

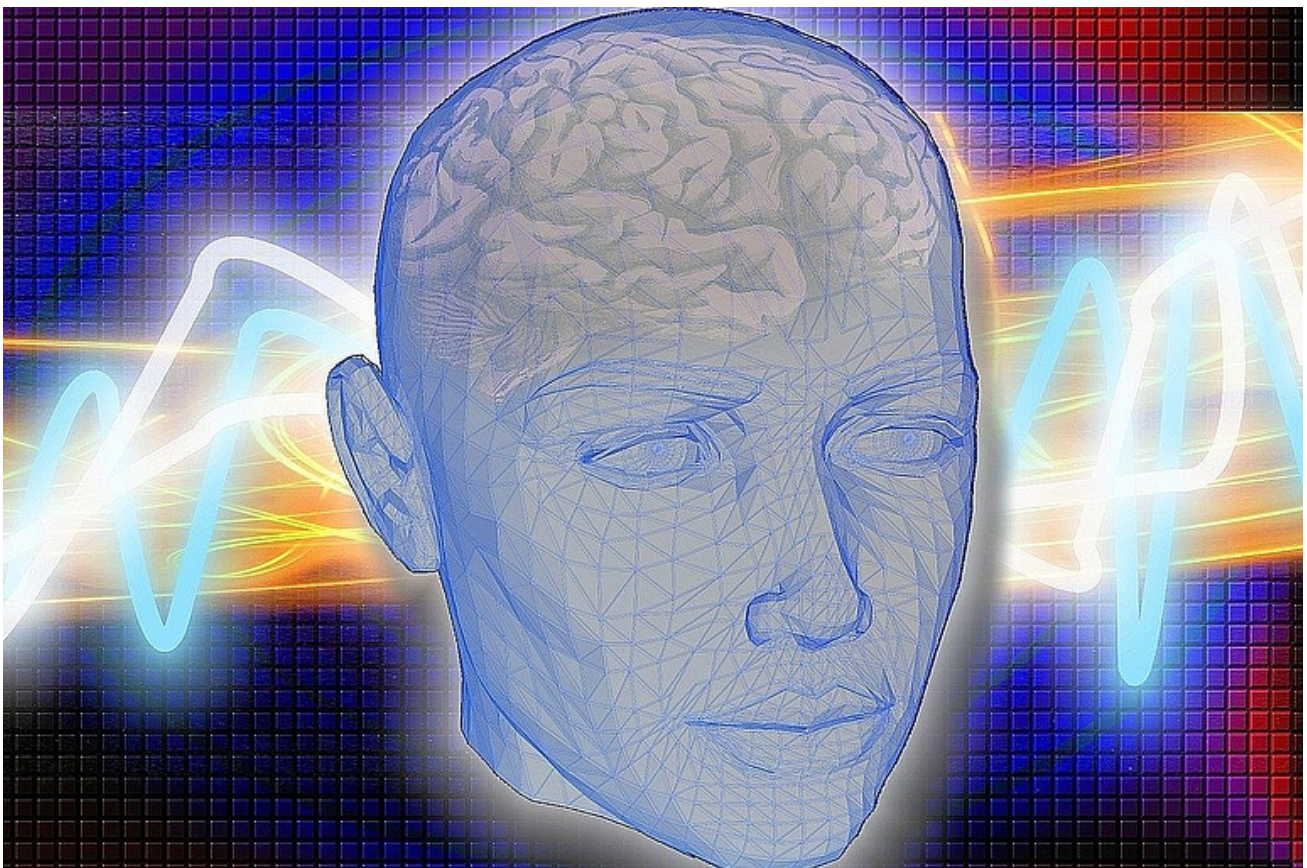


Advertorial

Top-Performance mit Intels® AI-Tools

Deep-Learning-Algorithmen zur Analyse menschlicher Organe beschleunigen

von Shailen Sobhee und Emanuel Rosenauer



Ein vielversprechendes Anwendungsfeld für „Artificial Intelligence“ ist die Medizin. Hier kann AI zum Beispiel den Ärzten bei der Diagnose von Krankheiten helfen. Das AI-Team von Intel® hat deshalb an einem Algorithmus gearbeitet, mit dem sich menschliche Organe analysieren und Tumoren erkennen lassen. Dieser Algorithmus wird derzeit bereits von einigen Krankenhäusern in Europa als „Proof of Concept“ in der Praxis ausprobiert.

