European Journal of Radiology 167 (2023) 111047

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Automated liver segmental volume ratio quantification on non-contrast T1–Vibe Dixon liver MRI using deep learning

Lukas Zbinden [a,b], Damiano Catucci [b,c], Yannick Suter [a], Leona Hulbert [b], Annalisa Berzigotti [d], Michael Brönnimann [b], Lukas Ebner [b], Andreas Christe [b], Verena Carola Obmann [b], Raphael Sznitman [a], Adrian Thomas Huber [b,*]

[a] ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland
[b] Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital, Bern University Hospital, Bern, Switzerland
[c] Graduate School for Health Sciences, University of Bern, Switzerland
[d] Hepatology, Department of Visceral Surgery and Medicine, Inselspital, Bern University Hospital, Bern, Switzerland

ARTICLE INFO

Keywords:
Magnetic Resonance Imaging
Liver
Cirrhosis
Biomarker
Artificial Intelligence

ABSTRACT

Purpose: To evaluate the effectiveness of automated liver segmental volume quantification and calculation of the liver segmental volume ratio (LSVR) on a non-contrast T1-vibe Dixon liver MRI sequence using a deep learning segmentation pipeline.

Method: A dataset of 200 liver MRI with a non-contrast 3 mm T1-vibe Dixon sequence was manually labeled slice-by-slice by an expert for Couinaud liver segments, while portal and hepatic veins were labeled separately. A convolutional neural network was trained using 170 liver MRI for training and 30 for evaluation. Liver segmental volumes without liver vessels were retrieved and LSVR was calculated as the liver segmental volumes I-III divided by the liver segmental volumes IV-VIII. LSVR was compared with the expert manual LSVR calculation and the LSVR calculated on CT scans in 30 patients with CT and MRI within 6 months.

Results: The convolutional neural network classified the Couinaud segments I-VIII with an average Dice score of 0.770 ± 0.03, ranging between 0.726 ± 0.13 (segment IVb) and 0.810 ± 0.09 (segment V). The calculated mean LSVR with liver MRI unseen by the model was 0.32 ± 0.14, as compared with manually quantified LSVR of 0.33 ± 0.15, resulting in a mean absolute error (MAE) of 0.02. A comparable LSVR of 0.35 ± 0.14 with a MAE of 0.04 resulted with the LSRV retrieved from the CT scans. The automated LSVR showed significant correlation with the manual MRI LSVR (Spearman $r = 0.97$, $p < 0.001$) and CT LSVR (Spearman $r = 0.95$, $p < 0.001$).

Conclusions: A convolutional neural network allowed for accurate automated liver segmental volume quantification and calculation of LSVR based on a non-contrast T1-vibe Dixon sequence.

1. Introduction

Liver cirrhosis is associated with a hypertrophy of the left liver segments (I-III) and a hypotrophy of the right liver segments (IV-VIII) [1,2], which correlates with the patient's individual prognosis [3], and helps differentiating cirrhosis from other types of liver diseases such as porto-sinusoidal vascular disorder (PSVD) [4]. The liver segmental volume ratio (LSVR) may be quantified on liver CT and MRI, as a quantitative non-invasive imaging biomarker to measure liver segmental remodeling [5]. LSVR describes the relative size of the different segments of the liver and is calculated as the volume of the left liver segments (I-III) divided by the right liver segments (IV – VIII). LSVR has been proven useful for liver fibrosis staging [6]. An LSVR threshold of ≥ 0.35 had a sensitivity and specificity for cirrhosis of 81.5% and 88.7% [7] and was able to

Abbreviations: LSVR, Liver segmental volume ratio; AI, Artificial Intelligence; DL, Deep learning; DSC, Dice similarity coefficient; ASSD, Average symmetric surface distance; HD95, Hausdorff distance 95; MAE, Mean absolute error; MAPE, mean absolute percentage error; SD, Standard deviation; CLD, Chronic liver disease; noCLD, No chronic liver disease; IQR, Interquartile range; GPU, Graphics processing unit; NAFLD, Nonalcoholic fatty liver disease; NASH, Nonalcoholic steatohepatitis; ARLD, Alcohol-related liver disease.

* Corresponding author at: Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital, Bern University Hospital, Freiburgstrasse 10, 3010 Bern, Switzerland.
E-mail address: adrian.huber@insel.ch (A.T. Huber).

https://doi.org/10.1016/j.ejrad.2023.111047
Received 4 April 2023; Received in revised form 29 July 2023; Accepted 13 August 2023
Available online 14 August 2023

