

Die Internationale Radiomics-Plattform – eine Initiative der Deutschen und Österreichischen Röntgengesellschaften

Erste Anwendungsbeispiele

The International Radiomics Platform – An Initiative of the German and Austrian Radiological Societies

First Application Examples

Autoren

Daniel Overhoff¹, Peter Kohlmann², Alex Frydrychowicz³, Sergios Gatidis⁴, Christian Loewe⁵, Jan Moltz², Jan-Martin Kuhnigk², Matthias Gutberlet⁶, H. Winter⁷, Martin Völker⁸, Horst Hahn², Stefan O. Schoenberg¹

Vorstandskommission Radiomics und Big data:

U. Attenberger, F. Bamberg, B. Baeßler, J. Barkhausen, M. Forsting, J. Furtner-Srajter, S. Gatidis, M. Gutberlet, H. K. Hahn, C. Herold, A. Kaindl, J. Kather, R. Kikinis, P. Kohlmann, B. J. Krause, G. Krombach, M. E. Ladd, G. Langs, S. Lohwasser, J. Lotz, D. Pinto dos Santos, H. Prosch, S. O. Schoenberg, M. Völker, H. Winther, M. Zimmermann

Vorstand der Deutschen Röntgengesellschaft:

G. Antoch, F. Anton, J. Barkhausen, A. Dörfler, F. Körber, G. Layer, S. Neumann, S. O. Schoenberg, J. Weßling, M. Wucherer

Präsidium der Österreichischen Röntgengesellschaft:

R. Forstner, E. R. Gizewski, K. A. Hausegger, G. Heinz, C. Herold, C. Loewe, H. Prosch, E. Sorantin, K. Wicke

Institute

- 1 Department of Radiology and Nuclear Medicine, University Medical Center Mannheim, Medical Faculty Mannheim, Heidelberg University, Germany
- 2 Fraunhofer Institute for Digital Medicine MEVIS, Bremen, Germany
- 3 Department of Radiology and Nuclear Medicine, University Hospital Schleswig-Holstein, Campus Lübeck, Germany
- 4 Department of Diagnostic and Interventional Radiology, University-Hospital Tübingen, Germany
- 5 Department of Radiology, Medical University of Vienna, Austria
- 6 Department of Diagnostic and Interventional Radiology, Leipzig Heart Centre University Hospital, Leipzig, Germany
- 7 Department of Diagnostic and Interventional Radiology, Hannover Medical School, Hannover, Germany
- 8 German Roentgen Society „Deutsche Röntgengesellschaft“, Berlin, Germany

Key words

artificial intelligence, myocarditis, magnetic resonance imaging, neural networks

eingereicht 02.06.2020

akzeptiert 17.07.2020

online publiziert 26.11.2020

Bibliografie

Fortschr Röntgenstr 2021; 193: 276–288

DOI 10.1055/a-1244-2775

ISSN 1438-9029

© 2020. Thieme. All rights reserved.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Korrespondenzadresse

Dr. Daniel Overhoff

University Medical Center Mannheim, Department of Radiology and Nuclear Medicine, Theodor-Kutzer-Ufer 1–3, 68167 Mannheim, Germany

Tel.: ++49/6 21/3 83 20 67

daniel.overhoff@umm.de

ZUSAMMENFASSUNG

Ziel Die DRG-ÖRG-IRP (Deutsche Röntgengesellschaft-Österreichische Röntgengesellschaft Internationale Radiomics-Plattform) stellt eine web-/cloudbasierte Radiomics-Plattform auf Grundlage einer öffentlich-privaten Partnerschaft dar. Sie bietet die Möglichkeit der gemeinsamen Nutzung von Daten, Annotation, Validierung und Zertifizierung auf dem Gebiet der künstlichen Intelligenz, Radiomics-Analyse und integrierter Diagnostik. In einer ersten Proof-of-Concept-Studie soll die automatisierte Myokardsegmentation sowie die automatisierte myokardiale Late-Gadolinium-Enhancement (LGE)-Detektion mittels radiomischer Bildmerkmale für Myokarditis-Datensätze evaluiert werden.

Material und Methoden Die DRG-ÖRG-IRP kann zur Erstellung qualitätsgesicherter, strukturierter Bilddaten in Kombi-

nation mit klinischen Daten und anschließender integrierter Datenanalyse genutzt werden und zeichnet sich durch die folgenden Leistungskriterien aus: Nutzungsmöglichkeit multizentrischer vernetzter Daten, automatisiert berechnete Qualitätsparameter, Bearbeitung von Annotationsaufgaben, Konturerkennung mittels herkömmlicher Verfahren sowie Verfahren der künstlichen Intelligenz und der Möglichkeit einer gezielten Einbindung von Algorithmen.

In einer ersten Studie wurde ein anhand kardialer CINE-Datensätze vortrainiertes neuronales Netz zur Segmentierung von PSIR-Datensätzen evaluiert. In einem zweiten Schritt wurden radiomische Bildmerkmale zur segmentalen Detektion von LGE der gleichen Datensätze, welche multizentrisch über die IRP zu Verfügung gestellt wurden, angewendet.

Ergebnisse Erste Ergebnisse zeigen die Vorteile (Datentransparenz, Zuverlässigkeit, breite Einbindung aller Mitglieder, kontinuierliche Evolution sowie Validierung und Zertifizierung) dieses plattformbasierten Ansatzes auf. In der Proof-of-Concept-Studie konnte im Vergleich zur Segmentierung des Myokards durch den Experten das neuronale Netzwerk einen Dice-Koeffizienten von 0,813 nachweisen. Im Rahmen der segmentbasierten myokardialen LGE-Detektion ergab sich ein AUC von 0,73 bzw. ein AUC von 0,79 nach Ausschluss von Segmenten mit unsicherer Annotation.

Die Auswertung und Bereitstellung der Daten findet auf der IRP unter Berücksichtigung der FAT-Kriterien (Fairness, Accountability, Transparency) sowie FAIR-Kriterien (findable, accessible, interoperable, reusable) statt.

Schlussfolgerung Es konnte gezeigt werden, dass die DRG-ÖRP-IRP als Kristallisationspunkt für die Generierung weiterer Einzel- und Verbundprojekte genutzt werden kann. Die Durchführung von quantitativen Analysen mit Verfahren der künstlichen Intelligenz wird hierbei durch den Plattformansatz der DRG-ÖRP-IRP erheblich erleichtert, da vorab trainierte neuronale Netze integriert und wissenschaftliche Gruppen vernetzt werden können.

In einer ersten Proof-of-Concept-Studie zur automatisierten Segmentierung des Myokards sowie zur automatisierten myokardialen LGE-Detektion konnten diese Vorteile erfolgreich angewendet werden.

Somit zeigt sich, dass sich mittels der DRG-ÖRP-IRP strategische Ziele interdisziplinär umsetzen, konkrete alltagstaugliche Proof-of-Concept-Beispiele aufzeigen sowie möglichst partizipativ unter Einbindung aller Gruppierungen eine Vielzahl an Einzel- und Verbundprojekten realisieren lassen.

Kernaussagen:

- Die DRG-ÖRG-IRP ist eine web-/cloudbasierte Radiomics-Plattform auf Grundlage einer öffentlich-privaten Partnerschaft.
- Die DRG-ÖRG-IRP ist nutzbar für die Erstellung qualitätsgesicherter, strukturierter Bilddaten in Kombination mit klinischen Daten und anschließender integrierter Datenanalyse.
- Erste Ergebnisse zeigen die Anwendbarkeit linksventrikulärer myokardialer Segmentierung mittels neuronalen

Netzwerks und segmentbasierte LGE-Detektion mittels radiomischer Bildmerkmale.

- Die DRG-ÖRG-IRP bietet die Möglichkeit der Integration vorab trainierter neuronaler Netze und Vernetzung wissenschaftlicher Gruppen.

Zitierweise

- Overhoff D, Kohlmann P, Frydrychowicz A et al. The International Radiomics Platform – An Initiative of the German and Austrian Radiological Societies. *Fortschr Röntgenstr* 2021; 193: 276–288

ABSTRACT

Purpose The DRG-ÖRG IRP (Deutsche Röntgengesellschaft-Österreichische Röntgengesellschaft international radiomics platform) represents a web-/cloud-based radiomics platform based on a public-private partnership. It offers the possibility of data sharing, annotation, validation and certification in the field of artificial intelligence, radiomics analysis, and integrated diagnostics. In a first proof-of-concept study, automated myocardial segmentation and automated myocardial late gadolinium enhancement (LGE) detection using radiomic image features will be evaluated for myocarditis data sets.

Materials and Methods The DRG-ÖRP IRP can be used to create quality-assured, structured image data in combination with clinical data and subsequent integrated data analysis and is characterized by the following performance criteria: Possibility of using multicentric networked data, automatically calculated quality parameters, processing of annotation tasks, contour recognition using conventional and artificial intelligence methods and the possibility of targeted integration of algorithms. In a first study, a neural network pre-trained using cardiac CINE data sets was evaluated for segmentation of PSIR data sets. In a second step, radiomic features were applied for segmental detection of LGE of the same data sets, which were provided multicenter via the IRP.

Results First results show the advantages (data transparency, reliability, broad involvement of all members, continuous evolution as well as validation and certification) of this platform-based approach. In the proof-of-concept study, the neural network demonstrated a Dice coefficient of 0.813 compared to the expert's segmentation of the myocardium. In the segment-based myocardial LGE detection, the AUC was 0.73 and 0.79 after exclusion of segments with uncertain annotation.

The evaluation and provision of the data takes place at the IRP, taking into account the FAT (fairness, accountability, transparency) and FAIR (findable, accessible, interoperable, reusable) criteria.

Conclusion It could be shown that the DRG-ÖRP IRP can be used as a crystallization point for the generation of further individual and joint projects. The execution of quantitative analyses with artificial intelligence methods is greatly facilitated by the platform approach of the DRG-ÖRP IRP, since pre-trained neural networks can be integrated and scientific groups can be networked.

