HHS Public Access

Author manuscript

Ann Surg Open. Author manuscript; available in PMC 2022 October 21.

Published in final edited form as:

Ann Surg Open. 2022 June ; 3(2): . doi:10.1097/as9.0000000000000155.

nnU-Net Deep Learning Method for Segmenting Parenchyma and Determining Liver Volume From Computed Tomography Images

Rowland W. Pettit, BS[*], **Britton B. Marlatt, MS**[†], **Stuart J. Corr, BE, MEE, MA, PhD**[‡,§,‖,¶,#], **Jim Havelka, MS, MBA**[†], **Abbas Rana, MD**[**]

[*]Department of Medicine, Baylor College of Medicine, Houston, TX

[†]InformAI, Houston, TX

[‡]Department of Innovation Systems Engineering, Houston Methodist, Houston, TX

[§]Department of Cardiovascular Surgery, Houston Methodist Hospital, Houston, TX

[‖]Department of Bioengineering, Rice University, Houston, TX

[¶]Department of Biomedical Engineering, University of Houston, Houston, TX

[#]Swansea University Medical School, Wales, United Kingdom

[**]Department of Surgery, Division of Abdominal Transplantation, Baylor College of Medicine, Houston, TX.

Abstract

Background: Recipient donor matching in liver transplantation can require precise estimations of liver volume. Currently utilized demographic-based organ volume estimates are imprecise and nonspecific. Manual image organ annotation from medical imaging is effective; however, this process is cumbersome, often taking an undesirable length of time to complete. Additionally, manual organ segmentation and volume measurement incurs additional direct costs to payers for either a clinician or trained technician to complete. Deep learning-based image automatic segmentation tools are well positioned to address this clinical need.

Objectives: To build a deep learning model that could accurately estimate liver volumes and create 3D organ renderings from computed tomography (CT) medical images.

Methods: We trained a nnU-Net deep learning model to identify liver borders in images of the abdominal cavity. We used 151 publicly available CT scans. For each CT scan, a board-certified radiologist annotated the liver margins (ground truth annotations). We split our image dataset into training, validation, and test sets. We trained our nnU-Net model on these data to identify liver borders in 3D voxels and integrated these to reconstruct a total organ volume estimate.

Reprints: Rowland W. Pettit, BS, Department of Medicine, Institute for Clinical and Translational Science, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030. Rowland.Pettit@bcm.edu.

R.W.P. and B.B.M. contributed equally.

Results: The nnU-Net model accurately identified the border of the liver with a mean overlap accuracy of 97.5% compared with ground truth annotations. Our calculated volume estimates achieved a mean percent error of 1.92% + 1.54% on the test set.

Conclusions: Precise volume estimation of livers from CT scans is accurate using a nnU-Net deep learning architecture. Appropriately deployed, a nnU-Net algorithm is accurate and quick, making it suitable for incorporation into the pretransplant clinical decision-making workflow.

Keywords

computed tomography; deep learning; liver; machine learning; nnunet; segmentation; solid organ transplantation; transplantation; volume

INTRODUCTION

Solid organ transplantations are medically and surgically complex procedures.[1] To achieve successful surgical outcomes, transplantation teams must appropriately consider numerous clinical, biochemical, and anthropometric features for both an organ donor and a patient recipient.[2] Of these considerations, the size and shape of an offered organ can be relevant. Size mismatches exist between donors and recipients and are often of little consequence.[3] However, if a donated organ is relatively large or a recipient has anatomical cavity limitations, a donor/recipient size mismatch may prevent a transplantation procedure from occurring or may lead to an adverse clinical outcome. In this way, estimation of donor organ volume and recipient abdominal cavity is essential to clinical decision-making in solid organ transplantation.[3]

Size mismatch limitations are prominent in the field of liver transplantation surgery[3] where abdominal cavity constraints, especially in pediatric populations,[4] exist and must be identified prior to organ harvesting. Interoperative realization of an organ size mismatch is problematic since it can lead to difficulty closing the abdomen or place excessive tension on the allograft leading to graft failure.[5] For this reason, size matching is a significant endeavor that requires accuracy. Transplantation teams take great care to estimate donor organ and recipient cavity volumes before initiating a transplantation procedure when necessary. At present, several methods for organ volume estimation and assessment of anatomical spatial constraints exist.[6-8] Organ size estimates may be made informally by viewing recipient and donor full computed tomography (CT) scans side-by-side prior to an operation. Total liver volume (TLV) estimates may also be estimated through general demographic-based estimation techniques[9] or with height and weight-based normalized calculators.[10] While these estimation techniques have some clinical utility, they are often inaccurate.[10] Alternatively, liver segmentation and volume estimation is conducted using semi-automatic computational methods.[11] In semi-automatic segmentation, probabilistic models are automatically generated from a manual seed point on the image identified by a user. The model then attempts to identify relevant boundaries per slice in the medical image, while human-guided pixel assignment is iteratively conducted to refine predictions. While accurate, these processes are iterative and require time, clinical expertise, and manual human guidance.

In light of recent advancements made with machine learning[12] and deep learning[13] based image segmentation tools,[14-19] it is time to reevaluate and create a top performing, and fully automated organ volume estimation model. The “no new U-net” or nnU-Net algorithm[20] is a deep learning convolutional neural network method that has recently been developed and validated as an accurate tool for automatically segmenting features from 2D or 3D image modalities.[20] The nnU-Net method is an extension of the recent, but highly impactful U-Net deep learning architecture,[21] which demonstrated impressive object identification[22] and image segmentation capabilities.[23,24] The problem that nnU-Net solves, which remains unaddressed in general U-Net architectures, is the problem of generalizability. Individual U-Net deep learning models are highly specific and require manual selection of model design elements. They are designed for optimized use on certain segmentation tasks, such as appreciating 2D or 3D image information, respectively. nnUNet optimizes these processes as an end-to-end solution. The nnU-Net method combines multiple U-Net approaches and automatically adapts its use of several U-Net architectures to appropriately model images of any given input geometry or information type.[20]

For this reason, the nnU-Net deep learning modality has the capability to optimally segment organs from CT scans, which relies on density or hypodensities revealed from x-ray images to differentiate between relevant biological features within a patient.[25] While the applications of nnU-Net are diverse, they have previously been successfully applied to estimate organ shapes, volumes, or other relevant biological features from multiple imaging modalities, including CT scans.[26-28] A gap in capability currently exists, however, determining the efficacy of utilizing nnU-Net to segment and specifically estimate intra-abdominal organ volumes prior to transplantation.

We sought to fill this gap in knowledge by creating and validating a nnU-Net model to segment and calculate accurate organ volume estimates acceptable for aiding in solid organ transplantation clinical decision-making. In this work, we will discuss the success of this initiative and present our results for the accurate 3D renderings and volume estimations of livers, which can be generated with nnU-Net from CT scans. These results can inform transplantation surgical decision-making and may have broader applications in the field of surgery and procedural-based specialties.

