrane Endothelial Keratoplasty*) [Treder et al. 2019].

### 3.3 SPALTLAMPENUNTERSUCHUNG

Die Spaltlampenuntersuchung ermöglicht die Untersuchung von sowohl anterioren als auch posterioren Augensegmenten und dient der allgemeinen Untersuchung des Auges und dessen Adnexe [Tong et al. 2020]. Ähnlich wie Fundus-Fotografien und OCT-Aufnahmen eignen sich auch Spaltlampenbilder für eine automatisierte Analyse. Huang et al. gelang es anhand eines 1.000 Bilder umfassenden Datensets einen Algorithmus zu entwickeln, mit dem sich mittels einer Ranking-Methode der Schweregrad von Katarakten bestimmen ließ [Huang et al. 2011]. Eine weitere Arbeitsgruppe nutzte Spaltlampenbilder, um typische Merkmale einer Katarakt zu extrahieren, und trainierte ein SVM- (*Support-Vector-Machine*-) Regressionsmodell für die automatische Klassifizierung von Katarakten nach Schweregrad [Li et al. 2010].



**Abbildung 5:** Mithilfe künstlicher Intelligenz lässt sich anhand quantitativer Drusenmerkmale das Risiko für den Übergang zu einer fortgeschrittenen AMD (in Form des *Hazard Ratio*) mit hoher Genauigkeit voraussagen; modifiziert nach [Schmidt-Erfurth et al. 2018a].

Besonders hilfreich ist die Spaltlampenuntersuchung bei der Diagnose angeborener Katarakte, die einen komplexeren Phänotyp aufweisen als senile Katarakte und einen der Hauptgründe für Blindheit im Kindesalter darstellen [Tong et al. 2020]. Hier kann mit DL-Methoden eine Schweregradeinteilung und Klassifizierung anhand

der Spaltlampenbilder vorgenommen werden [Liu et al. 2017]. Das System *CC-Cruiser* (Prototyp), das auf einem DCNN basiert, wurde speziell für das Screening von pädiatrischen Katarakten entwickelt und weist eine mit ophthalmologischen Experten vergleichbare Leistung auf [Long et al. 2017].

**Tabelle 5:** Studien zur Anwendung von Künstliche-Intelligenz-Systemen zur Diagnose und Überwachung der altersbedingten Makuladegeneration basierend auf Aufnahmen mittels optischer Kohärenztomographie; modifiziert nach [Ahmad et al. 2020].

Publikation	Modell/Ziel	Für Training und Test/ Validierung verwendete Datensets	Ergebnis
Bogunovic et al. 2017b	<i>Random Forest Machine Learning</i> , Verwendung von quantitativen OCT-Scan-Merkmalen, um den Injektionsbedarf von Anti-VEGF-Inhibitoren innerhalb von 20 Monaten in die Klassen hoch (> 15), mittel (6 – 15) und niedrig (< 6) einzuordnen	OCT-Bilder von 317 Patienten	AUC: 0,7 – 0,77
Prahs et al. 2018	<i>Deep Learning, Unsupervised CNN</i> , Vergleich von vorausgesagtem Injektionsbedarf mit tatsächlicher Injektionshäufigkeit (bestimmt durch den Arzt) im Zeitraum von 21 Tagen	OCT-Linien-Scans Training: n = 153.912 Validierung: n = 5.358	Sensitivität: 90,1 % Spezifität: 96,2 % AUC: 0,968
Aslam et al. 2018	<i>Deep Learning, Supervised NN</i> , Vergleich der vom System (auf Basis von OCT-Merkmalen wie SRF, SHRM) vorausgesagten mit der tatsächlich gemessenen Sehschärfe	OCT-Bilder von 1.210 Patienten	Mittlerer Regressionskoeffizient: 0,85 Mittlerer quadratischer Fehler: 8,2 Buchstaben
Schmidt-Erfurth et al. 2018b	<i>Machine Learning</i> Voraussagemodell für den Übergang zu einer fortgeschrittenen AMD in einem Zeitraum von 2 Jahren	OCT-Bilder von 495 Augen	AUC: 0,68 (CNV) 0,80 (GA)
Bogunovic et al. 2017a	<i>Machine Learning</i> , Voraussagemodell für Rückbildung von Drusen innerhalb von durchschnittlich 38 Monaten	OCT-Scans von 61 Augen	AUC: 0,75
Treder et al. 2018a	<i>Deep Learning</i> , automatische Detektion einer AMD	OCT-Scans Training: n = 1,2 Millionen (ImageNet) Validierung: n = 1.012 SD-OCT-Scans (AMD: 701, gesund: 311)	Score von 0,98 für Erkennung einer AMD

