Journal
Cover
Image

Automatic scan range for dose-reduced multiphase CT imaging of the liver utilizing CNNs and Gaussian models

Manh Ha Luu [a, b, c, *], Theo van Walsum [b], Hong Son Mai [d], Daniel Franklin [e], Thi Thu Thao Nguyen [f], Thi My Le [g], Adriaan Moelker [b], Van Khang Le [f], Dang Luu Vu [f], Ngoc Ha Le [d], Quoc Long Tran [h], Duc Trinh Chu [c], Nguyen Linh Trung [a]

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

ARTICLE INFO

ABSTRACT

Multiphase CT scanning of the liver is performed for several clinical applications; however, radiation exposure from CT scanning poses a nontrivial cancer risk to the patients. The radiation dose may be reduced by determining the scan range of the subsequent scans by the location of the target of interest in the first scan phase. The purpose of this study is to present and assess an automatic method for determining the scan range for multiphase CT scans. Our strategy is to first apply a CNN-based method for detecting the liver in 2D slices, and to use a liver range search algorithm for detecting the liver range in the scout volume. The target liver scan range for subsequent scans can be obtained by adding safety margins achieved from Gaussian liver motion models to the scan range determined from the scout. Experiments were performed on 657 multiphase CT volumes obtained from multiple hospitals. The experiment shows that the proposed liver detection method can detect the liver in 223 out of a total of 224 3D volumes on average within one second, with mean intersection of union, wall distance and centroid distance of 85.5%, 5.7 mm and 9.7 mm, respectively. In addition, the performance of the proposed liver detection method is comparable to the best of the state-of-the-art 3D liver detectors in the liver detection accuracy while it requires less processing time. Furthermore, we apply the liver scan range generation method on the liver CT images acquired from radiofrequency ablation and Y-90 transarterial radioembolization (selective internal radiation therapy) interventions of 46 patients from two hospitals. The result shows that the automatic scan range generation can significantly reduce the effective radiation dose by an average of 14.5% (2.56 mSv) compared to manual performance by the radiographer from Y-90 transarterial radioembolization, while no statistically significant difference in performance was found with the CT images from intra RFA intervention ($p = 0.81$). Finally, three radiologists assess both the original and the range-reduced images for evaluating the effect of the range reduction method on their clinical decisions. We conclude that the automatic liver scan range generation method is able to reduce excess radiation compared to the manual performance with a high accuracy and without penalizing the clinical decision.

© 20XX

1. Introduction

Liver cancer is the fourth-most common cause of cancer death globally, with approximately 1 million new cases per year worldwide (Bray et al., 2018). The rate of liver cancer is increasing significantly in devel-

* Corresponding author at: AVITECH, University of Engineering and Technology, VNU, Hanoi, Vietnam.

E-mail address: halm@vnu.edu.vn (M.H. Luu).

oping countries in East Asia, South East Asia and Africa (McGlynn and London, 2011). Although MRI is sometimes available as a low radiation exposure imaging modality for the diagnosis of liver cancer, multiphase CT scanning is the most typical choice due to its low costs and time efficiency. However, CT uses ionizing radiation, and thus CT imaging is associated with an increased risk of radiation-induced cancer to the patient (Lin, 2010; Shao et al., 2020). The United States Food and Drug Administration (FDA) suggested that a CT scan with an effective dose of 10 mSv is associated with a 1/2000 risk of the development of fatal cancers (FDA, 2018). For common minimally invasive liver cancer interventions such as radiofrequency ablation (RFA) and selective internal radiation therapy (SIRT), CT scanning is performed multiple times during the process of diagnosis and treatment. Consequently, the absorbed radiation doses to the patients may increase.

Awareness of the importance of the radiation dose associated with CT scanning is increasing (Shao et al., 2020; Brenner and Hall, 2007; Goldman and Maldjian, 2013; Raman et al., 2013). Lowering radiation dose by reducing the tube current voltage has been investigated in several studies (Goldman and Maldjian, 2013; Raman et al., 2013). However, the consequence of lowering radiation dose is a reduction in image quality, which may affect the clinical decision-making process of the radiologists. The accumulated radiation dose also relates to the scan range, where a larger scan range results in a higher absorbed dose to the patient (Raman et al., 2013; Zinsser et al., 2019). Several studies investigating scan coverage for several organs have been published (Goldman and Maldjian, 2013; Johnson et al., 2015). Goodman et al. (1979) is amongst the first to have raised the issue of optimizing liver scan range in CT imaging. Goldman and Maldjian (2013) suggested a CT scan protocol for liver transplant planning for patients with liver disease. Devapalasundaram et al. (2016) used a Liver Detection Algorithm, LDA, for reducing CT noise in the liver; however the paper did not mention the method details. Zhang et al. (2010) presented a hierarchical Markov network for automatically delimiting scan range via detecting anatomical landmarks in topogram images. McCollough and Leng (2020) first introduced that artificial intelligence algorithm can automatically delineate scan range which covers all of the lung anatomy. Demircioğlu et al. (2021) also applied deep learning to automatically determine scan range for multiphase CT imaging of the lung.

