Quantification of Liver-Lung Shunt Fraction on 3D SPECT/CT Images for Selective Internal Radiation Therapy of Liver Cancer using CNN-based Segmentations and Non-rigid Registration

Manh Ha Luu[a,b,c,*], Hong Son Mai[d], Xuan Loc Pham[b], Quoc Anh Le[a], Quoc Khanh Le[d], Theo van Walsum[c], Ngoc Ha Le[d], Daniel Franklin[e], Vu Ha Le[a,b], Adriaan Moelker[c], Duc Trinh Chu[b], Nguyen Linh Trung[a]

[a]*AVITECH, VNU University of Engineering and Technology, Hanoi, Vietnam*
[b]*FET, VNU University of Engineering and Technology, Hanoi, Vietnam*
[c]*Department of Radiology and Nuclear Medicine, Erasmus MC, Rotterdam, the Netherlands*
[d]*Department of Nuclear Medicine, Hospital 108, Hanoi, Vietnam*
[e]*School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Australia*

ARTICLE INFO

Keywords:
Liver-lung shunt
Liver cancer
SPECT/CT
CNNs
Segmentation
Registration

ABSTRACT

Purpose: Selective internal radiation therapy (SIRT) has been proven to be an effective treatment for hepatocellular carcinoma (HCC) patients. In clinical practice, the treatment planning for SIRT using ^{90}Y microspheres requires estimation of the liver-lung shunt fraction (LSF) to avoid radiation pneumonitis. Currently, the manual segmentation method to draw a region of interest (ROI) of the liver and lung in 2D planar imaging of ^{99m}Tc-MAA and 3D SPECT/CT images is inconvenient, time-consuming and observer-dependent. This study proposes and evaluates a nearly automatic method for LSF quantification using 3D SPECT/CT images, offering improved performance compared with the current manual segmentation method.

Methods: We retrospectively acquired 3D SPECT with non-contrast-enhanced CT images (nCECT) of 60 HCC patients from a SPECT/CT scanning machine, along with the corresponding diagnostic contrast-enhanced CT images (CECT). Our approach for LSF quantification is to use CNN-based methods for liver and lung segmentations in the nCECT image. We first apply 3D ResUnet to coarsely segment the liver. If the liver segmentation contains a large error, we dilate the coarse liver segmentation into the liver mask as a ROI in the nCECT image. Subsequently, non-rigid registration is applied to deform the liver in the CECT image to fit that obtained in the nCECT image. The final liver segmentation is obtained by segmenting the liver in the deformed CECT image using nnU-Net. In addition, the lung segmentations are obtained using 2D ResUnet. Finally, LSF quantitation is performed based on the number of counts in the SPECT image inside the segmentations.

Evaluations and Results: To evaluate the liver segmentation accuracy, we used Dice similarity coefficient (*DSC*), asymmetric surface distance (*ASSD*), and max surface distance (*MSD*) and compared the proposed method to five well-known CNN-based methods for liver segmentation. Furthermore, the LSF error obtained by the proposed method was compared to a state-of-the-art method, modified Deepmedic, and the LSF quantifications obtained by manual segmentation. The results show that the proposed method achieved a *DSC* score for the liver segmentation that is comparable to other state-of-the-art methods, with an

average of 0.93, and the highest consistency in segmentation accuracy, yielding a standard deviation of the *DSC* score of 0.01. The proposed method also obtains the lowest *ASSD* and *MSD* scores on average (2.6 mm and 31.5 mm, respectively). Moreover, for the proposed method, a median LSF error of 0.14% is obtained, which is a statically significant improvement to the state-of-the-art-method ($p = 0.004$), and is much smaller than the median error in LSF manual determination by the medical experts using 2D planar image (1.74% and $p < 0.001$).

Conclusions: A method for LSF quantification using 3D SPECT/CT images based on CNNs and non-rigid registration was proposed, evaluated and compared to state-of-the-art techniques. The proposed method can quantitatively determine the LSF with high accuracy and has the potential to be applied in clinical practice.

1. Introduction

Hepatocellular carcinoma (HCC) is a major healthcare challenge, with a global incidence rate that is increasing annually (Balogh et al., 2016). In the Asia-Pacific region, HCC is one of the most common causes of cancer-related deaths. HCC is most commonly staged using the Barcelona Clinic Liver Cancer (BCLC) system (EASL et al., 2018). Selective internal radiation therapy (SIRT) using ^{90}Y-loaded resin or glass microspheres is an optional treatment for unresectable HCC with promising outcomes (Salem et al., 2010; Sangro et al., 2011). The principle of SIRT is to deliver an effective radiation dose to the tumor while sparing the normal liver and lungs. SIRT is safe for patients when the delivery dose to the liver and lungs does not exceed 30 Gy and 50 Gy, respectively (Ho et al., 1997; Gil-Alzugaray et al., 2013). SIRT using ^{90}Y-loaded resin microspheres is commonly preceded by transarterial perfusion scintigraphy using ^{99m}Tc-macro aggregated albumin (MAA) for estimating the lung shunt fraction (LSF) and dosimetry. ^{99m}Tc–MAA is utilized as a planning tool to evaluate the microsphere distribution in tumors, normal liver tissue, and the lungs (see Figure 1). LSF is an important quantitative parameter that can help to predict damage to the lungs or other organs and to avoid radiation pneumonitis (Leung et al., 1995).

The LSF is calculated from ^{99m}Tc–MAA planar imaging using a simple geometric mean; however, the procedure for defining the regions of interest in the lungs and liver is often inaccurate. Planar imaging or 2D imaging (see Figure 2) may fail to differentiate the density and attenuation between the lungs and the dome of the liver due to the lack of anatomical references (Georgiou et al., 2021). In this scenario, ^{99m}Tc–MAA single-photon emission computed tomography (SPECT)/computed tomography (CT) is currently recommended in-

*Corresponding author at: AVITECH, VNU University of Engineering and Technology, Hanoi, Vietnam. *Email*: halm@vnu.edu.vn

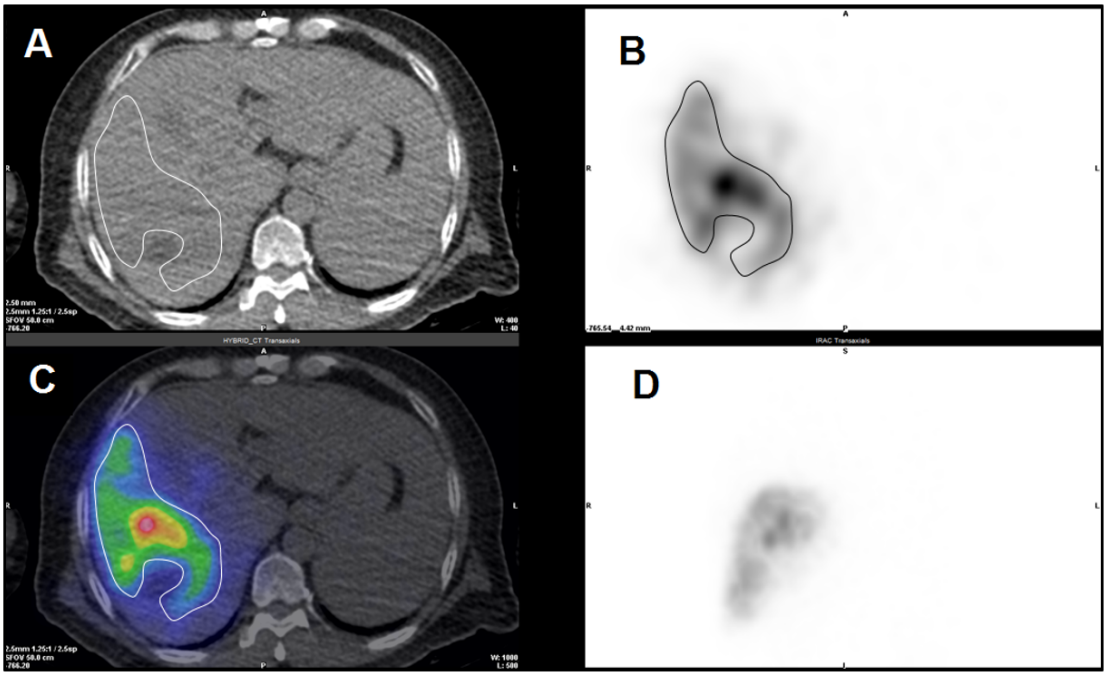

Fig. 1. An example of SPECT/CT images of the liver with ^{99m}Tc-MAA injection. Image A is the non-contrast CT image of the liver. Image B presents the SPECT image at the same axial slice. Image C shows the overlaid SPECT image on the CT image. Image D depicts the corresponding sagittal view of the SPECT image.

stead of relying on planar images (Kao et al., 2012). The advantages of SPECT/CT over planar imaging are that the images are volumetric imaging, enabling 3D visualization with higher spatial resolution, better contrast, and more precise quantitation (Son et al., 2021; Dittmann et al., 2018). Currently, performing slice-by-slice manual annotations of the lung and liver in the CT image required for such quantifications is a time-consuming, labor-intensive process, which is a significant practical limitation on SPECT/CT-based LSF calculation. Therefore, LSF quantification based on SPECT/CT images using an efficient (semi) automatic segmentation method would be a valuable improvement in the planning of SIRT treatment.

Computer-aided medical image analysis could be realized by either traditional image processing techniques or modern deep learning approaches. In the past, many traditional image processing algorithms for liver and lung segmentation, including region-growing, graph-cut, and active contour, have been developed (Heimann et al., 2009). In addition, machine learning methods have been applied to improve the accuracy of the segmentation task, from the traditional support vector machine (SVM) to active shape models or neural networks (Luo et al., 2014). However, relying on hand-crafted features made those traditional image processing techniques unreliable and inefficient in learning important features of the image (Hoang et al., 2019).

