GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Schlussbericht – Teil I

CompLS – RundFörderkennzeichen 031L0210A

LODE - Longitudinales Deep Learning für klinische Diagnosen, Merkmalsdarstellung und interventionelle Modellierung

Laufzeit des Projektes: 01.03.2020 - 30.11.2023

Projektpartner:

Prof. Dr. Dr. Fabian J. Theis
Helmholtz Munich
Ingolstädter Landstraße 1
D-85764 Neuherberg
fabian.theis@helmholtz-munich.de
Tel.: +49 3187 2926

Prof. Dr. Sach Mukherjee
Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE)
Venusberg Campus-1
53127 Bonn
sach.mukherjee@dzne.de
Tel.: +49 228 43302-853

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Neuherberg, 10.07.2024

I. Kurze Darstellung

1. Aufgabenstellung

Maschinelles Lernen (ML) im Allgemeinen und Deep Learning (DL) im Besonderen können einen tiefgreifenden Einfluss auf die medizinische Forschung und Praxis haben. Längsschnittdaten, die Veränderungen von Patientenphänotypen (im Laufe der Zeit und unter klinischen Eingriffen) erfassen, spielen in der klinischen Medizin eine Schlüsselrolle; allerdings stellen die spezifischen Merkmale von Längsschnittdaten, wie sie in realen klinischen Umgebungen gewonnen werden, Herausforderungen für das Lernen dar. Im biomedizinischen Umfeld besteht außerdem ein deutlicher Bedarf an der Entwicklung interpretierbarer Methoden und an der Entwicklung von Kausalmodellen über die Vorhersage hinaus. Unser Ziel für LODE war es, maßgeschneiderte Innovationen in der DL für große ophthalmologische Längsschnittdaten zu liefern.

Die übergreifenden Ziele von LODE waren (1) die Entwicklung neuer DL-Methoden für longitudinale bildbasierte klinische Daten, wobei der Schwerpunkt auf einer interpretierbaren Methodik lag, die auf groß angelegte klinische Routinebeobachtungen anwendbar ist, und (2) die Integration von Kausal- und Netzwerkmodellierungskonzepten mit DL in dieser Anwendung. Diese Ziele wurden in vier spezifischen Arbeitspaketen (WPs) behandelt, die im Folgenden erläutert werden.

2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

LODE war ein Projekt in enger Zusammenarbeit zwischen der Arbeitsgruppe von Fabian Theis am Helmholtz Zentrum München und der Arbeitsgruppe von Sach Mukherrje am DZNE Bonn durchgeführt. Die Hauptarbeit wurde von den Postdocs Dr. Hannah Spitzer (Helmholtz) und Dr. Johannes Wahle (DZNE) mit den Doktoranden Olle Holmberg (Helmholtz), Nastassya Horlova (Helmholtz) und Kai Lagemann (DZNE) durchgeführt. Herr Holmberg verließ das Helmholtz Zentrum während der Projektlaufzeit, aber die entsprechenden Aufgaben wurden von Frau Horlova und Dr. Spitzer bis zum Abschluss weitergeführt. Die fachliche Expertise und die Gesamtbetreuung wurde von Prof. Fabian Theis und Prof. Sach Mukherjee wahrgenommen. Unser klinischer Partner, Dr. Karsten Kortüm, leistete ebenfalls wertvolle Beiträge, insbesondere bei großen ophthalmologischen Datensätzen. Darüber hinaus unterstützte ein wissenschaftlicher Manager (kleinere) Aspekte der Projektkoordination (Sachleistung).