In a first proof-of-concept study on automated segmentation of the myocardium and automated myocardial LGE detection, these advantages were successfully applied.

Our study shows that with the DRG-ÖRP IRP, strategic goals can be implemented in an interdisciplinary way, that concrete

proof-of-concept examples can be demonstrated, and that a large number of individual and joint projects can be realized in a participatory way involving all groups.

Einleitung

In der modernen bildgebenden Diagnostik erfolgt derzeit ein Paradigmenwechsel. Bisher haben der radiologische Befund und die Beurteilung beispielsweise bei einer Auffälligkeit in der Lunge nur einen Rundherd beschrieben mit der wertenden Verdachtsdiagnose Lungenkarzinom. Heute ermöglichen maschinelles Lernen und die Extraktion statistischer Features Vorhersagen über Mutationen und ein Vorliegen von Mikrometastasen [1]. Neben dieser durch maschinelles Lernen und Extraktion statistischer Features gewonnenen Subebene existiert eine diagnostische Metaebene, in der durch interdisziplinäre Datenintegration Aussagen zu Therapieansprechen und Überleben möglich werden [2, 3]. Auch in der Onkologie vollzieht sich eine grundlegende Veränderung der bisherigen diagnostisch-therapeutischen Abläufe. Bisher wurde dort eine CT-Untersuchung mit dem Ergebnis mehrerer Metastasen von einer Gewebebiopsie an einer willkürlichen, in Abhängigkeit der Zugänglichkeit gewählten Lokalisation gefolgt. Darauf basierend wurde die Therapieentscheidung getroffen. Heute ist es bereits zumindest prinzipiell möglich, über erweiterte Radiomics-Analysen im gesamten Körper die Wahrscheinlichkeit für das Vorliegen einer spezifischen Mutation zu bestimmen. Erst darauf aufbauend kann eine durch die radiomische Analyse identifizierte Metastase gezielt interventionell-radiologisch biopsiert werden, die beispielsweise das höchste Risiko einer intratumoralen Heterogenität oder einer Sekundärmutation aufweist [4].

Die soeben beschriebene, häufig automatisierte Extraktion klinisch relevanter qualitativer und quantitativer Biomarker aus medizinischen Bilddaten wird als „Radiomics“ bezeichnet. Unter diesem Oberbegriff werden verschiedene algorithmische Vorgehensweisen zusammengefasst, z. B. klassische Verfahren der Texturanalyse, aber auch Verfahren der künstlichen Intelligenz und des maschinellen Lernens, einem Teilbereich der künstlichen Intelligenz. Grundlage der Radiomics-Verfahren ist die Interpretation medizinischer Bilddaten als Datenquelle weit über die traditionelle visuelle Betrachtung hinaus [5].

Basis für eine zuverlässige Radiomics-Analyse bilden eine reproduzierbare Datenaufbereitung und durchgängige Qualitätssicherung, wie mehrere Studien belegt haben [4]. Neben der Auswahl des korrekten Endpunkts ist eine Kaskade aus zuverlässiger Segmentierung, Test der Stabilität extrahierter Features und trainierter Modelle bedeutsam. Ebenfalls sind ein Bericht über die verwendeten Algorithmen sowie die Durchführung einer internen und externen Validierung, unter Aufführung der eingesetzten Methodologie, entscheidend. Die Ergebnisse bisheriger wissenschaftlicher Studien hinsichtlich der Adhärenz zu diesen Radiomics-Qualitätskriterien sind ernüchternd: Lediglich eine von ca. 20 Studien hält sich an die Kriterien zumindest zu 50 % [6]. Somit sind die trotz aller Euphorie

im Bereich der algorithmischen Bildanalyse gewonnenen Mehrinformationen aus bisher publizierten Studien mit Vorsicht zu betrachten. Unterstützt wird diese kritische Analyse durch aktuelle Studien zur Robustheit von Radiomics-Verfahren, insbesondere der Texturanalyse gegenüber verschiedenen Einflussfaktoren wie Messprotokollen und Rekonstruktionsverfahren. Zudem stellt die rasche Fortentwicklung der Methoden und gewonnen Erkenntnisse eine Herausforderung für den klinischen Einsatz dar [7]. Zusätzlich limitieren aktuell noch manuelle oder semiautomatisierte Segmentationen aufgrund des Zeitaufwands und der Interobserver-Variabilität den klinischen Routineeinsatz.

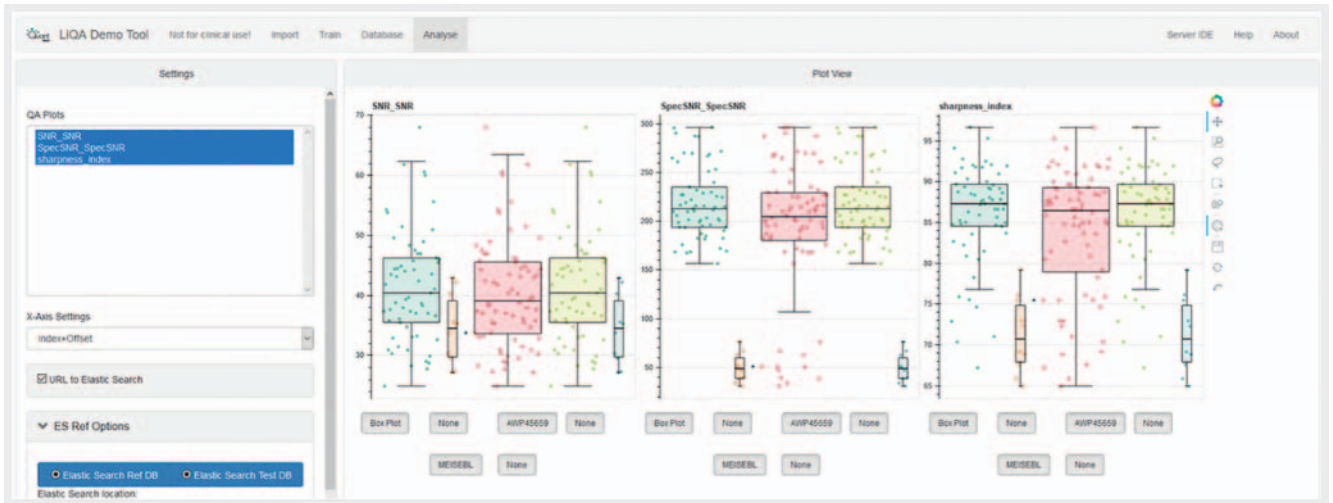
Deep-Learning-Verfahren neuester Generationen zeigen eine potenziell bessere Generalisierung und Adaptivität im Vergleich zu früheren Verfahren (klassische Feature-Extraktion in der Texturanalyse) und haben das Potenzial, bei den beschriebenen Limitationen Abhilfe zu schaffen [8]. Die Anwendung von Deep-Learning-Verfahren erfordert dabei eine vergleichsweise große Anzahl an Trainingsdaten, möglichst aus verschiedenen Standorten, und effiziente Werkzeuge zur Datenannotation.

Die damit verbundenen Chancen, aber auch erheblichen Herausforderungen waren Grundlage des durch die Deutsche Röntgengesellschaft initiierten internationalen Radiomics-Plattform-Projekts (IRPP). Das IRPP trägt die Besonderheit einer kooperativen öffentlich-privaten Partnerschaft, in Verbindung mit großen Universitätskliniken, Fraunhofer MEVIS und Industriepartnern im Bereich der Medizin- und Informationstechnologie. Ziel dieser Initiative ist die Generierung einer gemeinnützigen, cloud-basierten Analyseplattform (internationale Radiomics-Plattform, IRP) für die gemeinsame Nutzung von Daten, Annotation, Validierung und Zertifizierung auf dem Gebiet der künstlichen Intelligenz, Radiomics-Analyse und integrierter Diagnostik. Basis hierfür ist ein gemeinsamer Konsortialvertrag in Analogie zu den rechtlichen Strukturen für die Forschungscampusmodelle des Bundesministeriums für Bildung und Forschung [9].

Ziel der Arbeit ist es, Struktur, Eigenschaften und Möglichkeiten der IRP zu beschreiben und die Machbarkeit und erste Ergebnisse anhand erster Anwendungsbeispiele zur kardialen MRT zu demonstrieren.

Methoden

Im aktuell laufenden DRG-ÖRG-IRPP wird die web- bzw. cloudbasierte Radiomics-Plattform zur Erstellung qualitätsgesicherter und strukturierter Bilddaten in Kombination mit klinischen Daten und darauf folgender integrierter Datenanalyse genutzt [10]. Als Pilotprojekt wurde die retrospektive Analyse von Myokarditis-Patienten mit mindestens 2 MRT-Untersuchungen zum Zeitpunkt der Diagnose und in der Nachsorge nach 6 Monaten verwendet. Als Endpunkt wur-



► **Abb. 1** Demonstrator für die Visualisierung und Analyse der beim Datenimport automatisch berechneten Qualitätsparameter. Die für eine Auswahl von Bilddaten berechneten Parameter, wie beispielsweise das Signal-zu-Rausch-Verhältnis, können im Demonstrator untereinander oder, falls vorhanden, mit einer Referenzwerte-Datenbank verglichen werden.