AMD: altersbedingte Makuladegeneration; AUC: *Area Under the Receiver Operating Characteristic Curve*; CNN: *Convolutional Neural Networks*; CNV: choroidale Neovaskularisation; GA: geographische Atrophie; NN: *Neural Network*; OCT: optische Kohärenztomographie; SD: *Spectral Domain*; SHRM: *Subretinal Hyperreflective Material*; SRF: *Subretinal Fluid*; VEGF: *Vascular Endothelial Growth Factor*

### 3.4 HERAUSFORDERUNGEN UND AUSBLICK

Die oben beschriebenen Studien unterstreichen das große Potential von KI-basierten Systemen für die Verbesserung von Diagnostik und Therapiesteuerung in der Ophthalmologie. Die automatisierte Auswertung von Bilddaten mittels Computersoftware ermöglicht Geschwindigkeiten, die eine manuelle Auswertung bei Weitem übersteigen, und erlaubt es Zusatzinformationen aus den Bilddaten zu extrahieren, die für das menschliche Auge nicht erkennbar sind. Durch die Möglichkeit der Mustererkennung mittels automatisierter Verfahren können neuartige Krankheiten entdeckt und Biomarker für bereits bekannte Erkrankungen identifiziert werden.

Neben dem wissenschaftlichen Erkenntnisgewinn erleichtert die Anwendung von KI auch die Patientenversorgung. So kann das medizinische Personal durch die Automatisierung zeitaufwendiger Aufgaben entlastet werden, wodurch mehr Zeit für die Patienten zur Verfügung steht. Darüber hinaus ermöglichen KI-Anwendungen im Rahmen der Telemedizin auch die Diagnose und Behandlung von Patienten, die in großer Entfernung zu medizinischen Einrichtungen wohnen und keinen unmittelbaren Zugang zu Spezialisten haben [Korot et al. 2019]. Die meisten dieser Anwendungen konzentrieren sich gegenwärtig auf den Bereich des DR-Screenings. Neben der Zulassung des IDx-DR gibt es weitere erfolgreiche Modellprojekte [Keel et al. 2018, Walton et al. 2016], darunter auch eine Anwendung, die es den Patienten ermöglicht, mittels einer Smartphone-App zu Hause selbstständig Fundus-Fotografien ihrer Augen aufzunehmen und mit dem Arzt zu teilen [Natarajan et al. 2019]. Auch ein OCT-Gerät für die Anwendung zu Hause (von der Firma Notal Vision) befindet sich in der Testung und wird voraussichtlich bald verfügbar sein [Armstrong und Lorch 2020]. Es basiert auf einem von Chakravarthy et al. entwickelten ML-Algorithmus [Chakravarthy et al. 2016] und soll die Detektion von Flüssigkeitsansammlungen in der Netzhaut ermöglichen. Gedacht ist es für das Monitoring des Krankheitsverlaufs bei Hochrisiko-Patienten.

Bislang sind die Einsatzmöglichkeiten der KI-Systeme jedoch noch begrenzt. Noch gibt es ungelöste Probleme, die einer breiten Anwendung entgegenstehen.