METHODS

Data Acquisition

To implement nnU-Net, we obtained pre-annotated CT scan images from 2 databases. The full methods for data collection, annotation, and publication of these databases have been published previously, and we summarize them here.[29,30] The Liver Tumor Segmentation Challenge (LiTS)[29] is a publicly accessible competition consortium published in 2017 of 131 abdominal CT scans. The 3D Image Reconstruction for Comparison of Algorithm Database (IRCAD)[22] is also a publicly accessible dataset published in 2015 containing 20 abdominal CT scans. We accessed both databases on May 31, 2021, downloading the LiTS (https://www.kaggle.com/andrewmvd/liver-tumor-segmentation/version/1) and IRCAD images (https://www.ircad.fr/research/3dircadb/) completely. The CT scans housed in LiTS were segmented, annotated, and harmonized to single reference standard contour by 3 board-

certified radiologists,[29] demarcating the borders of the liver from other abdominal contents. All the CT scans in the training, validation, and test sets are uniformly 512 × 512 pixels per slice on their "*X*" and "*Y*" axes. The *Z* axes, or slice thickness, varied between CT samples. The average number of *z* axis slices across our CT samples is 142 slices with a standard deviation of 43 slices, and corresponding slice thicknesses ranged from 0.5 millimeters to 2.0 millimeters.

Preprocessing

Additional annotations exist within these datasets, including radiologist segmentations of bones, vasculature, intraparenchymal tumors, and other organs. To harmonize these data, we grouped and utilized only the radiologist annotations demarcating the contours of the liver, ignoring additional sub annotations for tumors, including intrahepatic tumors, vasculature, or other abdominal organs. Specifically, we did not retain any annotations denoting contours of the gallbladder, the intraparenchymal vasculature, the external biliary system, the hepatic portal vein, the proper hepatic artery, or the hepatic veins. These images are stored in these databases as Digital Imaging and Communications in Medicine (DICOM[31]). To facilitate our analysis with the nnU-Net deep learning method, and specifically for ease in converting these 3D images into a 3D array, we converted these DICOM files to Neuroimaging Informatics Technology Initiative (NIfTI[31]) image files using the python coding package dicom2nifti[32].

nnU-Net Architecture

We utilized the python package nnunet[33] (version 1.6.6) to train the deep learning algorithm. The *nnunet* python package is self-contained and optimizes a set 3D U-Net model architecture. As can be observed in Figure 1, the model takes the characteristic "U" shape. Three-dimensional arrays representing each input CT image are input into the model and undergo a series of convolutions, maximum pooling operations, up-convolutions, and finally, concatenation steps. Intrinsic to the U-Net design, information early on in the network architecture is leaked into the final stages of the network through horizontal concatenation steps (represented as green lines in Fig. 1). We represent the standard convolution steps as yellow forward arrows in the figure, which include batch normalization and a rectified linear unit activation function[34] implementation as these steps. Batch normalization was implemented for each layer of the nnU-Net through recentering/rescaling. This step standardized each batch of images trained per layer to speed up model training, likely through minimizing internal covariate shift.[35] The first half (left) of the U-shaped architecture in Figure 1 can be interpreted as conducting image analysis, while the second half synthesizes the segmentation image. Convolution steps were conducted with a 3 × 3 × 3 sliding voxel window. Maximum pooling steps condense a 2 × 2 × 2 voxel window. Up-convolution layers (also known as "deconvolution" or "transposed" convolution layers[36]) involve a 2 × 2 × 2 sliding voxel window, with a constrained step length of 2. The final convolution in the analysis layer is a 1 × 1 × 1 voxel convolution.

Model Training and Validation

After database harmonization and file conversion, 151 abdominal CT scans with annotated liver margins were available for analysis. We randomly split the 131 LiTS images into an

80% training set of 104 scans and a 20% validation set of 27 images. We reserved the remaining 20 independent images from the IRCAD database as a test set, including 10 male and 10 female CT scans. The sex or other demographic features of the LiTS database were not provided. With training, validation, and test sets sectioned, we initiated training of these segmented CT scans with nnU-Net. The nnunet python package has preprocessing, which occurs internally that automatically performs parameter tuning to optimize segmentation accuracy in the validation set during training and ultimately on the test set. Internal to the nnU-Net software is a unique loss function that combines the Dice coefficient[37] with cross-entropy and is utilized to measure model loss and update parameters during training of the segmentation tool. The Dice metric is a measure of overlap, quantifying the similarity between 2 segmented regions. We trained the nnU-Net model for 4000 epochs, where an epoch is defined as the case where the model has iteratively trained on each of our 151 images an additional time. Model performance was determined through a global Dice score, which averages the degree of overlap between the nnU-Net segmentation predictions of liver parenchyma compared to the known, ground truth, radiologist annotations. This evaluation metric was obtained at each step of model training from the test set; however, no feedback information from the test set was used in training the model. In this way, the test set remained an independent set to compare model performance on never-before-seen data objectively.

Volume Estimation and 3D Renderings

Our final trained and validated nnU-Net model segmented the CT scans by iteratively scanning 3D voxels across the entire image volume. A voxel is a standardized volume used as a base element during the convolutional processing of an image. In our model training, the pixel spacing and slice thickness for each CT scan was obtained through the NIfTI file header, standardizing voxel sizing across the unique scans. After identifying the borders of the liver with nnU-Net, we were able to create a separated, reconstructed 3D image representation of the liver by selecting only for the voxels that contained liver parenchyma or liver borders. Additionally, a colored region annotation was conducted for multiaxis 2D viewing of liver parenchyma predictions superimposed in space onto the original NIfTI CT image. Finally, by obtaining voxel dimensions per unique CT scan from the image meta-data, a volumetric estimate of liver volume was obtained by summing the nonoverlapping voxels that were predicted to comprise liver tissue. We compared our voxel-based predicted volume estimates of the IRCAD test CT scans against the ground truth volumes for these images with percent error metrics. These methods were conducted using 64 central processing units, 488 gigabyte random-access memory, and 8 Nvidia (Santa Clara, CA) V100 graphics processing units hosted on Amazon Web Services (Seattle, WA).

RESULTS

Training and Validation Outcomes

We successfully created a nnU-Net convolutional neural network to segment livers from the combined LiTS-IRCAD database of abdominal CT scans. The *nnunet* python package is self-contained and optimized the established 3D U-Net model architecture. In total, 19,069,955 independent parameters combined between the convolution, pooling, and up-

convolution steps were optimized to train the nnU-Net model. After 4000 epochs of training, an optimized nnU-Net model was established. A visualization of the training set, validation set, and final evaluation (test set) metrics can be visualized in Supplementary Figure 1 (http://links.lww.com/AOSO/A111) as blue, red, and green lines, respectively. A stochastic but steady decrease in the loss function for the training and validation sets was observed through the 4000 epochs. At 1000 epochs, a global Dice similarity score of 0.956 was obtained, with a Dice of 0.959 at 2000 epochs. We stopped model training at 4000 upon meeting a check-in thresh-old of <1% increase in global Dice similarity score within 2000 epochs. Ultimately, a training set loss of −0.7805, and validation set loss of −0.6785, and an evaluation metric of 0.966 were obtained at the conclusion of our model training.