In recent years, thanks to the significant advancement in computer hardware and algorithms, state-of-the-art deep learning models - specifically convolutional neural networks (CNNs) - were introduced and made crucial contributions to the field of medical image analysis. Ronneberger et al. (2015) introduced the U-Net architecture with its highly effective encoder-decoder hierarchical feature learning structure, being one of the

Fig. 2. Manual annotations of the liver and the lungs on the planar image drawn by a nuclear medicine doctor to estimate the LSF and ^{99m}Tc–MAA dosimetry.

first works to apply CNNs to the problem of medical image segmentation. Since then, CNNs are increasingly becoming the optimal solution in a wide range of medical applications. Lee et al. (2022a) introduced voxel-level Siamese representation learning to improve representation space and subsequently proposed an adversarial learning-based semi-supervised segmentation method for improving multi-organs segmentation in CT image (Lee et al., 2022b). Wang et al. (2023) proposed a cross-convolutional transformer to segment multi-organs in a variety of medical images. Li et al. (2022) presented multi-dimensional cascaded net (MDCNet) to segment multi-organs in abdominal CT images. Conze et al. (2021) successfully segmented abdominal multi-organs in both CT and MRI images using cascaded convolutional and adversarial deep networks. A comprehensive review of deep learning-based methods for multi-organ segmentation can be found at (Fu et al., 2021).

For liver segmentation applications, Christ et al. (2016) extended the U-Net architecture to create a cascaded network for segmenting the liver from CT and MRI images, reporting a Dice score of 94%. Li et al. (2018) proposed the H-DenseUNet structure by combining 2D and 3D Dense U-Net architectures; the modification improved the liver segmentation Dice score to 96.5%. Xu et al. (2019) introduced a framework for liver segmentation in CT image using ResUNet with 3D probabilistic and geometric post process. Recently, Isensee et al. (2021) leveraged the U-Net by releasing a generalized automatic training pipeline and segmentation platform, which demonstrated excellent performance metrics in various medical image segmentation tasks, including the liver. Despite the impressive Dice scores, there are few studies that specifically focused on the problem of liver segmentation in SPECT/CT images or non-contrast CT images.

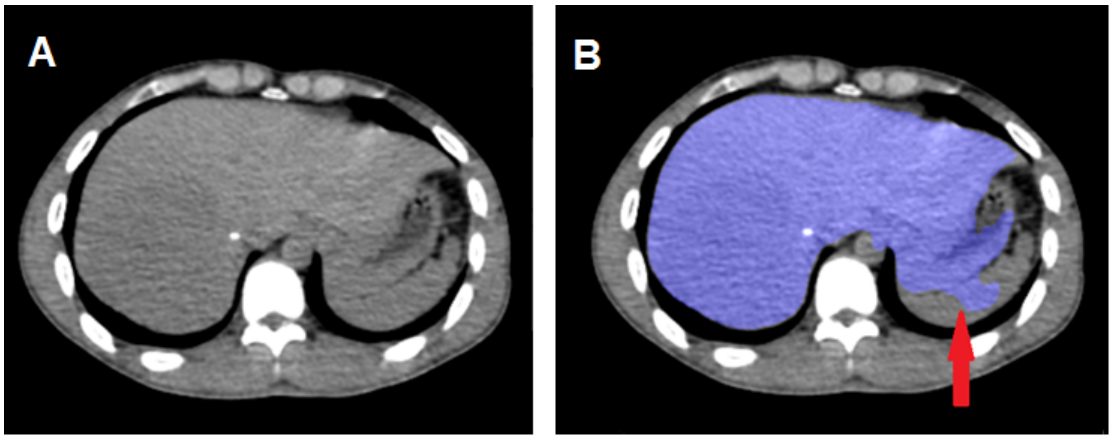

Fig. 3. Example of liver segmentation on nCECT image using 3D ResUnet. The arrow points out the incorrect segmentation due to low contrast between the liver and the other organs.

Recently, some studies on LSF quantification based on computer-aided methods have been introduced. Chaichana et al. (2021) demonstrated the superiority of (semi)automatic segmentation methods compared to traditional manual approaches in SIRT treatment planning by introducing LiverNet, achieving a median Dice score of 91% for liver segmentation. Tang et al. (2020) developed a modified Deepmedic model to automatically segment the lungs and liver in ^{99m}Tc-MAA SPECT/CT images, obtaining Dice scores of 98% and 91%, respectively. In addition, they were able to reduce the segmentation time to an average of about one minute per patient. Hence, it can be seen that computer-aided methods for LSF quantification have presented a great potential for application in clinical practice.

Although the methods achieved high accuracy in lung segmentation, both of the studies by Chaichana et al. (2021) and Tang et al. (2020) indicated that direct CNNs still may suffer from a failure when performing liver segmentation on non-contrast CT images (nCECT) with high low-dose noise, blurred boundary, or non-uniform intensity distribution (see Figure 3). In addition, none of the previous studies reports the accuracy of LSF quantification obtained by the automatic methods (only segmentation accuracy). Therefore, the aim of this study is to develop and assess an automatic method for liver segmentation in nCECT image. Subsequently, the accuracy of LSF quantification using the method will be assessed.

The remainder of this paper is organized as follows. Section 2 presents the materials and methods that we used to segment the liver and lung for LSF quantification. The experiments and results of the methods for liver segmentation and LSF quantification are described in detail in Section 3. We discuss the obtained results in Section 4. Finally, a summary of the findings and the conclusions of the study is presented in Section 5.

2. Material and methods

2.1. Data and ground truth annotation

We retrospectively collected data from 60 HCC patients who underwent SIRT treatment planning at Hospital 108 in Hanoi, Vietnam, from 2017 to 2021 [1]. This dataset (subsequently referred to as H108 dataset) includes diagnostic CECT, ^{99m}Tc–MAA SPECT, nCECT and planar image data. Of those patients, 36 subjects were eligible for SIRT treatment. The imaging data was anonymized before used in this study.

For each patient, a triple-phase abdominal CT was performed for staging and tumor vessel mapping. We obtained thirty-two available CECT images [2], acquired with either GE Healthcare or Siemens CT scanners, with in-plane resolution ranging from 0.67 mm to 0.98 mm, slice dimensions of 512×512 pixels, and the spacing between slices ranging from 0.5 mm to 3 mm. We used CT images in the portal venous phase as the CECT image to register to nCECT image. ^{99m}Tc–MAA with activity of 185 MBq was injected into the tumor-feeding artery for the patients. SPECT/CT and planar scans were obtained using an Optima NM/CT 640 scanner (GE Healthcare, Milwaukee, WI, USA) 1 hour after ^{99m}Tc–MAA injection. The planar imaging was acquired in anterior/posterior view and the range of scan is adjusted to ensure that it includes both the lung apex and the whole liver. Low-dose CT integrated with SPECT was acquired at 120 kVp and 30 mA and reconstructed with a slice thickness of 5 mm. SPECT was acquired with 60 projection images (6° rotation angle with step-and-shoot mode for 12 seconds per frame). nCECT and SPECT images were resampled to dimensions of 256×256×181 pixels, with an isotropic spatial resolution of 2.2 mm using Xeleris software. Before the registration, CECT images were down-sampled to 256×256 pixels to be the same as the slice dimensions of nCECT images.

For the ground truth of the lung segmentations of H108 dataset, a technician applied region growing to segment the lungs mask and then manually corrected the segmentation as in clinical routine. The manual segmentations were then manually corrected/verified by a senior expert using Mevislab 3.4.1. The segmentation label has a value of 0 for the background and 1 for the liver region. The lung segmentation ground truth will be used to evaluate performance of the lung segmentation method.

For the ground truth of the liver segmentations of H108 dataset, a technician first manually segmented the liver, followed by manual correction/verification by two medical experts. In this study, we also used the publicly availble CECT dataset in the liver and tumor segmentation task from the medical segmentation Decathlon (MSD dataset) (Antonelli et al., 2022). The MSD dataset consists of 131 abdomen CECT images in which experts have labeled the liver; it has been used in various studies on liver segmentation (Christ et al., 2016; Isensee et al., 2021; Tang et al., 2020). The CECT images were collected from numerous medical

[1] The local ethics committee approved the use of this dataset for our research purposes.
[2] Thirty-two CECT images were stored in the PACS and available for this study

facilities worldwide with image in-plane resolution ranging from 0.56 mm to 1 mm and the spacing between slices ranging from 0.45 mm to 6 mm. Of these images, nCECT images from thirty-five patients in H108 dataset and all CECT images from both MSD and H108 datasets were used for training the CNN models. Five and twenty nCECT images from the rest of H108 dataset were used as validating and testing images, respectively.

