



Abschlussbericht: Verbundprojekt: MDLMA - Multimodale Datenanalyse mit Multi-Task Deep Learning - Teilprojekt B - FKZ 031L0202B

Zuwendungsempfänger: Universität zu Lübeck

Verantwortlich: Prof. Dr. Mattias Heinrich

Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Einleitung

Projektziele

Die Hauptziele und Motivation dieses Teilvorhabens waren:

1. Das Lernen auf heterogenen, multimodalen Bilddaten und eine komplementäre Kombination von Deep Learning und mathematischer Modellierung.
2. Die Verbesserung der iterativen CT-Rekonstruktion und der multimodalen Artefaktreduktion (WP2).
3. Die Entwicklung neuartiger Algorithmen zur robusten multimodale Bildregistrierung mit Deep Learning: Vereinfachung von komplexen Netzwerken durch Entkopplung des Repräsentationslernens mittels eines geometrischen Transformationsteils (WP4).
4. Die Entwicklung eines Konzept für das Multi-Task Lernen, das insbesondere auf multiskalige und multimodale biomedizinische Daten anwendbar ist (WP5).

Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Das Forschungsprojekt MDLMA hat sich mit der Herausforderung befasst, heterogene und multimodale Bilddaten mittels fortschrittlicher lernbasierter und klassischer Bildverarbeitungsmethoden zu verknüpfen und zu analysieren. Die Notwendigkeit dieses Projekts ergibt sich aus der zunehmenden Komplexität medizinischer Bilddaten und der Vielfalt der Bildmodalitäten, die in der Diagnostik und Forschung verwendet werden. Die Integration und Analyse dieser vielfältigen Datenquellen erfordern innovative Ansätze, um die Genauigkeit diagnostischer Verfahren zu verbessern und neue Erkenntnisse über Krankheitsbilder zu gewinnen. Die geleistete Projektarbeit stellt in diesem Kontext einen bedeutenden Fortschritt dar, indem sie die Möglichkeiten herkömmlicher Bildverarbeitungsmethoden erweitert und die Effizienz sowie Effektivität medizinischer Bildanalysen signifikant verbessert.

Das Projektziel, eine komplementäre Kombination von Deep Learning und mathematischer Modellierung zu entwickeln, adressiert direkt die Bedürfnisse einer sich stetig weiterentwickelnden medizinischen Bildgebung. Durch die Verbesserung der CT-Rekonstruktion und Artefaktreduktion leistet das Projekt einen wesentlichen Beitrag zur Präzision und Qualität medizinischer Bilder. Die Entwicklung neuartiger Algorithmen zur robusten multimodalen Bildregistrierung unterstreicht die Fähigkeit des Projekts, komplexe Datenverarbeitungsprobleme zu lösen und die Anwendungsbereiche von Bildverarbeitungstechnologien zu erweitern.

Die Zusammenarbeit mit renommierten Instituten und Forschungseinrichtungen wie dem DESY, dem Helmholtz-Zentrum Hereon und der Syntellix AG sowie die Einbindung führender Experten aus dem Bereich der medizinischen Bildverarbeitung und -analyse, gewährleisten einen hohen wissenschaftlichen Standard und praktische Relevanz der Forschungsergebnisse. Die interdisziplinäre Ausrichtung des Projekts, die sowohl technische Innovationen als auch klinische Anwendungen umfasst, ist ein weiterer Beleg für die Angemessenheit der durchgeführten Arbeiten.

Durch die Entwicklung und Implementierung neuer Verfahren, die eine effiziente Verarbeitung und Analyse von heterogenen und multimodalen Bilddaten ermöglichen, trägt das Projekt zur Verbesserung der diagnostischen Genauigkeit und zur Erweiterung des Verständnisses komplexer Krankheitsbilder bei. Die erzielten Ergebnisse unterstreichen die Notwendigkeit und Angemessenheit der geleisteten Projektarbeit und legen den Grundstein für zukünftige Forschungsarbeiten in diesem wichtigen Bereich.

Projektplan

Im Folgenden wird der Projektplan zusammengefasst, dessen Arbeitspakete jeweils wie angegeben vom Institut für Medizinische Informatik (IMI) und Institut für Medizintechnik (IMT) der Universität zu Lübeck bearbeitet wurden.

WP1: Datenaufbereitung (IMT 3PM, IMI 2PM)

Lieferobjekte:

- **M1.1:** Gemeinsames Datenformat.
- **M1.2:** Datenbankintegration und einheitlicher Zugang zu den Daten.

WP2 Bildverbesserung (17 PM IMT, 2 PM IMI)

Teilpakete:

- **T2.1:** CNN-basierte Ansätze zur Reduzierung von Artefakten im Post-Processing.
- **T2.2:** Direkte DL-basierte CT-Rekonstruktion.
- **T2.3:** Iterativer Rekonstruktionsalgorithmus zur Artefaktreduktion in Kombination mit DL-basierter Bild- und Rohdatenverarbeitung.
- **T2.4:** Qualitätsbewertung von rekonstruierten CT-Scans mit adversen Netzwerken.
- **T2.5:** Integration einer nicht-lokalen, prior-bildbasierten Regularisierung und implantatabhängiger Gewichtung.
- **T2.6:** Transferlernen zwischen Modalitäten und Auflösungen.

Deliverables:

- **M2.1:** Erstes Werkzeug für Bildverbesserung mit CNN basierter Nachverarbeitung
- **M2.2:** Rekonstruktionsalgorithmus für die Reduktion von Artefakten / Rauschen für in ex und in vivo Bilddaten von labCT.
- **M2.3:** Verbesserte Auflösung und Qualitätskontrolle der Rekonstruktion in verschiedenen Modalitäten

WP3: Segmentierung (IMT 2PM, IMI 3PM)

Lieferobjekte:

- **M3.1:** Erfolgreicher Transfer der entwickelten Algorithmen zur Segmentierung über Modalitäts- und Domänengrenzen hinweg
- **M3.2:** Erfolgreiche Segmentierung der Daten mit Quality-Assessment, und Anwendung auf Projekt-externe Daten

WP4 Registrierung (1 PM IMT, 10 PM IMI)

Teilpakete:

- **T4.1:** Entwicklung von DL-basierten Registrierungsmethoden, die semantische Orientierungshilfen aus Segmentierungsnetzwerken und Landmark-Lokalisierungsaufgaben beinhalten.

- **T4.2:** Erlernen modalitätsinvarianter Merkmale für SR μ CT, SAXS, Histologie, In-vivo-MRT und CT.
- **T4.3:** Anwendung von Formkodierungsmethoden zur Definition einer kanonischen anatomischen Koordinate.

Lieferobjekte:

- **M4.1:** Lernbasiertes Registrierungswerkzeug für in- und ex-vivo Daten von SR μ CT, SAXS und Histologie
- **M4.2:** Evaluation der Registrierungsgenauigkeit und Verbesserungen für in- und ex-vivo Tierdaten (MRT, CT, Histologie)
- **M4.3:** Registrierungswerkzeug für zeitliche und multimodale CT/MRT Bilder und Transfer der Gewichte (für Multitask)

WP5 Multi-task learning (4PM IMT, 12 PM IMI)

Teilpakete:

- **T5.1:** Identifizierung von generischen Strukturen im neuronalen Netz aus WP2-4.
- **T5.2:** Verbesserte Segmentierung und Registrierung von Ex-vivo-Tierdaten durch iterative multimodale Segmentierung.
- **T5.2:** Entwicklung eines Software-Tools für die Multi-Task-Analyse zur gleichzeitigen Bildverbesserung, Segmentierung und Registrierung.
- **T5.3:** DL-basierte multimodale Segmentierung mit CNNs unter Verwendung registrierter (WP4) Daten, die aus verschiedenen experimentellen Methoden stammen. Wir werden sowohl die Verwendung der registrierten Rohdaten, d.h. der Bilder, als auch die Verwendung der einzelnen Segmentierungen aus WP3 untersuchen, um bessere Segmentierungen zu erhalten.

Lieferobjekte

- **M5.1:** Grundlage der Softwareplattform für Multi-Task-Lernen (Bildverbesserung, Segmentierung, Registrierung)
- **M5.2:** Einsatz der Multi-task Analyse (Bildverbesserung/Segmentierung) für ex vivo Tierdaten von SR μ CT, SAXS und Histologie
- **M5.3:** Einsatz der Multi-task Analyse für ex vivo Tierdaten von MRT und labCT sowie für labCT ex vivo SR μ CT und in / ex vivo MRT / labCT. 30
- **M5.4:** Verwendung des Multi-task Lernens (Bildregistrierung/Segmentierung) für longitudinale MRT Daten

WP6: Unified Framework (IMT 2PM, IMI 2PM)

Lieferobjekte

- **M6.1:** Entwicklung einer vereinheitlichten Plattform für die domänen- und modalitätsinvariante Multi-task Analyse.
- **M6.2:** Inbetriebnahme der vereinheitlichten Plattform am Maxwell Rechencluster am DESY.
- **M6.3:** Leistungsevaluierung der vereinheitlichten Plattform für in WP2-5 nicht betrachtete Domänen und Modalitäten.