Hierzu gehört z. B. die fehlende Standardisierung der Aufnahmen. Die Variabilität der Bildqualität, bedingt durch Unterschiede bei den eingesetzten Gerätetypen, verschiedene Aufnahmewinkel oder auftretende Bildreflexionen bzw. -Artefakte, erschwert das Training der KI und macht die automatische Detektion von Merkmalen fehleranfällig. Hinzu kommt, dass die Daten für die bisher entwickelten Trainingsmodelle auf unterschiedliche Weise vorbearbeitet sind, wodurch es zu Verzerrungen kommt. Für das Training der Systeme sind zudem große Datensätze erforderlich, die z. B. bei seltenen Erkrankungen und fortgeschrittenen Krankheitsstadien nicht zur Verfügung stehen. Aus diesem Grund können für die Erkennung dieser Erkrankungen meist (noch) keine entsprechenden Programme entwickelt werden.

Grundsätzlich besteht auch die Gefahr des blinden Vertrauens in die automatisch generierten Ergebnisse. Eine große Schwierigkeit ist hierbei der fehlende Einblick in die Datenverarbeitung. Wie die KI zu dem erhaltenen Ergebnis kommt, ist für das medizinische Fachpersonal bei den selbstlernenden Programmen in der Regel nicht mehr eindeutig nachvollziehbar. Hinzu kommt die Frage der Verantwortung und Haftung für eine auf Basis der Daten verordnete Behandlung. Auch wenn die KI-Systeme bereits eine hohe Leistungsfähigkeit und Präzision besitzen, bleibt die ärztliche Expertise weiterhin unverzichtbar. Daher sollten stets zusätzliche Untersuchungen für die klinische Beurteilung herangezogen werden. Dies stärkt auch das Vertrauen der Patienten in die durchgeführten Untersuchungen und fördert damit die Akzeptanz des Einsatzes von KI auf Patientenseite.

Um die Qualität von Studien zur Entwicklung und Anwendung von KI-Systemen besser bewerten zu können, sollte sich das Fachpersonal ein solides Basiswissen des technischen Hintergrunds aneignen. Ein wichtiges Merkmal guter Studien ist eine klare Nachvollziehbarkeit und Transparenz der Methodik sowie der Ergebnisse [Faes et al. 2020].

Nicht zuletzt spielen auch Fragen zum Datenschutz eine wichtige Rolle. Dies ist im Hinblick auf die Etablierung der elektronischen Patientenakte relevant, aber auch für die Verwendung von Patientenbildern in groß angelegten Bilddatenbanken.

## 4 FAZIT

Mit zunehmender Präzision und Leistungsfähigkeit KI-basierter Systeme finden diese immer häufiger Verwendung im Bereich der Diagnose und Therapieüberwachung in der Ophthalmologie. Durch das *Machine Learning*, und insbesondere das *Deep Learning*, lassen sich pathologische Merkmale automatisch erkennen und klassifizieren. Die automatisierte Auswertung von Bilddaten ermöglicht so die Früherkennung häufiger Augenerkrankungen wie z. B. der DR, der AMD oder dem Glaukom. Auch zur Therapieüberwachung und -steuerung werden die Anwendungen in Zukunft voraussichtlich vermehrt eingesetzt werden. So lässt sich beispielsweise die Wirksamkeit einer VEGF-Therapie bei der AMD anhand von OCT-Aufnahmen verfolgen,

wobei auch kleine Veränderungen mit hoher Präzision erkannt werden können. Die Leistungsfähigkeit vieler KI-basierter Systeme erreicht inzwischen die von medizinischem Fachpersonal oder kann diese sogar übersteigen. Dies betrifft einerseits die Geschwindigkeit der Auswertungen, da die Systeme große Datenmengen bewältigen können, andererseits auch die Erkennung von Bildmerkmalen, die für das menschliche Auge nicht sichtbar sind. Es ist daher zu erwarten, dass entsprechende Methoden zukünftig immer mehr Einzug in die ophthalmologische Diagnostik halten. Der Einsatz KI-basierter Methoden bildet hierbei jedoch keinen Ersatz der ärztlichen Expertise, sondern sollte vielmehr als eine hilfreiche Ergänzung betrachtet werden.