In clinical practice, a multiphase CT scan session of the liver often starts with a tomography for acquiring the field of view information; next, a non-contrast enhanced CT is performed, followed with the arterial phase scan at 15 to 40 s after the contrast agent is injected to the body of the patient (Lip-Pauwels et al., 2012). Subsequently, the portal-venous and delayed phase scans are performed at specific time points. The scanning process is performed by a radiographer who manually chooses the scan range based on a scout scan (either the topogram or the non-contrast enhanced CT) and an estimation of the liver motion due to the respiration (Demircioğlu et al., 2021; Zanca et al., 2012). The accuracy of the manually chosen scan range depends on experience of the radiographer. Also, the decision must be taken under time pressure, and thus the chosen scan range, often based on landmarks such as the iliac crest and the xiphosternum or diagram, is larger than the actual liver range (Goldman and Maldjian, 2013; Johnson et al., 2015). Therefore, an efficient automatic method for scanning the liver using CT images in liver interventions may benefit patients by reducing the absorbed radiation dose from CT scanning. It is the purpose of our work to develop such an automatic method.

Several organ localization methods have been proposed in the literature, which fall into two categories: conventional methods and deep learning-based methods. Conventional methods often use atlas-based approaches (Gauriau et al., 2015). Generally, medical experts define the information of anatomical structures; subsequently, these atlases are registered to an image with unknown anatomical structures for localizing/segmenting organs. In this approach, liver localization is commonly performed as a preprocessing step prior to the segmentation task

- a time-consuming procedure (Jimenez-del Toro et al., 2016; de Vos et al., 2017; Xu et al., 2019b; Navarro et al., 2020). Organ localization problems also have been investigated using various machine learning-based approaches (de Vos et al., 2017), including classification-based methods (Zhan et al., 2008; Zhou et al., 2012), marginal space learning (MSL) algorithms (Zheng et al., 2007; 2009), and regression analysis (Criminisi et al., 2013; Samarakoon et al., 2017). Classification-based approaches are utilized for estimating the bounding box of the organ target, which are based on various image features and heuristics (e.g., edge, intensity, texture, shape). Zhan et al. (2008) proposed a rigorous formula for solving of organ localization. One of the challenges that has been addressed by Zhou et al. (2012) is liver segmentation on PET-CT images, in which an ensemble learning strategy is used to determine 2D bounding boxes of the target organ from three orthogonal views, and then a 3D bounding box is determined by combining the 2D bounding boxes through collaborative majority voting. An MSL algorithm was used to estimate the 3D organ bounding box by using nine pose parameters from three consecutive stages (Zheng et al., 2007); these pose parameters include the localization, orientation, and size of the bounding box. Zheng et al. (2009) have proposed two novel MSL techniques, named constrained MSL and nonrigid MSL, for liver detection in CT scans. Criminisi et al. (2013) first utilized random regression forests (RRFs) based on texture features for solving multi-organ localization in CT scans. Later, Gauriau et al. (2015) improved the effectiveness of RRFs by applying a global-to-local cascade of RRFs for reducing organ localization errors. Samarakoon et al. (2017) proposed light RRFs for localizing the organ with the aim of speeding up the processing time and reducing computational cost.

With the recent explosion of interest in deep learning technologies, convolution neural networks (CNNs) have been proven to be effective in localizing objects when a sufficient amount of training data is available. In organ localization, it has been reported that CNNs have outperformed classical machine learning approaches (Xu et al., 2019b; Hussain et al., 2017). de Vos et al. (2016) trained three independent CNNs for performing organ localization in three orthogonal directions (axial, coronal, sagittal). The 3D bounding box is determined by combining organ location status (present or absent) in each orthogonal direction. Later, de Vos et al. (2017) utilized spatial pyramidal pooling (SPP) in single CNNs to allow the input to be compatible with multiple sizes. Humpire-Mamani et al. (2017, 2018) introduced CNNs for multiple organ localization in a 3D CT scan by simultaneously examining multiple adjacent slices of the data set in each of the three orthogonal image planes to take advantage of their mutual information. Zhou et al. (2019) utilized a CNN for localizing organs and achieved a significant improvement in computational efficiency over their previous machine-learning based approach, which employed ensemble-learning on 2D sections and 3D majority voting.

Several 3D CNNs have been proposed to take advantage of spatial contextual information for organ localization in 3D medical images. Xu et al. (2019b) proposed a novel backbone architecture of 3D CNNs based on 3D faster R-CNNs for localizing multiple organs in a CT scan. The modified 3D CNNs utilize a multi-candidate fusion block instead of the original faster R-CNN classifier; this block combines all predicted bounding boxes with the same label to achieve the final 3D bounding box. In a further study (Xu et al., 2019a), a 3D CNN with three separate branches was utilized in each of the three orthogonal directions, resulting in three-channel images as the input of the network. Navarro et al. (2020) first introduced deep Q-reinforcement learning for organ localization in CT scans. In this approach, an artificial agent utilizing a CNN-based estimator of expected future reward teaches itself to perform organ localization through iterated deformation of a bounding box, learning from both its successes and mistakes using a manually-segmented dataset as ground truth. This method achieves comparable performance to purely CNN-based methods but with a substantially smaller training dataset.