Die Unterstützung der Ärzte durch AI macht auf der einen Seite eine exakte und frühzeitige Diagnose möglich. Auf der anderen Seite kann AI aber auch möglichst wirksame Therapievorschlge unterbreiten und bei der Vor- und Nachsorge mithelfen. Wie segensreich AI an dieser Stelle sein kann, wird am Beispiel von Krebs deutlich; weltweit gab es allein 2018 laut [globaler Krebsstatistik \(Globocan\)](#) 18,1 Millionen neue Krebsflle – und 9,6 Millionen Menschen starben im Jahr 2018 an Krebs. Gezhlt hat man dabei 36 Krebsarten in 185 Lndern.

Krebs ist nach den Herz-Kreislauf-Erkrankungen in Deutschland [die zweithufigste Todesursache](#) [1]. Bei der Frherkennung der Krankheit kommt es viel auf die diagnostische Genauigkeit („accuracy“) an, also auf die Fhigkeit, sowohl einen pathologischen als auch einen nicht pathologischen Zustand als solchen richtig zu identifizieren und diese Flle voneinander zu unterscheiden. Speziell bei Hirntumoren lsst sich mit der Diagnose eines gutartigen Tumors unter Umstnden eine Operation vermeiden.

Selbst bsartige Tumore lassen sich erfolgreich durch andere Methoden als die riskante Operation am Gehirn behandeln – etwa durch Strahlen- oder Chemotherapie, aber auch Knochenmarks- und Stammzelltherapie oder neuerdings durch die gezielte Krebstherapie („[targeted therapy](#)“). Dank aktueller Erkenntnisse von Molekularbiologie und Biomedizin knnen bei der „targeted therapy“ Merkmale bsartiger Zellen fr die Entwicklung neuer, spezifischer Arzneistoffe genutzt werden, z. B. fr Antikrper gegen typische Proteine der Krebszellen oder zur Hemmung der krebsbedingten Neubildung von Blutgefen (Angiogenese). Da diese Merkmale bei gesunden Zellen meist kaum oder gar nicht vorkommen, ist eine gezielte Krebstherapie gleichermaen schonend und wirksam.

„Big Data“ und E-Health

Bei Frhdiagnose und Therapievorschlgen kann AI vor allem deshalb helfen, weil Forscherinnen und Forscher weltweit tagtglich Patientendaten aus Blutproben, bildgebenden Verfahren oder Genanalysen sammeln. Allein im Jahr 2013 wurden knapp 153 Exabyte an Gesundheitsdaten generiert – und dieses Datenvolumen wchst jhrlich um 48 Prozent. Es wird erwartet, dass in diesem Jahr 2.314 Exabyte erreicht werden [2]. Ein Exabyte, also eine Milliarde Gigabyte, entspricht 250 Mio. DVDs mit Informationen.

Speziell die vorliegenden Bilder aus Magnetresonanz-Tomografien (MRT) nutzt Intel® fr das Training eines knstlichen neuronalen Netzwerkes zur [Hirntumorsegmentierung mittels Deep Learning](#). Dabei kommen verschiedene Optimierungstechniken zum Einsatz, um die Laufzeit der Deep-Learning-Algorithmen auf einer modernen Intel®-Architektur mglichst stark zu verkrzen.

So beschleunigt zum Beispiel Intels® „Math Kernel Library“ (MKL) das Training des „Deep Neural Network“ mit TensorFlow, whrend OpenVINO™ beim konkreten Einsatz der KI bei der Diagnose die Inferenz verbessern. Durch das Skalieren auf mehrere Xeon-Knoten in Form verteilter Trainings lsst sich dann die beste Leistung fr Deep-Neural-Network-Trainings erzielen.

Zur Optimierung genutzt werden dabei die Intel® Distribution fr Python, Intels® optimierte TensorFlow-Implementierung mit MKL-DNN, ein Xeon-basierter Cluster fr verteilte AI sowie auch Intel® MPI, eine mit der Intel®-Compiler-Suite ausgelieferte Implementierung des „Message Passing Interface“-Standards (MPI) fr die Kommunikation zwischen den Knoten im Cluster.