2.2. *Liver lung shunt quantification pipeline*

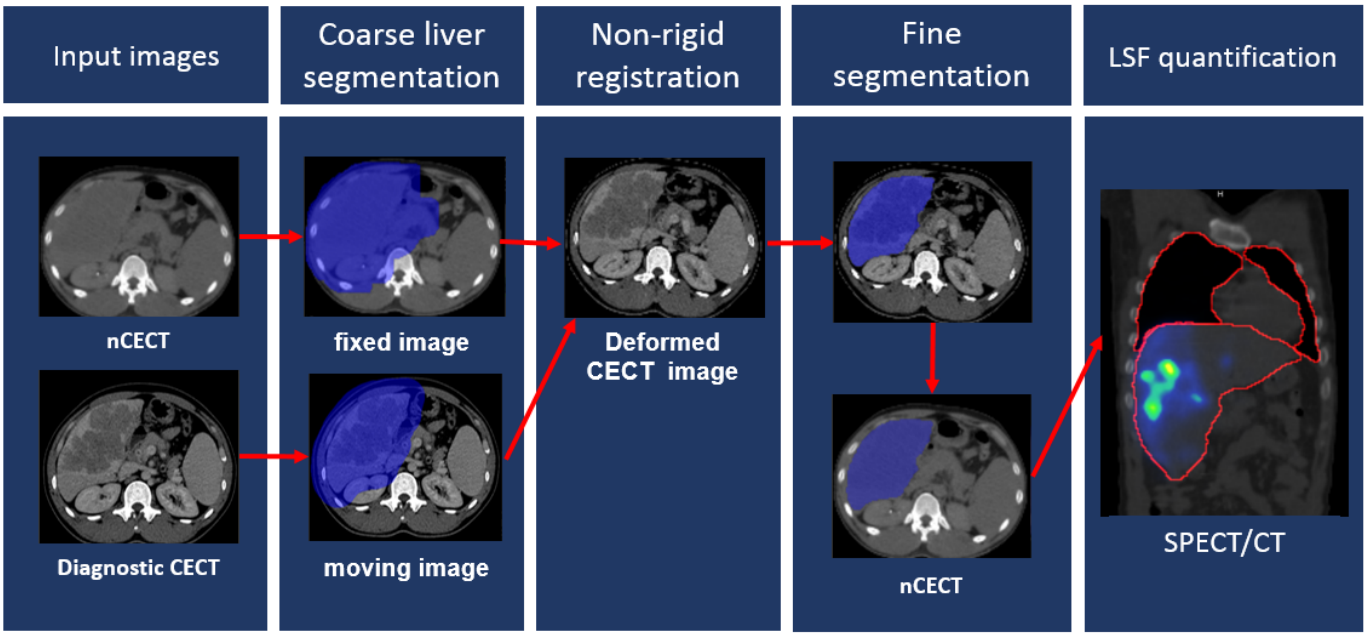

Fig. 4. Multi-stage framework for LSF quantification using SPECT/CT images.

In this study, we use a multi-stage framework for LSF quantification using SPEC/CT images, as depicted in Figure 4. In the nCECT image, the blurred liver boundary and uniform pixel intensity result in the liver being indistinguishable from other neighboring abdominal organs such as spleen, kidney or stomach. This poses challenges for segmentation, and may lead to substantial errors in a liver segmentation when segmentation methods are directly applied to nCECT images. Thanks to the contrast enhancement, CECT images (obtained from the diagnostic procedure) have clear contrast in the liver region, which enables better discrimination from other neighboring organs (see Figure 5). *Therefore, our key idea is to use the intra-patent contrast information of the liver from CECT in combination with nCECT to improve liver segmentation accuracy*. We propose a strategy that combines registration and deep learning techniques (CNNs) for segmenting the liver using the nCECT image. Firstly, coarse segmentation of the liver in nCECT image is obtained using a fast CNN. If the liver segmentation in the nCECT image is not sufficiently accurate (visually assessed by medical expert), we then segment the liver in CECT images utilizing the same CNN, and use the segmentations to generate the

regions of interest (ROIs) of the liver. Secondly, non-rigid image registration is applied to deform the liver in the CECT image to align with the liver in the nCECT image of the same patient. Subsequently, a CNN is applied to the deformed CECT image to achieve fine liver segmentation. In addition, lung segmentation in the nCECT image is also obtained by using a CNN. Finally, the LSF can be computed using the segmented liver and lung images. Details of each stage are described in the following section.

2.3. Proposed approach for liver segmentation

2.3.1. Coarse liver segmentation

Firstly, our aim is to coarsely segment the liver with fast processing time using a state-of-the-art CNN-based method. 3D ResUNet model (Xu et al., 2019) is applied to segment the liver from the SPECT/CT image. Compared to other state-of-the-art CNNs in the field, 3D ResUNet is a lightweight, fast, reliable and sufficiently accurate model with the classical encoder-decoder architecture. Inspired by the original ResNet, the residual component has been added to the U-Net structure to eliminate the vanishing/exploding gradient issue. Moreover, the model leverages the efficiency of feature extraction and representation. The performance of 3D ResUnet on liver segmentation will be evaluated in the experimental section 3. If the obtained segmentation contains a large error, it is then expanded by 10 mm in-plane using dilation morphology to guarantee the integrity of the liver boundary and used as the liver mask in the following stages. We also use 3D ResUNet to segment the liver in the CECT image, with a dilation of 10 mm applied to obtain the liver mask. The obtained liver masks are then used to locate the region of interest (ROI) in the images, which do not requires highly accurate liver segmentations. This ROI becomes auxiliary information for guiding the deformable liver registration process in the next stage. The training of 3D ResUNet model is described in details in Section 3.1.2.

2.3.2. Non-rigid registration of the liver

Here, we use non-rigid image registration to align the contrast enhanced liver in the CECT image to match to that in the nCECT image, thus enabling the liver boundary to be clear demarcated and separated from the surrounding organs (see Figure 5).

We denote the nCECT image of the liver as the fixed image $I_F(\mathbf{x}) : \Omega_F \subset R^D \rightarrow R$, and diagnostic CECT image as the moving image $I_M(\mathbf{x}) : \Omega_M \subset R^D \rightarrow R$. Accordingly, the transformation for mapping points from the fixed to the moving image domain can be defined as $T_\mu(\mathbf{x}) : \Omega_F \rightarrow \Omega_M$, where $\mu \in R^P$ is the parameter which represents the P-dimensional vector of the parameters of a B-spline transformation model (Rueckert et al., 1999).

The registration can be formulated as the following optimization problem:

$$\widehat{\mu} = \underset{\mu}{\operatorname{argmin}} \left(MI\left(T_\mu(\mathbf{x}), F(\mathbf{x})\right) + \lambda R\left(T_\mu(\mathbf{x})\right)\right), \tag{1}$$

where $MI(T; I_F, I_M)$ is the term for similarity measurement of the fixed and moving images; $R(T)$ is a regularization term (Fischer and Modersitzki, 2004) and λ is a weight which balances the terms.

We use mutual information (MI) (Viola and Wells III, 1997; Mattes et al., 2003; Thévenaz and Unser, 2000) as the similarity metric for the registration because CECT image contains the liver with enhanced contrast, while the liver in the nCECT image has lower contrast. In addition, due to the similar patient orientation with swallow respiratory during the diagnostic CECT and SPECT/CT scannings, only a small non-rigid deformation of the liver between the images is expected to be occurred. Therefore, we use bending energy (Rueckert et al., 1999) as the regulation term to constrain the non-rigid deformation.

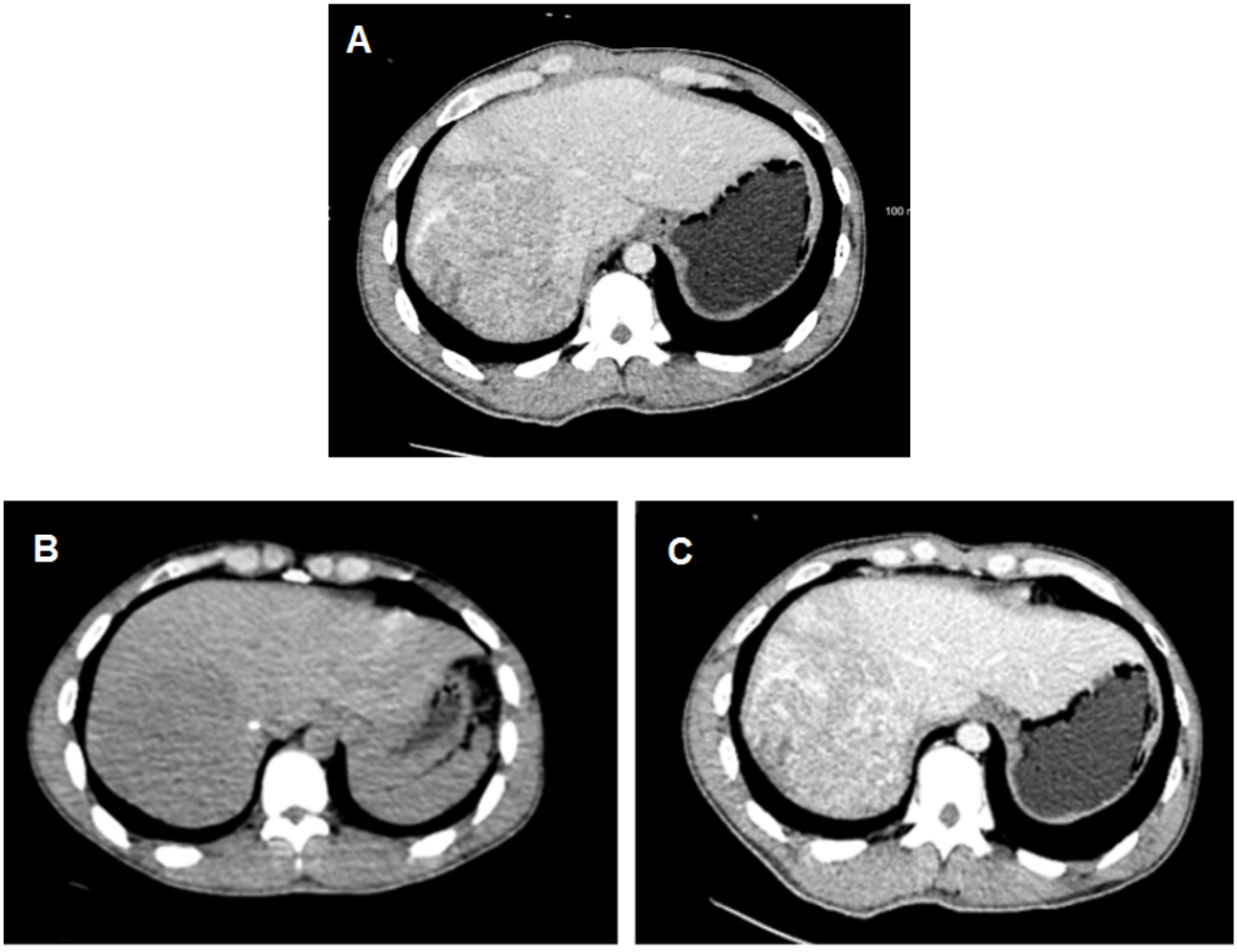

Fig. 5. Example illustrating the liver registration, which deforms a diagnostic CECT image (A) to align to an nCECT image (B) to obtain the deformed CECT image (C).