Whether the 3D CNN-based approach is superior to the 2D CNN-based approach is still not entirely clear, and 3D CNNs generally require much greater computational resources as compared to 2D CNNs. Recently, organ localization in CT via a series of 2D images has been investigated. Xia and Yin (2019) utilized a combination of a deep 2D CNN, DenseNet 121, and an edge-perception fusion network for localizing the liver in a range of different datasets with an improvement in detection efficiency as compared to many other competing network architectures, including the original DenseNet. Pang et al. (2019) proposed a novel Yolo-based model for identifying cholelithiasis and classifying gallstones in CT images, in which the liver localization was used as an prerequisite. Hammami et al. (2020) introduced a combined method for multi-organ detection in CT images. First, a Cycle Generative Adversarial Network (CycleGAN) was used to generate synthetic CT images from both MRI and CT images of the same patient. Then, the YOLOv3 detector was trained on the generated images, achieving a significant improvement in the accuracy of the organ localization over the YOLOv3 detection alone.

Although several liver detection methods have been investigated, to date there has been no study specifically addressing the problem of applying machine learning/deep learning to determine the liver scan range in multiphase CT images. Hence, the contribution of this study is to propose and assess an automatic method for the liver scan range generation based on 3D liver detection in multiphase CT images.

The main challenges of detecting liver range in multiphase CT images are:

1. The liver is an organ with varying size and shape along the slices;
2. The liver motion, caused by the reparatory effect; and
3. The performance of liver detection needs to be accurate and fast (should be completed within few seconds).

To deal with these challenges, we apply a CNN to detect the liver in 2D CT images. Subsequently, we propose a liver range search algorithm, namely LRS, to determine the whole liver range in a 3D CT volume, i.e. the scout scan. We then estimate the liver range positions in subsequent scans based on Gaussian models of the liver motion due to respiration. To assess the method, we compare the accuracy of the extracted liver range by the proposed method to the liver scan range estimation performed by radiographers in two independent hospitals for RFA and SIRT interventions, and subsequently estimate the potential excess radiation dose reduction that can be achieved.

The rest of this paper is organized as follows. The method section describes YOLOv4 as a CNN-based detector for the liver, the proposed algorithm and the Gaussian liver motion models. In the experiment and result section, we present in detail the data used for this study, and describe the implementation, training, testing, comparison to several well-known CNN-based liver detection methods, as well as the potential excess radiation reduction compared to clinician's performance. Next, the discussion section provides the implications and the limitations of the study. Finally, the conclusion section summarizes the most important findings in this work.

2. Methods

2.1. Data, annotations and preprocessing

For our study, we used 657 abdominal CT images from three sources. The first dataset is retrospectively reused from our previous study (Luu et al., 2018; 2021) which contains diagnostic (EMC_{diag} subset, 202 volumes) and intra-interventional (EMC_{intra} subset, 147 volumes) CT images in RFA liver cancer treatment of 51 patients in Erasmus MC, Rotterdam. The diagnostic scans were performed under either a three or four-phase protocol (Lip-Pauwels et al., 2012) while the radiographer, with the aim of minimizing the liver scan range, scanned the

interventional images manually. All images where the scan range did not cover the whole liver were excluded in the test phase of this study because the field of view of these images focuses on the ablator at the current tumor rather than on the whole liver. The second dataset contains 210 contrast enhanced CT volumes of the liver, portal venous phase, which are publicly available from the LiTS Challenge (*LiTS* dataset) (Bilic et al., 2019) and Mayo Clinic (*Mayo* dataset) (McCollough et al., 2017). The third dataset contains the three/four-phase contrast enhanced CT diagnostic and follow-up images from Y-90 SIRT treatment of 36 liver cancer patients in Hanoi's Hospital-108 (*H108* dataset), which was retrospectively collected in our previous study Mai et al. (2021). For 29 of the patients in the *H108* dataset, the diagnostics images are available ($H108_{diag}$ subset) and for 14 of the patients the follow-up, post treatment CT images ($H108_{post}$ subset) acquired one to three months after the intervention are available. All the data was anonymized before being used in this study. The dataset from *EMC*, *Mayo* and *H108* were acquired under a low radiation dose protocol (Mai et al., 2021; Hoang et al., 2019; Lip-Pauwels et al., 2012) while the *LiTS* dataset was acquired under regular dose protocol. The characteristics of the datasets are described in detail in Table 1.