Behandlung von Gliomen

- To assess the severity/for treatment of the tumour, segmentation is important for:
 - focusing on the tumour areas during radiotherapy
 - navigation during surgery

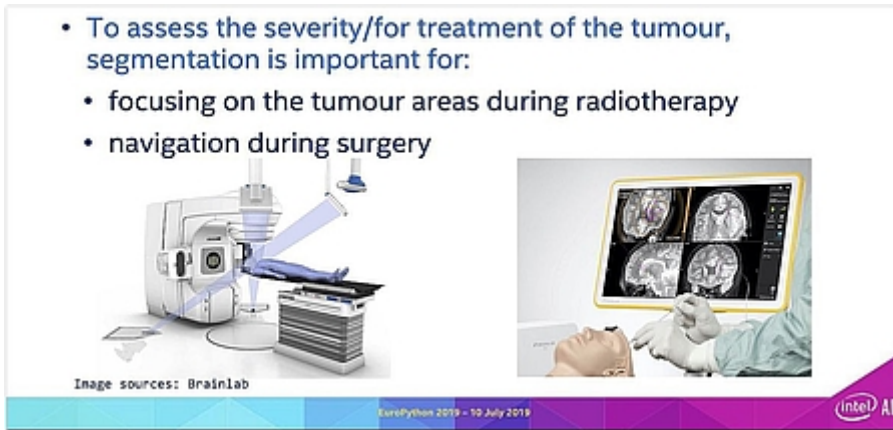


Image sources: Brainlab

EuroPython 2019 - 10 July 2019

intel AI

Segmentierung des Gehirntumors

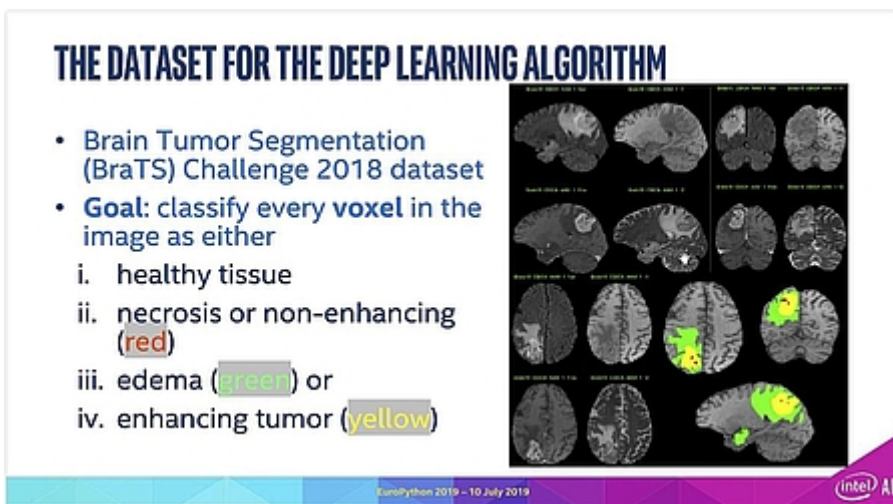
Im ersten Schritt steht bei diesem Algorithmus die Behandlung von **Gliomen** im Vordergrund. Gliom (griech. „glia“, Leim) ist ein Sammelbegriff für einige der häufigsten und bösartigsten Hirntumore bei Menschen. Gliome lassen sich durch eine Operation häufig nicht vollständig entfernen. Chemo- oder Strahlentherapie sind nur begrenzt wirksam, da der Tumor eine hohe Widerstandskraft besitzt. Neue und präzise validierte Behandlungsansätze werden daher dringend benötigt.

Die sogenannte Mehrsequenz-Magnetresonananz- oder Kernspintomografie ist die primäre Methode für das Screening und die Diagnose solcher Gliome. Ein MRT kann die Organe des Körpers – auch das Gehirn – detailliert darstellen und viele krankhafte Veränderungen sichtbar machen. Dieses bildgebende Untersuchungsverfahren erzeugt mithilfe eines starken Magnetfelds, zusätzlichen Wechselfeldern, Messantennen und eines Computers Schnittbilder des Körperinneren. Der Name ist daher abgeleitet vom altgriechischen Wort „tome“ für Schnitt.

Klassische MRTs sind als Röhre gebaut. Für die Untersuchung fährt eine Liege den Patienten so weit in den ringförmigen Magnettunnel, bis der zu untersuchende Körperbereich sich in der Röhrenmitte befindet. Die MRT-Bilder sind sehr detailgenau, sodass auch kleinste Veränderungen nachgewiesen werden können. Außerdem ermöglicht die MRT Aufnahmen in jeder gewünschten Körperebene – also nicht nur quer durch den Körper, sondern auch längs und schräg. Die Schnittebenen legt der Arzt fest, was das Verfahren zeitaufwendig und teuer macht.