2.3.3. Fine liver segmentation

In the fine liver segmentation stage, the same nnU-Net in Section 2.3.1 is applied to segment the liver from the deformed CECT image. We choose nnU-Net because it has previously been proved effective at segmenting the liver in CECT images (Isensee et al., 2021), and achieved high performance scores in the liver

segmentation task in Decathlon segmentation challenge Antonelli et al. (2022). Moreover, one of the practical features of nnU-Net is that it is an end-to-end framework which allows automatically configuring the whole pipeline including preprocessing, network architecture, training parameters and post-processing. Despite the favorable characteristics of the CECT image, liver segmentation by CNNs may suffer from the problems of non-connected components and indistinguishability among neighboring organs. Since the liver is the largest organ in the abdominal cavity, we select the largest connected component for the final liver segmentation selection in a post-processing step. The evaluation of performance of nnU-Net on the deformed CECT image and compared to other state-of-the-arts liver segmentation CNNs will be shown in Section 3.4.2.

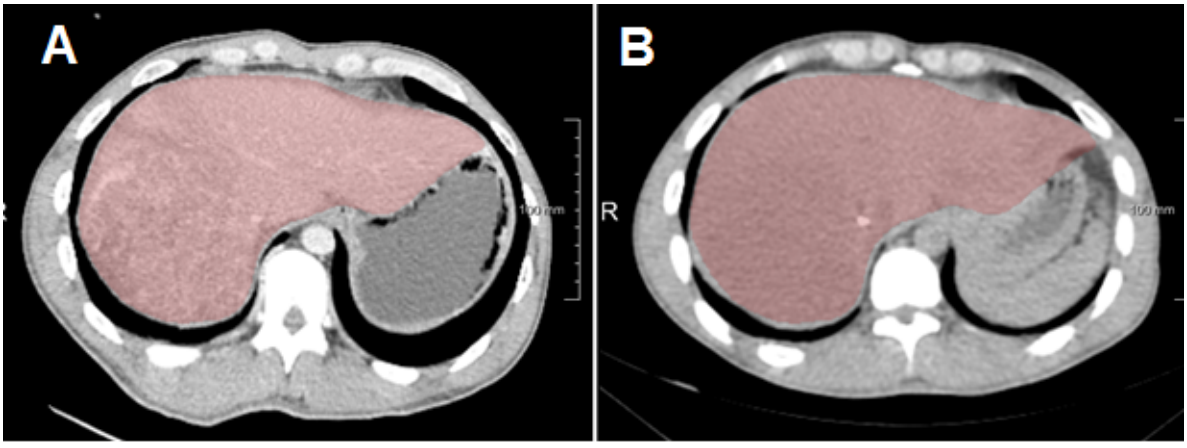

Fig. 6. Liver segmentation performed on the deformed CECT image by nnU-Net (left) and overlaid on the nCECT image (right).

2.4. Lung segmentation

In the nCECT image, the anatomical information of the lungs with separated pixel intensity range often makes it distinguishable from surrounding organs. Hence, the application of deep learning models in lung segmentation eliminates the problem of manual threshold-level searching, therefore reducing the segmentation time to seconds while maintaining a high level of lung segmentation accuracy (Hofmanninger et al., 2020; Chaichana et al., 2021). In this study, we reused the 2D ResUNet for segmenting the lungs with its trained model provided by Hofmanninger et al. (2020) using the data from Lung CT Segmentation Challenge (Yang et al., 2018). We choose the 2D ResUNet model because it has been proved effective in lung segmentation. In addition, the evaluation of the performance of the trained model on the dataset in this study will be described in Section 3.4.1.

3. Experiments and results

3.1. Implementation

3.1.1. Registration software and parameter settings

In this study, we utilized *Elastix*, which is an open source, itk-based registration tool created by Klein et al. (2009) for aligning CECT image to nCECT images. To reduce the registration time, both the nCECT and

CECT images were cropped such that the volumes-of-interest contain the liver in each image. To obtain the volumes-of-interest of the liver, we used the liver segmentations using the CNNs to detect the liver ranges and then expanded in both superior and inferior direction of 10 mm. Subsequently, multi-scale B-splines (Lester and Arridge, 1999) at four resolutions of [80, 40, 20, 10] mm were used for the transformation model, with the center of mass of the liver mask as the initial translation before starting the registration. We used adaptive stochastic gradient descent as the optimizer for finding the parameters $\widehat{\mu}$ of the B-spline transformation at each resolution. For each iteration, 2000 samples are randomly chosen inside the liver mask of nCECT to compute the similarity metric *MI* of the nCECT and CECT images. For each resolution, the number of iterations is set to 500 to guarantee convergence (Klein et al., 2009). The weights for *MI* and bending energy are set to 1 and 200, correspondingly. The other parameters were set to the default values from *Elastix*. The registration time on the test dataset is approximately 93 seconds for a case on average.

3.1.2. Implementation and training of CNNs

We utilized three CNN models in the proposed automatic LSF quantification pipeline, including 2D ResUNet for lung segmentation (Hofmanninger et al., 2020), 3D ResUNet for liver mask generation (Xu et al., 2019), and nnU-Net (Isensee et al., 2021) for fine liver segmentation. For 2D ResUNet, we reimplemented the network architecture and reused the trained model provided in the author's GitHub repository [3]. For 3D ResUNet, the implementation was obtained from 3D ResUNet's Github repository [4]. The implementation of nnU-Net was obtained from nnU-Net's GitHub repository [5].

We trained the 3D ResUNet model using the training data described in Section 2.1, and using Dice loss with an initial learning rate of 1×10^{-5}, decreasing to 80% of its previous value after every 25 epochs. The model was trained for 100 epochs, and the network parameters were learned using the Adam optimizer with the default parameters. Since processing the entire 3D CT input image is resource-consuming and possibly leads to an overflow in GPU VRAM, we split the original 181×256×256 pixel input into smaller patches of 48×256×256. For pre-processing of liver segmentation in the nCECT image, the pixel intensities were clipped to the liver intensity range of [-200, 200]. For nnU-Net, we followed the default preprocessing, postprocessing and training methods suggested by the authors. The source code and trained models of 3D ResUnet and nnU-Net are make publicly available at [6]

As a benchmark method, we reimplemented the modified Deepmedic network introduced by Tang et al. (2020), which also applied CNNs for segmentation of SIRT data. The modified Deepmedic contains three pathways instead of two pathways, enabling the network to capture a wider range of image resolution. In the

[3] https://github.com/JoHof/lungmask
[4] https://github.com/assassint2017/MICCAI-LITS2017
[5] https://github.com/MIC-DKFZ/nnUNet
[6] https://github.com/kennyha85/Liver-Lung-Shunt-project.

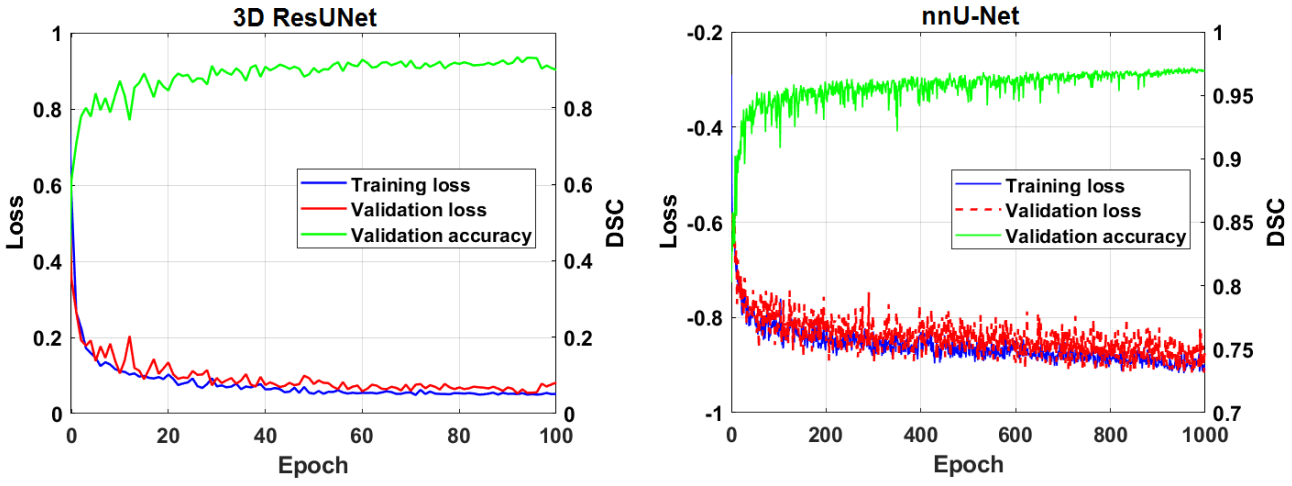

Fig. 7. The training curves of 3D ResUNet and nnU-Net using the dataset in this study. Dice score was used for validating the trained models.

training process, we train the modified Deepmedic with the same configuration as suggested by the authors. In addition, we also experimented two Transformer-based models, including TransUNet and Swin-Unet, which recently demonstrated their effectiveness in medical image segmentation (Cao et al., 2021; Chen et al., 2021). We used the source code published by the authors, with the implementation obtained from authors' GitHub repositories [7][8]. TransUNet and Swin-Unet models were trained using the Dice and CrossEntropy loss function with the training parameters set according to their respective authors' recommendations. We trained the two models with the same number of training/validation and testing datasets as for the proposed method. Since two Transformer-based models are 2D segmentation models and require pre-trained weights, we split the 3D image into multiple 2D images with the input size set of 256×256 for TransUNet and 224×224 for Swin-Unet.