For the *EMC*, *LiTS* and *Mayo* datasets, we used the manual liver segmentations which were performed by experts in the previous studies to compute the 2D bounding boxes of the liver in each slice. Subsequently, 3D bounding boxes can be determined by combining the boxes in the whole volume, and the liver range can be extracted from the top and bottom of the 3D liver bounding boxes. The 2D bounding boxes were used for training and evaluation of the liver CNN models, while the 3D bounding boxes were used to compare the liver detection methods in 3D (see Section 3.4). For all CT images in the *H108* dataset, a technician manually determined the upper and the lower extent of the liver in the Z axis, which then were verified by an expert. The liver range annotations were then used as the ground truth to evaluate the liver scan range generated by the proposed method.

CT images of 29 patients from the *EMC* dataset combined with *LiTS* and *Mayo* datasets were selected for training the CNN model, with a total of 335 volumes and 71,062 slices (51.2% of those containing the liver). The data from 22 patients (38 scan sessions, 155 CT volumes) from the *EMC* dataset and 69 test images in the *LiTS* dataset were used to evaluate liver detection methods, and the *H108* dataset (43 scan sessions) and the EMC_{intra} subset (10 scan sessions) were used for testing the proposed liver scan range determination method (46 patients in total). The details of the training and test datasets are listed in Table 2.

In the preprocessing step, we thresholded the image intensities from -100 to 400 HU since the intensities of the liver are within this range. Next, we scaled the threshold range to a numerical range of 0 to 255, and then converted the images from DICOM or NII into 8-bit TIFF.

Table 1
The detailed characteristics of each of the datasets.

Dataset	EMC	LiTS	Mayo	H108
Scanner	Siemens, Philips, Toshiba	-	Siemens	Toshiba, GE
Tube voltage (kVp)	80, 90, 100, 110, 120, 140, 150	-	100, 120	100, 120
Median X-ray Tube Current (mA)	215 (30 - 649)	-	418 (77 - 667)	143 (45 - 430)
Median Exposure (mAs)	157 (21 - 540)	-	322 (55 - 538)	50 (3 - 215)
Median CTDIvol (mGy)	6.3 (2–17.7)	-	17.4 (9.5 - 30.1)	7.3 (3.8– 17.5)
In-plane resolution (mm)	0.56 - 0.98	0.56 - 1	0.66 - 0.82	0.67 - 0.98
Spacing between slices (mm)	0.75 - 5	0.45 - 6	1	0.5 - 5
Number of slices per volume	24 - 892	42 - 1026	318 - 856	60 - 1601

Table 2

The number of CT images used for training CNN models for liver detection and testing/evaluation of the models and the proposed method in this study. The numbers in parentheses are the number of 2D images.

Dataset		Non-contrast	Arterial	Portal venous	Delay
Training	*LiTS*	-	-	131 (21707)	-
	EMC_{diag}	26 (3511)	34 (3650)	37 (4877)	23 (4079)
	EMC_{intra}	17 (778)	25 (3635)	26 (3770)	6 (1025)
	Mayo	-	-	10 (2323)	-
Test	*LiTS*	-	-	69 (26529)	-
	EMC_{diag}	13 (816)	28 (4066)	26 (5000)	15 (1362)
	EMC_{intra}	13 (809)	23 (3403)	28 (3985)	9 (1087)
	H108	43 (8911)	43 (12022)	-	12 (2936)

2.2. Proposed method for scan range determination

The proposed method for delineating the scan range in multiphase CT imaging of the liver is illustrated in Fig. 1. First, the liver range is detected in the scout scan (the non-contrast enhanced image) using a 2D CNN detector and LRS. Next, the liver motion in subsequent scans is estimated using the Gaussian models. This safety margin is then added to the estimated liver range determined from the scout image to obtain the scan range for the artery phase image and the delayed image as in Goldman and Maldjian (2013)'s protocol. We define the detection failure when either no *anchor* slice is found within the ground truth of the liver range, or the detected liver range does not cover the ground truth of the liver range. If the first attempt to detect the liver range in the non-contrast enhanced CT image is not successful, we apply the CNN detector and LRS on the artery phase image and the portal-venous phase image.

2.2.1. Liver detection in a 2D slice

The liver is an organ whose size and shape vary considerably along the slices. A 3D CNN-based liver detection approach may fail to detect small parts of the liver because the 3D convolutional kernels are not adept at dealing with the varying size of the upper and lower parts of the liver. Therefore, we have adopted a 2D approach as our core detection structure.

Here we describe YOLOv4 (Bochkovskiy et al., 2020) as a candidate for the 2D liver detector due to its two critical properties:

- YOLOv4 can detect objects at multiple scales; and
- YOLOv4 is accurate and has a fast inference time.