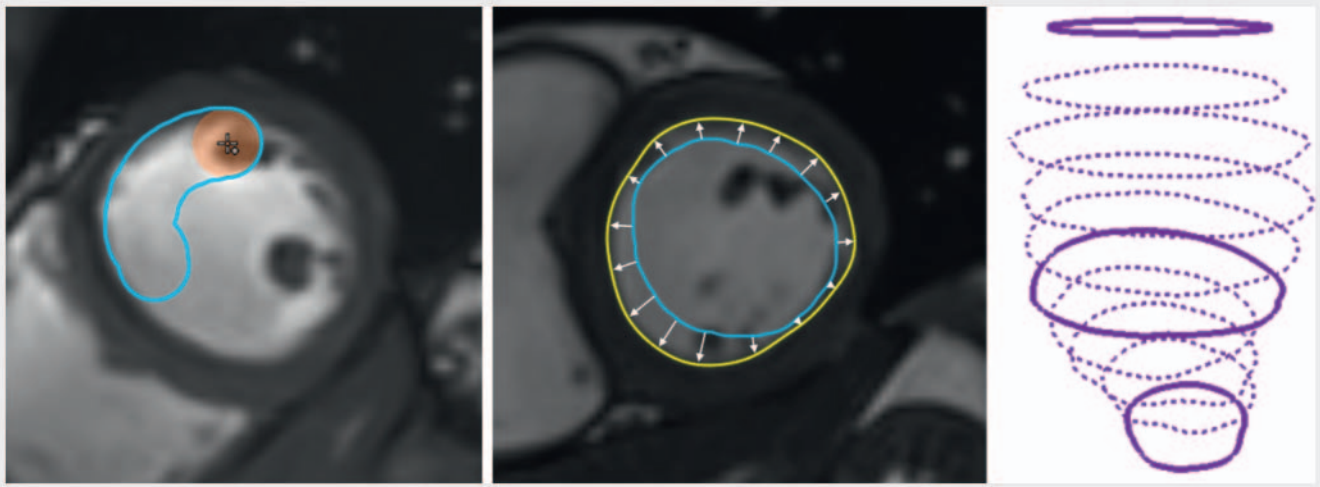
de hierbei die Erhaltung der myokardialen Funktion nach 6 Monaten definiert. Bisher nehmen 6 universitätsmedizinische Einrichtungen an der Studie teil, alle mit genehmigtem Votum der Ethikkommission, um eine initiale Zielgröße von ca. 200 Fällen zu erreichen.

Das IRPP zeichnet sich dabei durch folgende Leistungskriterien aus:

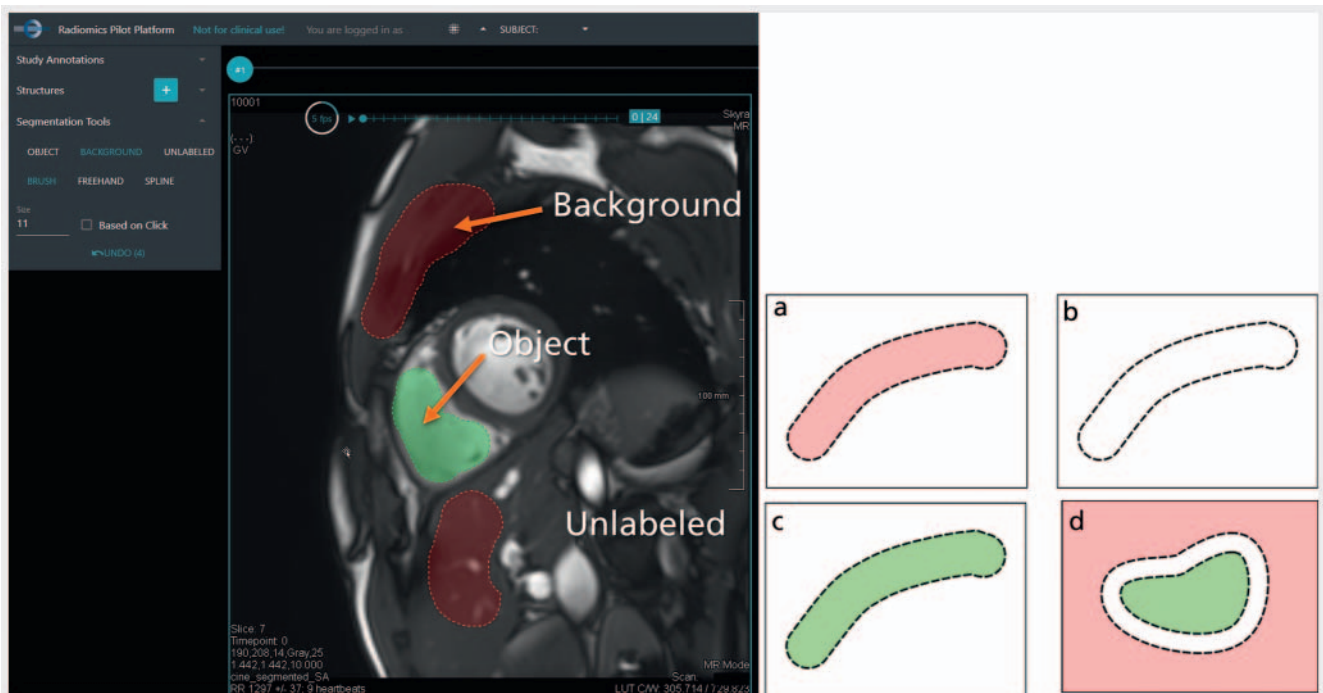
1. Bereitstellung einer cloudbasierten Plattform für die standardisierte und sichere Zusammenführung medizinischer Bilddaten und assoziierter Informationen aus verschiedenen Standorten. Aktuell ist ein Web-Upload von bereits anonymisierten Bilddaten per Drag-and-Drop verfügbar. In Kürze wird die IRP um eine integrierte Anonymisierung, die auf der Seite der Hochladenden stattfindet, erweitert. Dabei werden sowohl im DICOM-Standard definierte Deidentifikationsprofile als auch konfigurierbare Positivlisten unterstützt. Für den automatisierten Upload von Nichtbilddaten soll, wo es möglich ist, der HL7-FHIR-Standard verwendet werden.
2. Für die Qualitätsanalyse der Bilddaten können bereits beim Datenimport automatisiert Qualitätsparameter bzgl. Signalintensität, Homogenität und Artefakten berechnet werden. In einer separaten Webapplikation, die noch in die IRP integriert werden soll, werden die Datensätze untereinander oder mittels einer Referenzdatenbank verglichen und analysiert (► **Abb. 1**). Die resultierenden Daten können dafür genutzt werden, den Algorithmus für die Erkennung weiterer Artefakte in zukünftig hochgeladenen Datensätzen zu trainieren. Darüber hinaus können geladene Bilder visuell von den radiologischen Gutachtern bewertet werden (5-Sterne-Rating in Kombination mit strukturierten Eingabemasken).
3. Bündelung generischer Werkzeuge für die Bearbeitung vielfältiger Annotationsaufgaben. Die Annotationsergebnisse werden direkt für Radiomics- und Deep-Learning-Fragestellungen (z. B. als Trainingsdaten) verfügbar gemacht. Die Plattform ist hochgradig konfigurier- und erweiterbar, um schnell für den Einsatz in neuen Studien angepasst werden zu können. Für die effiziente

manuelle Segmentierung 2- und 3-dimensionaler Strukturen wurden ein dedizierter Werkzeugkasten und Konturierungsworkflow implementiert. Somit stehen verschiedene Tools (Freehand, Spline, Brush) zur Erstellung und Bearbeitung von Konturen sowie algorithmische Unterstützung wie A) Interpolation zwischen Konturen auf Schichten, auf denen die Struktur nicht händisch eingezeichnet wurde, und B) eine Verfeinerung von grob gezeichneten Konturen durch ein sich selbst optimierendes Verfahren (Snapping) (► **Abb. 2**) zur Verfügung. Darüber hinaus wurden innovative Interaktionskonzepte entwickelt, um mittels sogenannter Sparse-Labeling-Techniken bereits mit wenigen, aber hochqualitativen Labels das Training eines neuronalen Netzes zu starten und dieses iterativ zu optimieren (► **Abb. 3**).

4. Neben der Verwendung herkömmlicher Konturerkennungsalgorithmen werden spezielle Verfahren der künstlichen Intelligenz genutzt, um eine verbesserte Segmentierung auf Basis bereits annotierter Daten zu ermöglichen, z. B. für das Training des neuronalen Netzes zur automatisierten Konturanalyse des Endo- und Epikards. Hierfür ist die Radiomics-Plattform an das von Fraunhofer MEVIS entwickelte Deep-Learning-Framework RedLeaf [11] angebunden. Die IRP sendet neu erstellte oder korrigierte Segmentierungen an einen Trainingsserver für das Trainieren neuronaler Netze. Ein mit RedLeaf verbundenes Monitoring-Tool ermöglicht die Überwachung der Qualität der trainierten Netze. Ein Klassifikationsserver erlaubt es der IRP, automatische Segmentierungen für Strukturen, für die es bereits trainierte Netze gibt, abzurufen (► **Abb. 4**).
5. Es besteht die Möglichkeit einer gezielten Einbindung von Algorithmen, z. B. die Verwendung bereits trainierter Netze aus internationalen Algorithmen-Wettbewerben (Challenges). Hierzu hat das IRPP bereits die Kooperation mit der Grand-Challenge-Plattform [12] aufgenommen, welche algorithmische Vergleiche zu unterschiedlichsten Fragestellungen der medizinischen Bildgebung mit weltweiter Beteiligung organisiert. Hierbei können die zur Anwendung kommenden neuronalen Netze durch Algo-



► **Abb. 2** Verschiedene Werkzeuge (wie z. B. ein Brush) stehen zur Verfügung, um Konturen manuell einzuzeichnen (links). Die gezeichneten Konturen können optional automatisch optimiert werden. Ein Snapping-Algorithmus passt eine gezeichnete Kontur an hohe Gradienten in der Umgebung an. Die ungenau gezeichnete blaue Kontur wird dadurch in die gelbe Kontur umgewandelt (Mitte). Falls eine 3-dimensionale Struktur nicht auf allen Schichten eingezeichnet wird, werden Konturen auf Zwischenschichten durch Interpolation ergänzt (rechts).