Segmentation Accuracy

The ground truth whole liver volume segmentation mask for each CT scan is manually constructed by board-certified radiologists. To compare the complete liver volumes autosegmented via our trained 3D U-Net, we calculate the Dice coefficient, quantifying the overlap between the ground truth whole liver volume segmentation mask and the whole liver segmentation mask predicted by the trained nnU-Net model for each sample in the test set. For the male samples, we obtained a median global Dice similarity coefficient of 0.975. The standard deviation in global Dice coefficients among the male samples 0.006 with a maximum of 0.985 and a minimum of 0.965. For the female samples, we obtained a median global Dice similarity coefficient of 0.975. The standard deviation in global Dice coefficients among the female samples was 0.004 with a maximum of 0.980 and a minimum of 0.968. Combined, we achieved an average global Dice similarity coefficient of 0.974 across all samples in the test set with a standard deviation of 0.005. We also report the 95th percentile of the (symmetric) Hausdorff distance (HD) between predicted versus ground truth liver contours. Overall, we obtained an average 95th percentile HD of 2.458 ± 0.933. These model performance results on individual test samples and overall compiled summary statistics may be appreciated in Tables 1 and 2, respectively.

Volume Estimation Accuracy

The ground truth volumes for each of the CT scans obtained in our combined LiTS-IRCAD database are presented in Figure 2. It is visually appreciable through box plots the median and interquartile ranges of the TLVs (cm^3) captured between the training, validation, and testing image sets (Fig. 3). For the male samples, we obtained a median percent error of 2.055% in TLV (cm^3) estimation comparing our nnU-Net predictions to the ground truth for each CT. The standard deviation of percent error for male livers was 1.746%, with a maximum of 5.560% and a minimum of 0.223%. We obtained a median percent error of 0.911% in TLV (cm^3) estimation comparing our nnU-Net predictions to the ground truth for each CT for the female samples. The standard deviation of percent error for female livers was 1.149%, with a maximum of 3.127% and a minimum of 0.135%. Combined, we achieved an average percent error of 1.922% across all samples with a standard deviation of 1.537%. Our exact estimates of liver volume through summation of predicted liver parenchyma voxels via nnU-Net annotation on the test set are appreciable in Table 1. Predicting the segmentation mask with our trained nnU-Net model took on average 143.4 seconds for each test set CT scan with a standard deviation of 76 seconds.

3D Rendering

From the segmentation predictions obtained via our nnU-Net implementation, we successfully created both 2D and 3D renderings of the liver parenchyma for interactive viewing. Specifically, we compared 2 interactive image viewing platforms to appreciate the segmentation predictions for nnU-Net. The first is appreciable in Figure 4 and is a multiaxis 2D nnU-Net liver prediction overlay, which is superimposed onto the original CT image using the software platform ITK-SNAP.[38] This overlay tracks in standard DICOM or NIfTI image viewers, and it is possible to scroll, take in image point measurements, or utilize all the other image viewer features with this colored overlay intact. We were also successful in the second step using ITK-SNAP,[38] which afforded us the capability to separate out the predicted liver parenchyma and present a 3D interactive rendering of the liver in isolation. This 3D projection is freely rotatable and of high resolution, which can be appreciated in Figure 5.

DISCUSSION

Size matching is an essential step in the pretransplantation clinical workflow. When size discrepancies between a donated organ and recipient do occur, they can prevent a transplantation procedure from occurring. If a size mismatched transplant does go forward, it can lead to dangerous interoperative circumstances.[5] It is clear that population-based and demographic-based organ volume estimators are of some clinical utility.[6,7] However, the precision with which they estimate can be unsatisfactory for confident organ allocations. We set out to train and validate a deep learning model capable of segmenting out the liver parenchyma from CT medical images in this work. Implementing a nnU-Net deep learning architecture,[20] we created a computational solution that could accurately estimate liver volumes from CT scans with an average percent error of 1.922% ± 1.537%. Further, this successful implementation of nnU-Net affords for the expedient generation of flexible 2D or 3D interactive models representing the liver parenchyma. This advancement meets a clinical need to accurately size match donor and recipient before transplantation.

We view the similarity with which we segment liver volume compared to an average ground truth radiologist segmentation to be notable, with a global Dice similarity of 0.974 ± 0.005 and 95% HD of 2.46 ± 0.93 mm. This marks an improvement from other currently used organ volume estimators. For comparison, In 2020, Boers et al[39] were able to modify a 3D U-Net model to achieve a Dice similarity coefficient of 0.78 for automatic pancreas segmentation from CT scans. To automatically segment the humerus, a U-Net model has achieved a 0.946 Dice coefficient.[40] To segment out the liver, and training with our same LiTS database and testing on the IRCAD test set, Li et al[23] were able to achieve a global Dice score of 0.965 with their H-DenseUNet model, while Jin et al[41] developed a RA-UNet method that performed at 0.961 global Dice score. These can be compared to the segmentation accuracy of our nnUNet method that achieved an average global Dice score of 0.974. An improvement is demonstrated by our nnU-Net application when compared to previously constructed neural network-based segmentation techniques for the liver and other organs.

In addition to providing an increase in technical segmentation capabilities over other deep learning or approximation methods, developing this nnU-Net solution has practical benefits. The previously mentioned automatic methods are specific and require a certain level of data science expertise to meaningfully implement. This includes manual data preprocessing tasks and neural network architecture engineering per developed solution. As the nnU-Net method can automatically optimize key model tuning parameters based on unique input data constraints,[20] it is much easier to implement. We suggest its use to the reader as it requires less computational technical capabilities to build for this and other relevant image segmentation tasks.

Methods for TLV appraisal other than neural network models exist and can be highly accurate.[10] However, these methods often include additional imaging or data preprocessing steps that can be financially burdensome or clinically unjustifiable. For example, TLV estimators can be accurate with supplementary CT annotation steps,[7,42] or when combined with additional imaging modalities of such as ultrasound,[43] or magnetic resonance imaging (MRI).[44] The nnU-Net we present works with standard CT imaging and does not require any manual preprocessing steps. Linear regression TLV models also exist.[45-47] TLV may be estimated as a regression of total body surface area (TLV = 794.41 + 1267.28 × body surface area, $r^2 = 0.49$) or from body weight in kilograms (TLV = 191.80 + 18.51 × weight, $r2 = 0.49$).[48] These methods are only validated to predict TLV within 10% of the observed TLV 33% of the time.[33] The method we present, a trained nnU-Net model, is more accurate, yielding an average 2% error for our liver volume estimations. It also has a pronounced advantage over linear regression models as exact CT segmentations, unique liver representations, and ultimately total volume estimation are performed for each individual.

Our study has some limitations. Our nnU-Net is not currently deployable with other imaging modalities, including MRI. Our sample size is also limited, containing only 151 CT scan images. It is known, however, that rich information is stored in each image that is sufficient to produce a robust nnU-Net deep learning solution using few input patients.[20] For reference, in 2019, Ma et al[49] trained a modified U-Net algorithm to accurately perform bladder segmentation from CT scans. These authors were able to obtain an average percent error of 2.3% ± 21% for their model, training and testing with a total of 172 imaging studies. Our nnU-Net deep learning architecture currently only recognizes the general organ parenchyma of the liver. We aim to incorporate other relevant hepatic features such as vasculature or lobar anatomy in future model versions. We are additionally collecting and annotating prospective images to ensure model generalizability in varied clinical situations, disease states, and among varied demographic populations. Such developments, while helpful, are outside the scope of our initial aim, to precisely assess liver volume and shape before a transplantation procedure.