The registration and all CNN models in this study were trained on a workstation with an NVIDIA RTX8000 48GB VRAM GPU, an Intel Core i9-9900K 32GB RAM CPU running the Ubuntu Linux 20.04.3 LTS environment.

3.2. Evaluation metrics for CNN-based segmentation

Dice similarity coefficient (*DSC*): the *DSC* score evaluates the overlap between the predicted segmentation and the ground truth (label), which is generally used for segmentation evaluation. The *DSC* is defined as follows:

$$DSC(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|}, \tag{2}$$

where X and Y represent the ground truth segmentation and the predicted segmentation, respectively.

[7] https://github.com/Beckschen/TransUNet
[8] https://github.com/HuCaoFighting/Swin-Unet

False negative rate (*FNR*): *FNR* is used to measure the rate of negatively-predicted voxels that should be positive (i.e the under liver segmentation). A high *FNR* means the predicted segmentation missed a large portion of voxels as compared with the ground truth.

$$FNR(X, Y) = \frac{X - X \cap Y}{X \cup Y}, \tag{3}$$

False positive rate (*FPR*): In contrast to *FNR*, *FPR* indicates the rate of positively-predicted voxels that should be negative (i.e over segmentation). A high *FPR* means the predicted segmentation includes a large area that does not belong to the ground truth.

$$FPR(X, Y) = \frac{Y - X \cap Y}{X \cup Y}, \tag{4}$$

Average symmetric surface distance (*ASSD*): *ASSD* calculates the average of all distances from points on the surface of liver segmentation ground truth in one image to the predicted segmentation surface in the other image and vice versa.

$$ASSD(X, Y) = \frac{\sum_{\mathbf{x} \in X} d(\mathbf{x}, Y) + \sum_{\mathbf{y} \in Y} d(\mathbf{y}, X)}{X + Y}, \tag{5}$$

where $\mathbf{x}$ and $\mathbf{y}$ refer to an arbitrary voxel in X and Y respectively; $d(\mathbf{x}, Y)$ represents the shortest path from point $\mathbf{x}$ on the surface of X to the surface of Y and the same definition applied for $d(\mathbf{y}, X)$.

Maximum surface distance (*MSD*): *MSD* is the maximum distance between a point $\mathbf{x}$ on the surface of segmentation ground truth X to the surface of predicted segmentation Y and vice versa. Subsequently, the greater of the two distances is the value of the *MSD*.

$$MSD(X, Y) = \max\left(\max_{\mathbf{x} \in X} d(\mathbf{x}, Y), \max_{\mathbf{y} \in Y} d(\mathbf{y}, X)\right), \tag{6}$$

3.3. LSF evaluation

Clinically, the liver lung shunt fraction is calculated by the following formula (Georgiou et al., 2021):

$$\text{LSF} = \frac{C_{Lungs}}{C_{Lungs} + C_{Liver}} \times 100\%, \tag{7}$$

where C_{Lungs} and C_{Liver} refer to the the total percentage injected dose in the lung and liver regions, respectively.

The accuracy of LSF quantification for each method can be determined by taking the difference between the ground truth LSF quantification (with the label segmentations in nCECT image) and the obtained LSF.

Table 1. Median evaluation scores of performance of the five state-of-the-art methods on the deformed CECT images. The numbers in brackets are the lowest and the highest scores. The numbers in bold are the best scores for automatic liver segmentation methods.

Method	DSC	FNR	FPR	ASSD (mm)	MSD (mm)	Processing time
TransUNet	0.94±0.04 [0.81-0.95]	0.08±0.02 [0.05-0.11]	0.03±0.05 [0.01-0.22]	2±3.4 [1.3-14.1]	26.2±31.6 [17-122.6]	3.7 sec
Swin-Unet	0.89±0.04 [0.82-0.94]	0.12±0.08 [0.06-0.29]	0.04±0.04 [0.01-0.12]	4.1±1.4 [1.7-6.1]	39.1±27.1 [17.9-111.5]	2.8 sec
3D ResUNet	0.92±0.04 [0.84-0.96]	0.07±0.08 [0.01-0.23]	0.07±0.03 [0.03-0.1]	2.7±2.1 [1.1-7.1]	49.9±26.8 [11.3-91.9]	2.5 sec
modified Deepmedic	0.92±0.18 [0.3-0.96]	0.11±0.22 [0.03-0.82]	**0.02±0.01 [0-0.05]**	2.8±10.4 [0.9-39.8]	38.6±31.3 [11.9-134.7]	4.5 sec
nnU-Net	**0.96±0.01 [0.94-0.97]**	**0.03±0.02 [0.01-0.07]**	0.04±0.02 [0.02-0.07]	**1.1±0.3 [0.7-1.6]**	**15.9±5.9 [10.8-30.8]**	6.2 sec

3.4. Evaluation results

3.4.1. Lung segmentation

We evaluate performance of 2D ResUNet with the pretrained model for lung segmentation using 60 nCECT images in H108 dataset. The results show that the CNN obtains DSC of 0.972 on average with a standard deviation of 0.007, and a min of 0.953. The inference time is of 4.5 seconds on average. Moreover, the obtained means of *FNR*, *FPR*, *ASSD* and *MSD* are 0.015, 0.038, 0.160, 0.3 mm and 14.4 mm, respectively. To this end, two nuclear medicine doctors verify that the lungs segmentation has been performed with sufficient accuracy for the application.

3.4.2. Liver segmentation

In this experiment, for the proposed approach, we first apply 3D ResUNet on the test dataset. A nuclear medicine doctor visually assessed the results and determined that thirteen out of twenty test images contained large errors. We subsequently apply the proposed method to improve liver segmentation in nCECT image to those cases. In one out of the thirteen cases, the registration failed due to a large motion artifact during in the SPECT/CT images.

Firstly, we evaluate five state-of-the-art CNNs (TransUNet, Swin-Unet, 3D ResUNet, modified Deepmedic and nnU-Net) with their models trained with the same training dataset in this study and then test on the twelve deformed CECT images. The result is summarized in Table 1. It can be seen that nnU-Net performed the best in most of the evaluation scores, achieving the median *DSC*, *FNR*, *FPR*, *ASSD* and *MSD* scores of 0.96, 0.03, 0.04, 1.1 mm and 15.9 mm, respectively, with the inference time in order of second. Thus, we chose nnU-Net for fine liver segmentation stage in the proposed pipeline (see Figure 4).

Secondly, we also evaluate the performance of the five state-of-the-art CNNs directly on the test dataset (nCECT image). The resulting liver segmentations are also compared to manual liver correction performed by

a technician on 3D ResUNet's liver segmentation prediction. The liver segmentation accuracy of the methods is summarized Table 2. It can be seen that the liver segmentation from the manual interaction yields the highest accuracy, with an average *DSC* of 0.95 and a standard deviation of 0.01. Meanwhile, among the state-of-the-art CNN models, TransUNet shows the highest *DSC* score of 0.94 on average with a standard deviation of 0.03, approaching the manual method. The proposed method has a small lower margin with an average *DSC* of 0.93, with no statistically significant difference ($p = 0.58$, t-test). However, with a standard deviation of 0.01, the proposed method has shown the same consistent level as the manual interaction approach, and its performance is superior to all other CNN models. For *FNR* and *FPR* values, 3D ResUnet surprisingly shows the best *FNR* results with an average value as low as 0.02 and fluctuation of 0.01, while the proposed method achieves the worst *FPR* value of 0.06.

For boundary-distance-based evaluation metrics, segmentations with human intervention consistently produce the most reliable results compared to fully automatic predictions, with an average value of 1.5 mm and 27.4 mm obtained for the *ASSD* and *MSD* metrics, respectively. As compared to the five state-of-the-art CNN models, the proposed method has values that are statistically significantly smaller than those of the other methods, with mean values of 2.5 mm and 31.5 mm for *ASSD* and *MSD*, respectively.

Regarding execution time, 3D ResUNet is faster to the others with an average runtime of 2.3 seconds for each segmentation. With an average runtime of approximately 122 seconds, the proposed method requires a longer processing time than the other automatic segmentation models due to a composition of multiple steps in the pipeline; however, this remains a practical execution time in a clinical scenario. Lastly, as a trade-off for providing the highest accuracy, the manual correction method is the slowest, taking between 5 and 20 minutes. Note that all methods are experimented using the same hardware configurations as stated in Section 3.1.