Begründung der Abweichungen zum Projektplan:

WP2: Bildrekonstruktion und Verbesserung

Die Teilpakete T2.1, T2.3, T2.4, T2.5 und T2.6 konnten, wie bereits in den Zwischenberichten beschrieben, rechtzeitig abgeschlossen werden. Die direkte DL-basierte CT-Rekonstruktion (T2.2) wurde zu Beginn des Projekts auf einen späteren Bearbeitungszeitpunkt verschoben, da zwar plausible Rekonstruktionen erzeugt, aber die Konsistenz von Roh- und Rekonstruierten Daten durch etwaig eingebrachte Modell-Halluzinationen zum frühen Projektzeitpunkt (2020) nicht ausgeschlossen werden konnte. Das genannte Arbeitspaket wurde in 2022 mit der Möglichkeit zur Verwendung eines In-House-CT-Scanners begonnen, durch welchen ein Zugriff auf akquirierte Rohdaten möglich war. Eine abschließende Bewertung der DL-generierten Rekonstruktion wurde bis zum Abschluss des Projekts aus zeitlichen Gründen (siehe unten) nicht mehr durchgeführt.

Der Meilenstein M2.2 "Rekonstruktionsalgorithmus für die Reduktion von Artefakten / Rauschen für in ex und in vivo Bilddaten von labCT." wurde daher abweichend zum ursprünglichen Plan auf einem anderen Datensatz (humane CT Studie) durchgeführt und somit zumindest die technische Zielstellung des Arbeitspakets erreicht. Es steht durch Bereitstellung von Open-Source Tools auch eine neue generell einsetzbare und effektive Rekonstruktionsmethodik für weitere ex- und in-vivo Anwendungen zur Verfügung.

WP 3: Segmentierung

Die Meilensteine M3.1 und M3.2 wurden wie geplant im Jahr 2021 bearbeitet und als Vorarbeit zu WP4 umfassend ausgearbeitet, weshalb die Bearbeitung von WP4 erst nach Abschluss der Konzeption der multimodalen Segmentierungsalgorithmen begonnen wurde. M5.1 und M5.2 wurden 2022 erfolgreich abgeschlossen.

WP 4: Multimodale Bildregistrierung

Die Meilensteine M4.1 und M4.2 wurden 2020 erfolgreich abgeschlossen. Die Entwicklung des Registrierungswerkzeugs für zeitliche und multimodale CT/MRT Bilder M4.3 wurde verspätet begonnen, konnte jedoch 2022 erfolgreich unter Integration der multimodalen Segmentierungsmodelle abgeschlossen werden.

WP 5: Multitask-Lernen

Als Kombination von Segmentierung und Registrierung wurden die Teilpakete des Multitask-Lernen noch nicht Ende des Projektjahres 2021, sondern erst zum Anfang des Jahres 2022 begonnen.

Vorarbeiten für die Meilensteine "Einsatz der Multi-task Analyse (Bildverbesserung/Segmentierung) für ex vivo Tierdaten von SR μ CT, SAXS und Histologie" (M5.2) und "Einsatz der Multi-task Analyse für ex vivo Tierdaten von MRT und labCT sowie für labCT ex vivo SR μ CT und in / ex vivo MRT / labCT. 30" (M5.3) wurden ebenfalls in 2022 begonnen, konnten jedoch bis zum Laufzeitende aufgrund des Austritts von Christian Kruse und der notwendigen Einarbeitungszeit seiner Nachfolgerin Hanna Siebert abweichend vom ursprünglichen Projektplan nicht auf Tier- sondern Humanbilddaten durchgeführt. Es konnte somit die technische Zielstellung des Arbeitspakts erreicht werden und eine weitere Nutzung für neue biomedizinische Fragestellungen mit Hilfe der entwickelten kombinierten Registrierungs- und Segmentierungsalgorithmen ist durch die öffentliche Bereitstellung von Open Source Algorithmen sichergestellt.

WP 6: Unified Framework

Aufgrund der zeitlichen Schwierigkeiten und der Personalwechsel und der Abhängigkeiten zu den vorangegangenen Arbeitspaketen konnten die Meilensteine M6.1, M6.2 und M6.3 nicht alle entsprechend dem Plan erfolgreich erreicht werden (s.u.). Alternativ, wurde in Zusammenarbeit mit weiteren wissenschaftlichen Mitarbeitenden der Universität zu Lübeck ein neues Domain Generalisation Ansatz u.A. zur Bildsegmentierung entwickelt der teilweise ähnliche Ziele erreicht, wie das ursprünglich geplante "unified framework". Im Speziellen ermöglicht der Algorithmus, welcher Open Source zur Verfügung gestellt wurde, die robuste Anpassung von Deep Learning Modellen, welche auf einem mit Expertenlabel versehenen Datensatz trainiert wurden, auf neue ungesehene Datensätze und Problemstellungen.

Zeitliche Herausforderungen im Projektverlauf

Neben Verzögerungen durch die Covid-19 Pandemie in 2020 wurden weitere Verzögerungen durch eine elternzeitbedingte Auszeit sowie einen Personalwechsel innerhalb des Projektzeitraums erzeugt. Aus den genannten Gründen konnten manche Teilarbeitspakete aus WP6 nicht vollumfänglich wie im ursprünglichen Antrag bearbeitet werden. Allerdings konnten hierfür alternative Erkenntnisse erarbeitet und Veröffentlichungen umgesetzt werden, die teilweise über die Planungen hinausgingen.

Projektergebnisse

Im Rahmen des Teilvorhabens der Uni Lübeck im MDLMA Projekt konnten eine Reihe von neuen Erkenntnissen im Bereich des Multi-Task Lernens insbesondere für biomedizinisch Bilddaten erreicht werden. Eine grundlegende Herausforderung die durch die Entwicklung von neuen Deep Learning Verfahren in Bildregistrierung, Bildsegmentierung und Bildrekonstruktion gemeistert werden konnte ist eine erhöhte Zuverlässigkeit, des Trainings mit weniger annotierten Daten bzw. der Transfer von bestehendem Wissen einer Domäne (z.B. Bildmodalität) auf eine neue innerhalb der Modellentwicklung. Durch eine Verknüpfung von Bildsegmentierungs- und

Bildregistrierungsalgorithmen ist es uns gelungen Verfahren zu entwickeln die eine Kombination von heterogenen Datensätzen erleichtert. Es konnten damit Merkmale und Wissen von einer zur anderen Modalität transferiert werden (Transfer-Learning).

Multi-Task Lernen ermöglicht insbesondere dann Vorteile da biomedizinischer Datensätze häufig zu klein, unvollständig, unausgewogen oder unzureichend annotiert sind. Somit konnten unsere entwickelten Methoden gemeinsame und komplementäre Merkmale von verschiedenen anatomischen Regionen und über verschiedene bildgebenden Modalitäten hinweg erfassen. Durch die Entwicklung von neuen Konzepten, welche z.B. eine automatische Anpassung der Teilnetzwerke auf einem kleinen neuen (ungelabelten) Testdatensätze, wobei die Modelle auf einer anderen Modalität und Anatomie u.A. unter Maßgabe einer intrinsischen Generalisierung der Domänen (domain generalisation), vortrainiert wurden, konnten wir zeigen dass das ursprüngliche Netzwerk grundlegendes Verständnis von geometrischen, anatomischen Strukturen unabhängig von der Modalität erlernt und somit gut geeignet für neue Aufgaben ist.

WP2: Bildrekonstruktion und Verbesserung

Im Bereich der Bildrekonstruktion und -verbesserung wurden zahlreiche neue Methoden und Verfahren entwickelt und evaluiert, die im Nachfolgenden ausführlicher beschrieben werden.

T2.1 - CNN-based approaches for post-processing artifact reduction:

Für einen CNN-basierte Ansatz konnte eine Schwellwert-basierte Erkennung von Metallen im Projektionsbild umgesetzt werden. Mittels eines in-painting Ansatzes, bei dem die fehlenden Informationen durch eine U-Net Architektur aufgefüllt wird, werden die Projektionsbilder anschließend vorverarbeitet. Die korrigierten Daten werden darauf folgend an etablierte und eigens implementierte Rekonstruktions- algorithmen übergeben. Die Güte dieses Ansatzes wurde sowohl im Projektionsbild als auch im rekonstruierten Bild überprüft. Als Training wurden initial simulierte Daten des bekannten

X-Cat-Phantoms verwendet. Hierbei hat sich herausgestellt, dass die Anwendung von partiellen Faltungen eine Verbesserung der Ergebnisse im Vergleich zu normalen Faltungen erzielen kann. Die Ergebnisse weisen große Potential auf, liegen in der erzielten Bildqualität aber noch hinter etablierten Interpolationsansätzen zurück. Erste Ergebnisse wurden bei der Konferenz Medical Imaging with Deep Learning eingereicht. Eine Rückmeldung steht noch aus. Der unter M2.1 aufgeführte Meilenstein ist erfüllt und die erzielten Ergebnisse geben Anlass dazu, die eingeschlagene Strategie weiterzuentwickeln.