Advantages of this nnU-Net for liver organ segmentation and volume estimation remain despite these limitations. CT scan images are the most readily utilized image modality prior to transplantation procedures[50] and are often available to a physician for consideration. Ubiquitous pretransplantation access to CT scan images of the donor is on the immediate horizon, as United Network of Organ Sharing nears completion of their UNet Image Sharing pilot program to provide high resolution donor medical images directly in DonorNet.[51] This

nnU-Net application with a CT scan modality is robust providing exact volume estimates and precise multidimensional shape reconstructions. Various precisions of CT scans existed in these 151 standardized images, including differences in image granularity and width between image slices. Our model performed well despite this variance and was accurate across all CT scan qualities. Currently, it takes less than 20 seconds for our fully trained model to import a CT scan image, annotate the margins of the liver, provide a volume estimation, and return a modified CT scan image. Specifically, our model produces a masked image with color highlighting of the liver in a freely manipulable CT scan image, viewable on standard CT scan viewing platforms (Fig. 4). From this segmentation capture, a clinician can definitively extract specific dimensions, such as anterior-posterior distances, in addition to TLV.

Further, the 3D rendering capabilities of our model may be easily dually applied to both the donor and the recipient, allowing for a side-by-side, synchronized, direct organ comparison utilizing the same measurement scales. This capability may demonstrate utility when anatomical variation or complex partial allograft procedures are being considered. Broader applications of this technology within the field of surgery are abundant. Three-dimensional organ segmentation renderings from CT scans can be captured in series to track organ volume and other metrics over time and in parallel to disease state progression. Clear utility also exists for these technologies to aid in preoperative and preprocedural assessments, highlighting surgical access constraints that may guide clinical decision-making. The nnU-Net method is technically ready to provide even greater insight into these problems. Once sizable and accurate ground truth training datasets have been generated, the method will be well suited to automatically segment out vascular[52] and subsegmental organ anatomical features from CT imaging. More broadly, the nnU-Net method will also be useful for automatically segmenting tumors[53] and other anomalies[54] that exist across many organ systems. This method will have near-term application and impact in assisting in the clinical workflows for multiple specialties.[20]

In conclusion, our nnU-Net liver segmentation estimation model is accurate and robust across CT scan image qualities. This model’s deployment will allow for immediate incorporation into clinical practice workflows, without manual preprocessing steps, to provide accurate volumetrics and 3D reconstructions on donor organs for size matching. We envision particular utility for our model to be an aid in pre-liver transplantation clinical decisions. Future work is justified to produce similar nnU-Net models for other solid organs, to accurately measure the volume of kidneys, hearts, and lungs. Finally, we anticipate expanding our model to function appropriately across other imaging modalities, including MRI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Disclosure:

R.W.P. would like to thank the Baylor College of Medicine Medical Scientist Training Program and the Baylor Research Advocates for Student Scientists for their funding support. B.B.M. and J.H. received salary funding under

this award to conduct this investigation. The remaining authors declare that they have nothing to disclose. This work was funded as part of the National Science Foundation Small Business Technology Transfer Program Phase I (Award Number 2014827), in which InformAI was the award recipient.

The data that support the findings of this study are openly available. We have attached permanent links at which they can be accessed. For the LITS-Challenge Dataset, the training and testing files may be found at https://drive.google.com/drive/folders/0B0vscETPGI1-Q1h1WFdEM2FHSUE and https://drive.google.com/drive/folders/0B0vscETPGI1-NDZNd3puMlZiNWM, respectively. The IRCAD data is available for review at https://www.ircad.fr/research/3d-ircadb-01/ and for download at https://www.dropbox.com/s/8h2avwtk8cfzl49/ircad-dataset.zip?dl=0.

REFERENCES

1. Linden PK. History of solid organ transplantation and organ donation. Crit Care Clin. 2009;25:165–184, ix. [PubMed: 19268801]
2. Starzl TE. History of clinical transplantation. World J Surg. 2000;24:759–782. [PubMed: 10833242]
3. Fukazawa K, Nishida S. Size mismatch in liver transplantation. J Hepatobiliary Pancreat Sci. 2016;23:457–466. [PubMed: 27474079]
4. Starzl TE, Koep LJ, Schröter GP, et al. Liver replacement for pediatric patients. Pediatrics. 1979;63:825–829. [PubMed: 377201]
5. Reyes J, Perkins J, Kling C, et al. Size mismatch in deceased donor liver transplantation and its impact on graft survival. Clin Transplant. 2019;33:e13662. [PubMed: 31283049]
6. Herden U, Wischhusen F, Heinemann A, et al. A formula to calculate the standard liver volume in children and its application in pediatric liver transplantation. Transpl Int. 2013;26:1217–1224. [PubMed: 24118382]
7. Urata K, Kawasaki S, Matsunami H, et al. Calculation of child and adult standard liver volume for liver transplantation. Hepatology. 1995;21:1317–1321. [PubMed: 7737637]
8. Heinemann A, Wischhusen F, Püschel K, et al. Standard liver volume in the Caucasian population. Liver Transpl Surg. 1999;5:366–368. [PubMed: 10477836]
9. Chan SC, Liu CL, Lo CM, et al. Estimating liver weight of adults by body weight and gender. World J Gastroenterol. 2006;12:2217–2222. [PubMed: 16610024]
10. Pomposelli JJ, Tongyoo A, Wald C, et al. Variability of standard liver volume estimation versus software-assisted total liver volume measurement. Liver Transpl. 2012;18:1083–1092. [PubMed: 22532341]
11. Le DC, Chinnasarn K, Chansangrat J, et al. Semi-automatic liver segmentation based on probabilistic models and anatomical constraints. Sci Rep. 2021;11:6106. [PubMed: 33731736]
12. Giger ML. Machine learning in medical imaging. J Am Coll Radiol. 2018;15(3 pt B):512–520. [PubMed: 29398494]
13. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Annu Rev Biomed Eng. 2017;19:221–248. [PubMed: 28301734]
14. Minaee S, Boykov YY, Porikli F, et al. Image segmentation using deep learning: a survey [published online ahead of print February 17, 2021]. IEEE Trans Pattern Anal Mach Intell. doi: 10.1109/TPAMI.2021.3059968.
15. Wang K, Mamidipalli A, Retson T, et al. ; members of the NASH Clinical Research Network. Automated CT and MRI liver segmentation and biometry using a generalized convolutional neural network. Radiol Artif Intell. 2019;1:180022. [PubMed: 32582883]
16. Sengun KE, Cetin YT, Guzel MS, et al. Automatic liver segmentation from CT images using deep learning algorithms: a comparative study. arXiv. Preprint posted online January 25, 2021. doi: 10.48550/arXiv.2101.09987.
17. Guo X, Schwartz LH, Zhao B. Automatic liver segmentation by integrating fully convolutional networks into active contour models. Med Phys. 2019;46:4455–4469. [PubMed: 31356688]
18. Tang X, Jafargholi Rangraz E, Coudyzer W, et al. Whole liver segmentation based on deep learning and manual adjustment for clinical use in SIRT. Eur J Nucl Med Mol Imaging. 2020;47:2742–2752. [PubMed: 32314026]