## 5 LITERATUR

- Abbas** Q, Fondon I, Sarmiento A, et al. Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features. *Med Biol Eng Comput* 2017;55(11):1959 – 74
- Abràmoff** MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018;1:39
- Abràmoff** MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 2016;57(13):5200 – 6
- Ahmad** BU, Kim JE, Rahimy E. Fundamentals of artificial intelligence for ophthalmologists. *Curr Opin Ophthalmol* 2020;31(5):303 – 11
- Armstrong** GW, Lorch AC. A(eye): a review of current applications of artificial intelligence and machine learning in ophthalmology. *Int Ophthalmol Clin* 2020;60(1):57 – 71
- Aslam** TM, Zaki HR, Mahmood S, et al. Use of a neural net to model the impact of optical coherence tomography abnormalities on vision in age-related macular degeneration. *Am J Ophthalmol* 2018;185:94 – 100
- Barella** KA, Costa VP, Gonçalves Vidotti V, et al. Glaucoma diagnostic accuracy of machine learning classifiers using retinal nerve fiber layer and optic nerve data from SD-OCT. *J Ophthalmol* 2013;2013:789129
- Bellemo** V, Lim G, Rim TH, et al. Artificial intelligence screening for diabetic retinopathy: the real-world emerging application. *Curr Diab Rep* 2019;19(9):72
- Bizios** D, Heijl A, Hougaard JL, et al. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT. *Acta Ophthalmol* 2010;88(1):44 – 52
- Bogunovic** H, Montuoro A, Baratsits M, et al. Machine learning of the progression of intermediate age-related macular degeneration based on OCT imaging. *Invest Ophthalmol Vis Sci* 2017a;58(6):Bio141 – bio50
- Bogunovic** H, Waldstein SM, Schlegl T, et al. Prediction of anti-VEGF treatment requirements in neovascular AMD using a machine learning approach. *Invest Ophthalmol Vis Sci* 2017b;58(7):3240 – 8
- Burlina** PM, Joshi N, Pekala M, et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol* 2017;135(11):1170 – 6
- Chakravarthy** U, Goldenberg D, Young G, et al. Automated identification of lesion activity in neovascular age-related macular degeneration. *Ophthalmology* 2016;123(8):1731 – 6
- Chen** X, Xu Y, Wong DWK, et al. Glaucoma detection based on deep convolutional neural network. *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:715 – 8
- Cortes** C, Vapnik V. Support-vector networks. *Machine Learning* 1995;20(3):273 – 97
- Cover** TM, Hart PE. Nearest neighbor pattern classification. *IEEE Trans Inf Theory* 1967;13:21 – 7
- Cutler** A, Cutler DR, Stevens JR. Random Forests. 2012; In: Zhang C and Ma YQ, Eds., *Ensemble Machine Learning*, Springer, New York, 157 – 175
- Diener** R, Treder M, Eter N. Diagnostik von Erkrankungen des Sehnervenkopfes in Zeiten von künstlicher Intelligenz und Big Data. *Ophthalmologie* 2021; 10.1007/s00347-021-01385-6:1 – 6
- Du** XL, Li WB, Hu BJ. Application of artificial intelligence in ophthalmology. *Int J Ophthalmol* 2018;11(9):1555 – 61
- Eter** N. Deep Learning in der Augenheilkunde. *Ophthalmologie* 2018;115(9):712 – 3
- Faes** L, Liu X, Wagner SK, et al. A clinician's guide to artificial intelligence: how to critically appraise machine learning studies. *Transl Vis Sci Technol* 2020;9(2):7
- Gargeya** R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 2017;124(7):962 – 9