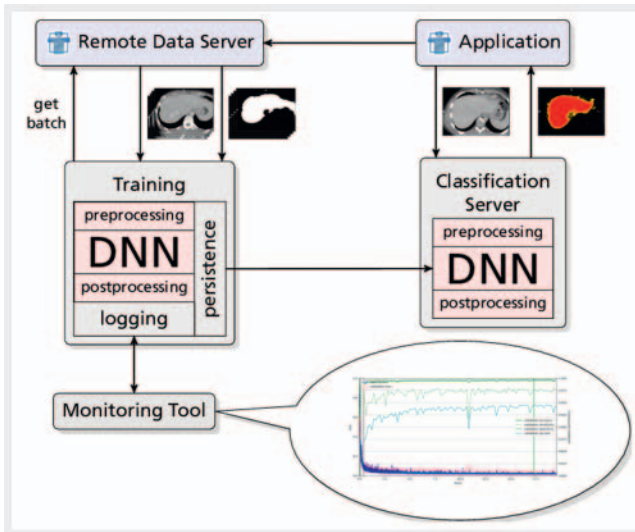


► **Abb. 3** Sparse-Labeling-Werkzeug, welches eine unterschiedliche und unvollständige Klassifikation von Objekten im Bild erlaubt: alles innerhalb einer Kontur ist **a** Hintergrund, **b** unbekannt, **c** Objekt oder **d** ein Unsicherheitsbereich. Die von der Kontur eingeschlossenen Bereiche sind das Objekt, und alles außerhalb der Kontur ist Hintergrund.

- rhythmen, die bereits auf externen Servern mit großen Datenmengen trainiert wurden, weiter optimiert werden.
- Integration von Bilddaten und korrespondierenden klinischen Daten sowie klinischen Endpunkten. Definierte Endpunkte können in Kombination mit Bilddaten und klinischen Daten als Zielvariablen für die Modellierung statistischer Modelle und maschineller Lernverfahren, z. B. zur Prädiktion des Therapieansprechens verwendet werden.

- Abschließend können aus den dominanten Features durch maschinelles Lernen mehrdimensionale Korrelationen zu den entsprechenden klinischen oder molekulargenetischen Endpunkten erstellt werden (Feature-Maps) (► **Abb. 5**).

Diese Leistungskriterien wurden im Rahmen erster Studien zur autonomen Segmentierung von Late-Gadolinium-Enhancement



► **Abb. 4** Anbindung einer Applikation (z. B. IRP) an einen Trainingsserver inkl. Monitoring-Tool und an einen Klassifikationsserver. Der Trainingsserver erhält originale Bilddaten zusammen mit dazugehörigen Segmentierungsmasken, die in der IRP erstellt wurden. Nach Erhalt dieser Daten kann mit einem Training begonnen werden, oder ein bereits trainiertes Netz (DNN) kann interaktiv weitertrainiert werden. Das an den Trainingsserver angebundene Monitoring-Tool überwacht die Klassifikationsgenauigkeit des trainierten Netzes in Bezug auf Validierungsdatensätze. Der Klassifikationsserver ermöglicht die Erstellung von Segmentierungen in Echtzeit aus der IRP basierend auf den trainierten Netzen.

(LGE)-Sequenzen sowie in einem zweiten Schritt zur Detektion von LGE mittels radiomischer Bildmerkmale angewandt.

Laufende Projekte

Derzeit laufen mit Unterstützung der Deutschen Röntgengesellschaft 2 multizentrische Analysen:

1. Eine MRT-Studie zur Prädiktion des kardialen Funktionsergebnisses im Verlauf (► **Abb. 6a, b**). Hierbei werden von Experten sorgfältig kuratierte Daten als Basis für gezielte Radiomics-Analysen und die (Weiter-) Entwicklung tiefer neuronaler Netze verwendet. Die multizentrisch erhobenen Daten werden bei dieser Studie in einem Konsensusverfahren ausgewertet.
2. Eine MRT-Studie bei 7 T zur Beurteilung normaler und pathologischer Veränderungen des Handgelenks im Vergleich zu 3 T (► **Abb. 7**). Bei dieser Studie geht es vorrangig um den radiologischen Bildvergleich zwischen beiden Feldstärken. Neben der Beurteilung unterschiedlicher Pathologien werden die Bildqualität und das Vorhandensein von Bildartefakten visuell beurteilt. Die monozentrisch erhobenen Daten werden bei dieser Studie von 7 Auswertenden unabhängig voneinander beurteilt.

Im Folgenden soll zur Illustration der Anwendbarkeit tiefer neuronaler Netze ausschließlich eine erste Analyse der myokardialen MRT-Studie präsentiert werden.

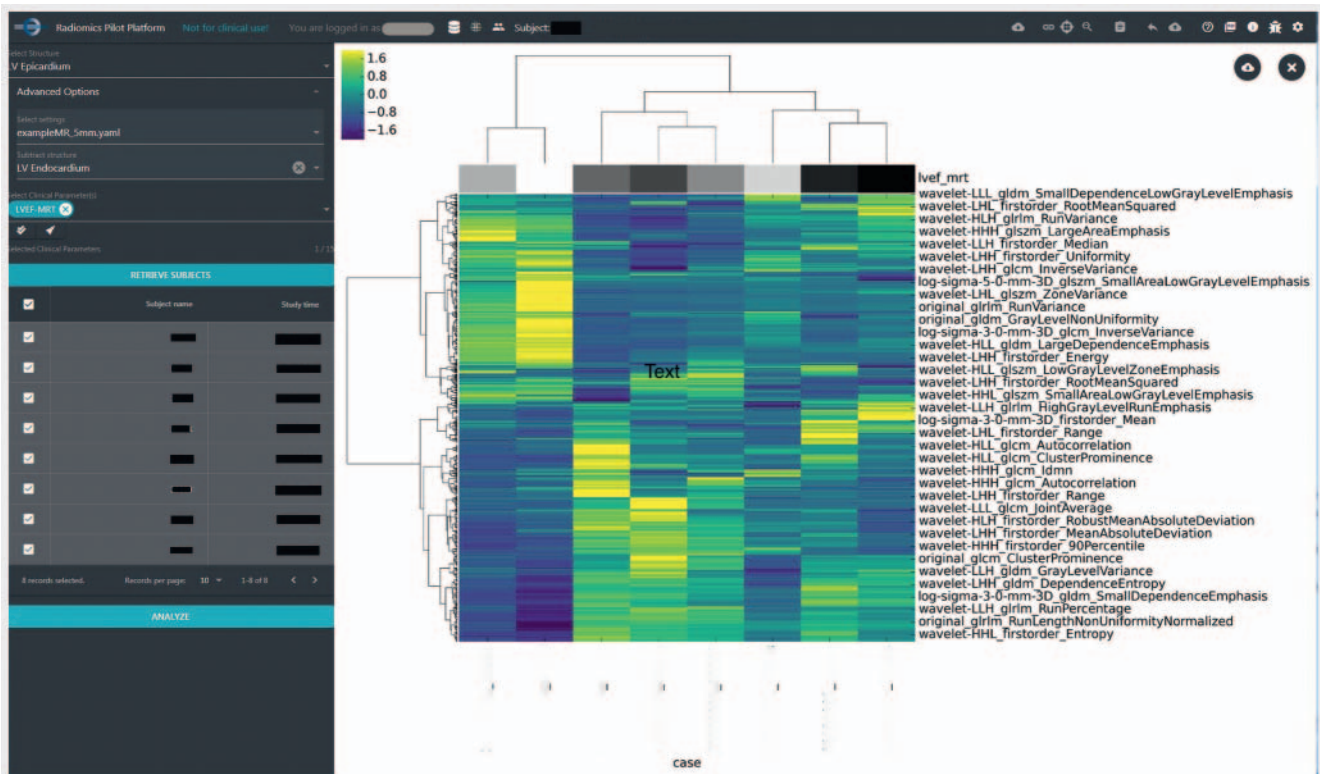
Erste Ergebnisse

Bei der Radiomics-Analyse konnten mittlerweile 47 multizentrisch erhobene kardiale MRT-Datensätze zu jeweils 2 Zeitpunkten auf den zentralen Server hochgeladen werden. An den teilnehmenden universitären Einrichtungen lag ein positives Votum der Ethikkommission für diese Studie vor. In die initiale Begutachtung der Daten flossen alle 47 Patienten ein, die zu einer endgültigen Datenbank von 992 zu analysierenden Segmenten (17-Segmente-Modell der American Heart Association (AHA)) führte (► **Abb. 8**).

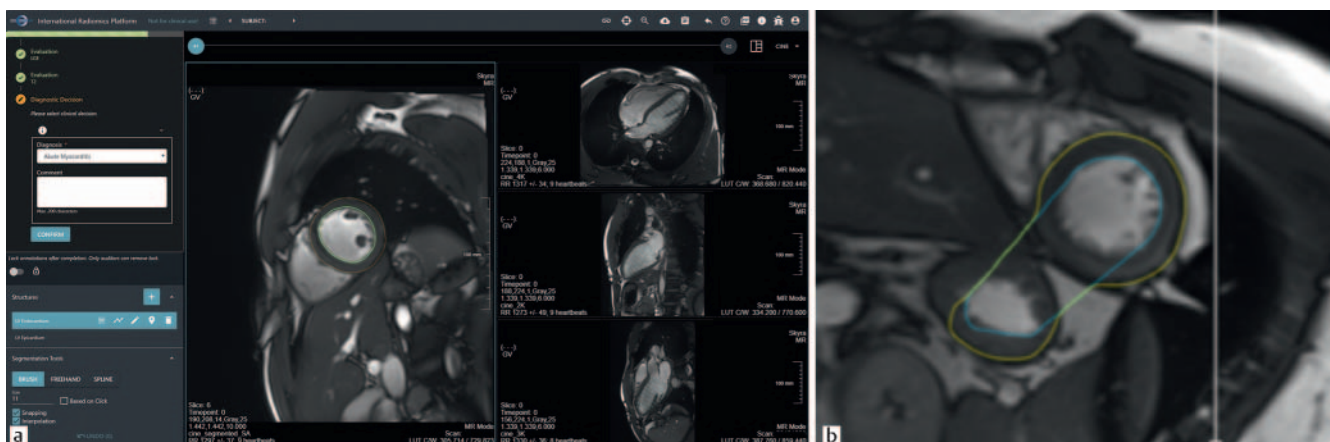
Die LGE-Bilder wurden 10–15 Minuten nach intravenöser Gabe von Gadolinium-haltigem Kontrastmittel unter Verwendung der „inversion recovery gradient echo“ (IR-GRE)-Pulssequenz akquiriert. Die Inversionszeit (TI) wurde pro Patient unter Verwendung einer TI-Scout-Sequenz optimiert und lag typischerweise zwischen 250 und 300 ms. Zur Auswertung wurde jeweils das Phasenbild der „phase-sensitive inversion recovery“ (PSIR)-Sequenz in der Kurzachse analysiert. Die Feldstärke der verwendeten MRT-Scanner betrug 1,5 und 3 Tesla und die Schichtdicke der PSIR-Sequenzen betrug 6–8 mm.