YOLO, introduced by Redmon and Farhadi (2018), is a well-known CNN-based method for automatic object detection. As a fully CNN, YOLO does not require handcrafted-feature selection for detecting objects. The key idea of YOLO is to detect bounding boxes of objects by classifying object parts in pre-defined grids and then combining candidate parts of the same object into a single box. The improved YOLOv4, introduced by Bochkovskiy et al. (2020) is the most recent evolution of the YOLO architecture. YOLOv4 uses CSPDarknet-53 as its **Backbone** to extract the CNN features; the **Neck** includes a Spatial Pyramid Pooling (SPP) (He et al., 2014) structure combined with PANet (Liu et al., 2018) which enables generalization of the size of the input image rather than limiting it to fixed dimensions; and finally, the **Head**, which is inherited from YOLOv3 (Redmon and Farhadi, 2018), employs detection kernels on feature maps of three different sizes in parallel, and thus enables the detection of the liver across a range of organ scales (Figs. 2).

2.2.2. Liver range search algorithm

In this section, we present the proposed algorithm, LRS, for liver range detection in a 3D CT image. The liver is a single organ in the abdomen, while the scan range in the scout image is often much larger than the liver range. As detecting the liver by sequentially processing the stack of 2D images may be time-consuming, the key idea of the algorithm is to detect a slice (LIVER-DETECTION) that intersects with the liver (the *anchor* slice) using an uniform-random search algorithm (GET-RANDOM) with a hopping step of HS, followed by detection of the upper extent and the lower extent by a hopping strategy. The value of HS depends on the slice spacing and the length of the liver (from 10 to 33.7 cm based on analysis of our data), which will be experimentally verified in Section 3.4.1). Once the hopping detections from the *anchor* slice are beyond the upper or lower extreme end of the liver, the liver detection (LIVER-DETECTION) runs towards the liver one slice per step until the upper and the lower extent are reached. Finally, because YOLOv4 still may fail to detect a very small liver section at the upper and lower extent, two margins (MLU and MLL, respectively) are added to the detected upper and the lower extent of the liver to guarantee the final detection range covers the entire liver. Values for these margins depend on quality of the scan, and were determined via experiment (see Section 3.4.3). Theoretically, the computational time mainly depends
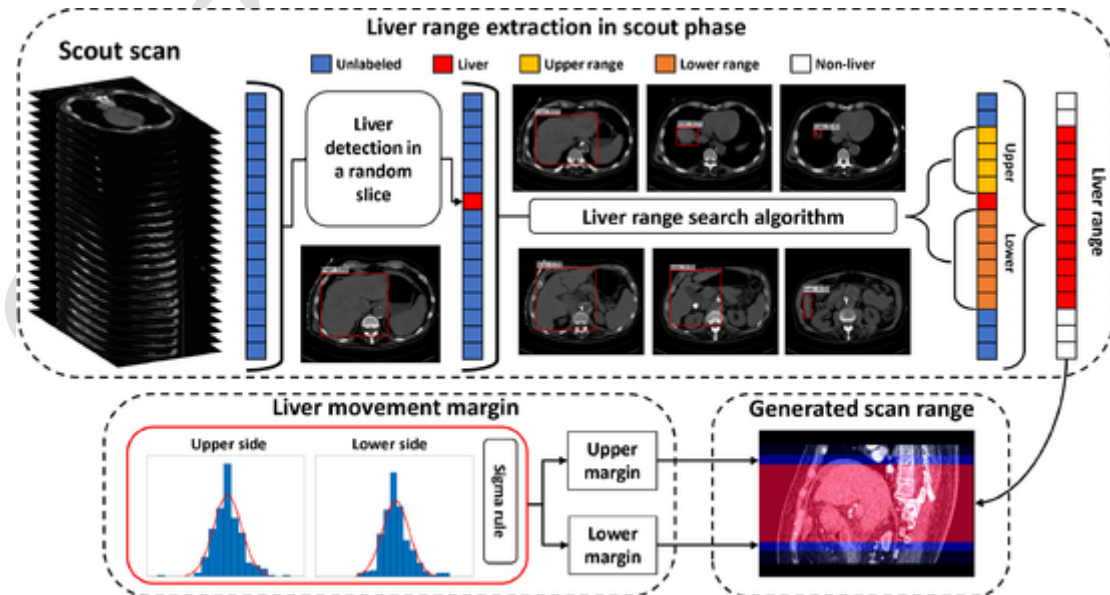

Fig. 1. Block diagram of the proposed method for delineating the scan range in multiphase CT imaging of the liver.
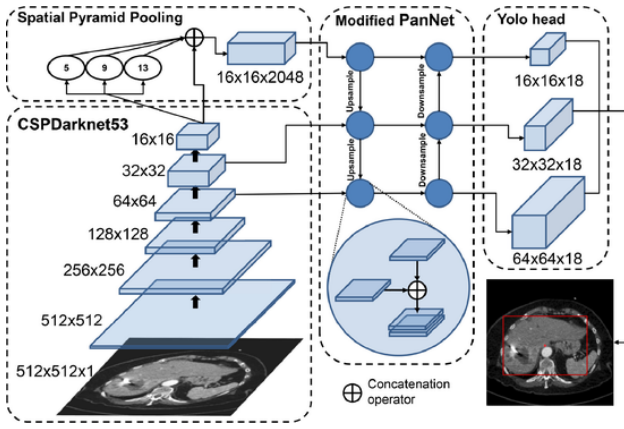
Fig. 2. Architecture of YOLOv4 for the task of liver detection in a 2D CT image.

on the *anchor* searching which relates to the ratio of the number of liver slice over the total number of slices in a volume. In this study, we investigate the computation time required by LRS and compare to a linear search in Section 3.4.3. The pseudo-code of the algorithm is shown in Algorithm 1.