- Grassmann F**, Mengelkamp J, Brandl C, et al. A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology* 2018;125(9):1410 – 20
- 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 – 10
- Haleem MS**, Han L, Hemert J, et al. Regional image features model for automatic classification between normal and glaucoma in fundus and scanning laser ophthalmoscopy (SLO) images. *J Med Syst* 2016;40(6):132
- Haleem MS**, Han L, van Hemert J, et al. Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: a review. *Comput Med Imaging Graph* 2013;37(7-8):581 – 96
- Heydon P**, Egan C, Bolter L, et al. Prospective evaluation of an artificial intelligence-enabled algorithm for automated diabetic retinopathy screening of 30 000 patients. *Br J Ophthalmol* 2020; 10.1136/bjophthalmol-2020-316594
- Höhn R**, Nickels S, Schuster AK, et al. Prevalence of glaucoma in Deutschland: results from the Gutenberg Health Study. *Graefes Arch Clin Exp Ophthalmol* 2018;256(9):1695 – 702
- Huang D**, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science* 1991;254(5035):1178 – 81
- Huang W**, Chan KL, Li H, et al. A computer assisted method for nuclear cataract grading from slit-lamp images using ranking. *IEEE Trans Med Imaging* 2011;30(1):94 – 107
- Kanagasingam Y**, Bhuiyan A, Abramoff MD, et al. Progress on retinal image analysis for age related macular degeneration. *Prog Retin Eye Res* 2014;38:20 – 42
- Kashim RM**, Newton P, Ojo O. Diabetic retinopathy screening: a systematic review on patients' non-attendance. *Int J Environ Res Public Health* 2018;15(1)
- Keel S**, Lee PY, Scheetz J, et al. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. *Scientific Reports* 2018;8(1):4330
- Korot E**, Wood E, Weiner A, et al. A renaissance of teleophthalmology through artificial intelligence. *Eye (Lond)* 2019;33(6):861 – 3
- Li A**, Cheng J, Wong DWK, et al. Integrating holistic and local deep features for glaucoma classification. *Annu Int Conf IEEE Eng Med Biol Soc* 2016;2016:1328 – 31
- Li H**, Lim JH, Liu J, et al. A computer-aided diagnosis system of nuclear cataract. *IEEE Trans Biomed Eng* 2010;57(7):1690 – 8
- Liu H**, Li L, Wormstone IM, et al. Development and validation of a deep learning system to detect glaucomatous optic neuropathy using fundus photographs. *JAMA Ophthalmol* 2019;137(12):1353 – 60
- Liu X**, Jiang J, Zhang K, et al. Localization and diagnosis framework for pediatric cataracts based on slit-lamp images using deep features of a convolutional neural network. *PLoS One* 2017;12(3):e0168606
- Long E**, Lin H, Liu Z, et al. An artificial intelligence platform for the multi-hospital collaborative management of congenital cataracts. *Nature Biomedical Engineering* 2017;1(2):0024
- Lowd D**, Domingos P. Naive Bayes models for probability estimation. *Proceedings of the 22nd International Conference On Machine Learning* 2005;529 – 536
- Lu W**, Tong Y, Yu Y, et al. Applications of artificial intelligence in ophthalmology: general overview. *J Ophthalmol* 2018;2018:5278196
- Matsuba S**, Tabuchi H, Ohsugi H, et al. Accuracy of ultra-wide-field fundus ophthalmoscopy-assisted deep learning, a machine-learning technology, for detecting age-related macular degeneration. *Int Ophthalmol* 2019;39(6):1269 – 75
- Moraes G**, Fu DJ, Wilson M, et al. Quantitative analysis of OCT for neovascular age-related macular degeneration using deep learning. *Ophthalmology* 2020; 10.1016/j.ophtha.2020.09.025
- Moraru AD**, Costin D, Moraru RL, et al. Artificial intelligence and deep learning in ophthalmology - present and future (Review). *Exp Ther Med* 2020;20(4):3469 – 73
- Natarajan S**, Jain A, Krishnan R, et al. Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA Ophthalmol* 2019;137(10):1182 – 8
- Olivera-Barrios A**, Heeren TF, Balaskas K, et al. Diagnostic accuracy of diabetic retinopathy grading by an artificial intelligence-enabled algorithm compared with a human standard for wide-field true-colour confocal scanning and standard digital retinal images. *Br J Ophthalmol* 2021;105(2):265 – 70
- Patel VL**, Shortliffe EH, Stefanelli M, et al. The coming of age of artificial intelligence in medicine. *Artif Intell Med* 2009;46(1):5 – 17
- Poplin R**, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* 2018;2(3):158 – 64
- Prahs P**, Radeck V, Mayer C, et al. OCT-based deep learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications. *Graefes Arch Clin Exp Ophthalmol* 2018;256(1):91 – 8
- Ramachandran N**, Hong SC, Sime MJ, et al. Diabetic retinopathy screening using deep neural network. *Clin Exp Ophthalmol* 2018;46(4):412 – 6
- Rokach L**, Maimon O. Data mining with decision trees: theory and applications. Second edition. World scientific. 2014
- Schmidt-Erfurth U**, Mulyukov Z, Gerendas B, et al. A comparison of the therapeutic response between brolicizumab and aflibercept in the HAWK & HARRIER trials using deep learning-based OCT analysis. *Invest Ophthalmol Vis Sci* 2020a;61(7):1159
- Schmidt-Erfurth U**, Reiter GS, Riedl S, et al. AI-based monitoring of retinal fluid in disease activity and under therapy. *Prog Retin Eye Res* 2022;86:100972
- Schmidt-Erfurth U**, Sadeghipour A, Gerendas BS, et al. Artificial intelligence in retina. *Prog Retin Eye Res* 2018a;67:1 – 29
- Schmidt-Erfurth U**, Vogl WD, Jampol LM, et al. Application of automated quantification of fluid volumes to anti-VEGF therapy of neovascular age-related macular degeneration. *Ophthalmology* 2020b;127(9):1211 – 9
- Schmidt-Erfurth U**, Waldstein SM, Klimescha S, et al. Prediction of individual disease conversion in early AMD using artificial intelligence. *Invest Ophthalmol Vis Sci* 2018b;59(8):3199 – 208
- Singh A**, Dutta MK, ParthaSarathi M, et al. Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image. *Comput Methods Programs Biomed* 2016;124:108 – 20
- Takahashi H**, Tampo H, Arai Y, et al. Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. *PLoS One* 2017;12(6):e0179790
- Tan TE**, Xu X, Wang Z, et al. Interpretation of artificial intelligence studies for the ophthalmologist. *Curr Opin Ophthalmol* 2020;31(5):351 – 6
- Tangelder GJ**, Reus NJ, Lemij HG. Estimating the clinical usefulness of optic disc biometry for detecting glaucomatous change over time. *Eye (Lond)* 2006;20(7):755 – 63
- Ting DSW**, Cheung CY, Lim G, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* 2017;318(22):2211 – 23
- Tong Y**, Lu W, Yu Y, et al. Application of machine learning in ophthalmic imaging modalities. *Eye Vis (Lond)* 2020;7:22