Der unter M2.1 aufgeführte Meilenstein in Bezug auf dieses Arbeitspaket konnten im weiteren Projektverlauf vollständig erfüllt werden und die erzielten Ergebnisse wurden unter anderem auf der Konferenz Medical Imaging with Deep Learning veröffentlicht (Blum, N., Buzug, T. and Stille, M.: Projection Domain Metal Artifact Reduction in Computed Tomography using Conditional Generative Adversarial Networks, In: Medical Imaging with Deep Learning, 2021. und Hellwege, L., Blum, N., Buzug, T. M. and Stille, M.: Partial Convolution Network for Metal Artifact Reduction in CT Preprocessing: Preliminary Results, In: Medical Imaging and Deep Learning, 2021.).

Obwohl die zu erreichenden Meilensteine in der Teilaufgabe T2.1 bereits erfüllt waren, wurde auf Grund des hohen Potentials weiter an der Methodik gearbeitet. Um das Fortschreiten anderer Arbeitspakete jedoch nicht zu gefährden, wurde hier der Weg eingeschlagen, die Problemstellung mit Hilfe von Abschlussarbeiten zu bearbeiten. Die Ergebnisse zweier Abschlussarbeiten befanden sich zum Projektabschluss noch in der Auswertung. Sie befassten sich mit einer CNN-basierten Reduktion von Metallartefakten im Bildbereich.

T2.2: Direkte DL-basierte CT-Rekonstruktion.

Die Bearbeitung der Aufgabe T2.2 wurde basierend auf der Geometrie eines Inhaus-CT-Scanners erforscht. Dies hat den großen Vorteil, dass der Zugang zu den gemessenen Rohdaten vorliegt. Bei Abschluss der Aufgabe kann die Methodik auf die Geometrie eines anderen Scanners übertragen werden. Hier bedarf es in dem Fall jedoch eines erneuten Trainings unter der Verwendung von dem Scanner zuzuordnenden Messdaten. Da es sich bei der Aufgabe, wie in früheren Berichten bereits dargestellt, um eine sehr rechenintensive Fragestellung handelt, konnte im Verlaufe des Projekts lediglich eine Machbarkeitsstudie als Ergebnis erreicht werden. Die Bearbeitung des Problems war auf Grund der numerischen Komplexität erschwert. Um die Zielstellung zu ermöglichen, musste das Sampling der verwendeten Daten derzeit an den daraus folgenden Rechenaufwand angepasst werden.

T2.3 - Iterative reconstruction algorithm for artifact reduction in combination with DL-based image and raw data processing:

Für das Arbeitspaket wurde ein völlig neuartiger Rekonstruktionsalgorithmus entworfen, welche auf einem iterativen Schema basiert. In diesem wird in jeder Iteration ein Zwischenergebnis mit Hilfe eines neuronalen Netzes bearbeitet. Hierbei werden insbesondere initial entstandene Artefakte gefiltert und unterdrückt. Das Ergebnis dieses neuronalen Netzes wird anschließend in Rohdaten umgewandelt und für eine weitere Iteration genutzt. Die Struktur des Algorithmus sieht es vor, dass das Netzwerk für jede Iteration separat trainiert wird, da in jeder Iteration eine andere Ausprägung der Artefakte erwartet wird. Zunächst konnte dies für drei verschiedene Iterationen umgesetzt werden, wobei jeweils mit echten klinischen Daten trainiert wurde. Erste Ergebnisse konnten auf der 6th International Conference on Image Formation in X-Ray Computed Tomography vorgestellt werden. Die vielversprechenden Ergebnisse gaben Anlass, den Ansatz auch im weiteren Verlauf des Projektes weiterzuentwickeln. Hierzu wurden dann neben der bereits durchgeführten Analyse von verschiedenen Loss-Funktionen weitere Netzstrukturen und ggf. Transferlearning mit berücksichtigt.

T2.4 - Quality assessment of reconstructed CT scans with adversarial networks:

In der ersten Phase wurde an einer geeigneten Netzwerkstruktur gearbeitet. Im Laufe der Bearbeitung dieses Arbeitspaketes hat sich eine große Synergie zwischen der Problemstellung des Quality assessment und dem Ansatz aus T2.3 herausgestellt. Die hier entwickelte neue datenbasierte Methode zur Reduzierung von Metallartefakten in CT-Bildern wendet generative

adversariale Netzwerke auf die verfälschten Daten an. Ein generatives Netzwerk wird direkt auf die Projektionen, welche durch die Metallobjekte verlaufen, angewendet, um die korrigierten Sinogrammdaten zu lernen. Außerdem wurden zwei Diskriminatornetzwerke entwickelt, um die Bildqualität der verbesserten Daten aus dem Generator zu bewerten. Die Methode wurde zunächst auf der Grundlage eines überwachten Ansatzes entwickelt. Allerdings gibt es in der Regel keine Grundwahrheit für tatsächliche klinische Daten ohne Artefakte, die für das Training der Netzwerke genutzt werden können. Daher wurde die Methode weiter verbessert, um ein unüberwachtes Netzwerk zu trainieren, d.h. ohne die Verwendung der Grundwahrheit. Zusätzlich wurden die Eingabedaten erweitert indem benachbarte Schichten und die stochastischen Komponenten des Bildes unter Verwendung des sogenannten Latent Space mit einbezogen wurden. Die Ergebnisse zeigen, dass das trainierte Generatorkomplex die fehlenden Projektionsdaten angemessen ersetzen und die Artefakte im rekonstruierten Bild reduzieren kann. Die Ergebnisse hier werden auf der 7th International Conference on Image Formation in X-Ray Computed Tomography vorgestellt.

T2.5 - Integration of a non-local Prior Image-based regularization and implant depended weighting:

In diesem Arbeitspaket werden einzelne Bildausschnitte einer zu rekonstruierenden Anatomie mit einem Prior-Bild verglichen. Die Regularisierung ermittelt dabei eine geeignete Gewichtung zwischen dem Prior, als Ausgleich bzw. Korrektur und den Artefakt-behafteten Bilddaten. Um anatomische Unterschiede zwischen dem untersuchten Patienten und den Referenzdaten des Priors zu berücksichtigen, wird hier ein nicht lokaler Ansatz verfolgt, welcher in der Lage ist Unterschiede zu berücksichtigen. Große Teile dieses Ansatzes wurden bereits vollständig implementiert. Der Regularisierungsansatz konnte bereits in eine vorhandene Bildrekonstruktion integriert werden, wobei einzelne Aspekte der Methode, wie z.B. die Suchfunktion, durch eine Parallelisierung auf der GPU beschleunigt werden konnte. Erste Ergebnisse konnten auf der 6th International Conference on Image Formation in X-Ray Computed Tomography vorgestellt werden. Im weiteren Verlauf des Projektes wird der Ansatz optimiert, auf klinische Daten übertragen und mit verschiedenen anderen Rekonstruktionsansätzen kombiniert.

Die weiteren methodischen Erkenntnisse und experimentellen Ergebnisse konnten auf der 6th International Conference on Image Formation in X-Ray Computed Tomography vorgestellt werden (Blum, N., Stille, M., Hofmann, C. and Buzug, T. M.: Metal Artifact Reduction for 3D Cone-Beam CT by Non-Local Prior Image Integration, In: CT Meeting 2020 (The 6th International Conference on Image Formation in X-Ray Computed Tomography), 396– 400, 2020.).

Seitdem wurde die Methodik weiterentwickelt. Diesbezüglich wird auch nach Projektabschluss weiter an einer Journalpublikation gearbeitet.

T2.6 - Transfer learning between modalities and resolutions.

Bei diesem Arbeitspaket lag der Fokus auf dem Transfer von trainierten Netzwerken mit Softwarephantomen hin zu der Anwendung mit klinischen Daten. Erste Ergebnisse in Bezug auf die Bildrekonstruktion und -verbesserung waren bereits aussichtsreich. Der im T2.6 geplante Transfer-Learning-Ansatz wurde in enger Verflechtung mit den Arbeitspaketen T2.3 und T2.4

umgesetzt. Der Übergang von Software-Phantom als Datengrundlage hin zu der Integration von klinischen Messdaten. Unser Ziel war es, die Leistungsfähigkeit des Transfer-Lernens zu nutzen, um Wissen von einem gut trainierten Netzwerk auf ein neues Netzwerk mit anderen Eingabedaten zu übertragen. Wir erreichten dies, indem wir das bereits bestehende Netzwerk mit neuen medizinischen Bildern feinabstimmten, so dass das Netzwerk neue, für den klinischen Scanner spezifische Merkmale erlernen konnte. Unsere Ergebnisse deuten an, dass Transfer-Learning die Leistung der Artefaktkorrektur erheblich steigern kann und eine kosteneffiziente Lösung für das Training von Modellen auf begrenzten Datensätzen darstellt. Die Arbeit unterstreicht an dieser Stelle das transformative Potenzial von KI-basierten Techniken bei der Verbesserung der Genauigkeit und Effizienz der medizinischen Bildgebung und

-analyse, was sich letztendlich in besseren Ergebnissen für die Patienten niederschlägt. Der Aspekt wird in der zuvor erwähnten, geplanten Journalpublikation mit aufgenommen.