19. Ahn Y, Yoon JS, Lee SS, et al. Deep learning algorithm for automated segmentation and volume measurement of the liver and spleen using portal venous phase computed tomography images. Korean J Radiol. 2020;21:987–997. [PubMed: 32677383]
20. Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18:203–211. [PubMed: 33288961]
21. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, et al., eds. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Vol. 9351. Springer Verlag; 2015:234–241.
22. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. Nat Methods. 2019;16:67–70. [PubMed: 30559429]
23. Li X, Chen H, Qi X, et al. H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans Med Imaging. 2018;37:2663–2674. [PubMed: 29994201]
24. Kang SK, Shin SA, Seo S, et al. Deep learning-based 3D inpainting of brain MR images. Sci Rep. 2021;11:1673. [PubMed: 33462321]
25. Schmidt CW. CT scans: balancing health risks and medical benefits. Environ Health Perspect. 2012;120:a118–a121. [PubMed: 22382352]
26. Zhong J, Bian Z, Hatt CR, et al. Segmentation of the thoracic aorta using an attention-gated u-net. In: Drukker K, Mazurowski MA, eds. Medical Imaging 2021: Computer-Aided Diagnosis. Vol. 2021. SPIE; 2021:19.
27. Xie J, Peng Y. The head and neck tumor segmentation using nnU-Net with spatial and channel 'squeeze & excitation' blocks. In: Andrearczyk V, Oreiller V, Depeursinge A, eds. Head and Neck Tumor Segmentation. HECKTOR 2020. Lecture Notes in Computer Science. Vol. 12603. Springer; 2021:28–36.
28. Graham-Knight JB, Djavadifar A, Lasserre P, et al. Applying nnU-Net to the KiTS19 Grand Challenge. Direct publication of results: Department of Computer Science, University of British Columbia, Kelowna BC, CA, 2019.
29. Bilic P, Christ PF, Vorontsov E, et al. The liver tumor segmentation benchmark (LiTS). arXiv preprint arXiv:1901.04056, 2019.
30. Soler L, Hostettler A, Agnus V, et al. 3D Image reconstruction for comparison of algorithm database: a patient specific anatomical and medical image database. IRCAD, Strasbourg, France, Tech Rep, 2010.
31. Larobina M, Murino L. Medical image file formats. J Digit Imaging. 2014;27:200–206. [PubMed: 24338090]
32. Brys A. GitHub - icometrix/dicom2nifti. Github. Published January 1, 2022. Available at: https://github.com/icometrix/dicom2nifti. Accessed March 18, 2022.
33. Isensee F, Jaeger PF, Kohl SAA, et al. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods. 2021;18:203–211. [PubMed: 33288961]
34. Agarap AF. Deep learning using rectified linear units (relu). arXiv preprint arXiv:1803.08375, 2018.
35. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: 32nd International Conference on Machine Learning, ICML 2015. Vol. 1. International Machine Learning Society (IMLS); 2015:448–456.
36. Dumoulin V, Visin F, Box GEP. A guide to convolution arithmetic for deep learning. arXiv. Preprint posted online March 23, 2016. doi: 10.48550/arXiv.1603.07285.
37. Rshamir R, Duchin Y, Kim J, et al. Continuous Dice coefficient: a method for evaluating probabilistic segmentations. arXiv. Preprint posted online June 26, 2019. doi: 10.48550/arXiv.1906.11031.
38. Yushkevich PA, Yang Gao, Gerig G. ITK-SNAP: an interactive tool for semi-automatic segmentation of multi-modality biomedical images. Annu Int Conf IEEE Eng Med Biol Soc. 2016;2016:3342–3345. [PubMed: 28269019]
39. Boers TGW, Hu Y, Gibson E, et al. Interactive 3D U-net for the segmentation of the pancreas in computed tomography scans. Phys Med Biol. 2020;65:065002. [PubMed: 31978921]

40. Mahdi FP, Tanaka H, Nobuhara K, et al. Automatic segmentation of the humerus region in 3-D shoulder CT images using U-Net. Int J Biomedical Soft Computing Hum Sci. 2020;25:67–74.
41. Jin Q, Meng Z, Sun C, et al. RA-UNet: a hybrid deep attention-aware network to extract liver and tumor in CT scans. Front Bioeng Biotechnol. 2020;8:605132. [PubMed: 33425871]
42. Heymsfield SB, Fulenwider T, Nordlinger B, et al. Accurate measurement of liver, kidney, and spleen volume and mass by computerized axial tomography. Ann Intern Med. 1979;90:185–187. [PubMed: 443650]
43. Rylance GW, Moreland TA, Cowan MD, et al. Liver volume estimation using ultrasound scanning. Arch Dis Child. 1982;57:283–286. [PubMed: 7082041]
44. Sahin B, Emirzeoglu M, Uzun A, et al. Unbiased estimation of the liver volume by the Cavalieri principle using magnetic resonance images. Eur J Radiol. 2003;47:164–170. [PubMed: 12880999]
45. Hashimoto T, Sugawara Y, Tamura S, et al. Estimation of standard liver volume in Japanese living liver donors. J Gastroenterol Hepatol. 2006;21:1710–1713. [PubMed: 16984594]
46. Yu HC, You H, Lee H, et al. Estimation of standard liver volume for liver transplantation in the Korean population. Liver Transpl. 2004;10:779–783. [PubMed: 15162473]
47. Olthof PB, van Dam R, Jovine E, et al. Accuracy of estimated total liver volume formulas before liver resection. Surgery. 2019;166:247–253. [PubMed: 31204072]
48. Vauthey JN, Abdalla EK, Doherty DA, et al. Body surface area and body weight predict total liver volume in Western adults. Liver Transpl. 2002;8:233–240. [PubMed: 11910568]
49. Ma X, Hadjiiski LM, Wei J, et al. U-Net based deep learning bladder segmentation in CT urography. Med Phys. 2019;46:1752–1765. [PubMed: 30734932]
50. Pannu HK, Maley WR, Fishman EK. Liver transplantation: preoperative CT evaluation. Radiographics. 2001;21(SPEC.ISS.):S133–S146. [PubMed: 11598253]
51. United Network of Organ Sharing. DonorNet® Imaging Study Pilot in 2019 - UNOS. 2019. Available at: https://unos.org/news/donornet-imaging-study-pilot-in-2019-2/. Accessed August 22, 2021.
52. Su Z, Jia Y, Liao W, et al. 3D attention U-Net with pretraining: a solution to CADA-Aneurysm segmentation challenge. Lect Notes Comput Sci. 2020;12643 LNCS:58–67.
53. Isensee F, Jäger PF, Full PM, et al. nnU-Net for brain tumor segmentation. Lect Notes Comput Sci. 2020;12659 LNCS:118–132.
54. Fidon L, Aertsen M, Mufti N, et al. Distributionally robust segmentation of abnormal fetal brain 3D MRI. Lect Notes Comput Sci. 2021;12959 LNCS:263–273.