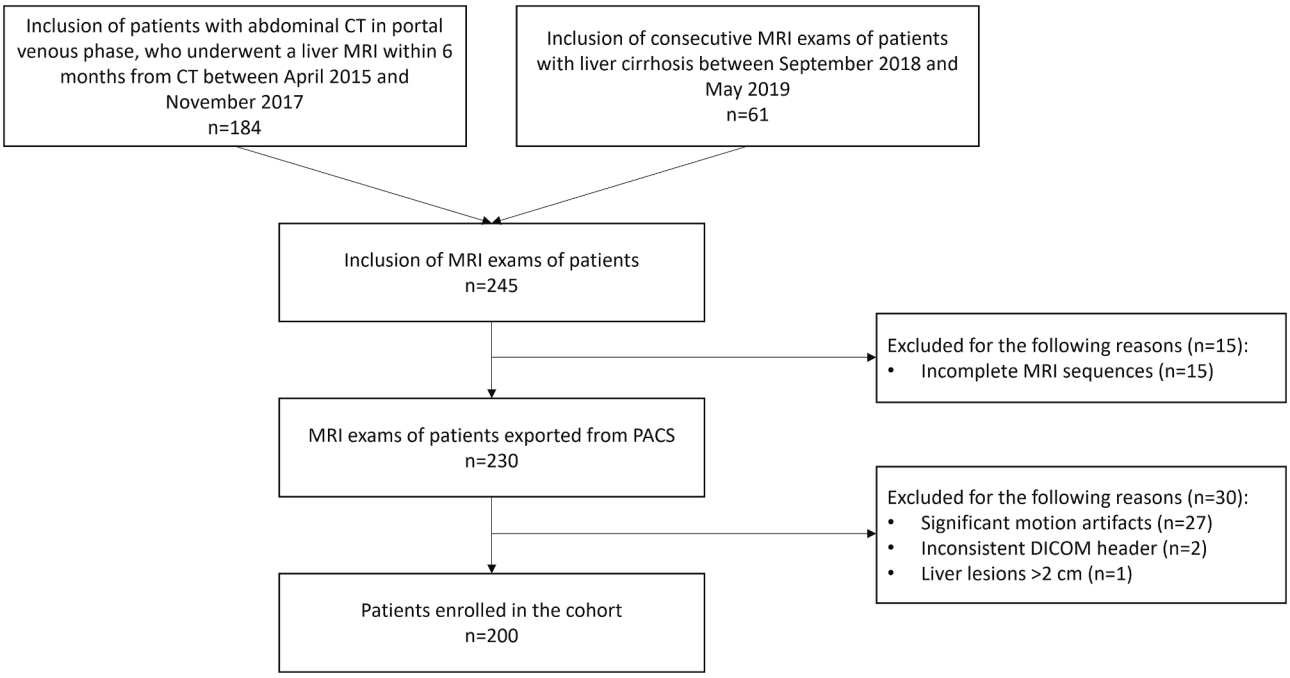

Fig. 1. Study participants inclusion flowchart.

predict significant liver fibrosis based on an MR elastography cutoff value $\geq$ 3.5 kPa [8]. As LSVR is calculated non-invasively, it may be used as a repetitive measurement during follow-up in patients with chronic liver disease.

The evaluation of liver volumes in MRI showed a high correlation with the measurement of the volume of the explanted liver in rats [9], and has also been validated in vivo in patients [10,11]. LSVR from MRI examinations is currently a complex procedure, as it requires time-consuming manual segmentation of the liver segments in the MR sequences by radiologists with dedicated software.

Deep learning (DL) is a branch of artificial intelligence that involves the use of neural networks to perform complex tasks such as image segmentation, voice recognition, and natural language processing. In the field of medical imaging, deep learning algorithms are increasingly being used to analyze and interpret medical images, such as X-rays, computer tomography (CT) scans, and magnetic resonance imaging (MRI) scans [12]. By automatically analyzing MRI scans, deep learning algorithms can provide clinicians with non-invasive, faster, and more accurate diagnoses than would be possible with human interpretation alone. In recent years, significant progress has been made in the development of deep learning algorithms for MRI [13,14].

Recently, a fully automated LSVR calculation has been developed on contrast-enhanced CT scans with a deep learning-based model [15], but up to our knowledge, there is no such model for liver MRI. As the algorithm employed in this study is applicable on non-contrast T1-vibe Dixon sequences, it is independent from the contrast medium and may be used as well on liver MRI without contrast medium administration, which is expensive and contraindicated in patients with kidney failure. In addition, it would also be available for abbreviated, non-contrast liver MRI protocols, as it is currently discussed for several indications, such as hepatocellular carcinoma screening [16].

The aim of our study was to evaluate the effectiveness of automated liver segmental volume quantification and calculation of the liver segmental volume ratio (LSVR) on a non-contrast T1-vibe Dixon liver MRI sequence using a deep learning segmentation pipeline.

2. Material and methods

This retrospective study was approved by the local ethics committee (Bern cantonal ethics committee, Bern, Switzerland) and was carried out in accordance with the principles of the Declaration of Helsinki. All patients gave written informed consent to participate in the study. The authors had full access to and take full responsibility for the integrity of the data.

2.1. Study participants

This was a retrospective analysis of a prospective study, including 184 patients with an abdominal CT in portal venous phase for any reason without malignant liver lesions or prior liver interventions who were invited to undergo a multiparametric liver MRI within 6 months from CT between 04/2015 – 11/2017. In addition, 61 consecutive patients with liver cirrhosis and clinically indicated liver MRI, without malignant liver lesions or prior liver intervention were included between 09/2018 – 05/2019, resulting in a dataset of 245 liver MRI (Fig. 1). Patients with incomplete MRI sequences (n = 15) were excluded. The remaining MRI exams from 230 patients were collected from the picture archiving system (PACS) of the hospital in Digital Imaging and Communications in Medicine (DICOM) format. Acquisitions with significant motion artifacts (n = 27), inconsistent DICOM header (n = 2), or liver lesions > 2 cm (n = 1) were excluded. The final study population thus included 200 patients. All MRI acquisitions included a non-contrast 3 mm T1-vibe Dixon in-phase sequence which was used in this study. Clinical characteristics were collected from the medical records and included the following: age, sex, weight, body mass index (BMI), and etiology of liver disease.

2.2. Magnetic resonance imaging

All magnetic resonance images were acquired at our university hospital on Siemens Magnetom scanners (154x Prisma^{fit} for 3 T, 46x Aera for 1.5 T; Siemens Healthineers, Erlangen, Germany). The acquisitions were made with a slice thickness of 3 mm and various pixel spacings ranging from 0.9375 × 0.9375 mm^2 to 1.5625 × 1.5625 mm^2. The sequences used in the study had axial dimensions ranging from 210 × 320 to 270 × 320 pixels with 60 to 104 axial slices. The choice of the non-contrast T1-vibe Dixon in-phase acquisition to build the dataset was based on previous work by Zbinden et al. [17], showing the best performance for parenchyma and vessel segmentation, while the addition