WP 3: Segmentierung

M3.1: Erfolgreicher Transfer der entwickelten Algorithmen zur Segmentierung über Modalitäts- und Domänengrenzen hinweg

Im Projektverlauf wurde zunächst das multimodale Lernen mittels generischer Merkmalsextraction untersucht. Dabei wurde sowohl das Lernen mittels Hilfsaufgaben als auch adversariales Lernen von gemeinsamen Merkmalen für verschiedene Domänen untersucht.

Hierbei konnten jedoch keine nennenswerte Vorteile gegenüber dem Stand der Technik erreicht werden. Im Weiteren wurde daher das Lernen auf klassischen Merkmalsextraktoren, insbesondere den "modality independent neighbourhood descriptors" (MIND) untersucht. Hierbei ließen sich große Fortschritte im Bereich der multimodalen Segmentierung erreichen. Es konnte gezeigt werden, dass für die Domänen CT und MRT die Segmentierung von Organen und Muskeln sehr gut funktioniert. Dabei werden keine Bilder aus einer spezifischen Zieldomäne verwendet, ein Modell funktioniert also zuverlässig auf MRT Bildern wenn im Training nur CT Bilder vorhanden waren und umgekehrt. Die Genauigkeit dieser Modelle reicht dabei an die von Modellen die auf der Zieldomäne trainiert wurden heran. Unsere Erkenntnisse wurden bei der SPIE medical imaging 2022 veröffentlicht: Christian N. Kruse and Mattias P. Heinrich "Bridging the domain gap for medical image segmentation with multimodal MIND features", Proc. SPIE 12032, Medical Imaging 2022: Image Processing, 1203231 (4 April 2022); <https://doi.org/10.1117/12.2612041>. Diese Methode wurde außerdem bei der MICCAI 2021 CrossMoDa Challenge eingesetzt bei der der Transfer von Gehirn-Segmentierungs Modellen zwischen verschiedenen MRT Kontrasten untersucht wurde. Dabei kam der MIND Ansatz als einziger ohne Daten in der Zieldomäne aus und konnte dennoch gut in die Zieldomäne übertragen werden. Die Ergebnisse der Challenge wurden bei Medical Image Analysis eingereicht (<https://arxiv.org/abs/2201.02831>).

M3.2: Erfolgreiche Segmentierung der Daten mit Quality-Assessment, und Anwendung auf Projekt-externe Daten

Desweiteren wurden neue Augmentierungsstrategien entwickelt, insbesondere die "randomfield augmentation" in der die Bilder zunächst mit einem geglätteten Zufallsfeld multipliziert werden und

auf dieses Ergebnis dann ein weiteres Zufallsfeld addiert wird. Durch diese Augmentierung geht im Training die Information von lokalen Intensitätsverhältnissen verloren und ein trainiertes Modell muss sich stärker auf die Bildkonturen fokussieren. Auch mit diesem Ansatz konnten wir Modelle trainieren die sich für den Transfer zwischen verschiedenen Domänen eignen. Diese Erkenntnisse wurden zur Veröffentlichung vorbereitet und sollen im März 2022 zur MICCAI eingereicht werden.

Sowohl das Lernen auf MIND Merkmalen und die Zufallsfeld-Augmentierung wurden in das nn-UNet (Isensee et al., 2020, Nature Methods, 1-9.) integriert, zur Zeit das beste öffentlich verfügbare Framework für medizinische Bildsegmentierung, um die bestmöglichen Ergebnisse zu erzielen. Zur Zeit werden beide Ansätze in diesem Framework für verschiedene Domänentransfer-Aufgaben getestet und mit state-of-the-art Baselines verglichen, wobei die ersten Ergebnisse vielversprechend waren.

Die Ergebnisse unserer zwei Beiträge zur MICCAI CrossMoDa Challenge für den unüberwachten Domain-Transfer wurden im sehr hochrangigen Journal Medical Image Analysis, Volume 83, (Impact Faktor 13.8) publiziert. <https://doi.org/10.1016/j.media.2022.102628>

WP 4: Multimodale Bildregistrierung

T4.1: Entwicklung von DL-basierten Registrierungsmethoden, die semantische Orientierungshilfen aus Segmentierungsnetzwerken und Landmark-Lokalisierungsaufgaben beinhalten.

Ausgehend von den Erfolgen für multimodale Segmentierung wurde im nächsten Schritt die multimodale Registrierung auf Basis der oben genannten Modelle untersucht. Dabei konnte in unseren laufenden Forschungsarbeiten gezeigt werden, dass sich z.B. die nn-UNet Merkmale gut für einen Registrierungsansatz mittels Convex-Adam Optimierer eignen (Siebert et al., MICCAI 2021, vol. 13166, https://doi.org/10.1007/978-3-030-97281-3_25). Dieser Ansatz wird im Weiteren auch in diesem Projekt integriert. Lern-basierte multi-task Registrierungs Ansätze wurden in der verbleibenden Projektlaufzeit untersucht.

Die direkter Kombination von Segmentierung und multimodaler Registrierung führten darüber hinaus zu einer weiteren hochrangigen Journal Publikation bei IEEE Transactions on Medical Imaging <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9925717> (Impact Faktor >11) bei der

Dieser Ansatz wurde und wird im Weiteren Verlauf des Projekts für die Anwendung auf den biomedizinischen Bilddaten der Projektbeteiligten (SR μ CT, SAXS, CT histology) untersucht und als Software Tool bei DESY integriert, sowie als Open-Source Bibliothek auf Github veröffentlicht: <https://github.com/multimodalllearning/convexAdam>

T4.2: Erlernen modalitätsinvarianter Merkmale für SR μ CT, SAXS, Histologie, In-vivo-MRT und CT und T4.3: Anwendung von Formkodierungsmethoden zur Definition einer kanonischen anatomischen Koordinate.

Für die multimodale Bildregistrierung wurde eine neue Methode entwickelt und erfolgreich auf 2D CT und MRT Bilddaten getestet. Die Methode basiert auf dem Konzept der "maximum classifier discrepancy" bei der zwei zufällig initialisierte Klassifizierungs Netzwerke verwendet werden um

ein Merkmalsextraktionsnetzwerk zu trainieren. Das Training ist dabei so gestaltet, dass die relevanten Merkmale aus beiden Bild Modalitäten einander angenähert werden (<https://arxiv.org/abs/1712.02560>). Im Rahmen dieses Projektes wurde dieser Ansatz um eine Metrik erweitert, die es erlaubt die Diskrepanz zwei Registrierung Netzwerke zu messen um so den Ansatz für medizinisch Bildregistrierung einsetzen zu können. Die Metrik basiert auf der Earth Movers distance, die auf 1D Linien projiziert wird um so die Lösung der Metrik zu stabilisieren und zu beschleunigen. Eine weitere Verbesserung der Registrierungsgenauigkeit konnte durch die Integration eines Optimierungsalgorithmus, des "Semi-Global matching", während der Inferenz erreicht werden. Erste Ergebnisse wurden bereits bei der nationalen Konferenz (Bildverarbeitung für die Medizin) veröffentlicht (<http://arxiv.org/abs/2005.14107>, https://link.springer.com/chapter/10.1007/978-3-658-33198-6_47).

Weiterhin wurde an einem neuen Ansatz zum selbstüberwachten Multitask-Lernen entwickelt und bei der renommierten MICCAI Konferenz 2023 eingereicht und zur Veröffentlichung angenommen (Akzeptanzrate 30%). Der Ansatz basiert auf der Kombination eines generischen Feature-Netzwerkes welches für zwei Eingabebilder (entweder aus der gleichen Modalität / Domäne oder unterschiedlichen) Merkmale extrahiert.

Zunächst werden die Netzwerkgewichte zufällig initialisiert, wobei jedoch bereits durch das induktive Bias der Faltungsschichten informative Bildinformationen ausgegeben werden. Diese werden anschließend in ein differenzierteres Registrierungs-Modul gegeben welches eine Transformation ausgibt. Auch wenn diese Transformation noch keine hohe Genauigkeit aufweist, so erlaubt sie zumindest schon eine bessere Angleichung der Bilder als deren initialer Zustand. Folglich können diese Transformationen über einen großen (ungelabelten) Trainingsdatensatz hinweg als verrauschte Pseudo-Label für einen überwachten Lernansatz genutzt werden der wiederum die Gewichte des Feature-Netzwerks verbessert und somit über mehrere Wiederholungen zu exzellenten Ergebnissen führt. Da dem Netzwerk mit Hilfe der Registrierungsaufgabe selbstüberwacht semantisches Wissen beigebracht wird kann es in einem weiteren Schritt auch z.B. zur Lösung von Segmentierungsproblemen mit sehr wenig Trainingsdaten (few shot multitask learning) eingesetzt werden.