Bisher wurde als erstes Ergebnis der abgeschlossenen ersten Anwendungsbeispiele die automatisierte Segmentierung der Late-Enhancement (LGE)-Daten etabliert. Diese erfolgte in 2 Schritten: (i) automatische Myokardsegmentierung mittels Deep Learning (2D-U-Net-Architektur [13] mit 4 Schichten) (► **Abb. 9**) und nachfolgend (ii) Detektion des LGE mittels eines Random-Forest-Klassifikators.

Zunächst wurden die LGE (PSIR)-Daten zur Generierung von Trainingsdaten durch einen Experten (Facharzt für Radiologie mit mehr als 5-jähriger Erfahrung in der kardialen Bildgebung) manuell segmentiert. Aufbauend auf einem bereits etablierten Algorithmus für die Segmentierung von „steady-state free-precession“ (SSFP)-CINE-Datensätzen wurde das neuronale Netz anhand der manuell segmentierten LGE-Daten weiter trainiert. Verwendet wurde ein neuronales Netz, welches zunächst mit SSFP-CINE-Daten und Realtime-Daten von 113 Patienten vortrainiert und evaluiert wurde. Im Anschluss wurden 80 % aus 75 Datensätzen von 41 Patienten (992 Frames) genutzt, um das vortrainierte neuronale Netzwerk an PSIR-Datensätzen zu adaptieren. Weitere 10 % der PSIR-Daten wurden für die Validierung und die übrigen 10 % für die Testung genutzt. Als Ergebnis entstanden 2 Klassifikatoren, die sequenziell auf die Daten angewendet werden können. Der erste Klassifikator identifiziert eine ungefähre Bounding Box um den linken Ventrikel, und der zweite Klassifikator segmentiert das Myokard (Bereich zwischen epikardialer und endokardialer Grenze) innerhalb dieser Bounding Box. Als Gütekriterium der Übereinstimmung zwischen der Segmentierung durch den Experten und der myokardialen Segmentierung durch das neuronale Netzwerk wurde der Dice-Koeffizient bestimmt. Dieser Ähnlichkeitskoeffizient gibt sowohl die räumliche Überlappung als auch die Reproduzierbarkeit an, wobei ein Dice-Koeffizient von 1 somit eine vollständige Überlappung bzw. Übereinstimmung darstellt und ein Wert von 0 eine fehlende Übereinstimmung anzeigt [14]. Hier konnte auf den Testdaten im Vergleich zur Segmentierung durch den Experten ein Dice-Koeffizient von 0,813 für das Myokard gezeigt werden. Der Dice-Koeffizient für den Blutpool (Bereich innerhalb der endokardialen Grenze) betrug 0,941.



► **Abb. 5** IRP-Radiomics-Analyse: Für eine ausgewählte Struktur (Myokard: Volumen innerhalb des Epikards abzüglich des Volumens innerhalb des Endokards) und ausgewählte Fälle werden radiomische Merkmale berechnet. Eine Heatmap stellt diese Merkmale und ggf. weitere klinische oder in IPR berechnete Parametern (hier exemplarisch LVEF-MRT: linksventrikuläre Ejektionsfraktion) nach erfolgter Clusterbildung dar. Aktuell können diese Ergebnisse für weiterführende Analysen heruntergeladen werden.



► **Abb. 6 a** IRP konfiguriert für die Myokarditis-Studie. **b** CINE-SSFP-Kurzachsenschnitt durch den linken und rechten Ventrikel. Beispiel für eine fehlerhafte Segmentierung des initialen ML-basierten Algorithmus, was die Notwendigkeit eines „supervised learnings“ auch bei vermeintlich einfachen Segmentierungsaufgaben unterstreicht. Flüssigkeitsgefüllter Magen mit dicker Muskulatur wird vom Algorithmus als linksventrikuläres Myokard fehlgedeutet.

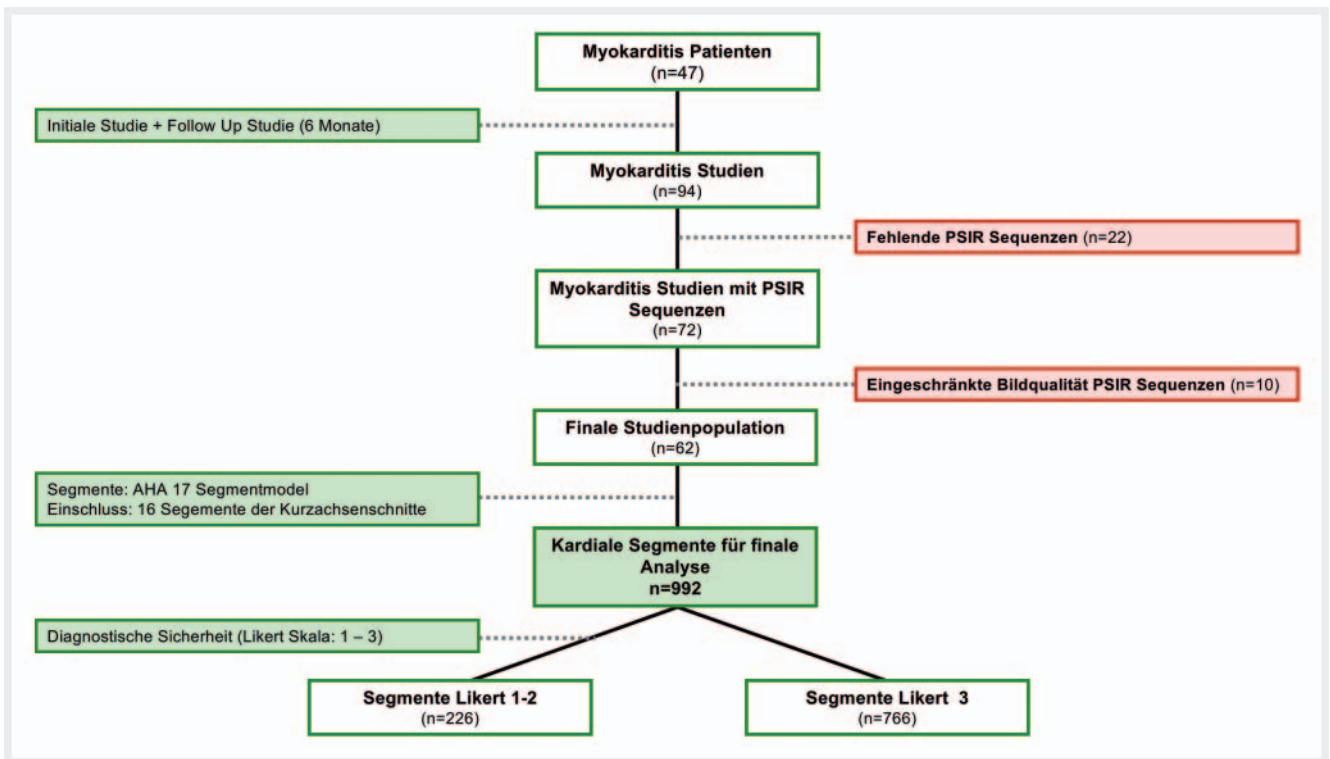
Im nächsten Schritt wurden jetzt die LGE-Areale durch den radiologischen Experten identifiziert, den in den Kurzachsenschnitten zu beurteilenden Segmenten des 17-Segmente-Modells der AHA-Klassifikation zugeordnet und die entsprechenden Segmente als LGE-positiv oder -negativ klassifiziert. Das LGE wurde visuell durch den radiologischen Experten für jedes Segment ent-

sprechend der Beurteilbarkeit in einer 3-gliedrigen Likert-Skala hinsichtlich der diagnostischen Sicherheit (1 = gering; 2 = mittel; 3 = hoch) ausgewertet.

Darauf basierend wurde untersucht, wie gut sich Late Enhancement in einzelnen Segmenten durch radiomische Bildmerkmale detektieren lässt. Dazu wurde mit der in die IRP integrierten Soft-



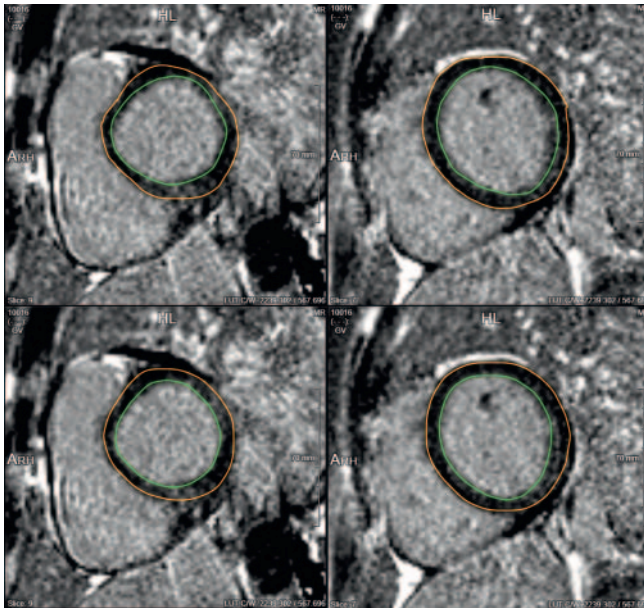
► **Abb. 7** IRP konfiguriert für die 3 T/7T-Handgelenk-Studie.