2.2.3. Liver motion margin estimation

Since no respiratory signal is acquired during multiphase CT scanning of the liver, it is infeasible to accurately predict the liver position the subsequent scans. Therefore, we model the difference between the upper and the lower extreme of the liver between two scans with Gaussian distributions, and estimate the maximum motion using a 3-sigma margin. We first determine individual differences of the upper extreme margin of the liver from the scans in a same scan sessions, then combine all of the differences between these extremes into a histogram. Furthermore, we repeat the similar experiment for the lower liver extreme. Note that within a scan session, while the liver moves due to respiration, the image coordinate remains the same (Fig. 3). Finally, we fit the parameter of the Gaussian models based on the histograms using the displacement of the upper and the lower liver extent from 132 multiphase CT scan sessions of 87 patients. The analysis shows that the standard deviations of the upper and the lower liver motion position are 5.3 mm and 5.1 mm, respectively, resulting in the upper margin and the lower margin of 15.9 mm and 15.3 mm, respectively. The Gaussian models of the liver motion are illustrated in Fig. 4.

3. Experiment and results

3.1. Implementation

The study was carried on an Windows 10 64-bit PC, with 02 CPU Intel®Xeon(R) Processor X5650 @ 2.67GHz, 36 GB DDR3, 1333 MHz bus and 512 GB M2 SSD storage. We used an NVIDIA GeForce GTX 1080 Ti/

PCIe 11GB VRAM and RTX 8000/PCIe 48GB VRAM with CUDA version 10.1 to train the CNN models and evaluate the the time consumption of the liver detection methods.

The source code for YOLOv4, which is written in C + + and uses the Darknet library, was taken from the authors' github repository at https://github.com/AlexeyAB/darknet, while the pre-trained model was trained with the MS COCO dataset (Lin et al., 2014). LRS was implemented in Python 3.8, including a Python wrapper to load the CNN model.

We also compare the performance of YOLOv4 with other well-known fast CNN-based detectors, such as SSD (Liu et al., 2016) and Faster R-CNN (Ren et al., 2017), for the 2D liver detection task. The implementation for SSD was obtained from https://github.com/kweisamx/SSD-PyTorch and the source code for Faster R-CNN was obtained from https://github.com/alankbi/detecto. In addition, we also reused the implementation of the detection network provided by Xu et al. (2019b) for the qualitative comparison of 3D liver detection task performance.

3.2. Parameters setting and training

The YOLOv4, SSD and Faster R-CNN models were initialized from pre-trained models and were fine-tuned with the training datasets over 50 epochs, using a batch size of 8 images, a learning rate of 1×10^{-3}, momentum of 0.9 and a decay of 5×10^{-4} as the default. The pre-trained model for YOLOv4 and Faster R-CNN were trained using MS COCO dataset (Lin et al., 2014) while those of SSD was trained on ImageNet dataset (Deng et al., 2009).

For the 3D liver detector provided by Xu et al. (2019b), we trained the model from scratch using the training dataset with the training parameters as suggested by the authors.

The hopping step of LRS and the confidence scores for the liver detection will be determined via the experiment in Section 3.4.1.

3.3. Evaluation metrics

3.3.1. Evaluation metrics for 2D liver detection

As liver detection in 2D images is an object detection problem, we used several common performance criteria taken from the COCO challenge (Lin et al., 2014; Hosang et al., 2016) for evaluating object detection methods, including average precision (AP), average precision at intersection over union (IoU) of 50% and 75% ($AP@50$ and $AP@75$, respectively), and average precision for small object (AP_S), medium object (AP_M) and large object (AP_L), which have a size smaller than 32×32 pixels, from 32×32 to 96×96 pixels, and larger then 96×96 pixels, respectively.