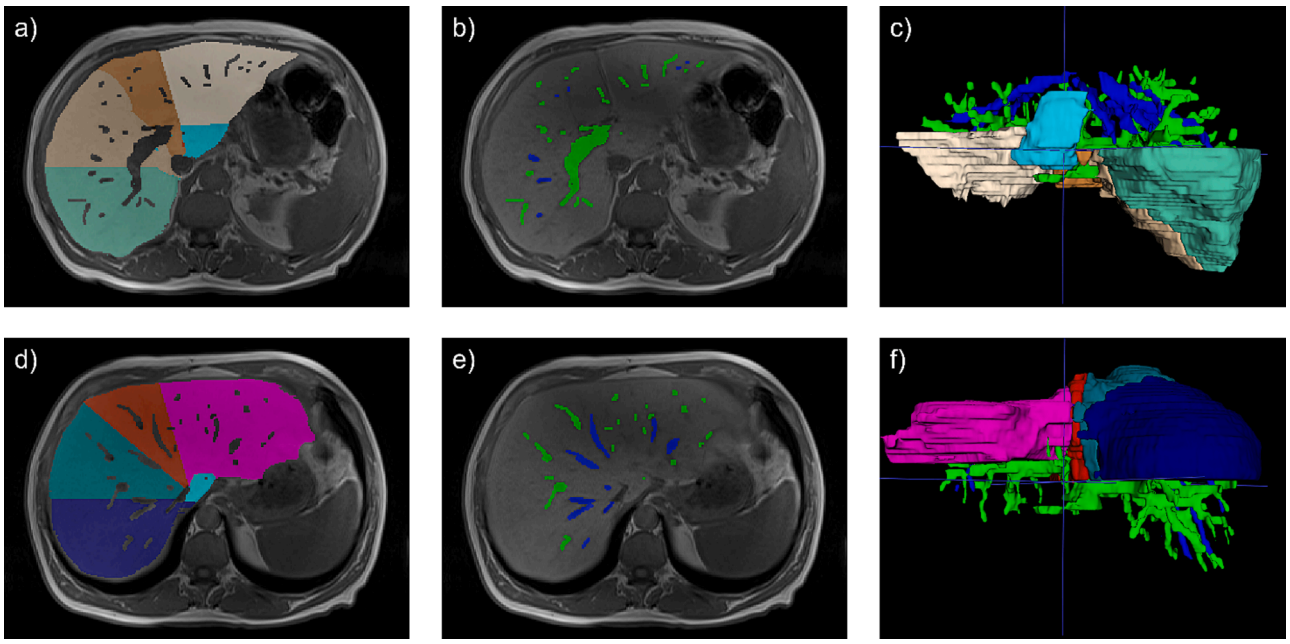

Fig. 2. Sample 63-year-old male patient without chronic liver disease. Manual annotations of 9 Couinaud segments are shown in a) and d), portal vessels and hepatic vessels are shown in b) and e), on two axial slices of a standard non-contrast T1-vibe Dixon in-phase sequence. The rightmost column c) and f) shows a 3D visualization of the complete manual segmentation.

of opposed-phase acquisition, as well as fat and water reconstructions, did not increase the accuracy of liver parenchyma and vessel segmentation.

2.3. Computer tomography

The CT scans of the patient cohort were acquired on Siemens Somatom Definition Flash (Siemens Healthineers, Erlangen, Germany), Siemens Somatom Definition Edge (Siemens Healthineers, Erlangen, Germany) and Philips Brilliance 64 (Philips, Best, Netherlands) scanners within 6 months of MRI, as previously published [18]. For CT scans, a pitch of 0.8 and a detector collimation of 0.6 were used. The acquisitions were automatically adapted based on a reference of 100 kVp and 150 mAs. Axial 1-mm slices were reconstructed with an increment of 1 mm in a liver parenchyma window, using the vendor-specific iterative reconstruction algorithm.

2.4. Preprocessing and manual segmentation

The liver MRI data sets were manually labeled slice-by-slice on in-phase images by a trained radiology resident with three years of experience in liver MRI (D.C.) and a second reader with one year of experience in liver MRI (L.H.) using ITK-SNAP (version 3.8.0) [19]. Liver segments and liver vessels (hepatic veins, portal veins) were manually segmented (Fig. 2). All results were reviewed by a board-certified abdominal radiologist with 7 years of experience (M.B.) and manually corrected, where necessary.

The Couinaud segment delineation on the CT scans was handled using the Philips IntelliSpace Portal software (version 10.1). The segmentation was done semi-automatically with an initial delineation of the liver parenchyma suggested by the system. Subsequently, a board-certified radiologist with 8 years of experience (V.C.O.) corrected the liver borders manually and provided key landmarks for the assignment of the segments by the system, while the borders of segments I-III and IV-VIII were manually corrected in every patient, where necessary. LSVR was calculated by dividing the liver segmental volumes I-III by the liver segmental volumes IV-VIII.

2.5. Deep learning algorithm

2.5.1. Segmentation network

The deep learning algorithm used for automated delineation of the Couinaud liver segments, and the liver vessels, was a 3D convolutional neural network with a U-Net architecture [20], the nnU-Net implementation [21]. The nnU-Net offers a fully automated machine learning pipeline including data preprocessing, data augmentation, network architecture optimization, and postprocessing. We employed it in its original form without modifying the pipeline. The nnU-Net neural network architecture has 30.7 million trainable parameters, enabling it to learn and represent complex patterns within the imaging data. The segmentation task for the DL model was to classify each MR imaging voxel into one of twelve categories, including 9 segments, portal vessels, hepatic vessels, and background. To be compatible with the neural network input file format, the T1-vibe Dixon in-phase sequences in Digital Imaging and Communications in Medicine (DICOM) format were converted to the Neuroimaging Informatics Technology Initiative (NIfTI) format. By employing nnU-Net without alterations, we aimed to capitalize on its state-of-the-art capabilities and demonstrate its effectiveness in the classification of MR imaging voxels, which holds significant promise for advancing medical image analysis and diagnosis.

2.5.2. Training and evaluation

The neural network was trained on the liver MRI with fivefold cross-validation using 170 patient MRI scans. Each fold had 136 liver MRI for training and 34 liver MRI for validation. The network 3D input was the single-modal T1-vibe Dixon in-phase sequence. 30 randomly selected patient MRI scans were not shown to the neural network for later out-of-sample hold-out testing. The nnU-Net network was trained with the default setup as published by Isensee, Jager, et al. [21]. This included as loss function the linear combination of Dice loss and cross-entropy loss, the Adam optimizer with an initial learning rate of 3×10^{-4}, and a learning rate scheduler that reduced the learning rate to at least 10^{-6} depending on the moving average of the training loss and the validation loss. For data augmentation, the following techniques were applied during training: random rotations, random scaling, random elastic deformations, gamma correction augmentation and mirroring. We did not

Table 1
Patient Characteristics.

	noCLD (n = 126)	CLD (n = 74)	p-value
Male (n, %)	64 (51%)	60 (81%)	<0.001
Age (years)	52 (38–62)	60 (54–67)	<0.001
Body mass index (kg/m^2)	24.5 (22.5–28.2)	27.6 (24.4–30.9)	0.001
Arterial Hypertension (n, %)	22 (17%)	35 (47%)	<0.001
Diabetes mellitus (n, %)	6 (5%)	28 (38%)	<0.001
Daily alcohol consumption (n, %)	1 (1%)	18 (24%)	<0.001
Pre-cirrhotic CLD (n, %)	0 (0%)	22 (30%)	<0.001
Liver cirrhosis (n, %)	0 (0%)	52 (70%)	<0.001
Child-Pugh-Groups (in patients with cirrhosis)			
- Child A (n, %)	0 (0%)	41 (55%)	<0.001
- Child B (n, %)	0 (0%)	11 (15%)	<0.001
- Child C (n, %)	0 (0%)	0 (0%)	1.0