► **Abb. 8** Flow Chart der Studienpopulation mit Einschluss- und Ausschlusskriterien.

ware-Bibliothek PyRadiomics [15] eine große Zahl standardisierter Merkmale berechnet. In die Analyse eingeflossen sind Merkmale, die aus den Intensitäten und entsprechenden Histogrammen abgeleitet wurden (Merkmale erster Ordnung), sowie Texturmerkmale,

für deren Berechnung die Beziehung zwischen mehreren Voxel berücksichtigt werden (Merkmale höher Ordnung). Für die Detektion des Late Enhancements anhand dieser Features wurde ein Random-Forest-Klassifikator verwendet. Die Analyse wurde als 10-fache



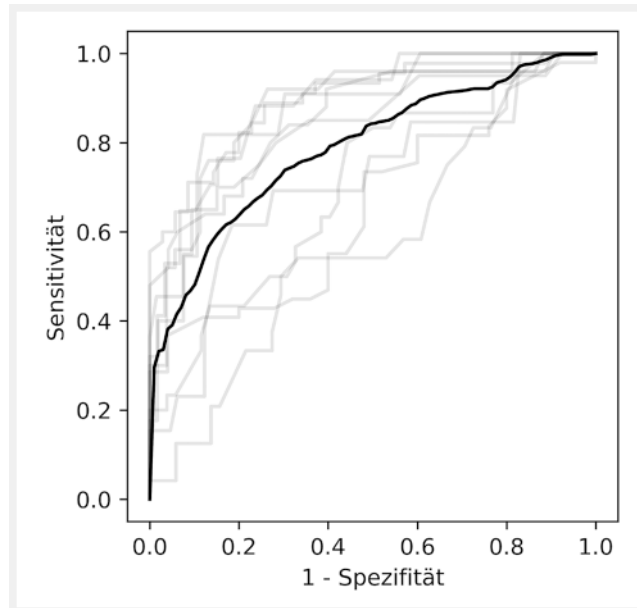
► **Abb. 9** Gegenüberstellung der Segmentierung durch den Experten (obere Zeile) und dem DNN (untere Zeile). Der dargestellte Datensatz wurde dem DNN im Training vorenthalten. In der linken Reihe erschwert ein geringer Perikarderguss die Segmentierung.

Kreuzvalidierung durchgeführt, wobei sichergestellt wurde, dass alle Segmente eines Patienten derselben Gruppe zugeordnet wurden. Bei Verwendung aller 992 Segmente, wovon 408 Late Enhancement aufwiesen, ergab sich dabei eine mittlere AUC von 0,73. Nach Ausschluss von 226 Segmenten mit unsicherer Annotation (Stufen 1 und 2 auf einer Likert-Skala von 1 bis 3 ► **Abb. 8**) betrug die mittlere AUC 0,79 (ROC-Kurve in ► **Abb. 10**). Hieraus resultierte bei einem optimalen Cut-off-Wert gemäß Youden-Index eine Sensitivität von 0,62 und eine Spezifität von 0,83. Für eine Sensitivität von 0,8 müsste die Spezifität auf 0,58 reduziert werden. Die für den Klassifikator wichtigsten Merkmale waren der Mittelwert des Originalbildes sowie Quantile nach Anwendung eines Laplacian-of-Gaussian-Filters zur Hervorhebung von Kanten. Weitere Experimente zeigten, dass eine vorgeschaltete automatisierte Merkmalsselektion (Univariate feature selection), welche auf einer Varianzanalyse basiert, keinen Mehrwert brachte.

Diskussion

Die IRP erfüllt in besonderer Weise die Aufgaben einer radiologischen Fachgesellschaft, strategische Ziele in einem sich besonders dynamisch entwickelnden Gebiet interdisziplinär umzusetzen, konkrete alltagstaugliche Proof-of-Concept-Beispiele und Ergebnisse aufzuzeigen sowie möglichst partizipativ unter Einbindung aller Gruppierungen eine Vielzahl an Einzel- und Verbundprojekten auf den Weg zu bringen.

Der initiale Einsatz der IRP für eine wichtige klinische Fragestellung zur kardialen Bildgebung zeigt die erheblichen grundsätzlichen Vorteile dieses plattformbasierten Ansatzes zur Erforschung und Entwicklung neuer Verfahren der künstlichen Intelligenz:



► **Abb. 10** ROC-Kurven des Random-Forest-Klassifikators zur LGE-Klassifikation (mittlere ROC-Kurve fett, ROC-Kurven der 10-fachen Kreuzvalidierung transparent).

1. **Datentransparenz:** Durch den Upload der Daten, die Qualitätssicherung und den Nutzerzugang über die wissenschaftliche Fachgesellschaft ist eine hohe Transparenz bei der Datenanalyse mit Verfahren der künstlichen Intelligenz gesichert.
2. **Zuverlässigkeit:** Die für die Verfahren der künstlichen Intelligenz wesentliche Annotation und damit eine „Kuratierung“ der Daten kann multizentrisch durch speziell ausgewählte Experten erfolgen und somit zu standardisierten, maximal qualitätsgesicherten Datensätzen führen.
3. **Breite Einbindung aller Mitglieder:** Durch die projektbasierte Zugänglichkeit über die wissenschaftliche Fachgesellschaft können sämtliche Fachgruppen mit unterschiedlichen wissenschaftlichen und klinischen Interessen Projekte und die dazugehörigen spezifischen Endpunkte definieren.
4. **Kontinuierliche Evolution:** Durch die Einbindung von Algorithmen-Wettbewerben (Challenges) auf der Plattform können die eingesetzten KI-Verfahren durch die internationale Community im freundschaftlichen Wettbewerb optimiert werden.
5. **Validierung und Zertifizierung:** Durch die qualitätsgesicherte Kontrolle der Punkte 1–4 kann die wissenschaftliche Fachgesellschaft die Algorithmen selbst validieren, die Ergebnisse zur Genauigkeit, Reproduzierbarkeit und insbesondere Generalisierbarkeit veröffentlichen und mit entsprechenden Institutionen (z. B. TÜV) für den klinischen Einsatz zertifizieren.

Dieser Ansatz entspricht auch den FAT-Kriterien (Fairness, Accountability and Transparency) der ethischen Aspekte beim Einsatz von künstlicher Intelligenz für die Datenanalyse, da die Schlüsselkriterien der Datentransparenz durch die multilaterale Konsortialpartnerschaft, der Verantwortlichkeit für die Daten durch die Annotation durch nominierte Experten und der gemeinschaftlichen Verfügbarkeit durch den gemeinnützigen Ansatz

gewährleistet sind [16]. Insbesondere unterstützt das IRPP die FAIR-Kriterien durch eine konsistente Identifizierbarkeit (findable), Zugänglichkeit (accessible), Austauschbarkeit (interoperable) und Wiederverwendbarkeit der Daten (reusable) [17]. Insbesondere zur Interoperabilität sind weitere Anstrengungen nötig, welche unter anderem auch im Kontext der nationalen Forschungs-Daten-Infrastruktur (NFDI) erfolgen sollen. Ein wesentlicher Aspekt hierbei ist auch die Erklärbarkeit der Ergebnisse in Bezug auf die tatsächliche Leistungsfähigkeit und die möglichen zufälligen oder systematischen Abweichungen, die durch den Einsatz von künstlicher Intelligenz bei spezifischen medizinischen Fragestellungen entstehen. Dies kann nur durch die genaue Definition der medizinischen Rahmenbedingungen und die sorgfältige Auswahl der resultierenden Endpunkte, gegen die das neuronale Netz getestet werden soll, erreicht werden.

Als erste Proof-of-Concept-Studie wurde bewusst die kardiale MRT bei Patienten mit Myokarditis ausgewählt. Die kardiale MRT ist für die weitergehende Datenanalyse mit der IRP besonders attraktiv, da sie die Möglichkeit der Akquisition standardisierter, untersucherunabhängiger Bilddaten und daraus die quantitative Berechnung von links- und rechtsventrikulären Funktionsparametern bietet [18, 19]. Hierzu ist eine Segmentierung des Endo- und Epikards in SSFP-CINE-Sequenzen unerlässlich. Diese Konturierung ist zeitaufwendig und untersucherabhängig. Teilautomatisierte und automatisierte Verfahren führen hier zu einer erheblichen Optimierung im täglichen Workflow sowie zu einer weiteren Standardisierung [20–23]. Diese Standardisierung ist gerade im Rahmen von weitergehenden Radiomics-Analysen ein wichtiges Qualitätskriterium [4, 24]. Zudem ist das Erkrankungsbild der Myokarditis im Besonderen für eine Analyse mit modernen Verfahren der kardialen MRT geeignet, da die quantitativen Parameter der Herzfunktion und entzündlichen Gewebeeränderung in Kombination mit klinischen Parametern bei standardisierter, systematischer Analyse eine hohe Genauigkeit hinsichtlich relativ eindeutige Diagnose und Differenzialdiagnose aufweisen [25, 26] sowie Aussagen zur Prognoseabschätzung des Patienten erlauben. In der vorliegenden Studie wurde ein neuronales Netz auf die automatisierte endo- und epikardiale Segmentierung von PSIR-Sequenzen nach Kontrastmittelapplikation trainiert und validiert. Als Referenz diente die manuelle Konturierung durch einen radiologischen Experten. Das neuronale Netz wurde zuvor mit einem anderen Sequenztyp, nämlich SSFP-CINE-Sequenzen für die Herzvolumetrie und Herzfunktionsanalyse, evaluiert. Als Maß für die Übereinstimmung wurde dabei der Dice-Koeffizient, eine häufig verwendete Kenngröße zum Vergleich des Überlappungsbereichs von Segmentierungen, berechnet.