The precision of liver detection (PR) is computed as follows:
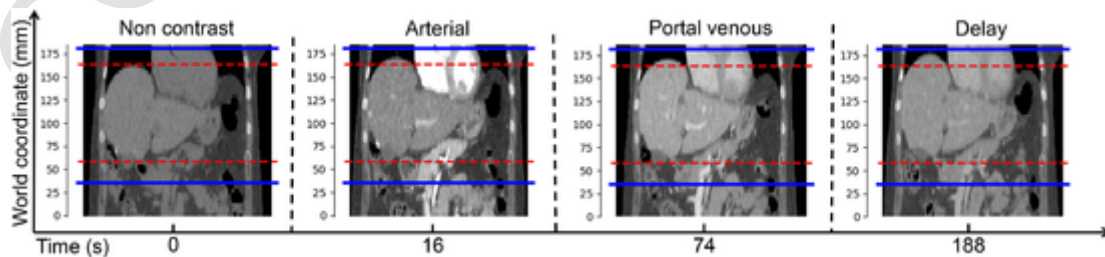
$$PR = \frac{TP}{TP + FP} \tag{1}$$

Fig. 3. The liver motion in a multiphase CT image of the liver: the scans were performed at four specific time points. The red dashed lines depict the live range in the first CT image and the blue solid lines indicate the delineated scan range. Note that in this study, we applied the delineated scan range on the arterial image and delayed image based on the Goldman's protocol. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

```
 1: Input: CT volume: Image, anchor theshold: AT, searching theshold: ST, hopping step: HS, margin: MLU and MLL
 2: Output: liver upper and lower extreme end: liver_upper, liver_lower
 3: N ← slice number of Image
 4: /*Extract a liver slice as the anchor*/
 5: index ← GET-RANDOM(N, HS)
 6: CS ← LIVER-DETECTION(Image[index])
 7: while CS < AT do
 8:     index ← GET-RANDOM(N, HS)
 9:     CS ←LIVER-DETECTION(Image[index])
10: end while
11: /* Find the upper extreme end */
12: current_upper ← index
13: step ← HS
14: while step > 0 do
15:     current_upper ← current_upper + step
16:     CS ←LIVER-DETECTION(Image[current_upper])
17:     if CS < ST then
18:         step ← step − 1
19:     end if
20: end while
21: /* Find the lower extreme end */
22: current_lower ← index
23: step ← HS
24: while step > 0 do
25:     current_lower ← current_lower − step
26:     CS ←LIVER-DETECTION(Image[current_lower])
27:     if CS < ST then
28:         step ← step − 1
29:     end if
30: end while
31: liver_upper ← current_upper + MLU
32: liver_lower ← current_lower + MLL
```

Algorithm 1. Liver Range Search.

where TP is the number of correctly-identified slices which contain the liver, and FP is the number of incorrectly-identified slices which contain the liver while there is no liver label inside.

The recall of the liver detection (RC) is computed as follows:

$$RC = \frac{TP}{TP + FN} \qquad (2)$$

where FN is the number of incorrectly-identified slices which do not contain the liver while there is/are liver label(s) inside.

3.3.2. Evaluation metrics for 3D liver detection

For evaluating the 3D bounding box liver detection, we use the $3D\,IoU$ criteria:

$$3D\,IoU = \frac{B_{GT} \cap B_P}{B_{GT} \cup B_P} \qquad (3)$$

where B_{GT} and B_P are the ground truth 3D bounding box and the predicted 3D bounding box, respectively.

Furthermore, we also evaluate the performance of the 3D liver detections using wall distance (WD), which is the absolute wall distance between the predicted bounding box and the ground truth bounding box, and centroid distance (CD), which is the absolute centroid distance between the predicted bounding box and the ground truth bounding box.

3.3.3. Evaluation metrics for liver range detection

The aim of our study is to develop an automatic method for liver scan range generation based on a liver detection/prediction approach; therefore, the liver range detection accuracy (Acc) can be evaluated as follows:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \qquad (4)$$

where TN is the number of the true negative liver slice detections.

In addition, for evaluating the liver range detection, the missing liver range (MLR) and the excess range (ER) can be computed using the difference between the actual end points of the liver range and the upper and the lower predicted liver range. Subsequently, the relative excess scan range reduction compared to the radiographer manual performance can be computed as:

$$RR = \frac{E_{Rad} - E_{Ma}}{E_{Rad}} \times 100\% \qquad (5)$$

where E_{Rad} and E_{Ma} are the excess range by the radiographer and the proposed method (machine generated).

3.4. Experimental results

3.4.1. Experiment for LRS parameters determination

In this experiment, we examine sensitivity of the parameters of LRS. For the confidence scores for the liver detection by YOLOv4, we verify the performance of the model by varying the confidence score w.r.t. the precision and recall metrics of the liver detections using the validation dataset. The result is illustrated in Fig. 5. The confidence score for the *anchor* liver detection (see Section 2.2.2) is set at 0.95 to ensure a high probability of correct liver detection (precision and recall of 1 and 0.64, respectively), and for the liver extend slice detections, a confidence score of 0.1 is chosen to guarantee a low liver detection miss rate (precision and recall of 0.98 and 0.98, respectively).