THE DATASET FOR THE DEEP LEARNING ALGORITHM

- Brain Tumor Segmentation (BraTS) Challenge 2018 dataset
- Goal: classify every voxel in the image as either
 - healthy tissue
 - necrosis or non-enhancing (red)
 - edema (green) or
 - enhancing tumor (yellow)



EuroPython 2019 - 10 July 2019

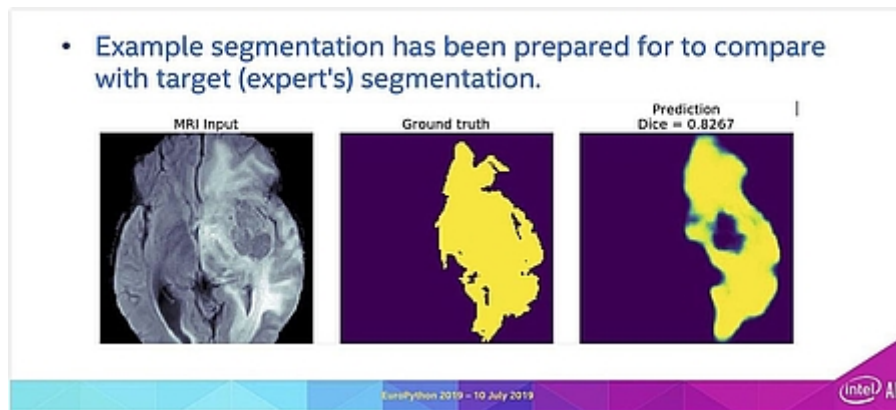
intel AI

Bei dieser sogenannten Segmentierung kann AI helfen, zumal es einen Mangel an entsprechend spezialisierten Ärzten gibt – durch den Zeitgewinn für den Radiologen. Genau an dieser Stelle setzt Intel® mit innovativen AI-Verfahren an. Das Ziel ist es, auf einem dreidimensionalen MRT-Bild einen Tumor einwandfrei zu identifizieren.

Dazu soll jedes Voxel im Bild entweder als gesundes Gewebe oder krankes Gewebe klassifiziert werden. Dabei ist „Voxel“ (zusammengesetzt aus dem englischen „volume“ und „elements“) ein Begriff aus der Computergrafik, der in einem dreidimensionalen Raster einen Gitterpunkt bzw. ein Datenelement bezeichnet – analog zu einem Pixel (Rasterpunkt) in einem

2D-Bild.

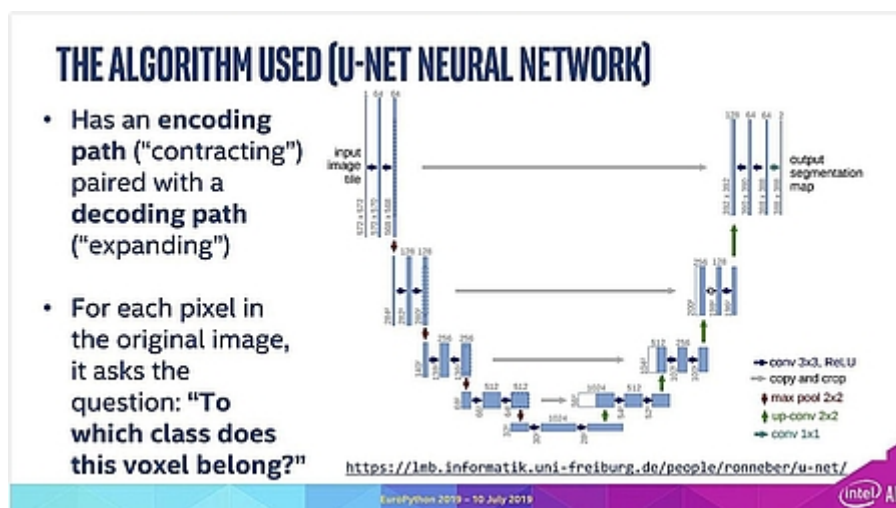
Intel® hat bei der Diagnose folgendes Verfahren ausgewählt. Krankes Gewebe wird vom KI-System unterschiedlich eingefärbt: Nekrosen oder nicht wuchernde Tumoren (rot), Ödem (grün) und wachsende Tumoren (gelb). Für das Training des „Deep Neural Network“ (DNN) wurde ein Open-Source-Datensatz verwendet, der im Jahr 2018 im Zuge der „Brain Tumor Segmentation“ – oder kurz BraTS-Challenge – gewonnen wurde.