3. Planung und Ablauf des Vorhabens

LODE war ursprünglich als 36-monatiges Projekt vom 01.03.2020 bis 28.02.2023 geplant. WP 1 wurde in diesem Zeitraum abgeschlossen. Aufgrund der verständlichen Verzögerungen durch die COVID-Pandemie wurde uns eine kostenneutrale Verlängerung um 9 Monate (bis 30.11.2023) gewährt, um auf den erzielten Fortschritten des Projekts aufzubauen und die weitere Dissemination unserer Arbeit auf wissenschaftlichen Konferenzen zu ermöglichen. Während des 9-monatigen Verlängerungszeitraums setzte das Theis Lab (i) die Arbeit an WP2 fort, insbesondere an WP2. 2, das sich mit der Stratifizierung von Patienten und der Vorhersage patientenspezifischer Verläufe im latenten Raum befasst. Weiterhin unterstützte das Theis-lab (ii) die Arbeiten an den WPs 3 (zeitliche Kausalmodelle) und 4 (gemeinsames tiefes kausales Lernen, beide geleitet durch das Mukherrje-Lab) durch die Bereitstellung von Vorhersagemodellen für das

Patientenschicksal (untersucht in WP2) und von Fachwissen bei der Datenanalyse und unterstützte (iii) die WPs 3 und 4 (), die auf den Ergebnissen der vorherigen WPs aufbauen.

4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

Das Hauptziel von LODE war die Entwicklung neuer DL-Ansätze für longitudinale klinische Daten, mit klinisch und wissenschaftlich interpretierbaren Ergebnissen und einem Schwerpunkt auf großen ophthalmologischen Daten. Im Rahmen des LODE-Projekts entwickelten, untersuchten, trainierten und bewerteten wir verschiedene Modelle des maschinellen Lernens (ML), darunter klassische ML-Ansätze (u. a. Lasso, elastisches Netz, Boosting) sowie Deep-Learning-basierte Ansätze (u. a. Standard-Feed-Forward-Netze, LSTM und neuartige Modelle, siehe unten). Es wurden Modelle entwickelt, um die Reaktionen der Patienten als kontinuierliche Variablen vorherzusagen. Diese Werte wurden dann verwendet, um den Reaktionstyp des Patienten auf der Grundlage medizinisch definierter Schwellenwerte zu bestimmen, wodurch die Interpretierbarkeit der Modelle demonstriert wurde. Einige der ML-Modelle zeigten im Vergleich zu klinischen Analysen eine gute Leistung bei der Vorhersage von Non-Respondern und schlechten Respondern auf die Medikation. Dies unterstreicht das Potenzial von ML zur Unterstützung von Klinikern bei der Identifizierung von Patienten, bei denen ein Ansprechen auf bestimmte Behandlungen unwahrscheinlich ist, und hilft so bei der Auswahl geeigneter Medikamente.

- **Angabe bekannter Konstruktionen, Verfahren und Schutzrechte, die für die Durchführung des Vorhabens benutzt wurden**

Die Projektziele wurden mit Hilfe von Open-Source-Software erreicht. Uns ist nicht bekannt, dass eines der Projektziele auf proprietären Softwarelösungen und/oder projektiertem geistigem Eigentum beruht.

- **Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste**

Es wurde primär auf die Online-Fachliteratur in der PubMed-Datenbank zurückgegriffen.

5. Zusammenarbeit mit anderen Stellen

Zu Beginn des Projekts wurde eine Kooperationsvereinbarung zwischen allen Partnern geschlossen. Mit den Teammitgliedern wurden regelmäßige Videokonferenzen organisiert, in denen im Vorfeld über eine gemeinsame Agenda verschiedene Schwerpunkte definiert wurden, die dann vorgestellt, diskutiert und weitere Schritte dokumentiert wurden. Auch unsere klinischen Partner nahmen regelmäßig an den Treffen teil, um ein bidirektionales Zusammenspiel zwischen der Modellierung und dem klinischen Anwendungsfall zu ermöglichen und um sicherzustellen, dass die KI-Aspekte mit den klinischen Anwendungsfällen, Denkweisen und Eingaben in Einklang gebracht werden.