- Bigalke, A., Hansen, L., Mok, T. C., & Heinrich, M. P. (2023, October). Unsupervised 3D registration through optimization-guided cyclical self-training. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 677-687). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43999-5_64

Die Meilensteine aus AP4 zum Einsatz von lernbasierten Registrierungsmodellen in Verbindung mit Segmentierungsmodellen M4.3 konnte abgeschlossen werden. Weiterhin wurden die Ziele für die Meilensteine M5.1 und M5.4 aus AP5 mit der Publikation und Code-Veröffentlichung der multimodalen Multitask (Registrierung+Segmentierung) Modelle erreicht werden.

WP 5: Multitask-Lernen

T5.1: Identifizierung von generischen Strukturen im neuronalen Netz aus

Es konnte in mehreren verschiedenen Anwendungen die Nutzung vergleichbarer Strukturen der neuronalen Netzwerke erfolgreich umgesetzt werden, z.B. eignen sich 3D U-Net Modelle sowohl in Registrierungs-, als auch Rekonstruktions- und Segmentierungs-Tasks. Für weitere Details sei auf die jeweiligen Beschreibungen der Unterarbeitspakete verwiesen.

T5.2: Verbesserte Segmentierung und Registrierung von Ex-vivo-Tierdaten durch iterative multimodale Segmentierung.

Da selbst mit generalistischen Multitask-Modellen in der Praxis weiterhin Probleme der fehlenden Domänen-Anpassung auftreten. Haben wir in der Arbeitsgruppe von Prof. Heinrich im weiteren Verlauf eine neue Methode zur iterativen Verbesserung von Segmentierungsergebnissen auf neuen Daten einer anderen Domäne/Modalität umgesetzt. Zur weiteren Verbesserung der Anpassung von (generalisierten) Segmentierungsmodellen auf ungesehene Testdaten kann dieses neu-entwickelte Verfahren mit Hilfe der Test-Time-Adaptation die Gewichte des neuronalen Netzwerks mit Hilfe eines Konsistenz-Maßes für unterschiedliche Augmentierungen der gleichen Eingabedaten anpassen. Das Verfahren wurde ausführlich auf unterschiedlichen Bilddaten der Biomedizin evaluiert. Diese Arbeit reduziert die Lücke in der Genauigkeit von vollständig überwacht und unüberwacht trainierten Modellen noch einmal deutlich und ist aktuell bei IEEE Transactions on Medical Imaging im Review:

- Weihsbach, C., Kruse, C. N., Bigalke, A., & Heinrich, M. P. (2023). DG-TTA: Out-of-domain medical image segmentation through Domain Generalization and Test-Time Adaptation. arXiv preprint arXiv:2312.06275. <https://arxiv.org/abs/2312.06275>

Das Verfahren steht außerdem als Source-Code und eigenständiges pip-Modul der Allgemeinheit zur Nutzung zur Verfügung: <https://github.com/multimodallearning/DG-TTA>

Außerdem wurden Arbeiten für M4.2, M5.2 und M5.3 umgesetzt. Eine erste Auswertung der Registrierungs- und Segmentierungsgenauigkeit auf Tierdaten (Mausmodell) wurde in der ersten Jahreshälfte 2023 umgesetzt jedoch nicht mehr in einem wissenschaftlichen Beitrag veröffentlicht.

T5.2: Entwicklung eines Software-Tools für die Multi-Task-Analyse zur gleichzeitigen Bildverbesserung, Segmentierung und Registrierung.

Der Ansatz zur Multimodalen Bildregistrierung zu WP 4 hat bisher den Nachteil, dass nur die für die Registrierung entscheidenden Merkmalsverteilungen angenähert werden. Für das in diesem AP angestrebte Multitask Lernen ist es daher notwendig die Merkmalsverteilungen insgesamt anzunähern. Das Lernen von Kontextinformationen aus Hilfsaufgaben wie zum Beispiel der Positionsschätzung von Bildausschnitten kann hier hilfreich sein. Im Projekt wurde der Ansatz verfolgt allgemeinere Merkmalsverteilungen zu lernen und anzunähern und diese ggf. in Kombination mit anderen Merkmalsverteilungen (eg. "Modality Invariant Neighbourhood descriptors") einzusetzen.

Eine Übersicht über die verschiedenen Software-Tools die erarbeitet und veröffentlicht wurden ist unter dem Kapitel *Veröffentlichungen* zusammenfasst.

In Zusammenarbeit beider Arbeitsgruppen der Uni Lübeck wurde außerdem die Kombination und Evaluation von DL-unterstützter Rekonstruktionsverfahren von CT Scans von Knochen mit

Bildgebungsartefakten und nachfolgenden Segmentierungsansätzen diskutiert. Damit kann in Zukunft der Vorteil beim Zusammenspiel verschiedener Multitask-Elemente noch stärker verdeutlicht werden.

T5.3: DL-basierte multimodale Segmentierung mit CNNs unter Verwendung registrierter (WP4) Daten, die aus verschiedenen experimentellen Methoden stammen.

Im Rahmen der oben erwähnten multimodalen CrossMoDA Segmentierungschallenge wurde die angedachten Strategie zur Verwendung registrierter Daten für die Segmentierung von Daten einer ungesesehenen Modalität ausführlich untersucht und evaluiert. Der vorgeschlagene Ansatz basiert auf einer robusten deformierbaren multimodalen Multi-Atlas-Registrierung zur Überbrückung der Bereichslücke zwischen T1- (Quelle) und T2- (Ziel) gewichteten MRT-Scans. Der Quell- und der Zielbereich werden mit einer isotropen Auflösung von 1 mm neu abgetastet und mit einem ROI der Größe $64 \times 64 \times 96$ Voxel innerhalb der linken und rechten Hemisphäre beschnitten. 30 Trainingsbilder werden nach dem Zufallsprinzip ausgewählt und automatisch mit einer Teilmenge der Trainingsscans des Ziels sowohl linear als auch nicht starr registriert. Die Registrierung erfolgt mit dem diskreten Optimierungsverfahren deeds (Heinrich et al., 2013) mit multimodalen Merkmalsdeskriptoren (MIND-SSC (Heinrich et al., 2012)). Die propagierten Label werden mit dem bekannten STAPLE-Algorithmus fusioniert. Für eine schnelle Inferenz wurde dann anschließend ein nnU-Net-Modell auf den rauschbehafteten Segmentierungen in der Zielmodalität trainiert. Auf der Grundlage der vorhergesagten Segmentierungen wird ein automatischer Zentrumsausschnitt von $48 \times 48 \times 48$ mm mit einer Auflösung von 0,5 mm gewählt. Der beschriebene Prozess (Multi-Atlas-Registrierung, Label-Propagation und Fusion mit STAPLE, nnU-Net-Training) wurde dann für die verfeinerten Ausschnitte wiederholt. Das Verfahren erreichte eine sehr gute Genauigkeit für Tumour und gesunde Anatomien von ca. 75-80%.

WP 6: Unified Framework

M6.1: Entwicklung einer vereinheitlichten Plattform für die domänen- und modalitätsinvariante Multi-task Analyse.

Die einzelnen Aspekte der domänen- und modalitätsinvariante Multi-task Analyse wurden bereits in den vorangegangenen Abschnitten im Detail erläutert. Es lässt sich zusammenfassend schlussfolgern, dass zwar bestimmte komplementäre Aufgaben (Registrierung/Segmentierung) ein großes Potential haben von einem gemeinsamen Training zu profitieren. Jedoch die Performanz spezifischer Einzellösungen mit entsprechendem Fine-Tuning in Vergleichswettbewerben meist noch höher ist. Die Vorteile von Multi-Task Methoden zeigen sich jedoch in ihrer hohen Robustheit und guten Qualität der erreichten Ergebnisse beim Training mit wenigen Daten mit Expert-Labels.

M6.2: Inbetriebnahme der vereinheitlichten Plattform am Maxwell Rechencluster am DESY.

Diese Teilaufgabe lag maßgeblich bei den Projektpartnern des DESYs und ist im entsprechenden Abschlussbericht detailliert erläutert.

M6.3: Leistungsevaluierung der vereinheitlichten Plattform für in WP2-5 nicht betrachtete Domänen und Modalitäten.

In den Detailbeschreibungen von WP3-WP5 wurden bereits eine Vielzahl der über den ursprünglichen Projektantrag hinausgehenden und eingesetzten Datensätze für die Leistungs-Evaluierung unserer Algorithmen erwähnt und näher diskutiert.

Verwertung

Bisher wurden keine Erfindungen und Patente im Rahmen des Projekts angemeldet bzw. in Anspruch genommen. Wie weiter oben erwähnt, wurden jedoch zahlreiche Softwaretools entwickelt und als Open Source Code der wissenschaftlichen Community und Öffentlichkeit zur Verfügung gestellt. Die Verwertung als Universität erfolgte, wie im Antrag angegeben, primär über die Veröffentlichung von hochrangigen Publikationen und Präsentationen auf wissenschaftlichen Konferenzen. Somit gab es keine wesentlichen Änderungen des Verwertungsplans.