3.4.3. LSF quantification

In this experiment, we first investigate the accuracy of LSF quantification using 2D planar images analyzed by the human-expert manual drawing technique (the conventional method). Two nuclear medicine doctors with 5 years' and 3 years' experience (denoted as doctor 1 and doctor 2, respectively) manually draw the lung and the liver regions on the planar images in the test dataset. Subsequently, the LSF values can be obtained using Formula 7. The result shows that the LSF quantification by doctor 1 has the median error of 1.7% while the number for doctor 2 is 1.8%, compared to the LSF quantification using manual liver segmentation in the nCECT image (see Section 3.3). We apply the paired t-test on the LSF by the two doctors, yielding a p-value of 0.023, which shows a statistically significant difference between the two doctor's performances. The results are illustrated in Figure 9.

We also evaluate the accuracy of LSF quantification using the proposed method, the modified Deepmedic

Table 2. Quantitative evaluation scores of the proposed liver segmentation method and state-of-the-art methods on nCECT images. The numbers in brackets are the lowest and the highest scores. The numbers in bold are the best scores for automatic liver segmentation methods.

Method	DSC	FNR	FPR	ASSD (mm)	MSD (mm)	Processing time
TransUNet	**0.94**±0.03 [0.84-0.97]	0.05±0.02 [0.02-0.1]	0.06±0.06 [0.02-0.25]	4.4±4.9 [0.8-15.4]	91.7±40.0 [30.0 - 159.3]	7.9 sec
Swin-Unet	0.88±0.03 [0.81-0.93]	0.07±0.04 [0.03-0.15]	0.14±0.07 [0.05-0.27]	6.4±3.2 [2.5-13.9]	117.9±48.6 [33.2-187.7]	6.4 sec
3D ResUNet	0.92±0.03 [0.84-0.96]	**0.02±0.01** [0-0.05]	0.12±0.04 [0.06-0.25]	10.4±5.8 [2.5-26.4]	211.4±23.6 [180.5-266.5]	**2.3 sec**
modified Deepmedic	0.91±0.03 [0.86-0.96]	0.03±0.01 [0.01-0.04]	0.14±0.05 [0.06-0.23]	7.9±5.8 [2.1-29.4]	143.2±61.7 [53.3-243.2]	27.8 sec
nnU-Net	0.93±0.05 [0.81-0.98]	0.03±0.01 [0.02-0.05]	0.1±0.08 [0.02-0.29]	4.9±4.7 [0.7-13.4]	70.6±44.6 [13.4-142.0]	19.6 sec
Proposed method	0.93±**0.01** [0.9-0.96]	0.07±0.02 [0.03-0.1]	**0.06±0.03** [0.03-0.13]	**2.5±1.7** [1.7-7.9]	**31.5±26.0** [15.9-116.3]	~122 sec
Manual correction	0.95±0.01 [0.93-0.96]	0.03±0.01 [0.02-0.04]	0.07±0.01 [0.06-0.09]	1.5±0.5 [1.0-2.4]	27.4±12.3 [16.2-52.9]	~5-20 min

model (Tang et al., 2020), which is one of state-of-the-art CNN-based methods, and finally, the manual liver segmentation correction using the same test dataset (Figure 9). The modified Deepmedic model achieves a median error of 0.42%, which is less than one quarter of the error obtained using the conventional method. The proposed method achieves the best fully-automatic LSF quantification accuracy, with a median error of 0.14%, while the LSF value obtained by the manual correction approach achieves the best median error of 0.03%. A paired t-test between the LSF errors obtained using the proposed method and modified Deepmedic shows the statistical significance of the improvement obtained by the proposed method ($p = 0.004$). Additionally, a paired t-test demonstrates that the manual correction approach outperformed the proposed fully automatic method ($p = 0.03$).

Regarding consistency, errors with the conventional LSF measurement method are widely distributed, with a standard deviation of around 6% for the two doctors – and one outlier reaching approximately 11%. Consistency is improved by using the modified Deepmedic model, for which the errors have a standard deviation of approximately 1% and a maximum error of 1.44%. The proposed method reduces the error range even further, with a maximum error of 0.38% and minimum error as low as 0.02%, and with a standard deviation of approximately 0.12%. Lastly, the manual correction method shows the least fluctuation with a standard deviation of approximately 0.11% and one outlier case of 0.31%.

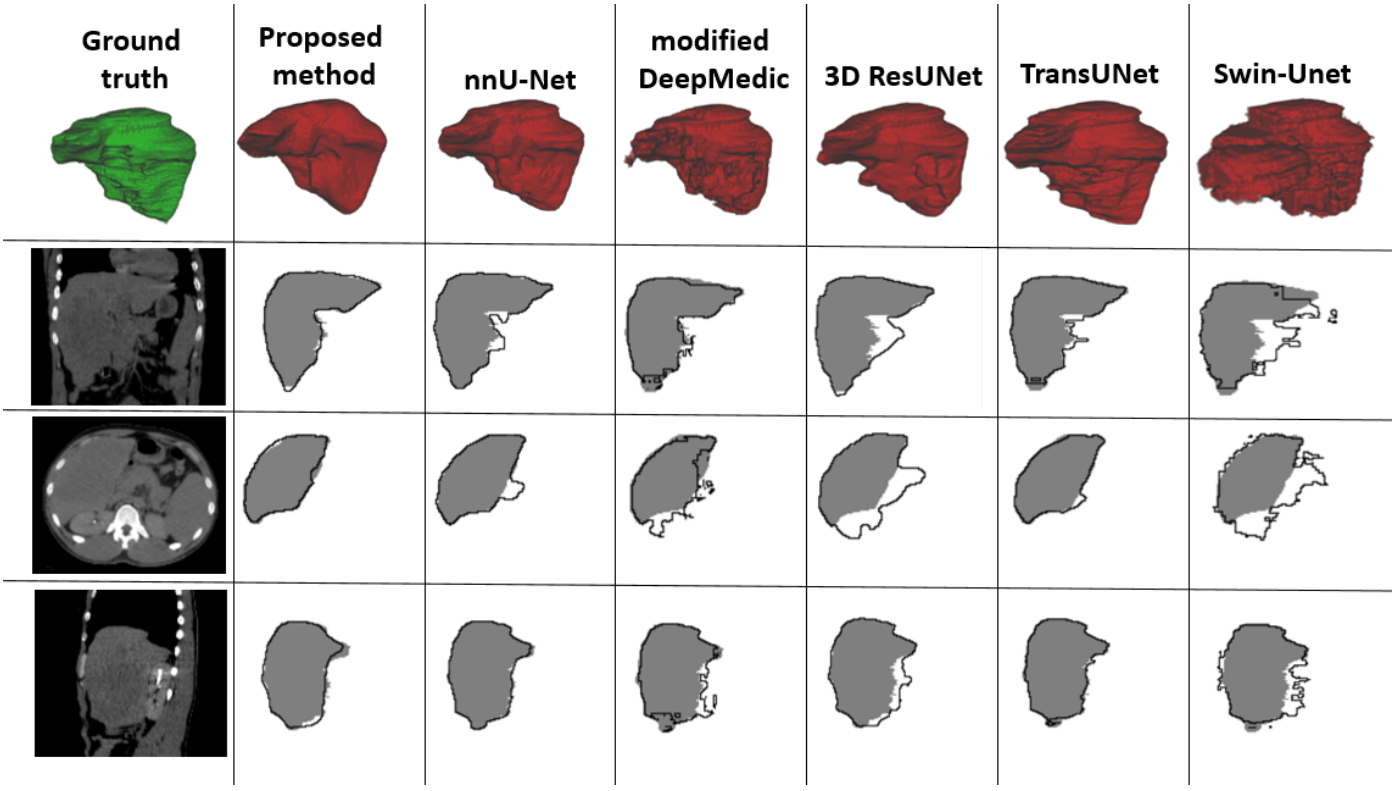

Fig. 8. Example of liver segmentations by the proposed method and other state-of-the-art method of an nCECT image. The outer lines are the border of the segmentation of each method, and the solid brown segmentations are the corresponding ground truths.

4. Discussion

In this study, we have proposed and evaluated a LSF quantification method using SPECT/CT image based on automatic liver segmentation by CNN-based methods and non-rigid registration. The liver segmentation accuracy for nCECT images and the processing time of the proposed method were compared to five state-of-the-art methods and a manual correction approach. Moreover, the LSF quantification error for the proposed method was also investigated and compared to state-of-the-art method as well as the manual methods.

The evaluation of 2D ResUnet for the lung segmentation on H108 dataset shows that the model achieved the mean *DSC* score of 0.972, which is equivalent to the numbers reported by Hofmanninger et al. (2020); Chaichana et al. (2021) and Tang et al. (2020). The inference time of 4.5 seconds on average is also sufficient fast for the application in this study. This result also suggests that the whole lung segmentation seems to be not a challenging problem with our dataset.