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Schlussbericht – Teil II

CompLS – RundFörderkennzeichen 031L0210A

LODE - Longitudinales Deep Learning für klinische Diagnosen, Merkmalsdarstellung und interventionelle Modellierung

Laufzeit des Projektes: 01.03.2020 - 30.11.2023

Projektpartner:

Prof. Dr. Dr. Fabian J. Theis
Helmholtz Munich
Ingolstädter Landstraße 1
D-85764 Neuherberg
fabian.theis@helmholtz-munich.de
Tel.: +49 3187 2926

Prof. Dr. Sach Mukherjee
Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE)
Venusberg Campus-1
53127 Bonn
sach.mukherjee@dzne.de
Tel.: +49 228 43302-853

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Neuherberg, 10.07.2024

II. Eingehende Darstellung

II.1 Verwendung der Zuwendung und des erzielten Ergebnisses im Einzelnen, mit Gegenüberstellung der vorgegebenen Ziele

WP 1: Static feature extraction in multi-modal real-world medical images using convolutional networks (Statische Merkmalsextraktion in multimodalen medizinischen Bildern unter Verwendung von Faltungsnetzen)

Für die Meilensteine dieses Arbeitspakets wurde ein U-Netz-basiertes Modell entwickelt und eingesetzt, um OCT-Scans zu segmentieren. Das Modell wurde so konzipiert, dass es medizinische Merkmale (Biomarker) aus diesen Scans erfasst und relevante klinische Merkmale effektiv identifiziert und annotiert. Diese Segmentierung wurde an einem großen und vielfältigen Datensatz von Patientenproben durchgeführt, um die Robustheit und Verallgemeinerbarkeit des Modells bei verschiedenen demografischen Merkmalen und Erkrankungen von Patienten zu gewährleisten.

WP2: Predicting time-resolved clinical outcomes and patient fates (Vorhersage von klinischen Ergebnissen und Patientenschicksalen im Zeitverlauf)

Um ein umfassendes Verständnis des Krankheitsverlaufs zu erlangen, umfasst die betrachtete Datenbank Aufnahmen der optischen Kohärenztomographie (OCT), die zu mehreren Zeitpunkten gemacht wurden. Darüber hinaus wird der longitudinale Verlauf jedes Patienten durch begleitende klinische Informationen (Patientencharakteristika, Krankengeschichte, Sehschärfemessungen usw.) angereichert.

Wir verwendeten den U-Net-Segmentierungsalgorithmus, um das Volumen der Biomarker für jeden Patienten zu erfassen. Dadurch wurde ein nahezu automatisierter Arbeitsablauf geschaffen, der eine interpretierbare Merkmalsdatenbank über die Patientenpopulation bereitstellte, einschließlich zeitaufgelöster und interpretierbarer Merkmale. Unter Verwendung dieser Daten haben wir Faltungsschichten zur Merkmalsextraktion aus Bildgebungsdaten und verschiedene Ansätze zur Behandlung des zeitlichen Aspekts der Längsschnittdaten integriert.

Die Modelle wurden auf Zeitserien-Patientendaten in drei verschiedenen Situationen trainiert: mit Daten von einer einzigen Untersuchung, von zwei Untersuchungen und von drei Buntersuchungen, um das Fortschreiten der Krankheit im Laufe der Zeit zu erfassen. Um die Analyse und den Vergleich der Ergebnisse zu erleichtern, haben wir eine Shiny-App entwickelt. Mit dieser App können die Benutzer die Ergebnisse der Experimente in verschiedenen Einstellungen untersuchen, einschließlich der Anzahl der Untersuchungen und der verschiedenen Reaktionstypen. Dies bietet eine interaktive Plattform, um die Leistung der ML-Modelle unter verschiedenen Bedingungen zu visualisieren und zu verstehen.