Für einen wissenschaftlichen und wirtschaftlichen Anschluss bieten sich aus unserer Sicht folgende Möglichkeiten:

- 1) wissenschaftliche Weiterentwicklung im Rahmen anderer Forschungsvorhaben
- 2) Verwertung der Verfahren im Rahmen von Dienstleistungen oder Ähnlichem
- 3) Lizenzierung bzw. Produktentwicklung im Rahmen einer Ausgründung

Zu Punkt 1): Es gibt bereits mehrere laufende Forschungsprojekte an der Universität zu Lübeck, die zumindest teilweise auf den erarbeiteten Ergebnissen und Erkenntnissen aufgebaut haben und somit einen wissenschaftlichen Anschluss ermöglichen. Jedoch gab es bisher nach unserem Wissen keine geeignete Ausschreibung für Folgeprojekte. Sofern sich dies in Zukunft ändert werden wir eine erneute Antragsstellung in Erwägung ziehen.

Zu Punkt 2): Derzeit ist noch nicht absehbar, inwieweit sich die entwickelten Softwaretools in der wissenschaftlichen Community (Biomedizin) durchsetzen. Sofern bzw. sobald eine wesentliche Anzahl an Nutzer:innen die Algorithmen einsetzt, würde sich die Möglichkeit von einer wirtschaftlichen Verwertung durch Dienstleistungen (Training und Services) der Tools anbieten. Dies ist jedoch aktuell noch nicht konkret geplant.

Zu Punkt 3): Auf Grund des erfolgreichen Projektverlaufs ist es grundsätzlich denkbar, die entwickelten Registrierungs-, Segmentierungs- und Rekonstruktionsalgorithmus im Rahmen einer Multi-Task-Learning-Plattform z.B. durch eine EXIST-gestützten Initiative zu kommerzialisieren ist. Universität zu Lübeck ist EXIST-Gründerhochschule und hat eine exzellente Infrastruktur zur Begleitung solcher Technologie-orientierten Ausgründungen. Jedoch ist es uns bisher noch nicht gelungen ein passendes Gründerteam zu finden. Die Projektbeteiligten arbeiten intensiv mit der Fraunhofer IMTE Einrichtung für personalisierte Medizintechnik zusammen, hierdurch ist in Zukunft ggf. eine Lizenzierung der entwickelten Techniken durch Industriepartner denkbar.

Softwaretools welche eine niederschwellige Nutzung der entwickelten Methoden für neue Anwendungsprobleme ermöglichen wurden öffentlich bereitgestellt:

<https://gitlab.com/maik.stille/partconvmar>

<https://gitlab.com/maik.stille/ganmar>

<https://github.com/multimodalllearning/DG-TTA>

<https://github.com/multimodalllearning/convexAdam>

Darüber hinaus gibt es die Möglichkeit Videos einiger wissenschaftlicher Präsentationen anzuschauen:

Blum et al. <https://www.youtube.com/watch?v=etLZHYVtkU4>

Hellwege et al. <https://www.youtube.com/watch?v=XD9THahFQrY>

Heinrich et al. https://www.youtube.com/watch?v=yfj_hAg89rg

Heinrich et al. <https://cloud.imi.uni-luebeck.de/s/z7zJasARZRrYAZE>

Für weitere Nutzung durch informatik-ferne Anwendergruppen aus den Lebenswissenschaften sei auf die Projektpartner Hereon und DESY verwiesen.

Wichtigsten Positionen des zahlenmäßigen Nachweises

Auf Seiten des Instituts für Medizintechnik haben zur gesamten Projektlaufzeit Maximilian Wattenberg, MSc. und Dr. rer. nat. Maik Stille mit jeweils einer 50% TV-L E13 Stelle gearbeitet. Am Institut für Medizinische Informatik hat hauptsächlich Dr. Christian Kruse zu 100% TV-L E13 als wissenschaftlicher Mitarbeiter gearbeitet. Zeitweise waren Hellena Hempe, M.Sc. und Hanna Siebert, M.Sc. in Teilzeit beschäftigt, um Elternzeiten und verspätete Einstellung aufzufangen. Es wurden wie im Projektantrag erläutert Mittel für Konferenzreisen (Medical Imaging with Deep Learning, MIDL) sowie Open Access Publikationen verausgabt. Für die technische Entwicklung wurden mehrere Grafikkarten angeschafft, die in einem aus anderen Mitteln beschafften GPU-Server installiert wurden.

Konkurrierende Ansätze

Unserer Kenntnis nach gab es im Verlauf des Projekts keine Veröffentlichung von neuen Erkenntnissen, die eine substantielle Änderung des Projektvorhabens nötig gemacht hätten. Wir verfolgten jedoch aufmerksam die Veröffentlichungen in internationalen Konferenzen und Fachzeitschriften und haben einige der dort veröffentlichten Verfahren bereits im Kontext dieses Projektes ausprobiert (z.B. Chen, C., et al. (2021). MICCAI 2021, vol 2903. Springer, https://doi.org/10.1007/978-3-030-87199-4_14). Wir konnten jedoch keine essentiellen Vorteile gegenüber unseren Ansätzen erkennen, sodass wir keine weitreichenden Änderungen des Projektvorhabens für nötig befunden haben.

Es lässt sich jedoch allgemein feststellen, dass der Bereich des Multitask-Lernens nochmals deutlich an Bedeutung gewonnen hat und für andere Aufgabenstellungen, zum Beispiel im Bereich der Computer Vision oder dem selbstüberwachten Lernen, inzwischen das Training von

Netzwerken, welche verschiedene Tasks und Domänen gemeinsam betrachten, häufiger untersucht wurde. Als Beispiel sei hier u.A. zu nennen:

Sun, Ximeng, Rameswar Panda, Rogerio Feris, and Kate Saenko. "Adashare: Learning what to share for efficient deep multi-task learning." *Advances in Neural Information Processing Systems* 33 (2020): 8728-8740.

Hu, Ronghang, and Amanpreet Singh. "Unit: Multimodal multitask learning with a unified transformer." In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1439-1449. 2021.

Veröffentlichungen

Im Rahmen des Teilprojekts konnten zahlreiche technische und wissenschaftliche Erfolge erzielt werden, die zu verschiedenen Publikationen und Präsentationen bei internationalen Konferenzen geführt haben. Es wurden Ergebnisse zu neuen CNN-basierten Ansätzen zur Artefaktreduktion mit Hilfe der Nutzung partieller Faltungen in der Bildrekonstruktion bei der Konferenz Medical Imaging with Deep Learning als peer-review Proceedingsartikel veröffentlicht:

- Hellwege, L., Blum, N., Buzug, T. M. and Stille, M.: Partial Convolution Network for Metal Artifact Reduction in CT Preprocessing: Preliminary Results, In: *Medical Imaging and Deep Learning*, 2021.)

Weiterhin wurden wie geplant Ansätze von konditionierten generativen Modellen zur Regularisierung der artefakt-reduzierenden Bildrekonstruktion direkt in der Domäne der physikalischen Bildaufnahme untersucht und als weiterer Konferenzartikel publiziert:

- Blum, N., Buzug, T. and Stille, M.: Projection Domain Metal Artifact Reduction in Computed Tomography using Conditional Generative Adversarial Networks, In: *Medical Imaging with Deep Learning*, 2021.

Für beide Ansätze wurde Open Source Code zur Steigerung sekundärer Verwertung und besserer Reproduzierbarkeit der wissenschaftlichen Ergebnisse und Erkenntnisse entsprechend des Open Science Grundsatzes veröffentlicht:

<https://gitlab.com/maik.stille/partconvmar>

<https://gitlab.com/maik.stille/ganmar>

Zur multi-domänen Bildsegmentierung wurden mehrere Methoden entwickeln, welche es ermöglichen ein Deep Learning Modell auf einer Modalität mit Expertenannotationen ohne vorherige Kenntnis der Zieldomäne zu trainieren. Durch eine geeignete Generalisierungs-Methode erlaubt es dann den Einsatz dieses Modells auf einer neuen Modalität. Arbeiten hierzu wurden einerseits bei der SPIE medical imaging 2022 veröffentlicht:

- Christian N. Kruse and Mattias P. Heinrich "Bridging the domain gap for medical image segmentation with multimodal MIND features", *Proc. SPIE 12032, Medical Imaging 2022: Image Processing*, 1203231 (4 April 2022); <https://doi.org/10.1117/12.2612041>.