Values are presented as median with interquartile range (25–75%) or n (%). P-values were calculated using the Mann-Whitney *U* test or Fisher's Exact test as appropriate.
noCLD = no chronic liver disease; CLD = chronic liver disease; PDFF = Proton-density fat fraction; Alcohol consumption was defined as average consumption of 2 alcoholic units per day for men and 1 alcoholic unit per day for women; Focal liver lesion was defined as actual focal liver lesion or treated focal liver lesion.

make any changes to the architecture of the nnU-Net.

The network optimization was performed on a NVIDIA GeForce RTX 3090 GPU for 150 epochs with a batch size of two, using 5-fold cross-validation following the approach by Isensee, Jaeger, et al. [21]. Subsequently, experiments using the hold-out test data were carried out. Liver segmental volumes were retrieved on the unseen test patients and the LSVR was calculated automatically as the liver segment volumes I-III divided by the liver segmental volumes IV-VIII. The LSVR was then compared with LSVR obtained from the MRI manual segmentation volumes and the LSVR obtained from the CT manual segmentation volumes.

2.6. Statistical analysis

The performance of the convolutional neural network to segment liver segments, portal veins, and hepatic veins was assessed with respect to the MRI manual annotations. We used the standard metrics Dice similarity coefficient (DSC), average symmetric surface distance (ASSD), and Hausdorff distance 95 (HD95) to quantitatively measure the segmentation quality of the DL model. DSC is used to measure the voxel overlap between the prediction and the ground truth. The value of DSC ranges from 0 to 1, with 1 indicating perfect overlap and 0 none. ASSD is the average of all the surface-based distances between the prediction and the ground truth. HD95 quantifies the maximum surface-based difference between the prediction and the ground truth. HD95 expresses this distance as the 95th percentile of all distances. DSC, ASSD, and HD95 values are reported as mean ± standard deviation (SD).

The LSVR was compared between the calculations by the DL model, the MRI manual method, and the CT manual method. We conducted statistical significance tests using the Kruskal-Wallis test with Dunn's multiple comparison post-hoc test for between-group differences. Further, the Spearman correlation coefficient was used to analyze the correlation between manual LSVR calculation and LSVR automated calculation. In addition, linear regression and Bland-Altman plots were used to further compare the manual and automated calculations. The Shapiro-Wilk test was performed to assess the patient parameters for normal distribution in the groups. The values were not normally distributed; therefore, nonparametric tests were used for all analyses. Parameters among the patient groups were compared using the Mann-Whitney *U* test for continuous variables or the Fisher's Exact test for categorical variables. Significance level was chosen to be $\alpha = 0.05$. To demonstrate inter-reader and inter-modality reliability of volumetry and LSVR measurements, intraclass correlation coefficients (ICCs) were calculated between the MRI readers and the CT reader. ICC classifications from 0.4 to 0.59 are considered as fair, those from 0.6 to 0.74 as good, and those from 0.75 to 1.00 are considered excellent [22]. All statistical analysis was conducted with the Python ecosystem (Python 3.6.12, SciPy 1.5.4, scikit-learn 0.24.2, dicom2nifti 2.4.6, SimpleITK 2.0.2, Matplotlib 3.3.3).

Table 2
Etiology liver disease.

Etiology	noCLD (n = 126)	CLD (n = 74)
NAFLD / NASH (n, %)	0 (0%)	29 (39%)
Viral Hepatitis (n, %)	0 (0%)	12 (16%)
ARLD (n, %)	0 (0%)	27 (37%)
Inflammatory / Autoimmune (n, %)	0 (0%)	5 (7%)
Drug-induced liver injury (DILI)	0 (0%)	1 (1%)

Values are presented as n (%).
noCLD = no chronic liver disease; CLD = chronic liver disease; NAFLD = nonalcoholic fatty liver disease; NASH = nonalcoholic steatohepatitis; ARLD = alcohol-related liver disease.

Table 3
Deep learning segmentation performance on Couinaud liver segments and liver vessels.

	DSC	ASSD	HD95
Portal vessels	0.663 ± 0.07	2.168 ± 0.64	10.706 ± 3.81
Hepatic vessels	0.559 ± 0.08	3.391 ± 1.35	18.194 ± 7.27
Segment I	0.744 ± 0.10	2.584 ± 1.40	8.729 ± 4.53
Segment II	0.801 ± 0.08	2.956 ± 1.85	11.114 ± 13.66
Segment III	0.754 ± 0.16	3.099 ± 1.95	9.089 ± 7.29
Segment IVa	0.743 ± 0.06	3.255 ± 0.96	11.717 ± 5.56
Segment IVb	0.726 ± 0.13	3.746 ± 2.90	13.821 ± 14.62
Segment V	0.810 ± 0.09	3.201 ± 1.43	11.852 ± 7.93
Segment VI	0.746 ± 0.20	3.797 ± 3.54	12.751 ± 13.94
Segment VII	0.805 ± 0.11	2.974 ± 1.55	10.396 ± 7.60
Segment VIII	0.796 ± 0.10	3.505 ± 1.91	12.02 ± 8.82
Segments I-III	0.868 ± 0.05	2.260 ± 1.48	9.420 ± 12.14
Segments IV – VIII	0.942 ± 0.01	1.444 ± 0.25	5.112 ± 1.27
All Segments	0.770 ± 0.03	3.235 ± 0.37	11.276 ± 1.56

Results are presented as mean ± SD.
DSC, Dice score; ASSD, Average Symmetric Surface Distance; HD95, Hausdorff. Distance 95.