Vorausgegangene Studien zur automatisierten Segmentierung des linken Ventrikels nutzten zumeist CINE-Sequenzen. Eine Metaanalyse früherer „deep learning“ neuronaler Netzwerke wies einen Mittelwert für den Dice-Koeffizienten von 0,965 für die endokardiale linksventrikuläre Konturerkennung auf, wobei Isensee et al. mit einem Wert von 0,968 den besten Dice-Koeffizienten aller vorliegenden Arbeiten generieren konnten [27]. Das hier verwendete neuronale Netzwerk wies hier ähnliche Werte auf. Im Vergleich zu schwellenwertbasierten Verfahren konnte jetzt schon eine verbesserte Überlappung der generierten Segmentie-

rung mit der „Ground Truth“, der manuellen Segmentierung durch einen erfahrenen Radiologen, festgestellt werden. Hier liegen die Werte für die endokardiale Segmentierung des linken Ventrikels bei SSFP-CINE-Sequenzen exemplarisch bei 0,88–0,89 [28, 29].

Tao et al. konnten 2015 mit einer automatisierten Segmentierungsmethode für kontrastmittelverstärkte Datensätze einen Dice-Koeffizienten von 0,81 erreichen [30], was dem unserer Studie entspricht. Generell muss berücksichtigt werden, dass in unserer Studie Myokarditiden untersucht wurden, bei denen das Verteilungsmuster des LGE typischerweise epikardial oder intramyokardial verortet ist. Dies führt im Vergleich zu vorangegangenen Studien unweigerlich zu einer erschwerten epikardialen Konturerkennung. Die niedrigere Übereinstimmung für das Myokard im Vergleich zu vorangegangenen Studien ist auch in der für neuronale Netze kleinen Testpopulation begründet. Hier ist insbesondere durch den Aufbau der internationalen Radiomics-Plattform mit einer erheblichen Verbesserung zu rechnen.

In einem zweiten Ansatz sollte die automatisierte, segmentale LGE-Detektion (17-Segmente-Modell der American Heart Association (AHA)) für PSIR-Sequenzen durch einen Random-Forrest-Klassifikator analysiert werden. Der Random-Forrest-Klassifikator zeigte eine AUC von 0,71 in der ROC-Kurven-Analyse für die segmentale LGE-Detektion. Unter Berücksichtigung der Segmente, die mittels Likert-Skala mit einer hohen diagnostischen Sicherheit durch den Experten beurteilt werden konnten (Likert-Skala = 3), zeigte sich in der weiteren Analyse eine verbesserte AUC von 0,77. Weitergehende Analysen des Autorenteam zielen gegenwärtig darauf ab, das trainierte neuronale Netz und den Random-Forrest-Klassifikator für die automatisierte Diagnosestellung einer Myokarditis und Abgrenzung gegenüber Normalbefunden sowie anderen myokardialen Pathologien einzusetzen. Die dafür erforderliche systematische Annotation aller Bilddaten mit den korrespondierenden klinischen Daten ist derzeit im Abschluss befindlich.

Ein potenzieller Vorteil eines neuronalen Netzes besteht darin, dass möglicherweise die „Kontamination“ durch helle Pixel an der Segmentgrenze zwischen dem endokavitären Blutpool und dem Endokard als typische Problematik bei schwellenwertbasierten Verfahren vermieden werden kann und es somit möglicherweise seltener zu einer Fehlklassifikation von LGE-Arealen kommt. Dies scheinen auch aktuelle Publikationen zu dem Thema zu belegen. So hatte in einer im Jahr 2019 publizierten Studie die Analyse unter Verwendung von „deep neural learning“ in Bland-Altman-Plots gegenüber der „Ground Truth“ die geringste Varianz von weniger als 10%, während die Varianz bei schwellenwertbasierten Verfahren bei über 20% lag [31]. Eine aktuelle Publikation in Radiology belegt außerdem, dass Texturanalysen auf Basis bereits vorprozessierter Daten, beispielsweise mit T1- und T2-Mapping, eine deutlich verbesserte Klassifikation von Patienten mit Myokarditis ermöglichen können [32].

Um die IRP als Kristallisationspunkt für die Generierung weiterer Einzel- und Verbundprojekte zu nutzen, muss eine Routinefunktionalität gewährleistet sein, die die Akkumulation, Annotation und Auswertung der Daten einerseits möglichst einfach gestaltet, andererseits technische Innovationen kontinuierlich und unkompliziert integrieren kann. Die Durchführung von quan-

titativen Analysen mit Verfahren der künstlichen Intelligenz wird hierbei durch den Plattformansatz der IRP erheblich erleichtert, beschleunigt und präzisiert. Dies liegt insbesondere an 4 Aspekten:

1. Bereits vorab trainierte neuronale Netze auf Basis anderer Datensätze können auf verwandte Fragestellungen der internationalen Radiomics-Plattform übertragen werden und somit die Zeitdauer für eine weitere, iterative Verfeinerung bis zu einer maximalen Präzision deutlich verkürzen (transfer learning).
2. Dadurch können sich ähnlich ausgerichtete wissenschaftliche Gruppen vernetzen und somit interaktive Wertschöpfungsketten bilden, um möglichst viele wissenschaftliche Fragestellungen aufeinander aufbauend zu bearbeiten.
3. Durch die (inter-) nationale Vernetzung und damit Erhöhung der Datengrundlage für die Validierung neuronaler Netze steigt deren Präzision weiter an und kann, wie von internationalen Fachgesellschaften mittlerweile gefordert, kontinuierlich optimiert werden [33]. Dafür muss neben der retrospektiven Analyse bereits existierender Datensätze auch ein prospektiver Ansatz für eine standardisierte Akquisition und Auswertung entlang der zu erwartenden wichtigsten Parameter möglich sein.
4. Die IRP bietet weiterhin auch Schnittstellen zu anderen Datenplattformen von großen nationalen und internationalen Konsortien, beispielsweise der Joint-Imaging-Plattform des Deutschen Konsortiums für translationale Krebsforschung zu gezielten Radiomics-Analysen im Bereich von onkologischen Studien.

Um möglichst viele Nachwuchswissenschaftler sowie etablierte Radiologen im Gebiet der Verfahren zur künstlichen Intelligenz weiterzubilden, wurde auf Basis eines im Jahr 2016 interdisziplinär abgehaltenen Workshops ein White Paper durch die Deutsche Röntgengesellschaft veröffentlicht [34]. Hierbei müssen auch gezielt neue Allianzen zwischen Radiologen, Informatikern, Mathematikern und benachbarten klinischen Fachdisziplinen aufgebaut werden [35]. Ein wesentlicher Baustein hierfür ist die geplante Nachwuchsakademie zur Aus- und Weiterbildung im Bereich der Verfahren künstlicher Intelligenz.

KLINISCHE RELEVANZ

- Das DRG-ÖRG-IRP-Projekt ist eine web- bzw. cloudbasierte Radiomics-Plattform auf Grundlage einer öffentlich-privaten Partnerschaft.
- Die DRG-ÖRG-IRP erleichtert, beschleunigt und präzisiert die Durchführung quantitativer Analysen mittels künstlicher Intelligenz.
- Erste Studienergebnisse der DRG-ÖRG-IRP zur automatisierten myokardialen Segmentierung mittels neuronalen Netzwerks und zur automatisierten Detektion von myokardialem LGE auf Grundlage radiomischer Features zeigen die Nutzbarkeit der Plattform.
- Die DRG-ÖRG-IRP bietet die Möglichkeit, vorab trainierte neuronale Netze auf die Plattform zu übertragen und wissenschaftliche Gruppen zu vernetzen.

Glossar

Cloud

IT-Ressourcen zur Speicherung oder Verarbeitung von Daten werden ortsfern/ortsunabhängig zur Verfügung gestellt und stehen dem Nutzer über Internet oder Intranet zur Verfügung.

Bounding Box

Die Bounding Box bezeichnet eine Möglichkeit der Objektdetektion. Es wird dabei ein Rechteck (2D) oder eine Box (3D) um das zu identifizierende Objekt generiert. Innerhalb des Rechtecks/der Box kann darauf basierend eine genauere (automatisierte) Segmentierung erfolgen.

Brush

Werkzeug für eine Segmentierung/Konturierung anatomischer Strukturen. Über einen grafisch dargestellten Kreis (Pinsel) mit anpassbarem Durchmesser wird entlang des äußeren Pinselrandes die Struktur segmentiert (► **Abb. 2**, links).

Deep Learning

Deep Learning ist eine Unterform des maschinellen Lernens, bei dem tiefe neuronale Netze genutzt werden.

Dice

Der Dice-Koeffizient (DSC) ist ein Ähnlichkeitskoeffizient und gibt sowohl die räumliche Überlappung als auch die Reproduzierbarkeit an, wobei ein Dice-Koeffizient von 1 somit eine vollständige Überlappung bzw. Übereinstimmung darstellt und ein Wert von 0 eine fehlende Übereinstimmung anzeigt. Der Dice-Koeffizient wird wie folgt berechnet: $DSC(A,B) = 2(A \cap B) / (A + B)$, wobei \cap die Schnittmenge bildet [14].

Feature-Heat-Map

Grafische Darstellung radiomischer Bildmerkmale, nachdem darauf ein Clustering-Verfahren angewendet wurde. Mit einer geeigneten Farbkodierung repräsentiert die Feature-Heat-Map hohe Korrelationen von radiomischen Merkmalen (und ggf. klinischen Parametern) in visueller Form.

Freehand

Werkzeug für eine Segmentierung/Konturierung anatomischer Strukturen. Die Segmentierung erfolgt mittels freier Zeichnung entlang der zu segmentierenden Struktur.

Snapping

Korrekturmodus für manuell gezeichnete Konturen. Die eingezeichnete Kontur passt sich per Algorithmus optional automatisch an in der Nähe befindliche signifikante Objektgrenzen der zu segmentierenden Struktur an (► **Abb. 2**, Mitte).