The evaluation of the five state-of-the-art CNNs on the deformed CECT images (see Table1) indicates that nnU-Net has the best performance with the median *DSC* score of 0.96, which is similar to the result reported in Isensee et al. (2021) (evaluation on MSD dataset). TransUNnet and modified Deepmedic also obtained reasonable results with the median *DSC* scores of 0.94 and 0.92, respectively, which are also similar to those

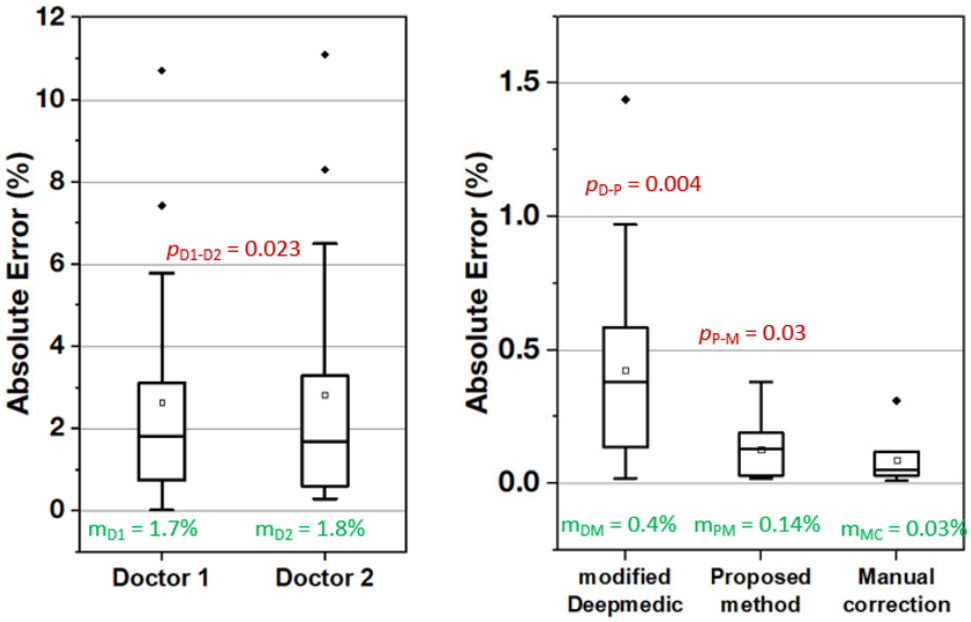

Fig. 9. Boxplots of LSF quantification errors by two medical experts using 2D planar images (left) and by CNN-based methods on SPECT/CT image (right). The green numbers are median values, and the red numbers are p values.

reported by Chen et al. (2021) and Tang et al. (2020). The other models seem to do not perform as high as reported in the original studies. The differences may come from the reason that we did not optimize the training parameters for our dataset rather than using the defaults values.

From Table 2 (Section 3), we can see that for liver segmentation accuracy, the proposed method achieved the mean *DSC* score of 0.93 which is comparable to the best state-of-the-art method, TransUNet (mean *DSC* of 0.94). The obtained *DSC* score for the proposed method is also equivalent to those reported by Tang et al. (2020) and Chaichana et al. (2021) (median *DSC* scores of 0.94 and 0.91, respectively) on their private dataset with similar experiment. Moreover, the manual segmentation obtained the mean *DSC* core of 0.95, which demonstrates that liver segmentation in nCECT image dataset is challenging even for human annotator. We suppose this is also the reason for the lower performance of the five state-of-the-art CNNs than those on the deformed CECT images. The proposed method achieves a consistently high *DSC* score compared to other methods, ranging from 0.90 to 0.96. However, the error in registration may contribute to a small part of the final segmentation error. It can be seen from both Figure 6 and Figure 8, the plain CNNs for liver segmentation on the nCECT image may have large errors at local liver boundaries with unclear contrast among the organs, while the use of CT images from same patient with similar acquisition condition enables sufficiently accurate registration at the boundaries.

Table 2 also shows that the proposed method obtained the best mean *ASSD* and *MSD* scores compared to other state-of-the-art methods, yielding the scores of 2.5 mm and 31.5 mm, respectively. Thanks to the registration, the liver segmentation on the deformed CECT image obtained better liver surface accuracy on

average (see Figure 8). In addition, observation on our dataset shows that high ^{99m}Tc–MAA concentration occurs at regions deeper inside the liver compared to the liver boundary. Hence, the segmentation errors at swallow liver boundary regions may not affect the accuracy of the LSF quantification. Yet, compared to the manual correction method, which the obtained mean *ASSD* and *MSD* scores of 1.5 mm and 27.4 mm, respectively, the proposed method is less accurate. Therefore, we can infer from these results that the (nearly) automatic methods have not yet reached human-equivalent performance levels.

Liver segmentation performance does not explicitly demonstrate the accuracy of LSF quantification. The experiments in Section 3.4.3 demonstrated the accuracy of each of the different methods, compared to conventional human tracing methods. The LSF quantification error by the experts using the convectional method with 2D planar image has the median error of 1.74% and may be as high as 10% (see Figure 9). On average, the error is not significant compared to the threshold of 20% for the criteria of SIRT intervention (Son et al., 2021; Dittmann et al., 2018); however with the high LSF cases, a large error can pose a risk to the patient. Moreover, the t-test performed on the two experts yielded a p-value of 0.023, indicating that the difference between the two experts is statistically significant. The evaluation on the test dataset shows that the proposed method obtained the LSF error of 0.14% and the largest error is less than 0.5%, which is statistically significantly better than modified Deepmedic method ($p = 0.004$) as well as the conventional method, and is sufficiently small for use in clinical practice.

Processing time is also an important factor that we need to consider in clinical practice. It has been reported that manual segmentation of the liver and lung in nCECT image may requires up to one hour to complete (Chaichana et al., 2021). The proposed method requires an average of approximately two minutes for computing LSF for a case. Therefore, the method has the potential to save time for nuclear medicine physicians performing LSF quantification. Compared to the current protocol, two minutes for a case on average is a substantial improvement in clinical practice. From Table 2, it can be seen that the most time-consuming stage of the proposed method is the registration process (93 seconds on average). Recently, newly-introduced CNN-based registration techniques are capable of reducing the registration time to the order of a few seconds with similar accuracy (Balakrishnan et al., 2019; Fu et al., 2020). Future studies may consider CNN-based registration approaches to speed up registration while maintaining accuracy when sufficient training data available.

Our study still has several limitations. Firstly, this study was carried out on data from 60 patients from a single medical center. CNN-based approaches may obtain better results if more training data can be included. However, collecting large amount of data with manual labelling is challenging in nuclear medicine (Visvikis et al., 2022). We suppose the segmentations performed using the proposed method can be a candidate for the data preparation for a larger study. Secondly, we only assessed *Elastix* for registration of the liver in this study, and did not experiment with other registration software. However, the accuracy of the liver segmentation is

comparable to the current state-of-the-art, suggesting that the registration framework is sufficiently reliable. Finally, in this study, the SPECT/CT scanning requires patient maintain shallow breathing during the scan to minimize motion artefacts. It should be noted that deep breathing motion may cause issues not only for the registration but also for the segmentation methods. We did not manage to deal with this exception situation using automatic approach. We suggest for those cases, nuclear medicine doctor may consider conventional method based on 2D planar images as a reference.

5. Conclusion

This study has proposed an automatic CNN-based LSF quantification method for liver cancer SIRT treatment planning with high accuracy and consistency using an auxiliary non-rigid registration component. Regarding segmentation quality, the proposed method is compatible or superior to the state-of-the-art CNN models in almost all metrics, while achieving the median LSF quantification error of 0.14%. The obtained LSF quantification accuracy is also better than that of conventional manual approaches used by medical experts using 2D planar image and modified Deepmedic CNN model on SPECT/CT images, and is sufficient small to be used in clinical practice. In conclusion, this study proves the potential of the nearly automatic methods in supporting nuclear medicine doctor in treatment planning prior to the use of SIRT for liver cancer.

Conflicts of interest

None.

Acknowledgments

Luu Manh Ha, ID VNU.2021.TTS.13, thanks The Development Foundation of Vietnam National University, Hanoi for sponsoring this research. Le Quoc Anh was funded by Vingroup JSC and supported by the Master, PhD Scholarship Programme of Vingroup Innovation Foundation (VINIF), Institute of Big Data, code VINIF.2021.ThS.95. We would like to thank NVIDA Corporation for supporting a GPU RTX 8000 for this study. We also would like to thank Professor Huu Tue Huynh for the consulting in this study. We would like to thank nuclear medicine staffs of Hospital 108 for the support in this study.

References

Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B.A., Litjens, G., Menze, B., Ronneberger, O., Summers, R.M., et al., 2022. The medical segmentation decathlon. Nature communications 13, 1–13. doi:10.1038/s41467-022-30695-9.

Balakrishnan, G., Zhao, A., Sabuncu, M.R., Guttag, J., Dalca, A.V., 2019. Voxelmorph: a learning framework for deformable medical image registration. IEEE transactions on medical imaging 38, 1788–1800. doi:10.1109/TMI.2019.2897538.

Balogh, J., Victor III, D., Asham, E.H., Burroughs, S.G., Boktour, M., Saharia, A., Li, X., Ghobrial, R.M., Monsour Jr, H.P., 2016. Hepatocellular carcinoma: a review. Journal of hepatocellular carcinoma 3, 41. doi:10.2147/JHC.S61146.

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M., 2021. Swin-unet: Unet-like pure transformer for medical image segmentation. doi:10.48550/arXiv.2105.05537.

Chaichana, A., Frey, E.C., Teyateeti, A., Rhoongsittichai, K., Tocharoenchai, C., Pusuwan, P., Jangpatarapongsa, K., 2021. Automated segmentation of lung, liver, and liver tumors from tc-99m maa spect/ct images for y-90 radioembolization using convolutional neural networks. Medical physics 48, 7877–7890. doi:10.1002/mp.15303.

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y., 2021. Transunet: Transformers make strong encoders for medical image segmentation. doi:10.48550/arXiv.2102.04306.

Christ, P.F., Elshaer, M.E.A., Ettlinger, F., Tatavarty, S., Bickel, M., Bilic, P., Rempfler, M., Armbruster, M., Hofmann, F., D'Anastasi, M., et al., 2016. Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields, in: International conference on medical image computing and computer-assisted intervention, Springer. pp. 415–423. doi:10.1007/978-3-319-46723-8_48.