3. Results

3.1. Patient characteristics

The clinical patient characteristics of the final cohort (n = 200) are shown in Table 1. There were 126 patients without chronic liver disease (noCLD) and 74 with chronic liver disease (CLD). Median age was 52 years (interquartile range (IQR) 38 – 62 years) for patients without and 60 years (IQR 54 – 67 years) for patients with CLD ($p < 0.001$). Patients without CLD showed a lower median BMI of 24.5 (IQR 22.5 – 28.2) than patients with CLD ($p = 0.001$) with a median of 27.6 (IQR 24.4 – 30.9). Etiology of chronic liver disease was nonalcoholic fatty liver disease or nonalcoholic steatohepatitis (NAFLD/NASH) in 29 patients, viral hepatitis in 12 patients, alcohol-related liver disease (ARLD) in 27 patients, and other causes in 6 patients (Table 2). Of the 74 patients with chronic liver disease, 52 patients had liver cirrhosis.

3.2. Liver segmental volume evaluation

The Couinaud liver segments and vessel segmentation results are summarized in Table 3. The neural network model classified the Couinaud segments I-VIII with an average DSC score of 0.770 ± 0.03,

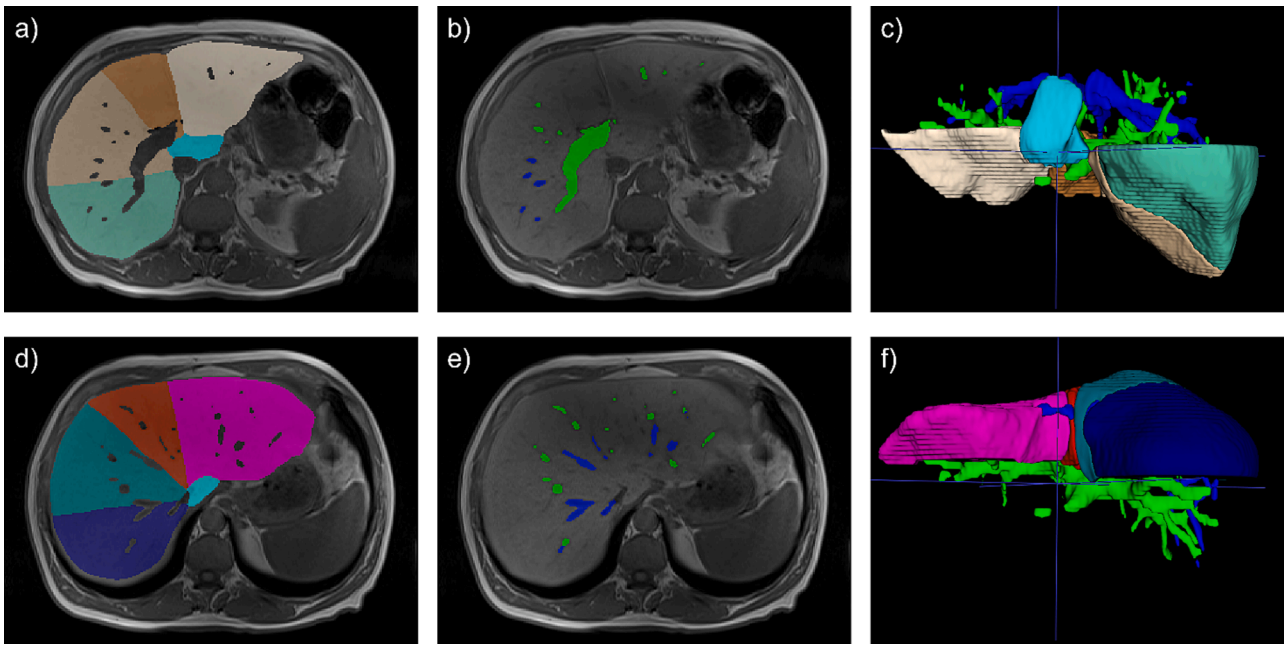

Fig. 3. AI automated segmentation of the same patient as in Fig. 2. 9 Couinaud segments are shown in a) and d), portal vessels and hepatic vessels are shown in b) and e), on two axial slices of a standard non-contrast T1-vibe Dixon in-phase sequence. The rightmost column c) and f) shows a 3D visualization of the complete automated segmentation.

Table 4
Liver segmental volume ratio (LSVR) as calculated by the MRI deep learning model (DL), MRI manual and CT manual on the validation data.

	MR DL	MR manual	CT manual	p-value	MR DL vs. MRI manual	MR DL vs. CT manual	MR manual vs. CT manual
LSVR	0.323 ± 0.14	0.328 ± 0.15	0.352 ± 0.14	0.652			
LSVR MAE					0.024	0.039	0.032
LSVR MAPE					7.19%	10.34%	12.24%

Results are presented as mean ± SD and percentage.
The p-value was calculated using the Kruskal-Wallis test with Dunn's multiple comparison post-hoc test.
MAE, mean absolute error; MAPE, mean absolute percent error. MR DL, deep learning-automated LSVR on MR; MR manual, manual LSVR on MR; CT manual, manual LSVR on CT.

ranging from 0.726 ± 0.13 (segment IVb) to 0.810 ± 0.09 (segment V). The liver vessels were segmented with a DSC of 0.663 ± 0.07 for portal vessels and a DSC of 0.559 ± 0.08 for hepatic vessels. A qualitative segmentation result is shown in Fig. 3.

3.3. Mean LSVR comparison between neural network and manual segmentation

The mean LSVR of the test liver MRIs as calculated by the neural network, the MR manual, and the CT manual is presented in Table 4. The calculated mean LSVR with test liver MRI unseen by the neural network was 0.323 ± 0.14, as compared with MRI manually quantified LSVR of 0.328 ± 0.15, resulting in a mean absolute error (MAE) of 0.024 and a mean absolute percentage error (MAPE) of 7.19%, respectively. A comparable mean LSVR of 0.352 ± 0.14 with a MAE of 0.039 and a MAPE of 10.34% resulted with the LSVR retrieved from the CT scans in the same patients. There was no statistically significant difference between the three mean LSVRs as shown by a Kruskal-Wallis test ($p = 0.652$). The variance of the test LSVRs was identical in all three methods (from ± 0.14 to ± 0.15).

In Fig. 4, the scatterplots showcase the strong correlation between MR DL and MR manual, and MR DL and CT manual. The Bland-Altman plot between MR manual and MR DL shows only a slight mean bias (0.01) with almost no outliers, underlining that the MR DL calculates the LSVR in close approximation to the MR manual. Moreover, the scatterplots show that the LSVR measurements occur on both sides of the cirrhosis threshold of 0.35 [1], evidencing that the DL model can robustly and reliably predict the LSVR in pathological cases as well.