- Treder M**, Diener R, Eter N. Künstliche Intelligenz zum Management von Makulaödemen: Chancen und Herausforderungen. *Ophthalmologie* 2020;117(10):989 – 92
- Treder M**, Eter N. Chancen von künstlicher Intelligenz und Big Data für die Diagnostik und Behandlung der altersabhängigen Makuladegeneration. *Klin Monbl Augenheilkd* 2019;236(12):1418 – 22
- Treder M**, Lauer mann JL, Alnawaiseh M, et al. Using deep learning in automated detection of graft detachment in descemet membrane endothelial keratoplasty: a pilot study. *Cornea* 2019;38(2):157 – 61
- Treder M**, Lauer mann JL, Eter N. Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning. *Graefes Arch Clin Exp Ophthalmol* 2018a;256(2):259 – 65
- Treder M**, Lauer mann JL, Eter N. Deep learning-based detection and classification of geographic atrophy using a deep convolutional neural network classifier. *Graefes Arch Clin Exp Ophthalmol* 2018b;256(11):2053 – 60
- Tufail A**, Rudisill C, Egan C, et al. Automated diabetic retinopathy image assessment software: diagnostic accuracy and cost-effectiveness compared with human graders. *Ophthalmology* 2017;124(3):343 – 51
- Vujosevic S**, Aldington SJ, Silva P, et al. Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol* 2020;8(4):337 – 47
- Walton OB**, Garoon RB, Weng CY, et al. Evaluation of automated tele-retinal screening program for diabetic retinopathy. *JAMA Ophthalmol* 2016;134(2):204 – 9
- Wu M**, Leng T, de Sisternes L, et al. Automated segmentation of optic disc in SD-OCT images and cup-to-disc ratios quantification by patch searching-based neural canal opening detection. *Opt Express* 2015;23(24):31216 – 29