Außerdem wurde ein Beitrag zur CrossMoDA Challenge erarbeitet, bei der der Transfer von Gehirn-Segmentierungs Modellen zwischen verschiedenen MRT Kontrasten untersucht wurde. Dabei kam der genannte MIND Ansatz als einziger ohne Daten in der Zieldomäne aus und konnte dennoch gut in die Zieldomäne übertragen werden. Die Ergebnisse der Challenge wurden bei Medical Image Analysis veröffentlicht:

- Dorent, Reuben, et al. "CrossMoDA 2021 challenge: Benchmark of cross-modality domain adaptation techniques for vestibular schwannoma and cochlea segmentation." Medical Image Analysis 83 (2023): 102628. (<https://arxiv.org/abs/2201.02831>)

Zur weiteren Verbesserung der Anpassung von (generalisierten) Segmentierungsmodellen auf ungesehene Testdaten wurde ein Verfahren der Test-Time-Adaptation entwickelt und ausführlich evaluiert. Diese Arbeit reduziert die Lücke in der Genauigkeit von vollständig überwacht und unüberwacht trainierten Modellen noch einmal deutlich und ist aktuell bei IEEE Transactions on Medical Imaging im Review:

- Weihsbach, C., Kruse, C. N., Bigalke, A., & Heinrich, M. P. (2023). DG-TTA: Out-of-domain medical image segmentation through Domain Generalization and Test-Time Adaptation. arXiv preprint arXiv:2312.06275. <https://arxiv.org/abs/2312.06275>

Das Verfahren steht außerdem als Source-Code und eigenständiges pip-Modul der Allgemeinheit zur Nutzung zur Verfügung: <https://github.com/multimodallearning/DG-TTA>

Die Integration eines nichtlokalen Priors zur bild-basierten Regularisierung der Bildrekonstruktion und einer Gewichtung der Messwerte unter Berücksichtigung von Metallimplantaten führte zu Ergebnissen die auf der 6th International Conference on Image Formation in X-Ray Computed Tomography vorgestellt wurden

- Blum, N., Stille, M., Hofmann, C. and Buzug, T. M.: Metal Artifact Reduction for 3D Cone-Beam CT by Non-Local Prior Image Integration, In: CT Meeting 2020 (The 6th International Conference on Image Formation in X-Ray Computed Tomography), 396– 400, 2020.

Im Bereich der multimodalen Bildregistrierung konnte eine Reihe von weiteren hochrangigen Publikationen durch die Erkenntnisse aus diesem Projekt realisiert werden. Zum einen wurde ein Verfahren der domain adaptation auf Grundlage des Maximum Classifier Discrepancy Ansatzes entwickelt und auf medizinischen Bilddaten evaluiert:

- Kruse, C.N., Hansen, L., Heinrich, M.P. (2021). Multi-modal Unsupervised Domain Adaptation for Deformable Registration Based on Maximum Classifier Discrepancy. In: Palm, C., Deserno, T.M., Handels, H., Maier, A., Maier-Hein, K., Tolxdorff, T. (eds) Bildverarbeitung für die Medizin 2021. Informatik aktuell. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-33198-6_47

Außerdem wurde das Erlernen von Merkmalen zur Bildregistrierung in einem selbstüberwachten Multi-Task Setting in der MICCAI Publikation von Alexander Bigalke untersucht:

- Bigalke, A., Hansen, L., Mok, T. C., & Heinrich, M. P. (2023, October). Unsupervised 3D registration through optimization-guided cyclical self-training. In International Conference on

Medical Image Computing and Computer-Assisted Intervention (pp. 677-687). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43999-5_64

Weiterhin konnten unseren Forschungsarbeiten zeigen, dass sich die Verknüpfung von Segmentierungsnetzen mit optimierungsbasierten Registrierungsansätzen, z.B. mittels des Convex-Adam Optimierers, sehr gut umsetzen lassen. Die Ergebnisse führten zu Bestwerten in einer internationalen Vergleichs-Challenge, welche sowohl als Konferenzbeitrag als auch Journalpublikation bei IEEE Transactions on Medical Imaging (Impact Faktor 11.1) veröffentlicht wurden:

- Siebert, H., Hansen, L., Heinrich, M.P. (2022). Fast 3D Registration with Accurate Optimisation and Little Learning for Learn2Reg 2021. In: Aubreville, M., Zimmerer, D., Heinrich, M. (eds) Biomedical Image Registration, Domain Generalisation and Out-of-Distribution Analysis. MICCAI 2021. Lecture Notes in Computer Science(), vol 13166. Springer, Cham. https://doi.org/10.1007/978-3-030-97281-3_25
- Hering, A., Hansen, L., Mok, T. C., Chung, A. C., Siebert, H., Häger, S., ... & Heinrich, M. P. (2022). Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. IEEE Transactions on Medical Imaging, 42(3), 697-712. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9925717>

Die Softwaremethoden wurden außerdem als Open Source Code bereitgestellt: <https://github.com/multimodalllearning/convexAdam>



MDLMA - Multimodale Datenanalyse mit Multi-Task Deep Learning - Teilprojekt B

Aufgabenstellung Ziel des Projektes MDLMA war die Verknüpfung und das Lernen von heterogenen und multimodalen Bilddaten mit Hilfe von lernbasierten Verfahren und klassischen Verfahren der Bildverarbeitung. Hierzu gehörten unter anderem die Untersuchung von bestimmten Klassen von Biomaterialien mittels Labor-Röntgen-Computertomografie (LabCT), Sychrotronstrahlungs-Mikrocomputertomografie (SR μ CT), Magnetresonanztomografie (MRT), Kleinwinkel-Röntgenstreuung (SAXS) und histologische Bilddaten, sowie die Analyse von Phantomen und klinischen Daten durch 3D-MRT und 3D-CT Daten. Um die Herausforderung unterschiedlicher räumlicher und zeitlicher Auflösungen, Dimensionalitäten und modalitätsspezifischer Erscheinungsbilder meistern zu können, mussten unterschiedlichste Teilaspekte gelöst werden. Dies umfasst die Analyse und Extraktion relevanter Information, durch eine Qualitätsverbesserung und Objektsegmentierung. Sowie die Kombination mehrerer separat verarbeiteter Datensätze, z.B. Bilder verschiedener Modalitäten, Segmentierungen und Ergebnisse der Bildverbesserung durch Verfahren der Bildregistrierung.

Arbeitsplan und Zielstellungen:

Die Hauptziele und Motivation dieses Projekts waren:

1. Das Lernen auf heterogenen, multimodalen Bilddaten und eine komplementäre Kombination von Deep Learning und mathematischer Modellierung.
2. Die Verbesserung der iterativen CT-Rekonstruktion und der multimodalen Artefaktreduktion.
3. Die Entwicklung neuartiger Algorithmen zur robusten multimodalen Bildregistrierung mit Deep Learning: Vereinfachung von komplexen Netzwerken durch Entkopplung des Repräsentationslernens mittels eines geometrischen Transformationsteils.
4. Die Entwicklung eines Konzeptes für das Multi-Task Lernen, das insbesondere auf multiskalige und multimodale biomedizinische Daten anwendbar ist.

Neben dem Institut für Medizinische Informatik (IMI, Prof. Dr. Mattias Heinrich) und dem Institut für Medizintechnik (IMT, Prof. Dr. Thorsten Buzug und Dr. Maik Stille) an der Uni Lübeck waren im Verbund auch DESY (Dt. Elektronen-Synchrotron), das Helmholtz-Zentrum Hereon und die Syntellix AG vertreten.

Wissenschaftlicher und technischer Stand an dem angeknüpft wurde

Das IMT und das IMI der Universität zu Lübeck hatten bereits in vorherigen Projekten Expertise im Feld des maschinellen Lernens, der Bildsegmentierung und -registrierung sowie bei Rekonstruktion von medizinischen Bilddaten aufgebaut. Insbesondere für Techniken und Methoden des Deep Learnings für Anwendungen auf inverse Probleme und Bildanalyse mit Computertomographie und Magnetresonanztomographie

Die Projektpartner am Hereon deckten Kenntnisse der Abteilung metallischer Biomaterialien sowie Legierungsdesign, Produktion, Abbaustudien, Zellkultur und Tierversuche ab. Schwerpunkt ist die Korrelation von Materialeigenschaften, Umwelteinflüssen und zellulären Reaktionen, mit Expertise in Bildgebung und Implantatforschung. Das DESY hat in den letzten Jahren seine Aktivitäten in der Entwicklung und Implementierung von lernbasierten Algorithmen erhöht. Dazu gehört die Segmentierung von SR μ CT Daten mit Hilfe von neuronalen Netzwerken. Maschinelles Lernen ermöglicht u.A. eine automatisierte Segmentierung, indem es die optimale Entscheidungsgrenze im hochdimensionalen Raum der diskriminanten Merkmale lernt und dabei gleichzeitig Merkmale extrahiert, die die Daten für das jeweilige Problem optimal repräsentieren. CNNs können die Bildsegmentierung in einer sehr effizienten Weise durchführen. Das U-Net ist eine der bekanntesten Architekturen für die Segmentierung, die für End-to-End-Lernen von Gewebesegmentierungen und -grenzen geeignet ist. Die größte Herausforderung liegt oft in der Verfügbarkeit ausreichend hochwertiger und gut annotierter Trainingsdaten, die oft durch intensive Datenaugmentation angegangen werden.

Im Vergleich zu anderen Bereichen wurde bisher relativ wenig Forschung zur bildbasierten Registrierung mit Deep Learning betrieben. Die meisten dieser Untersuchungen wurden in den letzten beiden Jahren veröffentlicht. Die vorgestellten Methoden zielen hauptsächlich darauf ab, eine Funktion in Form eines CNNs zu lernen, die eine räumliche Deformation vorhersagt, welche ein sogenanntes bewegliches Bild zu einem

festen Bild verformt. Je nach Art des Trainings der Netzwerke lassen sich diese Ansätze in vollständig überwachte, unüberwachte und schwach überwachte Methoden kategorisieren.