Das Ergebnis des Deep-Learning-Algorithmus

Das Ergebnis des Deep-Learning-Algorithmus sieht so aus: Der Scan eines Gehirnsegmentes (linkes Bild) dient als Input für den Algorithmus. Das mittlere Bild zeigt den MRT-Scan dieses Segmentes, bei dem vom Radiologen die erkrankten Bereiche eingefärbt wurden. Diese Bilder werden als „Labeled Data“ bezeichnet und für das Training des DNN verwendet. Das trainierte DNN liefert dann auf Basis des aktuellen MRT-Scans eines Patienten das rechte Bild, also die „Prognose“ bzw. das „inferred Image“ mit den Einfärbungen aufgrund der Schlussfolgerungen des DNN.

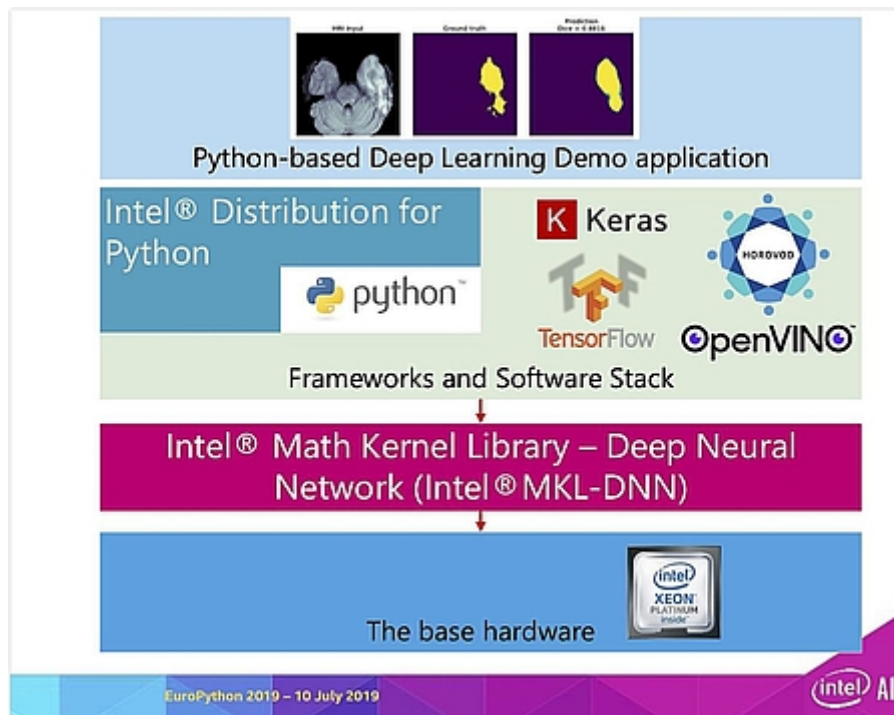
Auf diese Weise kann das AI-System bei neuen Patienten schnell erkennen, ob diese gesund oder krank sind, wo sich die Tumoren befinden und welcher Art diese Tumoren sind. Als Algorithmus wurde dabei das im Gesundheitswesen populäre künstliche neuronale Netzwerk „U-Net“ verwendet. Es verfügt über einen Codierungspfad („Contracting“), der mit einem Decodierungspfad („Expanding“) gekoppelt ist. Für jedes Voxel im Originalbild wird die Frage gestellt: „Zu welcher Klasse gehört es? Muss es eingefärbt werden – und falls ja: wie?“



Zunächst wird das U-Net-Modell mit den vorbereiteten „Labeled Data“ trainiert. Dieses trainierte Modell ist dann als „Inferenz-Maschine“ bereit für die Analyse neuer MRT-Scans von Patienten. Dabei ist zu berücksichtigen, dass MRT-Aufnahmen qualitativ hochwertige Bilder mit detailliertester Auflösung sind. Es geht also um riesige Datenmengen, bei deren Verarbeitung die Computer extrem gefordert sind, sowohl zunächst beim Training des DNN als auch später beim Inferencing.