**Bildnachweise:**

**Abbildung 4:** Reprinted from *Ophthalmology: Journal of the American Academy of Ophthalmology*, 127 (9), Schmidt-Erfurth U, Vogl WD, Jampol LM, Bogunović H, Application of Automated Quantification of Fluid Volumes to Anti-VEGF Therapy of Neovascular Age-Related Macular Degeneration, Pages 1211 – 1219, © (2020), with permission from Elsevier.

**Abbildung 5:** Reprinted from *Progress in Retinal and Eye Research*, Volume 67; Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunović H. Artificial intelligence in retina, Pages 1 – 29, © (2018), with permission from Elsevier.

**Icons:** © Flaticon/Freepik/Smashicons

## LERNKONTROLLFRAGEN

Bitte kreuzen Sie jeweils nur **eine** Antwort an.

### 1. Welche der folgenden Aussagen in Bezug auf künstliche Intelligenz (KI) trifft zu?

- a) Im medizinischen Bereich wird das *Machine Learning* seit den 90er Jahren angewandt.
- b) *Machine Learning* ist eine Unterkategorie des *Deep Learning*.
- c) Das *Supervised Learning* und das *Unsupervised Learning* sind Unterkategorien des *Machine Learning*.
- d) Beim *Supervised Learning* wird die Maschine mit unbearbeiteten Daten trainiert.
- e) Das *Unsupervised Learning* erfordert eine zeitaufwendige Vorbereitung der Daten.

### 2. Welcher Algorithmus findet keine Anwendung beim *Machine Learning*?

- a) Entscheidungsbäume (*Decision Trees*)
- b) Naiver Bayes-Algorithmus
- c) *Random-Forest*-Verfahren
- d) Bayes-Marshall-Algorithmus
- e) *K-Nearest-Neighbor*(KNN)-Algorithmus

### 3. Welche der folgenden Aussagen zu Instanzen ist falsch?

- a) Die beim *Machine Learning* angewandten Datensets umfassen eine Sammlung von Instanzen.
- b) Instanzen stehen repräsentativ für einzelne Objekte (z. B. Patienten).
- c) Instanzen beschreiben ein Objekt näher (z. B. Alter, Geschlecht).
- d) Instanzen werden durch Variablen charakterisiert.
- e) Instanzen werden in Klassen zusammengefasst.

### 4. Welche der folgenden Aussagen zu neuronalen Netzwerken ist richtig?

- a) Bei den dem menschlichen Gehirn nachempfundenen neuronalen Netzwerken sind alle Verbindungen zwischen den „Neuronen“ gleich stark gewichtet.
- b) Während des Trainingsvorgangs eines neuronalen Netzwerks muss manuell geprüft werden, ob die getroffenen Voraussagen korrekt sind oder nicht.
- c) Die Qualität des Algorithmus wird anhand des Trainingsdatensets überprüft.
- d) Um die Qualität des Vorhersagemodells zu bewerten, wird häufig die *Area Under the Receiver Operating Characteristic Curve* (AUC) herangezogen.
- e) Die AUC von effektiven Modellen bewegen sich im Bereich zwischen 0,3 – 0,5.