Kreuzvalidierung

Kreuzvalidierung bezeichnet eine Methode, in der eine Teilmenge der Datensätze für das Training eines Modells genutzt wird und eine weitere Teilmenge der Datensätze zur Evaluation des trainierten Netzes zur Verfügung gestellt wird. Es gibt hierzu unterschiedliche Ansätze wie K-Fold-Cross-Validation oder LOOCV (Leave-One-Out-Cross-Validation), auf die nicht näher eingegangen werden soll.

Laplacian-of-Gaussian-Filter (LoG-Filter)

Es handelt sich um einen Filter in der Bildbearbeitung, welcher der Hervorhebung und Detektion von Kanten dient.

Sparse-Labeling-Technik

Diese ermöglicht es, statt einer exakten Konturierung von Strukturen lediglich lokale Hinweise darüber zu geben, was und was nicht zur gewünschten Struktur gehört (► **Abb. 3**).

Random-Forest-Klassifikator

Es handelt es sich um eine maschinelle Lernmethode, bei der mehrere Entscheidungsbäume aus Teilmengen des Trainingsdatensatzes generiert werden. Die Entscheidungen jedes einzelnen Entscheidungsbaums werden aggregiert und die Klassifikation mit den meisten Stimmen als endgültige Klassifikation definiert.

PyRadiomics

PyRadiomics stellt ein OpenSource-Software-Paket dar und bietet die Möglichkeit, radiomische Merkmale aus medizinischen Bildern zu extrahieren.

Spline

Werkzeug für eine Segmentierung/Konturierung anatomischer Strukturen. Es werden Stützpunkte entlang der zu segmentierenden Struktur gesetzt. Ein Algorithmus interpoliert zwischen den Stützpunkten.

RedLeaf

RedLeaf (Remote Deep Learning Framework). Es handelt sich um eine von Fraunhofer MEVIS entwickelte Softwarelösung zum Trainieren, Testen und Anwenden tiefer neuronaler Netze.

Interessenkonflikt

Die Autorinnen/Autoren geben an, dass kein Interessenkonflikt besteht.

Danksagung

Teile dieser hier vorgestellten Arbeit wurden durch ein Forschungsstipendium (2018_01: 7T MRT des Handgelenks) der Deutschen Röntgengesellschaft DRG) finanziert. Die Förderstelle war nicht beteiligt am Studiendesign; bei der Erhebung, Analyse und Interpretation von Daten; beim Verfassen des Berichts und bei der Entscheidung, den Artikel zur Veröffentlichung einzureichen.

Literatur

- [1] Coroller TP, Grossmann P, Hou Y et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol* 2015; 114: 345–350. doi:10.1016/j.radonc.2015.02.015
- [2] Vaidya P, Bera K, Gupta A et al. CT derived radiomic score for predicting the added benefit of adjuvant chemotherapy following surgery in stage I, II resectable non-small cell lung cancer: a retrospective multicohort study for outcome prediction. *The Lancet Digital Health* 2020; 2: e116–e128
- [3] Coroller TP, Agrawal V, Narayan V et al. Radiomic phenotype features predict pathological response in non-small cell lung cancer. *Radiotherapy and Oncology* 2016; 119: 480–486
- [4] Lambin P, Leijenaar RTH, Deist TM et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017; 14: 749–762. doi:10.1038/nrclinonc.2017.141
- [5] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016; 278: 563–577
- [6] Sanduleanu S, Woodruff HC, de Jong EEC et al. Tracking tumor biology with radiomics: A systematic review utilizing a radiomics quality score. *Radiother Oncol* 2018; 127: 349–360. doi:10.1016/j.radonc.2018.03.033
- [7] Baessler B, Weiss K, Pinto Dos Santos D. Robustness and Reproducibility of Radiomics in Magnetic Resonance Imaging: A Phantom Study. *Invest Radiol* 2019; 54: 221–228. doi:10.1097/rli.0000000000000530
- [8] Hahn HK. Radiomics & Deep Learning: Quo vadis? *Forum* 2020; 35: 117–124. doi:10.1007/s12312-020-00761-8
- [9] Bundesministerium für Bildung und Forschung. Forschungscampus. Im Internet (Stand: 20.05.2020): <https://www.forschungscampus.bmbf.de>
- [10] Klein J, Wenzel M, Romberg D et al. QuantMed: Component-based deep learning platform for translational research. In: *Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications: International Society for Optics and Photonics* 2020; 113180U
- [11] MEVIS Fraunhofer. Deep Learning in Medical Imaging. Im Internet (Stand: 20.05.2020): <https://www.mevis.fraunhofer.de/en/solutionpages/deep-learning-in-medical-imaging.html>
- [12] grand-challenge.org. Grand Challenges in Biomedical Image Analysis. Im Internet (Stand: 20.05.2020): <https://grand-challenge.org/>
- [13] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer-assisted intervention* Springer. 2015: 234–241
- [14] Zou KH, Warfield SK, Bharatha A et al. Statistical validation of image segmentation quality based on a spatial overlap index. *Acad Radiol* 2004; 11: 178–189. doi:10.1016/s1076-6332(03)00671-8
- [15] van Griethuysen JJM, Fedorov A, Parmar C et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017; 77: e104–e107. doi:10.1158/0008-5472.Can-17-0339
- [16] Schönberger D. Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *International Journal of Law and Information Technology* 2019; 27: 171–203
- [17] Wilkinson MD, Dumontier M, Aalbersberg IJ et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*. 2016; 3
- [18] Grothues F, Moon JC, Bellenger NG et al. Interstudy reproducibility of right ventricular volumes, function, and mass with cardiovascular magnetic resonance. *American heart journal* 2004; 147: 218–223
- [19] Maceira A, Prasad S, Khan M et al. Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic resonance. *Journal of Cardiovascular Magnetic Resonance* 2006; 8: 417–426
- [20] Avendi M, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Medical image analysis* 2016; 30: 108–119
- [21] Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Medical image analysis* 2017; 35: 159–171
- [22] Codella NC, Weinsaft JW, Cham MD et al. Left ventricle: automated segmentation by using myocardial effusion threshold reduction and intravoxel computation at MR imaging. *Radiology* 2008; 248: 1004–1012
- [23] Mahnken AH, Muhlenbruch G, Koos R et al. Automated vs. manual assessment of left ventricular function in cardiac multidetector row computed tomography: comparison with magnetic resonance imaging. *European radiology* 2006; 16: 1416–1423. doi:10.1007/s00330-006-0226-1
- [24] Zwanenburg A, Vallières M, Abdalah MA et al. The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 2020; 295: 191145

- [25] Lurz P, Luecke C, Eitel I et al. Comprehensive Cardiac Magnetic Resonance Imaging in Patients With Suspected Myocarditis: The MyoRacer-Trial. *J Am Coll Cardiol* 2016; 67: 1800–1811. doi:10.1016/j.jacc.2016.02.013
- [26] Ferreira VM, Schulz-Menger J, Holmvang G et al. Cardiovascular Magnetic Resonance in Nonischemic Myocardial Inflammation: Expert Recommendations. *J Am Coll Cardiol* 2018; 72: 3158–3176. doi:10.1016/j.jacc.2018.09.072
- [27] Bernard O, Lalonde A, Zotti C et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? *IEEE Trans Med Imaging* 2018; 37: 2514–2525. doi:10.1109/TMI.2018.2837502
- [28] Huang S, Liu J, Lee LC et al. An image-based comprehensive approach for automatic segmentation of left ventricle from cardiac short axis cine MR images. *J Digit Imaging* 2011; 24: 598–608. doi:10.1007/s10278-010-9315-4
- [29] Liu H, Hu H, Xu X et al. Automatic left ventricle segmentation in cardiac MRI using topological stable-state thresholding and region restricted dynamic programming. *Acad Radiol* 2012; 19: 723–731. doi:10.1016/j.acra.2012.02.011
- [30] Tao Q, Piers SR, Lamb HJ et al. Automated left ventricle segmentation in late gadolinium-enhanced MRI for objective myocardial scar assessment. *J Magn Reson Imaging* 2015; 42: 390–399. doi:10.1002/jmri.24804
- [31] Zabihollahy F, White JA, Ukwatta E. Convolutional neural network-based approach for segmentation of left ventricle myocardial scar from 3D late gadolinium enhancement MR images. *Med Phys* 2019; 46: 1740–1751. doi:10.1002/mp.13436
- [32] Baessler B, Luecke C, Lurz J et al. Cardiac MRI and Texture Analysis of Myocardial T1 and T2 Maps in Myocarditis with Acute versus Chronic Symptoms of Heart Failure. *Radiology* 2019; 292: 608–617. doi:10.1148/radiol.2019190101
- [33] Langlotz CP, Allen B, Erickson BJ et al. A Roadmap for Foundational Research on Artificial Intelligence in Medical Imaging: From the 2018 NIH/RSNA/ACR/The Academy Workshop. *Radiology* 2019; 291: 781–791. doi:10.1148/radiol.2019190613
- [34] Deutsche Röntgengesellschaft (DRG) e.V. Radiomics in der Radiologie. Im Internet (Stand: 22.05.2020): <https://www.drg.de/de-DE/3601/radiomics/>
- [35] Recht MP, Dewey M, Dreyer K et al. Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. *European radiology* 2020. doi:10.1007/s00330-020-06672-5

ERRATUM

Erratum 22.04.2021: Die Internationale Radiomics-Plattform – eine Initiative der Deutschen und Österreichischen Röntgengesellschaften. Overhoff D, Kohlmann P, Frydrychowicz A et al. *Fortschr Röntgenstr* 2021; 193: 276–288

Dieser Artikel wurde gemäß Erratum am 22.04.2021 korrigiert.

Danksagung

Teile dieser hier vorgestellten Arbeit wurden durch ein Forschungsstipendium (2018_01: 7T MRT des Handgelenks) der Deutschen Röntgengesellschaft DRG) finanziert. Die Förderstelle war nicht beteiligt am Studiendesign; bei der Erhebung, Analyse und Interpretation von Daten; beim Verfassen des Berichts und bei der Entscheidung, den Artikel zur Veröffentlichung einzureichen.