3.4. Intra-patient correlation between neural network and manual segmentation

Pairwise LSVR measurements between neural network and manual MR and CT segmentation are shown in Fig. 4. LSVR calculated automatically by the neural network correlated very well with manual MR segmentation ($r = 0.97$, $p < 0.001$), as well as manual CT segmentation ($r = 0.95$, $p < 0.001$). There was also a good correlation between manual MR and CT segmentation ($r = 0.97$, $p < 0.001$). Bland-Altman plots showed small mean biases, ranging between 0.01 and 0.03 for all three pairwise comparisons.

3.5. Segmentation time of the neural network

Liver segments and vessel segmentation inference in one test liver MRI and calculation of the LSVR was performed in 70 s using an NVIDIA GeForce RTX 3090 GPU, an AMD EPYC 7302 16-Core Processor CPU, and an IBM Spectrum Scale-based file system. A manual segmentation took 35 min on average for a single liver MR.

3.6. Interobserver reliability

Interobserver reliability between manual CT and MRI liver volume

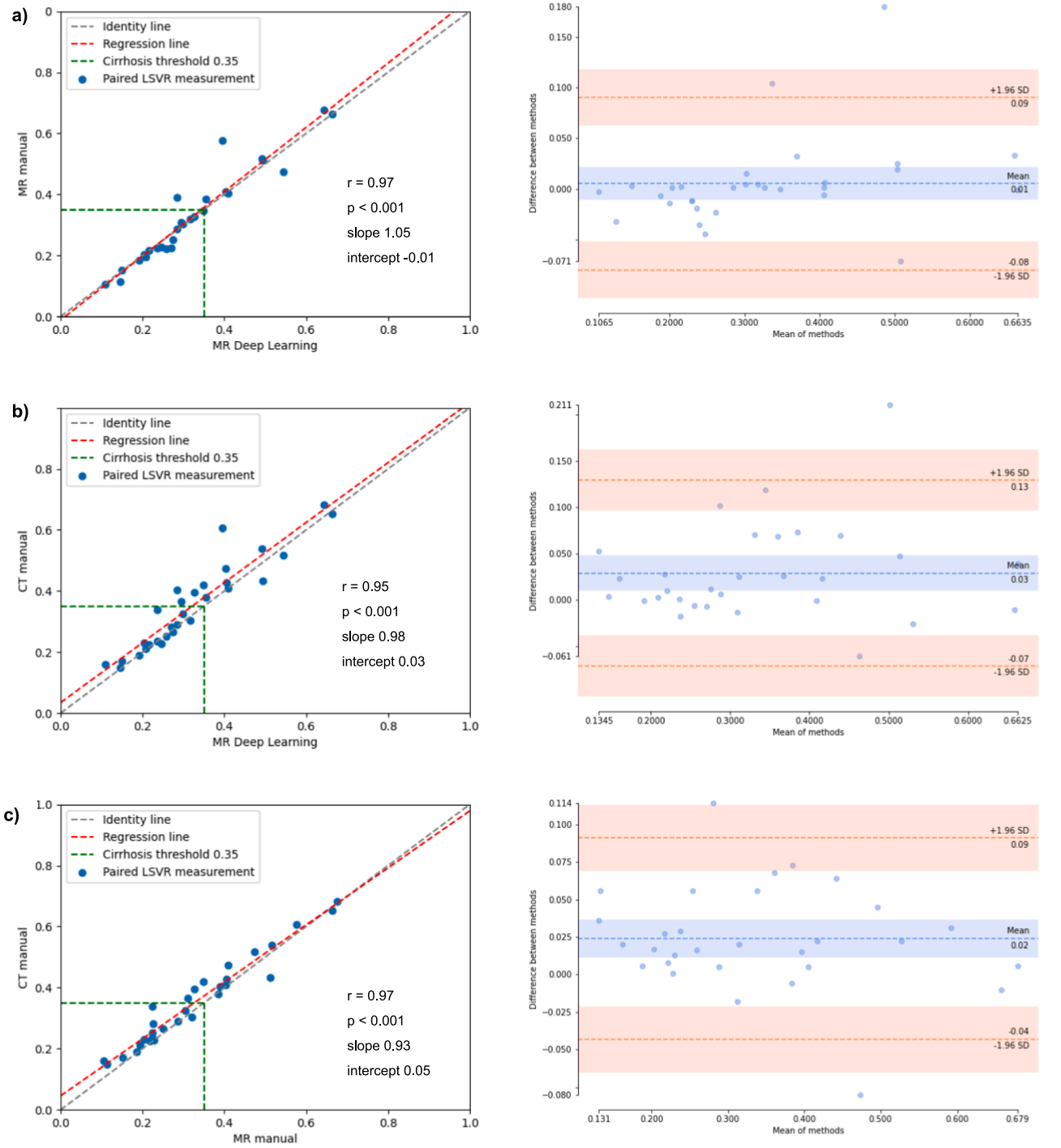

Fig. 4. Scatterplots of pairwise LSVR measurements (left column) with corresponding Bland-Altman plots (right column). Row a) compares MR deep learning with MR manual. Row b) compares MR deep learning with CT manual. Row c) compares MR manual with CT manual. All pairwise comparisons show a significant Spearman correlation ($r > 0.94$, $p < 0.001$).

segmentation was excellent, with an ICC of 0.93 (95% CI 0.44, 0.98) for Couinaud segments I-III, and an ICC of 0.95 (95% CI 0.84, 0.98) for Couinaud segments IV-VIII. In addition, the manual LSVR measurements were found with an ICC of 0.96 (95% CI 0.86, 0.99), demonstrating an excellent inter-reader and inter-modality reliability.

4. Discussion

We showed that a deep learning-based algorithm for automated 3D segmentation of Couinaud liver segments and liver vessels allows accurate calculation of the liver segmental volume ratio (LSVR). To our knowledge, this is the first work to present a neural network-based automated calculation of the LSVR using non-contrast T1-vibe Dixon

Table 5
Quantitative performance comparison between MRI-based liver segments and veins segmentation methods.