### 5. Welcher Algorithmus wird bei der Bilderkennung im Rahmen des *Deep Learning* am häufigsten eingesetzt?

- a) *Deep Belief Networks*
- b) *Convolutional Neural Networks*
- c) *Recurrent Neural Networks*
- d) *Supervised Learning*
- e) *Reinforcement Learning*

### 6. Welches klinische Merkmal lässt sich mittels automatischer Auswertung von Fundus-Fotografien (noch) nicht detektieren?

- a) Augeninnendruck
- b) *Cotton-Wool*-Herde
- c) Makulaödeme
- d) Mikroaneurysmen
- e) *Cup-to-Disc-Ratio*



- 7. Wie hoch** liegt die Treffsicherheit für die Erkennung einer altersbedingten Makuladegeneration (AMD) mittels KI-basierter Auswertung von *Spectral-Domain*-Bildern der optischen Kohärenztomographie (OCT)?
- a) 60 %
  - b) 65 %
  - c) 73 %
  - d) 82 %
  - e) 98 %
- 8. Wofür eignet sich die automatische Auswertung von OCT-Bildern (noch) nicht?**
- a) Detektion intraretinaler Flüssigkeit und subretinaler Flüssigkeit
  - b) Detektion von Drusen
  - c) Voraussage des Ansprechens auf eine anti-VEGF-(*Vascular Endothelial Growth Factor*-) Therapie
  - d) Überwachung des Therapieerfolgs einer anti-VEGF-Therapie
  - e) Bestimmung des anti-VEGF-Injektionsintervalls
- 9. Bei welcher Indikation** ist die automatisierte Auswertung von **Spaltlampenbildern** am weitesten fortgeschritten?
- a) AMD
  - b) Katarakt
  - c) Diabetische Retinopathie
  - d) Diabetisches Makulaödem
  - e) Glaukom
- 10. Was gehört nicht zu den größten Herausforderungen bei der KI-basierten Bildauswertung?**
- a) Geringe Geschwindigkeit der Bildauswertung
  - b) Fehlende Standardisierung der Bildaufnahmen
  - c) Bedarf an großen Datensätzen für das KI-Training
  - d) Gefahr des blinden Vertrauens in die generierten Ergebnisse
  - e) Fragen zum Datenschutz

## IMPRESSUM

### AUTORIN

**Prof. Dr. med. Nicole Eter**

Klinik für Augenheilkunde  
Universitätsklinikum Münster, Münster

### INTERESSENKONFLIKTE DER AUTORIN

Advisory Boards und Vorträge: Novartis, Novartis Global, Bayer, Bayer Global, Allergan, Roche, Appellis Pharmaceuticals, Alcon

Sonstiges: Ärztlicher Herausgeber der Pro Compliance Aufklärungsbögen beim Thieme Verlag, Aufsichtsratsmitglied am Universitätsklinikum Aachen, Weiterbildung und Facharztprüfungen sowie Planung und Durchführung von MFA-Weiterbildungen für die Ärztekammer Westfalen-Lippe, Mitglied der QS-Kommission und Teilnahme an Sitzungen und Vorträgen der Kassenärztlichen Vereinigung Westfalen-Lippe

### REDAKTION & LAYOUT

Dr. Barbara Knapp & Stefanie Blindert  
KW MEDIPOINT, Bonn

Die Zertifizierung dieser Fortbildung durch die Bayerische Landesärztekammer wurde von CME MEDIPOINT, Neusäß organisiert.

Diese Fortbildung wurde von der Novartis Pharma GmbH mit insgesamt 17.087,- € finanziert. Die Ausarbeitung der Inhalte der Fortbildung wird dadurch nicht beeinflusst.

### BEGUTACHTUNG

Diese Fortbildung wurde von zwei unabhängigen Gutachtern auf wissenschaftliche Aktualität, inhaltliche Richtigkeit und Produktneutralität geprüft. Jeder Gutachter unterzeichnet eine Konformitätserklärung.

Diese Fortbildung ist auf [www.cmemedipoint.de](http://www.cmemedipoint.de) online verfügbar.