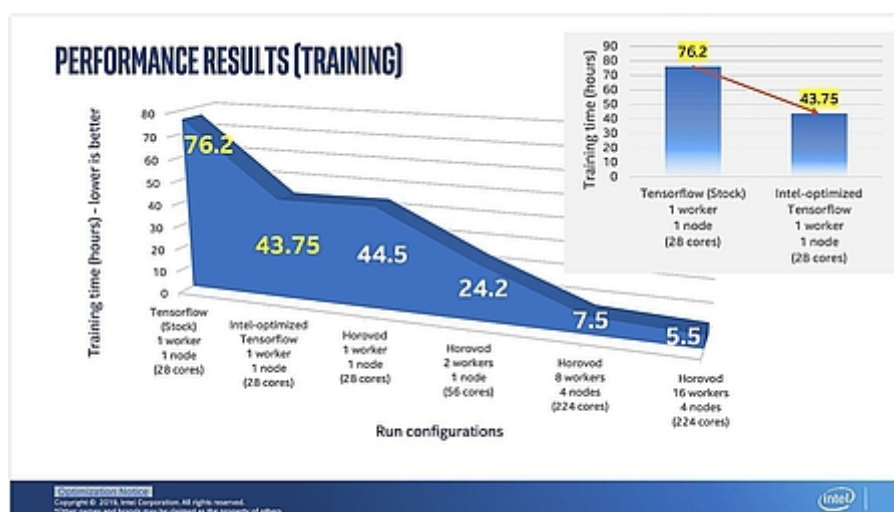
Deshalb ist Optimierung an jeder Stelle gefragt, um Speicherplatz und Rechenzeit zu sparen. Genau deshalb ist aber auch ein sorgfältig aufeinander abgestimmtes Hardware- und Software-System als Basis des AI-Systems entscheidend für akzeptable

Antwortzeiten bzw. einen vernünftigen Durchsatz. Andernfalls würde der Klinikbetrieb aufgehoben – und die Ärzte und Patienten würden keine Zeit sparen.



Für diese Optimierung benutzt Intel® als Software die eigene Distribution der Programmiersprache [Python](#), die auf die Hardware der Xeon-Prozessoren hin optimiert ist. Das DNN wurde mithilfe von TensorFlow dargestellt; weil [TensorFlow](#) aber etwas „sperrig“ und kompliziert sein kann, kommt als API-Layer [Keras](#), ebenfalls aus dem Hause Google, ins Spiel. Keras bietet dem Entwickler die Möglichkeit, das „Deep Neural Network“ quasi wie bei Legobausteinen einfach zusammenzustecken.

Mithilfe von [Horovod](#), dem Open-Source-basierten Deep-Learning-Framework für TensorFlow von Uber, werden Rechenaufgaben auf verschiedene Prozessoren verteilt, sodass ein verteiltes System oder ein Cluster für das Training genutzt werden kann, um diesen Prozess zu beschleunigen. Um das Inferencing ebenfalls zu beschleunigen und zu vereinfachen, kommt schließlich auch noch das Intel®-Toolkit [OpenVINO™](#) zum Einsatz. Die Basis dafür bildet eine Xeon-basierte Hardware-Infrastruktur mit Intels® Math Kernel Library – Deep Neural Network (Intel® MKL-DNN).



Wie sehr sich die Optimierung lohnt, macht ein Blick auf die notwendigen Trainingszeiten deutlich. Mit dem klassischen TensorFlow von Google dauerte das Training des DNN auf einem Rechner mit 28 Prozessor-Cores 76,2 Stunden (siehe Grafik 6); zugrunde liegen dabei jeweils 30 Epochen.