		Han et al. [23]	Arya et al. [24]	**Ours**
Model		3D CNN	LPD	nnU-Net
All segments		0.902	0.802 ± 0.08	0.770 ± 0.03
Portal veins		–	–	0.663 ± 0.07
Hepatic veins		–	–	0.559 ± 0.08
Dataset	Contrast enhanced	yes	no	no
	Liver vessels excluded from parenchyma	no	no	yes
	Caudate lobe segmented separately	no	yes	yes
	Patients with liver cirrhosis	no	n/a	yes
	MRI characteristics	T1 portal venous phase, 1.5 T/3T	T1-weighted MRI, 1.5/3T	T1 Dixon in-phase, 1.5 T/3T
	Name	private	HepaT1ca	private
	Training size, n	367	85	170
	Testing size, n	158	23	30

All scores are in Dice (DSC) metric. Results are shown as mean ± SD. The MRI dataset used by each method is summarized. Our study is the first to. evaluate a 3D convolutional neural network (nnU-Net) for the measurement of the LSVR on non-contrast T1-vibe Dixon liver MRIs.
MRI, magnetic resonance imaging; 3D CNN, 3-dimensional convolutional neural network; LPD, landmark prediction derived plus predicted liver.

liver MRI. LSVR was performed in<2 min with the neural network model, as compared to manual segmentation which took more than half an hour on average. The results of the automated LSVR calculation on the test data were very accurate and correlated strongly with the MRI manual calculation (Spearman $r = 0.97$, $p < 0.001$), as well as with the manual segmentation on CT scans of the same patients as an intermodal cross-validation (Spearman $r = 0.95$, $p < 0.001$). As CT and MRI manual LSVR calculation was performed by different radiologists, this comparison showed a high inter-reader and inter-modality reliability of the automated LSVR (ICC = 0.96). As the automated segmentation did not rely on contrast-enhanced MRI, it may be used on any standard MRI, even without contrast medium administration, as it is currently discussed in several indications [16].

Our liver segment results are comparable with two other similar studies in existing literature, as shown in Table 5. The larger study of Han et al. [23] used contrast-enhanced MRI scans and excluded patients with liver cirrhosis. In addition, they did not differentiate liver vessels that were counted as liver segmental volume and did not distinguish segment IV into IVa and IVb. This may explain the slightly higher Dice score (0.902) for the liver segmentation than we found in our study (0.770). However, the exclusion of patients with liver cirrhosis is not helpful in a realistic setting and the additional exclusion of the liver vessels in our study allows for more accurate segmentation of the liver segmental volume. In addition, our approach has the advantage to be applied to liver MRI without contrast medium administration. The smaller study of Arya et al. [24] excluded the anatomically difficult liver segment I and just analyzed segments II-VIII. As in the study of Han et al., liver vessels were counted as liver parenchyma, which made the segmentation task easier. Nevertheless, they achieved a comparable Dice score (0.802) to the one we found in our study.

Mojtahed et al. [25] studied a deep learning-based Couinaud segment volume measurement tool for MR images but did not report DSC scores on the segments. The only study investigating LSVR calculation with a neural network was recently performed in contrast-enhanced CT scans, reporting an average DSC of 0.834 ± 0.08 [15].

The main limitations of this study were the single-center study design and that all MRI scans were performed on MRI scanners from the same manufacturer. Therefore, the study results should be externally validated on different scanners. Another limitation is that the dataset did not include focal liver lesions or postoperative livers.

In conclusion, a convolutional neural network allowed for accurate automated liver segmental volume quantification and calculation of LSVR based on a non-contrast T1-vibe Dixon sequence.

CRediT authorship contribution statement

Lukas Zbinden: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing, Validation. **Damiano Catucci:** Formal analysis, Investigation, Methodology, Writing – review & editing. **Yannick Suter:** Data curation, Investigation, Methodology, Writing – review & editing. **Leona Hulbert:** Data curation, Methodology, Writing – review & editing. **Annalisa Berzigotti:** Data curation, Investigation, Methodology, Writing – review & editing. **Michael Brönnimann:** Data curation, Investigation, Methodology, Writing – review & editing. **Lukas Ebner:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing, Validation. **Andreas Christe:** Data curation, Methodology, Validation, Writing – review & editing. **Verena Carola Obmann:** . **Raphael Sznitman:** Conceptualization, Data curation, Investigation, Methodology, Supervision, Writing – review & editing. **Adrian Thomas Huber:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was funded by the Swiss National Science Foundation (SNSF) #320030_188591. Calculations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

References

[1] W.E. Torres, L.F. Whitmire, K. Gedgaudas-McClees, M.E. Bernardino, Computed Tomography of Hepatic Morphologic Changes in Cirrhosis of the Liver, J. Comput. Assist. Tomogr. 10 (1986) 47–50, https://doi.org/10.1097/00004728-198601000-00009.
[2] A. Huber, L. Ebner, J.T. Heverhagen, A. Christe, State-of-the-art imaging of liver fibrosis and cirrhosis: A comprehensive review of current applications and future perspectives, Eur. J. Radiol. Open. 2 (2015) 90–100, https://doi.org/10.1016/j.ejro.2015.05.002.
[3] M. Romero-Cristóbal, A. Clemente-Sánchez, M. Peligros, E. Ramón, A. Matilla, A. Colón, S. Alonso, M. Catalina, A. Fernández-Yunquera, A. Caballero, R. García, J.Á. López-Baena, M. Salcedo, R. Bañares, D. Rincón, Liver and spleen volumes are associated with prognosis of compensated and decompensated cirrhosis and parallel its natural history, United European, Gastroenterol. J. 10 (2022) 805–816, https://doi.org/10.1002/ueg2.12301.
[4] A. de Gottardi, C. Sempoux, A. Berzigotti, Porto-sinusoidal vascular disorder, J. Hepatol. 77 (2022) 1124–1135, https://doi.org/10.1016/j.jhep.2022.05.033.
[5] O.M. Furusato Hunt, M.G. Lubner, T.J. Ziemlewicz, A. Muñoz del Rio, P. J. Pickhardt, The Liver Segmental Volume Ratio for Noninvasive Detection of