Conze, P.H., Kavur, A.E., Cornec-Le Gall, E., Gezer, N.S., Le Meur, Y., Selver, M.A., Rousseau, F., 2021. Abdominal multi-organ segmentation with cascaded convolutional and adversarial deep networks. Artificial Intelligence in Medicine 117, 102109. doi:10.1016/j.artmed.2021.102109.

Dittmann, H., Kopp, D., Kupferschlaeger, J., Feil, D., Groezinger, G., Syha, R., Weissinger, M., La Fougère, C., 2018. A prospective study of quantitative spect/ct for evaluation of lung shunt fraction before sirt of liver tumors. Journal of Nuclear Medicine 59, 1366–1372. doi:10.2967/jnumed.117.205203.

EASL, E.A.F.T.S.O.T.L., et al., 2018. Easl clinical practice guidelines: management of hepatocellular carcinoma. Journal of hepatology 69, 182–236. doi:10.1016/j.jhep.2018.03.019.

Fischer, B., Modersitzki, J., 2004. A unified approach to fast image registration and a new curvature based registration technique. Linear Algebra and its Applications 380, 107–124. doi:10.1016/j.laa.2003.10.021.

Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X., 2020. Deep learning in medical image registration: a review. Physics in Medicine & Biology 65, 20TR01. doi:10.1088/1361-6560/ab843e.

Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X., 2021. A review of deep learning based methods for medical image multi-organ segmentation. Physica Medica 85, 107–122. doi:10.1016/j.ejmp.2021.05.003.

Georgiou, M.F., Kuker, R.A., Studenski, M.T., Ahlman, P.P., Witte, M., Portelance, L., 2021. Lung shunt fraction calculation using 99mtc-maa spect/ct imaging for 90y microsphere selective internal radiation therapy of liver tumors. EJNMMI research 11, 1–10. doi:10.1186/s13550-021-00837-z.

Gil-Alzugaray, B., Chopitea, A., Iñarrairaegui, M., Bilbao, J.I., Rodriguez-Fraile, M., Rodriguez, J., Benito, A., Dominguez, I., D'Avola, D., Herrero, J.I., et al., 2013. Prognostic factors and prevention of radioembolization-induced liver disease. Hepatology 57, 1078–1087. doi:10.1002/hep.26191.

Heimann, T., Van Ginneken, B., Styner, M.A., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., et al., 2009. Comparison and evaluation of methods for liver segmentation from ct datasets. IEEE transactions on medical imaging 28, 1251–1265. doi:10.1109/TMI.2009.2013851.

Ho, S., Lau, W., Leung, T., Chan, M., Johnson, P., Li, A., 1997. Clinical evaluation of the partition model for estimating radiation doses from yttrium-90 microspheres in the treatment of hepatic cancer. European journal of nuclear medicine 24, 293–298. doi:10.1007/BF01728766.

Hoang, H.S., Pham, C.P., Franklin, D., van Walsum, T., Luu, M.H., 2019. An evaluation of cnn-based liver segmentation methods using multi-types of ct abdominal images from multiple medical centers, in: 2019 19th international symposium on communications and information technologies (ISCIT), IEEE. pp. 20–25. doi:10.1109/ISCIT.2019.8905166.

Hofmanninger, J., Prayer, F., Pan, J., Röhrich, S., Prosch, H., Langs, G., 2020. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. European Radiology Experimental 4, 1–13. doi:10.1186/s41747-020-00173-2.

Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H., 2021. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods 18, 203–211. doi:10.1038/s41592-020-01008-z.

Kao, Y.H., Tan, A.E.H., Burgmans, M.C., Irani, F.G., Khoo, L.S., Lo, R.H.G., Tay, K.H., Tan, B.S., Chow, P.K.H., Ng, D.C.E., et al., 2012. Image-guided personalized predictive dosimetry by artery-specific spect/ct partition modeling for safe and effective 90y radioembolization. Journal of Nuclear Medicine 53, 559–566. doi:10.2967/jnumed.111.097469.

Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P., 2009. Elastix: a toolbox for intensity-based medical image registration. IEEE transactions on medical imaging 29, 196–205. doi:10.1109/TMI.2009.2035616.

Lee, C.E., Chung, M., Shin, Y.G., 2022a. Voxel-level siamese representation learning for abdominal multi-organ segmentation. Computer Methods and Programs in Biomedicine 213, 106547. doi:10.1016/j.cmpb.2021.106547.

Lee, C.E., Park, H., Shin, Y.G., Chung, M., 2022b. Voxel-wise adversarial semi-supervised learning for medical image segmentation. Computers in Biology and Medicine 150, 106152. doi:10.1016/j.compbiomed.2022.106152.

Lester, H., Arridge, S.R., 1999. A survey of hierarchical non-linear medical image registration. Pattern Recognition 32, 129–149. doi:10.1016/S0031-3203(98)00095-8.

Leung, T., Lau, W.Y., Ho, S., Ward, S.C., Chow, J., Chan, M., Metreweli, C., Johnson, P.J., Li, A., 1995. Radiation pneumonitis after selective internal radiation treatment with intraarterial 90yttrium-microspheres for inoperable hepatic tumors. International journal of radiation oncology, biology, physics 33, 919–924. doi:10.1016/0360-3016(95)00039-3.

Li, C., Mao, Y., Guo, Y., Li, J., Wang, Y., 2022. Multi-dimensional cascaded net with uncertain probability reduction for abdominal multi-organ segmentation in ct sequences. Computer Methods and Programs in Biomedicine 221, 106887.

Li, X., Chen, H., Qi, X., Dou, Q., Fu, C.W., Heng, P.A., 2018. H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes. IEEE transactions on medical imaging 37, 2663–2674. doi:10.1109/TMI.2018.2845918.

Luo, S., et al., 2014. Review on the methods of automatic liver segmentation from abdominal images. Journal of Computer and Communications 2, 1. doi:10.4236/jcc.2014.22001.

Mattes, D., Haynor, D.R., Vesselle, H., Lewellen, T.K., Eubank, W., 2003. PET-CT image registration in the chest using free-form deformations. IEEE transactions on medical imaging 22, 120–128. doi:10.1109/TMI.2003.809072.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241. doi:10.1007/978-3-319-24574-4_28.

Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE transactions on medical imaging 18, 712–721. doi:10.1109/42.796284.

Salem, R., Lewandowski, R.J., Mulcahy, M.F., Riaz, A., Ryu, R.K., Ibrahim, S., Atassi, B., Baker, T., Gates, V., Miller, F.H., et al., 2010. Radioembolization for hepatocellular carcinoma using yttrium-90 microspheres: a comprehensive report of long-term outcomes. Gastroenterology 138, 52–64. doi:10.1053/j.gastro.2009.09.006.

Sangro, B., Carpanese, L., Cianni, R., Golfieri, R., Gasparini, D., Ezziddin, S., Paprottka, P.M., Fiore, F., Van Buskirk, M., Ignacio Bilbao, J., et al., 2011. Survival after yttrium-90 resin microsphere radioembolization of hepatocellular carcinoma across barcelona clinic liver cancer stages: a european evaluation. Hepatology 54, 868–878. doi:10.1002/hep.24451.

Son, M.H., Ha, L.N., Bang, M.H., Bae, S., Giang, D.T., Thinh, N.T., Paeng, J.C., 2021. Diagnostic and prognostic value of 99mtc-maa spect/ct for treatment planning of 90y-resin microsphere radioembolization for hepatocellular carcinoma: comparison with planar image. Scientific reports 11, 1–9. doi:10.1038/s41598-021-82887-w.

Tang, X., Jafargholi Rangraz, E., Coudyzer, W., Bertels, J., Robben, D., Schramm, G., Deckers, W., Maleux, G., Baete, K., Verslype, C., et al., 2020. Whole liver segmentation based on deep learning and manual adjustment for clinical use in sirt. European journal of nuclear medicine and molecular imaging 47, 2742–2752. doi:10.1007/s00259-020-04800-3.

Thévenaz, P., Unser, M., 2000. Optimization of mutual information for multiresolution image registration. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society 9, 2083–2099. doi:10.1109/83.887976.

Viola, P., Wells III, W.M., 1997. Alignment by Maximization of Mutual Information. International Journal of Computer Vision 24, 137–154. doi:10.1023/A:1007958904918.

Visvikis, D., Lambin, P., Beuschau Mauridsen, K., Hustinx, R., Lassmann, M., Rischpler, C., Shi, K., Pruim, J., 2022. Application of artificial intelligence in nuclear medicine and molecular imaging: a review of current status and future perspectives for clinical translation. European journal of nuclear medicine and molecular imaging , 1–12.doi:10.1007/s00259-022-05891-w.

Wang, J., Zhao, H., Liang, W., Wang, S., Zhang, Y., 2023. Cross-convolutional transformer for automated multi-organs segmentation in a variety of medical images. Physics in Medicine and Biology doi:10.1088/1361-6560/acb19a.

Xu, W., Liu, H., Wang, X., Qian, Y., 2019. Liver segmentation in ct based on resunet with 3d probabilistic and geometric post process. 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP) , 685–689.doi:10.1109/SIPROCESS.2019.8868690.

Yang, J., Veeraraghavan, H., Armato III, S.G., Farahani, K., Kirby, J.S., Kalpathy-Kramer, J., van Elmpt, W., Dekker, A., Han, X., Feng, X., et al., 2018. Autosegmentation for thoracic radiation treatment planning: a grand challenge at aapm 2017. Medical physics 45, 4568–4581. doi:10.1002/mp.13141.