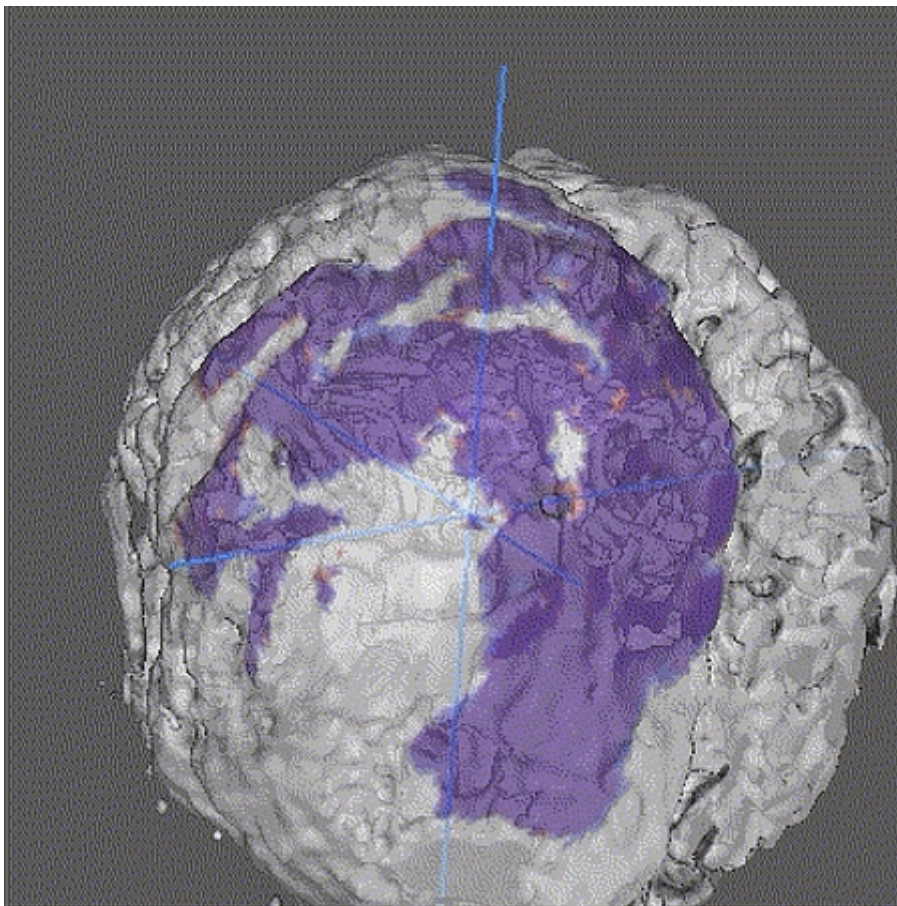
Allein schon die Umstellung auf die Intel®-optimierte Version von TensorFlow bringt einen Performance-Gewinn von mehr als 43 Prozent. Bei dieser Version handelt es sich nicht um eine eigene Implementierung, sondern um die ursprüngliche Codebasis von Google, die lediglich mit aktivierten Intel®-MKL-DNN-Flags neu kompiliert wurde, sodass TensorFlows mathematische Operationen innerhalb von MKL-DNN ausgeführt (oder an MKL-DNN ausgelagert) werden. Weiter steigern lässt sich die Performance durch „Distributed Training“; werden mehrere Maschinen genutzt, lässt sich die Trainingsdauer von rund 44 Stunden (1 Bearbeiter, 1 Knoten mit 28 Cores) auf 7,5 Stunden (8 Bearbeiter, 4 Knoten mit 224 Cores) bzw. 5,5 Stunden (16 Bearbeiter, 4 Knoten mit 224 Cores) drücken.

Optimierung lohnt sich - und kann einfach sein

Das ist praktikabel, denn so lassen sich die 30 Epochen für das Training des DNN an einem einzigen Arbeitstag bewältigen. Ohne Optimierung dauerte das 30-malige Durchspielen der Bildsequenzen zum Training des DNN fast zehn Arbeitstage, viel zu lange für den Klinikalltag. Die Botschaft hinter diesen Messergebnissen liegt auf der Hand: Optimierung lohnt sich - und die Verteilung der Workload auf mehrere Bearbeiter und Rechner ebenso.

Aber vor allem die Halbierung der Trainingsdauer allein schon durch die Verwendung optimierter Software hat Aufmerksamkeit verdient, zumal diese Performance-Steigerung praktisch überhaupt keine Mühe macht. Auch dafür wird Googles TensorFlow verwendet, wobei Intel® zur Optimierung in dieses Original einzig und allein seine MKL-DNN eingebaut hat, die dann die vielen mathematischen Operationen für das U-Net abwickelt. Weil diese Operationen dank MKL-DNN sehr viel schneller abgewickelt werden, beschleunigt sich das Training des U-Net deutlich.

Von 2D zu 3D



Visualisierung des Endergebnisses in 3D

Noch ist allerdings die Diagnose nicht ganz abgeschlossen, denn bisher wurden nur die einzelnen zweidimensionalen Schichten analysiert, aus denen dann das dreidimensionale Bild des Gehirns zusammengesetzt wird. Man kann sich vorstellen, wie mühsam

das für den Arzt ist, der dabei auf kleinste Details achten muss – und wie lange der Patient dafür in dem MRT-Gerät bleiben muss, bis die 3D-Aufnahme des Gehirns endlich komplett fertiggestellt ist.