WP3: Temporal causal inference based on neural network-derived patient state features (Zeitliche Kausalschlüsse auf der Grundlage von durch neuronale Netze abgeleiteten Patientenzustandsmerkmalen)

Verschiedene Patienten können aufgrund patientenspezifischer Faktoren unterschiedlich auf dieselbe Intervention reagieren. Während der durchschnittliche kausale Effekt in einer Population als durchschnittlicher Behandlungseffekt bezeichnet wird, kann der Effekt für einen bestimmten Patienten vom Durchschnitt abweichen, d. h. ein bestimmter Patient kann mehr oder weniger gut ansprechen. Wir versuchten, die Längsschnittmodelle zur Vorhersage von Effekten auf Patientenebene zu nutzen, indem wir die Veränderung unter der Behandlung als die Reaktion behandelten, die anhand von Merkmalen (auf Patientenebene) vorhergesagt werden sollte. Zu diesem Zweck verglichen wir eine breite Palette verschiedener Modelle, darunter klassische ML-Methoden (einschließlich Lasso, elastisches Netz, Boosting und andere) sowie auf Deep Learning basierende Ansätze (einschließlich Standard-Feed-Forward-Netze, LSTM und andere). Wir fanden heraus, dass einige der Modelle in der Lage waren, die Vorhersage von Behandlungseffekten (Ansprechen auf die Medikation) mit den Klinikern abzustimmen, was das Potenzial des maschinellen Lernens zur Unterstützung der klinischen Entscheidungsfindung im Zusammenhang mit interventionellen Behandlungen verdeutlicht. Modelle, die für die Vorhersage von Patientenreaktionen als kontinuierliche Werte trainiert wurden, wurden verwendet, um diskrete Patientenreaktionstypen auf der Grundlage von medizinisch definierten Schwellenwerten zu bestimmen, was die Vorstellung unterstützt, dass die ML-Modelle klinisch interpretierbare Ergebnisse liefern können.

WP4: Joint end-to-end learning of the DL and temporal causal models (Gemeinsames Ende-zu-Ende-Lernen von DL und zeitlichen Kausalmodellen)

Wir haben neuartige Architekturen entwickelt, die eine Art kausaler induktiver Verzerrung nutzen, um ein zeitliches Aufmerksamkeitsnetzwerk durchgängig zu lernen, einschließlich des Lernens von Repräsentationen und der Vorhersage von Trajektorien. Diese Modelle sind sehr flexibel in Bezug auf die Handhabung verschiedener Zeitpunkte oder der Anzahl von Zeitpunkten und nutzen modernste Transformer-Architekturen. Diese Architekturen wurden mit den bestehenden Ansätzen verglichen und scheinen die bestehenden Werkzeuge im Großen und Ganzen zu übertreffen. Dies ist ein wichtiger Schritt, da er das Potenzial von transformatorbasierten Architekturen für solche Probleme mit komplexen und heterogenen Eingabedaten aufzeigt, jedoch sind weitere Arbeiten erforderlich, um das volle Potenzial in verschiedenen Anwendungen zu verstehen.

II.2. Wichtigste Positionen des zahlenmäßigen Nachweises

Die Finanzierung umfasste ein Gesamtbudget von 566.713 €, wovon 394.092,89 € für Personal, 85.000 € für Fremdleistungen, 5.980 € für Reisen, 14.000 € für sonstige Projektkosten und 67.640,11 € für Verwaltungsgemeinkosten bestimmt waren. Die Mittel wurden weitgehend wie geplant verwendet (insgesamt wurden 534.678 € ausgegeben), mit Ausnahme der Fremdleistungen (die geringer ausfielen als erwartet) sowie einiger Reise- und sonstiger

Projektkosten (z. B. Kosten für die Organisation eines CompLS-Treffens/Workshops am Helmholtz Munich, der aufgrund der COVID-Pandemie nicht stattfinden konnte).