In den letzten Jahren hat Multi-Task-Learning (MTL) im Bereich Computer Vision zunehmend an Bedeutung gewonnen, da tiefe neuronale Netzwerke mit ihren vielfältigen Architekturen das visuelle Verständnis als zentrales Element nutzen können, um mehrere Bildanalyseaufgaben gleichzeitig zu bearbeiten. In der medizinischen Domäne sind umfassende Multi-Task-Deep-Learning-Ansätze noch selten. Transfer-Learning für Domänenanpassung wurde intensiv für die Pathologiedetektion und -segmentierung untersucht. MTLs wurden u. A. angewendet, um Einzel- und Kreuzmodalitätsbildanalysen durchzuführen, z. B. um die Gadoliniumdosis in kontrastverstärkten MRTs zu reduzieren und CTs aus MRTs zu synthetisieren.

Planung und Ablauf des Vorhabens

Lern-Basierte Verfahren in Bezug auf inverse Probleme Im Verlauf des Vorhabens haben wir uns auf die Weiterentwicklung von Deep-Learning-Techniken im Bereich der CT-Rekonstruktion konzentriert, welche zu vielversprechenden Ergebnissen geführt hat. Insbesondere bei Niedrigdosis- und supraauflösenden Bildern konnten wir feststellen, dass Convolutional Neural Networks (CNNs) das Potenzial zur Reduzierung von Artefakten und zur Verbesserung der Bildqualität aufweisen. Da die direkte Nutzung von Deep-Learning-basierten Rekonstruktionsalgorithmen für CT-Bildern in der Umsetzung problematisch und qualitativ herausfordernd war haben wir auch Ansätze bearbeitet die iterative Optimierungsmethoden mit trainierten neuronalen Netzwerken (z.B. Autoencodern) für die Bildrekonstruktion verbinden. Der Fokus lag hier insbesondere auf der Artefaktkorrektur, wobei sowohl überwachte, als auch unüberwachte Lernverfahren entwickelt wurden. Als technische Weiterentwicklung wurden CNN-basierten Ansätzen zur Artefaktreduktion mit Hilfe von partiellen Faltungen in der Bildrekonstruktion veröffentlicht [Hellwege 2021]. Weiterhin wurden Ansätze von konditionierten generativen Modellen zur Regularisierung der Bildrekonstruktion direkt in der Domäne der physikalischen Bildaufnahme untersucht und publiziert [Blum 2021]. Die Integration eines nicht-lokalen Priors zur bild-basierten Regularisierung der Bildrekonstruktion und einer Gewichtung der Messwerte bei Metallimplantaten führte zur einer weiteren Publikation [Blum 2020].

Segmentierung Zur Multi-Domänen-Bildsegmentierung wurden Methoden entwickelt, welche ein DL-Modell auf einer Modalität mit Expertenannotationen ohne Kenntnis der Zieldomäne trainieren. Durch eine geeignete Generalisierungs-Methode funktioniert dieses Modells auf einer neuen Modalität [Kruse 2022]. Außerdem wurde ein Beitrag zur CrossMoDA Challenge zum Transfer von Segmentierungs-Modellen zwischen MR-Kontrasten erarbeitet und veröffentlicht [Dorent 2023]. Zur Verbesserung der Anpassung auf ungesehene Testdaten wurde ein Verfahren der Test-Time-Adaptation entwickelt und evaluiert. Diese Arbeit reduziert die Lücke in der Genauigkeit deutlich und ist derzeit im Review [Weihsbach 2023].

Registrierung Für multimodalen Bildregistrierung konnten mehrer Publikationen realisiert werden. Zum einen wurde ein neues Verfahren der Domain Adaptation entwickelt [Kruse 2021], welches die Entscheidungsgrenze schrittweise von einer auf die andere Domäne überträgt. Das Erlernen von Merkmalen zur Bildregistrierung in einem selbstüberwachten Multi-Task Setting in einer MICCAI Publikation untersucht [Bigalke 2023].

Multitask-Learning Mit den vorliegenden Arbeiten konnten auch wichtige Schritte zur Multi-Task Analyse erreicht werden, da z.B. die CNNs für die Merkmalsextraktion in [Bigalke 2023] anschließend sowohl für die Registrierung als auch Segmentierung genutzt werden können. Weiterhin konnten wir zeigen, dass sich Segmentierungsnetzwerke gut mit optimierungsbasierten Registrierungsansätzen verknüpfen lassen. Dies führte zu Bestwerten in einer großangelegten Vergleichsstudie (Challenge) und wurde als Konferenz- und im Rahmen einer Journalpublikation veröffentlicht [Siebert 2022, Hering 2022].

Wesentliche Ergebnisse: Grundlegenden Herausforderungen in der Bildverarbeitung, insbesondere in den Bereichen der Bildregistrierung, Bildsegmentierung und Bildrekonstruktion, konnten durch die Entwicklung neuer Deep-Learning-Verfahren gemeistert werden. Die Fortschritte haben zur erhöhten Zuverlässigkeit der Modelle geführt, insbesondere in Bezug auf das Training mit einer begrenzten Menge an annotierten Daten. Darüber hinaus ermöglichte der Transfer von bestehendem Wissen aus einer Domäne in eine neue eine effizientere und effektivere Modellentwicklung. Durch die Kombination von Segmentierung und Registrierung haben wir Methoden entwickelt, die das Zusammenführen heterogener Datensätze vereinfachen, wodurch Merkmale und Wissen zwischen verschiedenen Modalitäten transferiert werden können. Somit konnten unsere entwickelten Methoden gemeinsame und komplementäre Merkmale von verschiedenen anatomischen Regionen und über verschiedene bildgebenden Modalitäten hinweg erfassen. Durch die Entwicklung innovativer Konzepte, die eine automatische Anpassung an kleine, ungelabelte Testdatensätze ermöglichen, wurden signifikante Fortschritte gemacht. Diese Modelle, vorab auf Datensätzen anderer Modalitäten trainiert, zeigen eine intrinsische Generalisierungsfähigkeit über verschiedene Domänen hinweg. Die Ergebnisse zeigen, dass diese Netzwerke ein tiefgehendes, modalitätsunabhängiges Verständnis für geometrische und anatomische Strukturen entwickeln, was sie ideal für neue medizinische

Bildgebungsaufgaben macht. Die entwickelten Algorithmen wurden größtenteils öffentlich verfügbar gemacht¹ welche vielseitig über verschiedene Datensätze, einschließlich Human-CTs und Kleintier-MRT einsetzbar sind. Inhaltlich wurden sowohl konzeptionell neue Deep Learning Ansätze, als auch der mathematischen Optimierung sowie der Regularisierung und Generalisierung von Verfahren auf nur teilweise überwachten Trainingsdaten untersucht.

Literaturverzeichnis

Publikationen die im Rahmen des MDLMA Projekts entstanden sind

- Hellwege, L., .. Buzug, T. M. and Stille, M.: Partial Convolution Network for Metal Artifact Reduction in CT Preprocessing, In: Medical Imaging and Deep Learning, 2021.
- Blum, N., Buzug, T. and Stille, M.: Projection Domain Metal Artifact Reduction in Computed Tomography using Conditional Generative Adversarial Networks, In: MIDL 2021.
- Kruse and Heinrich, MP: "Bridging the domain gap for medical image segmentation with multimodal MIND features", Proc. SPIE 12032, Medical Imaging 2022 <https://doi.org/10.1117/12.2612041>.
- Dorent, et al. "CrossMoDA 2021 challenge: Benchmark of cross-modality domain adaptation techniques for vestibular schwannoma and cochlea segmentation." Medical Image Analysis 83 (2023): 102628.(<https://arxiv.org/abs/2201.02831>)
- Weihsbach, C., ..Heinrich, M. P. (2023). DG-TTA: Out-of-domain medical image segmentation through Domain Generalization and Test-Time Adaptation. <https://arxiv.org/abs/2312.06275>
- Blum, N., .. Buzug, T. M.: Metal Artifact Reduction for 3D Cone-Beam CT by Non-Local Prior Image Integration, In: CT Meeting 2020 396– 400, 2020.
- Kruse, C.N., .. Heinrich, M.P. (2021). Multi-modal Unsupervised Domain Adaptation for Deformable Registration Based on Maximum Classifier Discrepancy. Bildverarbeitung für die Medizin 2021. https://doi.org/10.1007/978-3-658-33198-6_47
- Bigalke, A., ..Heinrich, M. P.. Unsupervised 3D registration through optimization-guided cyclical self-training.MICCAI 2023 https://doi.org/10.1007/978-3-031-43999-5_64
- Siebert, ..Heinrich, M.P. (2022). Fast 3D Registration with Accurate Optimisation and Little Learning for Learn2Reg 2021. MICCAI-W 2021https://doi.org/10.1007/978-3-030-97281-3_25
- Hering, A., ... & Heinrich, M. P. (2022). Learn2Reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. IEEE TMI <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=9925717>

¹ <https://gitlab.com/maik.stille/partconvmar>, <https://gitlab.com/maik.stille/ganmar>,
<https://github.com/multimodallearning/convexAdam>, <https://github.com/multimodallearning/DG-TTA>