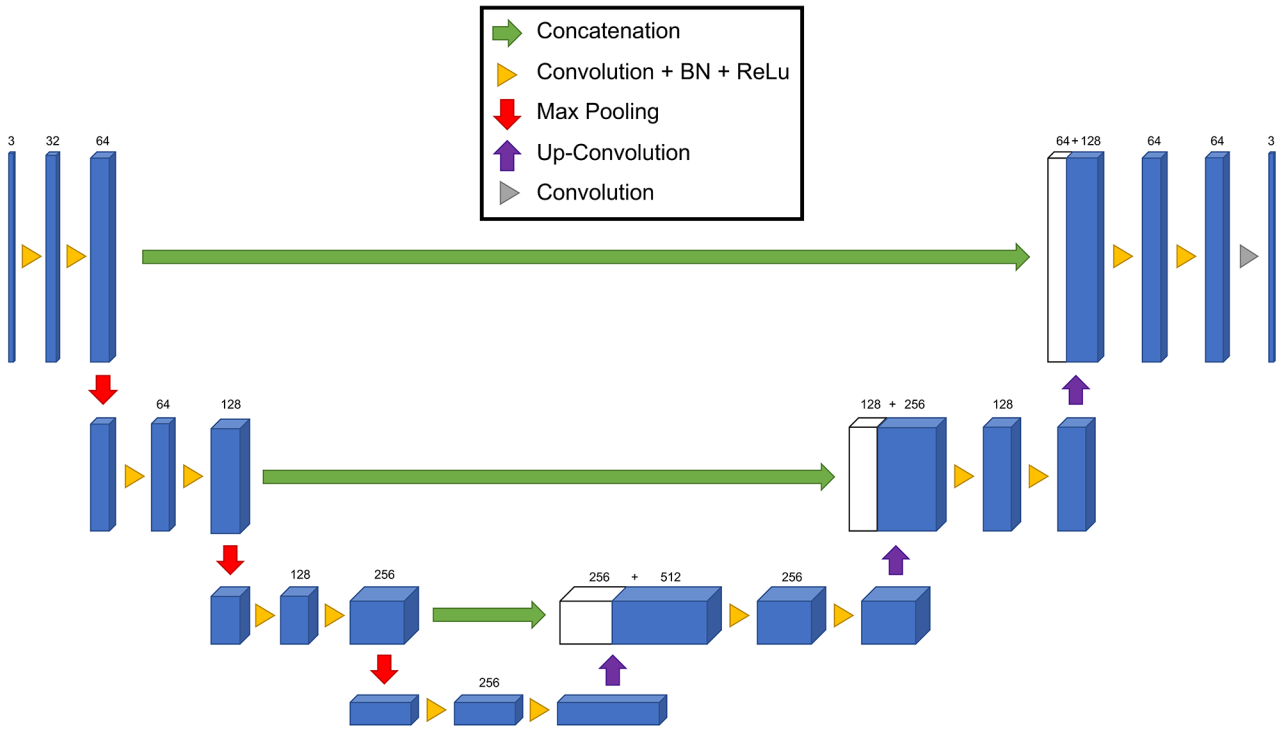

FIGURE 1.

nnU-Net convolutional neural network architecture for the segmentation of liver tissue from abdominal CT scans. BN indicates batch normalization; CT, computed tomography; ReLU, rectified linear unit activation function.

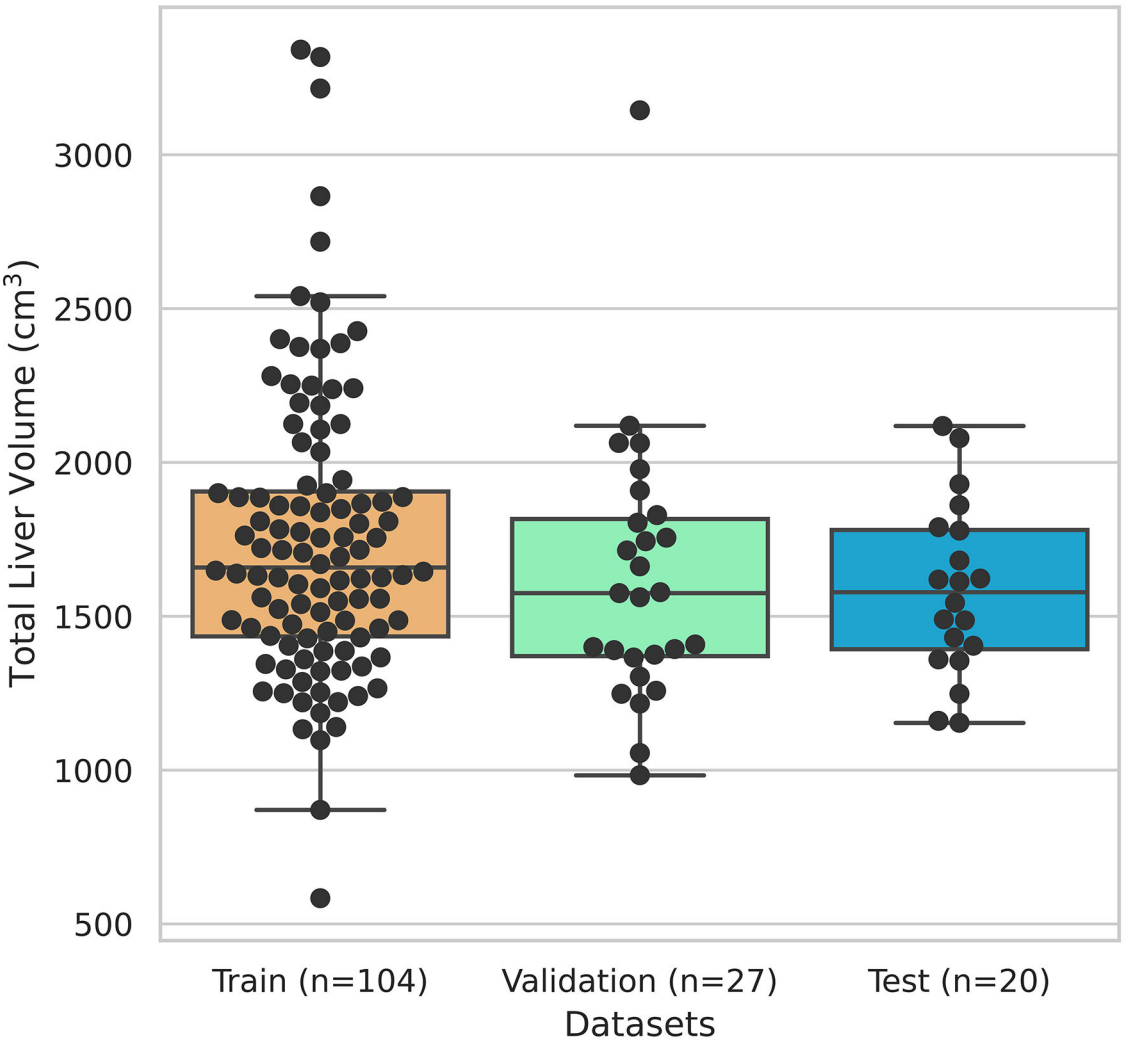

FIGURE 2.

Box plots representing median and interquartile range of liver volumes captured in training, validation, and test sets. Volume estimations (cm^3) were annotated slice by slice in CT scans by board-certified radiologists. One-hundred four training and 27 validation CT images were obtained from LiTS database, and 20 test images were obtained from IRCAD. CT indicates computed tomography; IRCAD, The 3D Image Reconstruction for Comparison of Algorithm Database; LiTS, The Liver Tumor Segmentation Challenge.