2.1. Personalkosten

Bewilligte Stellen im Rahmen der regulären Projektlaufzeit:

Eine Aufschlüsselung der Personalkosten für Helmholtz Munich finden Sie in der folgenden Tabelle

Name	**Einrichtung**	**Position**	**Kosten (in €)**			
			2020	**2021**	**2022**	**2023**
Dr. Hannah Spitzer	Helmholtz Munich	Postdoc	64.393,13	47.358,13	60.035,58	32.950,92
Olle Holmberg	Helmholtz Munich	Doktorand	29.178,34	24.899,18	-	-
Nastassya Horlova	Helmholtz Munich	Doktorandin	-	11.726,06	48.447,62	48.705,04
Gesamt (inkl. 10% Gemeinkosten)			**404.463,13**			

II.3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die im Rahmen des Projekts untersuchten Daten deuten darauf hin, dass die Fähigkeit der derzeitigen klinischen Ansätze zur Vorhersage des Ansprechens auf die Behandlung begrenzt bleibt. Das bedeutet, dass es in diesem Bereich ein großes Potenzial für ML oder ML-gestützte Arbeitsabläufe gibt.

Es ist wichtig, frühzeitig vorherzusagen, ob ein Patient nicht auf die Behandlung anspricht, damit die Wahl des Medikaments geändert werden kann, bevor es zu spät ist und irreversible Schäden entstanden sind.

In Zukunft wird die Einbeziehung von ML-Modellen in die klinische Entscheidungsfindung dazu beitragen, die Vorhersagegenauigkeit zu verbessern. Dies wird bei der Festlegung von Behandlungsstrategien helfen, die für den Einzelnen geeignet sind, und so die Gesamtwirksamkeit der Therapie erhöhen. Die Fähigkeit, Non-Responder frühzeitig zu identifizieren, kann zu besser personalisierten und schnelleren Interventionen führen, die letztlich die Ergebnisse für die Patienten verbessern.

II.4. Voraussichtlicher Nutzen, insbesondere die Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans

Da LODE auf früheren erfolgreichen Kooperationen zwischen den Labors von Mukherjee und Theis aufbaut, wurden die komplementären Stärken der beiden Labors kombiniert, um kritische Fragen der computergestützten Analytik für DL in der Augenheilkunde anzugehen. Die spezifischen Ziele, die in vier Arbeitspaketen erreicht wurden, umfassten die gemeinsame Entwicklung neuer ML-Methoden und -Software sowie die Verbreitung in gemeinsamen Projekten.

Das LODE-Projekt hat durch die Entwicklung und Anwendung fortschrittlicher Modelle des maschinellen Lernens (ML) erhebliche Fortschritte bei der Bewältigung wichtiger Herausforderungen in der Biomedizin und der Augenheilkunde gemacht. Ein grundlegendes Problem in diesen Bereichen ist der Mangel an hochwertigen Datensätzen. Durch die Nutzung einer umfassenden Datenbank von optischen Kohärenztomographie-Scans (OCT) und die enge Zusammenarbeit mit klinischen Partnern, hat LODE dazu beigetragen, wichtige Fragen in diesem Bereich zu definieren. Eine weitere kritische Herausforderung ist die Nichtinterpretierbarkeit von ML-Modellen. LODE hat dieses Problem durch die Entwicklung von Modellen angegangen, die nicht nur die Reaktionen der Patienten vorhersagen, sondern auch interpretierbare Ergebnisse liefern, die Kliniker verstehen und interpretieren können. Wir stellten Verbindungen zum kausalen Lernen her und entwickelten neue Modelle, die tiefes Repräsentationslernen und kausales Lernen integrieren. Darüber hinaus befasste sich das Projekt mit dem Problem der ungenauen Definitionen in der klinischen Forschung. Insbesondere das Fehlen einer klaren Definition von Reaktionstypen bei der Behandlung von nAMD stellte eine große Hürde dar. LODE arbeitete daran, präzise Kriterien für die Definition von Reaktionstypen festzulegen, was die Genauigkeit und Zuverlässigkeit der ML-Modelle verbesserte.