Durch die Komposition der vielen 2D-Segmente entsteht die 3D-Aufnahme des Gehirns, mit der dann zum Beispiel der Chirurg die Tumoren für seine Eingriffe exakt lokalisieren kann. Auch der Radiologe muss ja für eine etwaige Strahlentherapie möglichst exakt wissen, wo genau im Gehirn die Tumoren platziert sind. Wer den Source-Code ausprobieren möchte, [findet ihn zum Download auf GitHub](#).

Bei der Beurteilung der Wirksamkeit einer Therapie kann AI ebenfalls helfen, denn hier ist eines der wesentlichen Kriterien die Wachstumsdynamik, die gleichfalls über die MRT-Bildgebung ermittelt wird. Doch das manuelle Messen der Tumorausdehnung in zwei Ebenen in den kontrastverstärkten MRT-Aufnahmen ist, wie geschildert, aufgrund der manuellen Bearbeitung fehleranfällig und führt leicht zu abweichenden Ergebnissen, was die Präzision bei der Beurteilung der Wirksamkeit der Therapie negativ beeinflusst. Auch in der radiologischen Diagnostik gibt es also viel Potenzial für maschinelle Lernverfahren, an denen auch Intel® intensiv arbeitet. Das AI-Team von Intel® hat neuronale Netzwerke entwickelt, um computerbasiert die Heilungserfolge bei der Behandlung von Hirntumoren standardisiert zu beurteilen und klinisch zu validieren.

Intels® AI-Strategie basiert auf Open Source

Das Schöne daran: Sowohl der komplette Software-Stack für das DNN als auch die für das Training verwendeten Daten sind Open Source, was typisch ist für Intels® AI-Engagement. Auch das hier beschriebene U-Net-DNN zur biomedizinischen Bildsegmentierung ist [vollständig auf GitHub publiziert](#), inklusive Quellcode und Anweisungen zum Abrufen der medizinischen Datensätze. Das Ziel ist es letztendlich, durch lernende AI-Systeme die Behandlung von Patienten und ihre Prognosen deutlich zu verbessern. Dabei gibt es folgende Ansatzpunkte:

Behandlungserfolg: Frühzeitige Diagnosen und individuellere Therapien verbessern Ergebnisse bei der Behandlung.

Sicherheit: Ärztinnen und Ärzte können sich bei ihrer Entscheidung auf zusätzliche patientenindividuelle sowie weltweit vorhandene medizinische Daten stützen und so besser entscheiden.

Erkenntnisgewinn: AI ermöglicht Ärztinnen und Ärzten den Zugang zu exponentiell wachsendem medizinischem Wissen. Dadurch gelangen neueste Erkenntnisse rascher in die breite Versorgung.

Effizienz: Durch den Zugriff auf die elektronische Patientenakte und AI-Unterstützung bei Diagnose und Therapie können Ärztinnen und Ärzte schneller die richtigen Entscheidungen treffen.

Die Algorithmen erlernten unter dem Einsatz künstlicher neuronaler Netzwerke, die verschiedenen Varianten von Hirntumoren automatisch zu erkennen und zu lokalisieren. Das kann über Leben und Tod entscheiden, denn für Präzisionstherapien ist eine standardisierte und verlässliche Beurteilung der Effektivität der Therapie von herausragender Bedeutung.

Literatur & Links

[1] Destatis: Todesursachen: Zahl der Todesfälle im Jahr 2017 um 2,3 % gestiegen
<https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Todesursachen/todesfaelle.html>

[2] Corbin K. How CIOs Can Prepare for Healthcare „Data Tsunami“ [Internet]. CIO. 2014 [cited 8 FEB 2019].

[3] Fenton SH, Low S, Abrams KJ, Butler-Henderson K.: Health Information Management: Changing with Time. IMIA Yearbook of Medical Informatics 2017.

[4] Stanford Medicine. 2017 Health Trends Report: Harnessing the Power of Data in Health. Accessed online 8 FEB 2019.

[5] Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, Ahmedin Jemal: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, <https://onlinelibrary.wiley.com/doi/full/10.3322/caac.21492>

[6] Deborah Abrams Kaplan: A Physician Shortage, But Too Many Radiologists, www.diagnosticimaging.com/residents/physician-shortage-too-many-radiologists



Shailen Sobhee

ist AI Software Technical Consultant Engineer bei Intel®.



Emanuel Rosenauer

ist Verlagsleiter bei der SIGS DATACOM GmbH.

Bildnachweise

Intel®, kalhh / Pixabay

AI Trendletter

Impressum

|

Kontakt & Anfrage