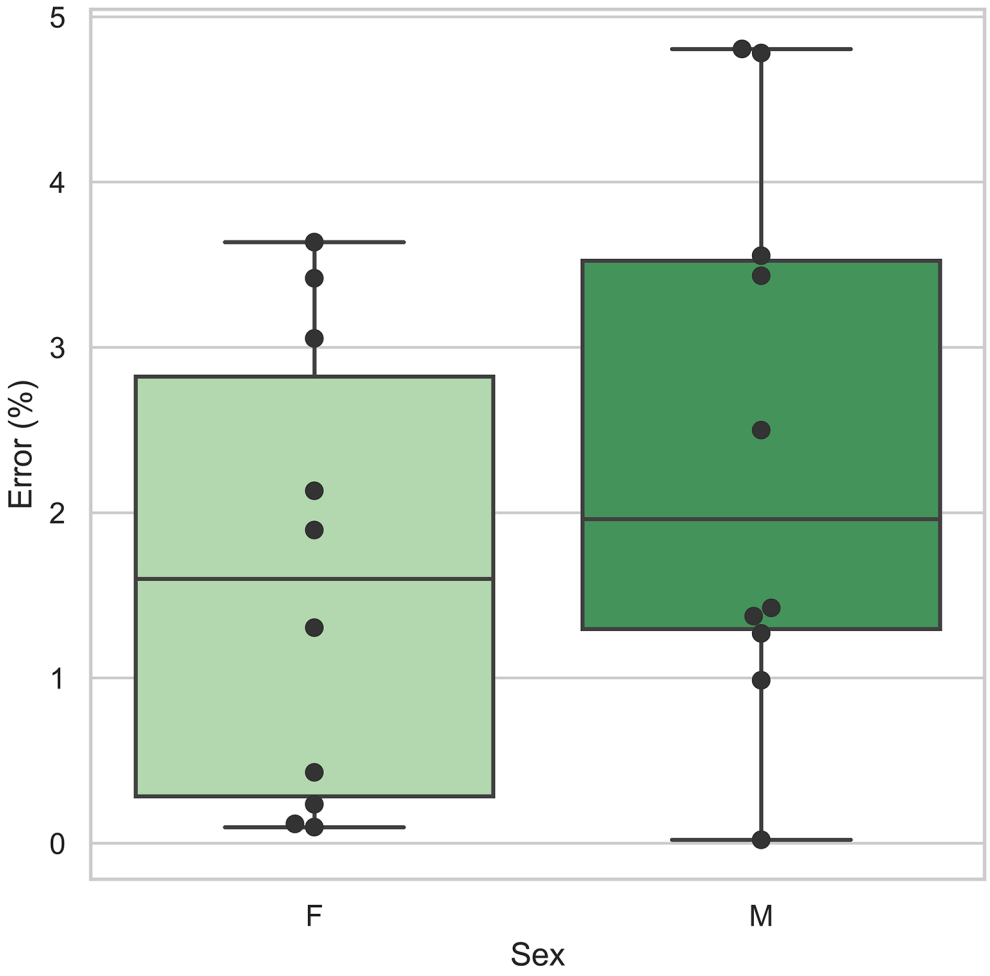

FIGURE 3.

Box plots visualizing the median and interquartile ranges for percent error when comparing nnU-Net predicted volumes to ground truth radiologist annotations. Percent error on the test set is stratified by patient sex. F indicates female; M, male.

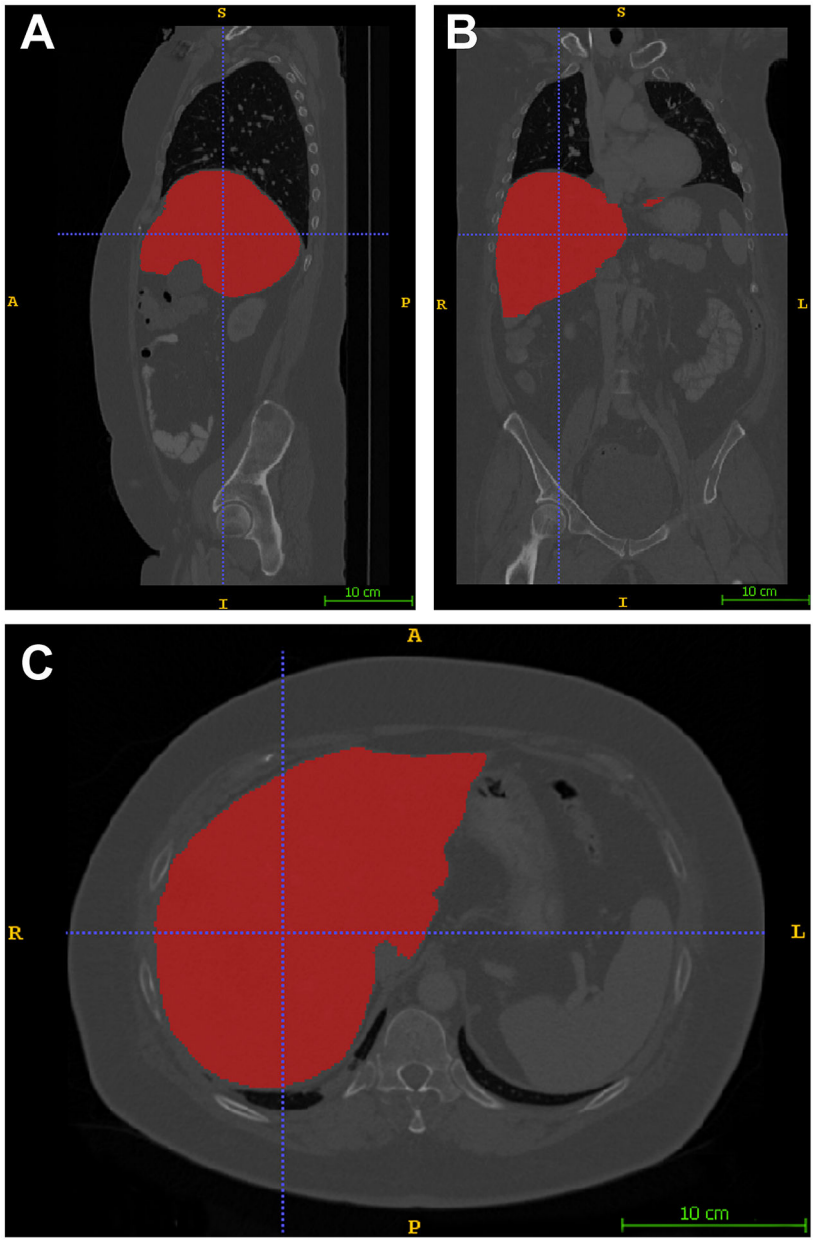

FIGURE 4.

nnU-Net segmentation and color annotation of liver parenchyma superimposed onto original computed tomography image. Areas identified by nnU-Net as part of the liver are noted in red in sagittal cross-section (A), coronal cross-section (B), and transverse cross-section (C).