II.5. Während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen

Im Laufe des LODE-Projekts wurden mehrere Fortschritte bei den Techniken und Kenntnissen auf diesem Gebiet beobachtet. Eine bemerkenswerte Entwicklung ist die Erforschung von selbstüberwachten Lernmodellen auf großen Datensätzen von OCT-Scans. Dieser Ansatz geht die Herausforderung kleiner beschrifteter Datensätze an, indem er die riesige Menge an unbeschrifteten Daten nutzt, um die Modellschulung und -leistung zu verbessern.

Darüber hinaus hat sich die Integration von genetischen Daten mit klinischen und bildgebenden Daten als vielversprechender Weg erwiesen. Durch die Kombination genetischer Informationen mit OCT-Scans und anderen Patientendaten ist es möglich, umfassendere Modelle zu entwickeln, die nicht nur die Patientenergebnisse genauer vorhersagen können, sondern auch genetische Zusammenhänge mit dem Krankheitsverlauf und dem Ansprechen auf die Behandlung aufdecken.

Diese Fortschritte unterstreichen das Potenzial für eine weitere Verbesserung der Wirksamkeit und Interpretierbarkeit von ML-Modellen in der Augenheilkunde und Biomedizin und bieten neue Wege für die Forschung und Anwendung in der personalisierten Medizin und beim Verständnis von Krankheiten. In Zukunft wird die Integration von genetischen Daten und selbstüberwachten Lerntechniken in die untersuchten Ansätze einen weiteren Beitrag zur personalisierten Medizin und einer evidenzbasierten Gesundheitspolitik leisten. Durch die Entwicklung neuer fortschrittlicher Techniken ebnet das Projekt den Weg für präzisere und wirksamere Behandlungen in der Zukunft und stellt sicher, dass Gesundheitsdienstleister ihren Patienten die bestmögliche Versorgung bieten können.

II.6. Erfolgte oder geplante Veröffentlichungen des Ergebnisses

Nachfolgend finden Sie eine Liste der Veröffentlichungen, die LODE anerkannt haben:

- Ben Asani, Olle Holmberg, Johannes B Schiefelbein, et al. Evaluation of OCT biomarker changes in treatment-naive neovascular AMD using a deep semantic segmentation algorithm. https://doi.org/10.1101/2022.06.16.22276342 [provisionally accepted – Eye]
- Hannah Spitzer, Scott Berry, Mark Donoghoe, et al. Learning consistent subcellular landmarks to quantify changes in multiplexed protein maps. Nature Methods. 2023; 20(7), 1058–1069. https://doi.org/10.1101/2022.05.07.490900.
- Lukas Heumos, Philipp Ehmele, Tim Treis, et al. Exploratory electronic health record analysis with ehrapy. https://doi.org/10.1101/2023.12.11.23299816 [provisionally accepted – Nature Medicine]
- Johannes Schiefelbein, Nastassya Horlova, Hannah Spitzer, et al. New OCT based definition of treatment response to anti-VEGF in treatment naive neovascular AMD patients using an Deep Learning Algorithm [Abstract, presented at the 2023 ARVO Annual Meeting, held in New Orleans, LA, April 23-27, 2023]
- Lagemann, K., Lagemann, C., Taschler, B. et al. Deep learning of causal structures in high dimensions under data limitations. Nature Machine Intelligence 5, 1306–1316 (2023)
- Lagemann K., Lagemann, C. & Mukherjee, S. Invariance-based Learning of Latent Dynamics. Proceedings of The Twelfth International Conference on Learning Representations (ICLR 2024).
- Giovanni Palla, Hannah Spitzer, Michal Klein, et al. Squidpy: a scalable framework for spatial omics analysis. Nature Methods. 2022; 19:171-178. https://doi.org/10.1038/s41592-021-01358-2.

Darüber hinaus hat Helmholtz München das LODE-Projekt zu Doktor Whatson als wissenschaftliche Verbreitungsmaßnahme diskutiert (siehe Link: https://www.youtube.com/watch?v=gxjjAV-oFMA).