Cirrhosis, J. Comput. Assist. Tomogr. 40 (2016) 478–484, https://doi.org/10.1097/RCT.0000000000000389.
[6] P.J. Pickhardt, K. Malecki, O.F. Hunt, C. Beaumont, J. Kloke, T.J. Ziemlewicz, M. G. Lubner, Hepatosplenic volumetric assessment at MDCT for staging liver fibrosis, Eur. Radiol. 27 (2017) 3060–3068, https://doi.org/10.1007/s00330-016-4648-0.
[7] O.M.F. Hunt, M.G. Lubner, T.J. Ziemlewicz, A.M. del Rio, P.J. Pickhardt, The liver segmental volume ratio for noninvasive detection of cirrhosis: Comparison with established linear and volumetric measures, J. Comput. Assist. Tomogr. 40 (2016) 478–484, https://doi.org/10.1097/RCT.0000000000000389.
[8] V.C. Obmann, C. Marx, J. Hrycyk, A. Berzigotti, L. Ebner, N. Mertineit, C. Gräni, J. T. Heverhagen, A. Christe, A.T. Huber, Liver segmental volume and attenuation ratio (LSVAR) on portal venous CT scans improves the detection of clinically significant liver fibrosis compared to liver segmental volume ratio (LSVR), Abdominal, Radiology 46 (2021) 1912–1921, https://doi.org/10.1007/s00261-020-02834-7.
[9] Y. Qin, M. van Cauteren, M. Osteaux, G. Willems, Determination of liver volume in vivo in rats using MRI, Eur. J. Radiol. 11 (1990) 191–195, https://doi.org/10.1016/0720-048X(90)90054-F.
[10] A.M. Roloff, P. Heiss, T.P. Schneider, A. Quadrat, M.L. Kromrey, F. Zeman, C. Stroszczynski, B. Mensel, J.P. Kühn, Accuracy of simple approaches to assessing liver volume in radiological imaging, Abdominal, Radiology 41 (2016) 1293–1299, https://doi.org/10.1007/s00261-016-0672-4.
[11] M. Mazonakis, J. Damilakis, T. Maris, P. Prassopoulos, N. Gourtsoyiannis, Comparison of two volumetric techniques for estimating liver volume using magnetic resonance imaging, J. Magn. Reson. Imaging 15 (2002) 557–563, https://doi.org/10.1002/jmri.10109.
[12] M. Puttagunta, S. Ravi, Medical image analysis based on deep learning approach, Multimed. Tools Appl. 80 (2021) 24365–24398, https://doi.org/10.1007/s11042-021-10707-4.
[13] A.S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on MRI, Z. Med. Phys. 29 (2019) 102–127, https://doi.org/10.1016/j.zemedi.2018.11.002.
[14] S. Gassenmaier, T. Küstner, D. Nickel, J. Herrmann, R. Hoffmann, H. Almansour, S. Afat, K. Nikolaou, A.E. Othman, Deep Learning Applications in Magnetic Resonance Imaging: Has the Future Become Present? Diagnostics. 11 (2021) 2181, https://doi.org/10.3390/diagnostics11122181.
[15] S. Lee, D.C. Elton, A.H. Yang, C. Koh, D.E. Kleiner, M.G. Lubner, P.J. Pickhardt, R. M. Summers, Fully Automated and Explainable Liver Segmental Volume Ratio and Spleen Segmentation at CT for Diagnosing Cirrhosis, Radiol Artif Intell. 4 (2022), https://doi.org/10.1148/ryai.210268.
[16] M.V. Chan, S.J. McDonald, Y.-Y. Ong, K. Mastrocostas, E. Ho, Y.R. Huo, C. Santhakumar, A.U. Lee, J. Yang, HCC screening: assessment of an abbreviated non-contrast MRI protocol, Eur Radiol Exp. 3 (2019) 49, https://doi.org/10.1186/s41747-019-0126-1.
[17] L. Zbinden, D. Catucci, Y. Suter, A. Berzigotti, L. Ebner, A. Christe, V.C. Obmann, R. Sznitman, A.T. Huber, Convolutional neural network for automated segmentation of the liver and its vessels on non-contrast T1 vibe Dixon acquisitions, Sci. Rep. 12 (2022) 22059, https://doi.org/10.1038/s41598-022-26328-2.
[18] V.C. Obmann, N. Mertineit, A. Berzigotti, C. Marx, L. Ebner, R. Kreis, P. Vermathen, J.T. Heverhagen, A. Christe, A.T. Huber, CT predicts liver fibrosis: Prospective evaluation of morphology- and attenuation-based quantitative scores in routine portal venous abdominal scans, PLoS One 13 (2018), https://doi.org/10.1371/journal.pone.0199611.
[19] P.A. Yushkevich, J. Piven, H.C. Hazlett, R.G. Smith, S. Ho, J.C. Gee, G. Gerig, User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability, Neuroimage 31 (2006) 1116–1128, https://doi.org/10.1016/j.neuroimage.2006.01.015.
[20] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, (2015). http://arxiv.org/abs/1505.04597.
[21] F. Isensee, P.F. Jaeger, S.A.A. Kohl, J. Petersen, K.H. Maier-Hein, nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, Nat. Methods 18 (2021) 203–211, https://doi.org/10.1038/s41592-020-01008-z.
[22] D.V. Cicchetti, Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology, Psychol. Assess. 6 (1994) 284–290, https://doi.org/10.1037/1040-3590.6.4.284.
[23] X. Han, X. Wu, S. Wang, L. Xu, H. Xu, D. Zheng, N. Yu, Y. Hong, Z. Yu, D. Yang, Z. Yang, Automated segmentation of liver segment on portal venous phase MR images using a 3D convolutional neural network, Insights Imaging. 13 (2022), https://doi.org/10.1186/s13244-022-01163-1.
[24] Z. Arya, G. Ridgway, A. Jandor, P. Aljabar, Deep Learning-Based Landmark Localisation in the Liver for Couinaud Segmentation, in: Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial intelligence and Lecture Notes in Bioinformatics), Springer Science and Business Media Deutschland GmbH, 2021, pp. 227–237, https://doi.org/10.1007/978-3-030-80432-9_18.
[25] A. Mojtahed, L. Núñez, J. Connell, A. Fichera, R. Nicholls, A. Barone, M. Marieiro, A. Puddu, Z. Arya, C. Ferreira, G. Ridgway, M. Kelly, H.J. Lamb, F. Caseiro-Alves, J. M. Brady, R. Banerjee, Repeatability and reproducibility of deep-learning-based liver volume and Couinaud segment volume measurement tool, Abdominal, Radiology 47 (2022) 143–151, https://doi.org/10.1007/s00261-021-03262-x.