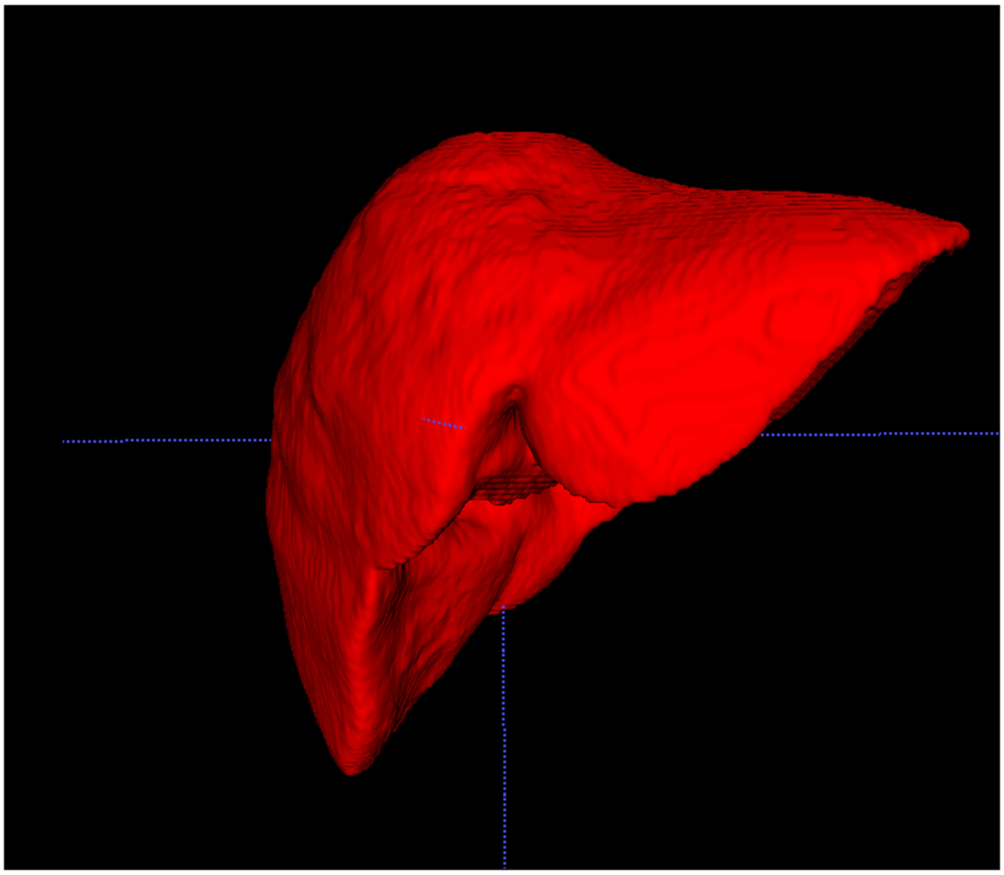

FIGURE 5.
A 3D rendering of liver parenchyma with nnU-Net. This rendering may be rotated freely in space and is used for final organ volume estimation.

Table 1.

Side by side liver volume estimations predicted by nnU-Net compared to the ground truth volumetric annotations by board certified radiologists in the test dataset.

Index	Sex	Total Liver Volume from Radiologist Annotations (cm3)	Total Liver Volume from Our 3D UNet Predictions (cm3)	Percent Error in Liver Volume Prediction (%)	Dice Similarity Coefficient	Hausdorff Distance (mm)	95th Percentile of Hausdorff Distance (mm)
1	F	1,489.410	1,480.438	0.602	0.980	18.176	1.802
2	F	1,612.490	1,604.298	0.508	0.974	16.210	2.346
3	F	1,618.970	1,611.261	0.476	0.976	17.324	2.253
4	F	1,622.318	1,586.608	2.201	0.979	8.429	1.646
5	F	1,543.661	1,541.581	0.135	0.976	18.868	1.819
6	F	1,355.007	1,371.519	1.219	0.971	13.642	2.688
7	F	1,359.611	1,323.460	2.659	0.969	18.204	2.945
8	F	1,680.848	1,675.151	0.339	0.972	23.576	2.500
9	F	1,153.317	1,122.455	2.676	0.968	24.828	4.000
10	F	1,778.006	1,833.598	3.127	0.976	27.989	2.572
11	M	1,159.707	1,153.143	0.566	0.985	12.702	1.398
12	M	1,247.579	1,316.949	5.560	0.965	29.297	5.507
13	M	2,078.697	2,023.016	2.679	0.979	8.051	1.749
14	M	1,789.068	1,721.530	3.775	0.971	12.013	2.212
15	M	1,430.422	1,484.953	3.812	0.969	24.883	1.945
16	M	1,929.034	1,954.936	1.343	0.976	17.827	2.619
17	M	1,404.006	1,350.078	3.841	0.974	32.235	1.755
18	M	1,486.234	1,482.924	0.223	0.979	20.326	1.860
19	M	2,118.118	2,145.089	1.273	0.978	21.027	2.792
20	M	1,860.623	1,833.989	1.431	0.972	32.794	2.744

Dice Similarity Coefficient: Computes the Dice coefficient (also known as Sorensen index) between the binary objects in two images. It is defined as the harmonic mean of the precision and recall or a global F1 score.

Hausdorff Distance (mm): Computes the (symmetric) Hausdorff Distance (HD) between the binary objects in two images. It is defined as the maximum surface distance between the objects.

95th Percentile of Hausdorff Distance (mm): Computes the 95th percentile of the (symmetric) Hausdorff Distance (HD) between the binary objects in two images. Compared to the Hausdorff Distance, this metric is slightly more stable to small outliers and is commonly used in biomedical segmentation.

Table 2.

nnU-Net Performance Statistics on Test Set

	F (N=10)	M (N=10)	Overall (N=20)
Total Liver Volume from Radiologist Annotations (cm3)			
Mean ± Standard Deviation	1521.36 ± 185.634	1650.35 ± 346.787	1585.86 ± 278.688
Median	1578.076	1637.651	1578.076
Min, Max	1153.317, 1778.006	1159.71, 2118.118	1153.317, 2118.118
Total Liver Volume from Our 3D nnU-Predictions (cm3)			
Mean ± Standard Deviation	1515.037 ± 200.773	1646.661 ± 336.484	1580.849 ± 278.001
Median	1122.455, 1833.598	1153.143, 2145.089	1122.455, 2145.089
Min, Max	1122.455, 1833.598	1153.143, 2145.089	1122.455, 2145.089
Percent Error in Liver Volume Prediction (%)			
Mean ± Standard Deviation	1.394 ± 1.149	2.450 ± 1.746	1.922 ± 1.537
Median	0.911	2.055	1.387
Min, Max	0.135, 3.127	0.223, 5.560	0.135, 5.560
Dice Similarity Coefficient			
Mean ± Standard Deviation	0.974 ± 0.004	0.975 ± 0.006	0.974 ± 0.005
Median	0.975	0.975	0.975
Min, Max	0.968, 0.980	0.965, 0.985	0.965, 0.985
Hausdorff Distance (mm)			
Mean ± Standard Deviation	18.725 ± 5.650	21.116 ± 8.673	19.920 ± 7.229
Median	18.190	20.677	18.536
Min, Max	8.429, 27.989	8.051, 32.794	8.051, 32.794
95th Percentile of Hausdorff Distance (mm)			
Mean ± Standard Deviation	2.457 ± 0.686	2.458 ± 1.170	2.458 ± 0.933
Median	2.423	2.079	2.300
Min, Max	1.646, 4.000	1.398, 5.507	1.398, 5.507