For the hopping step, we fix the confidence score of 0.95 for the *anchor* detection and 0.1 for the liver extend detections, and vary the hopping step from one slice to 100 mm with increments of 5 mm. The experiment is carried out on ten randomly selected CT volumes from the test dataset. The result of the experiment is illustrated in Fig. 6. It can be seen that with the smallest hopping steps (i.e. a single slice), it takes more than 3 s on average to complete a liver detection in 3D. The processing time reduces when the hopping step increases. In addition, for the small hopping steps (≤ 20 mm), the accuracy scores of the liver detection ($3D\,IoU$ and WD) remain unchanged, whereas they decrease for the large hopping steps. Therefore, the hopping step of LRS is chosen at 15 mm, yielding an average of processing time, $3D\,IoU$ and WD of 0.55 s, 87.6% and 6.1 mm, respectively.

3.4.2. Evaluation of 2D liver detection accuracy

In the first experiment, we evaluate three well-known fast 2D object detectors, i.e., SSD, Faster R-CNN and YOLOv4 on the same hardware (Section 3.1) with the same training and test datasets in this study (see Table 3). The result shows that YOLOv4 achieved the highest AP score of 71% with lowest inference time of 0.02. Furthermore, YOLOv4 also has the best performance with the AP at IoU of 50% and 75%, yielding the scores of 93.1% and 80.1%, respectively. YOLOv4 also obtains the best performance in detecting large livers (> 96x96 pixels) with an AP_L score of 75.9%, while Faster R-CNN performs better in detecting small
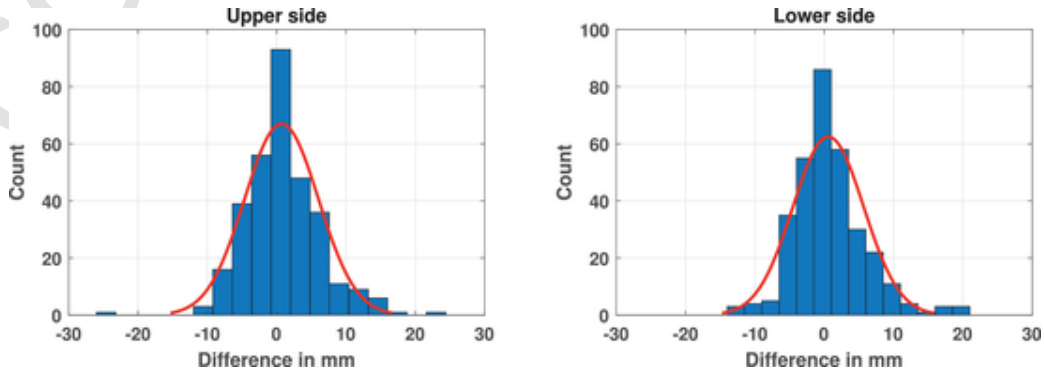

Fig. 4. The distribution and Gaussian fitting of the motion w.r.t. the upper and the lower liver among the phases of the CT scanning.
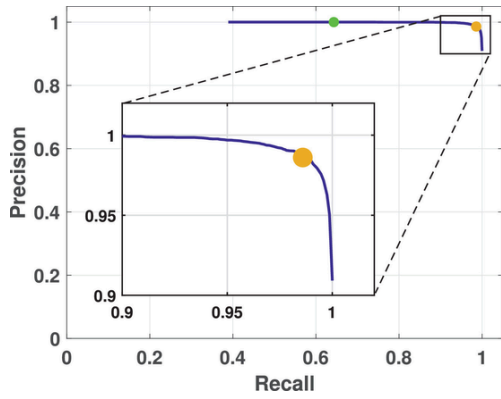
Fig. 5. Precision-Recall curve w.r.t. confidence scores of the liver detection using YOLOv4 on the validation dataset. The blue point is at confidence score of 0.95 (precision and recall of 1 and 0.64, respectively) while the yellow point is at confidence score of 0.1 (precision and recall of 0.98 and 0.98, respectively). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and medium liver sizes. Based on the fact that the smallest liver sections are at the top and bottom of the organ, incorrect small part liver detection can be compensated with an additional margin for the whole liver detection. Therefore, we choose YOLOv4 as the core detector of our liver detection.

Table 4 summarizes RC scores of YOLOv4 on the test dataset as compared to the RC scores reported in the state-of-the-art 2D liver detectors. It can be seen that YOLOv4 tested on the $LiTS$ dataset achieves a RC score of 96.2% on the $LiTS$ dataset which is similar to that of YOLOv3 tested on another dataset. Furthermore, the results of YOLOv4 tested on the EMC dataset are comparable with a combination of VGG16 and DenseNet (Xia and Yin, 2019). Note that the EMC dataset was

acquired with a low radiation dose protocol (see Table 1), with noisy CT images (Hoang et al., 2019), which may be responsible for the lower performance of YOLOv4 on this dataset.

3.4.3. Evaluation of 3D liver detection accuracy

In this section, we empirically verify the computational burden of Modified 3D faster R-CNN, YOLOv4 + linear search, and YOLOv4 + RLS on the same test dataset from this study using two GPUs: 1080 Ti and RTX 8000. Here, YOLOv4 + linear search was evaluated using the model trained with the training dataset. The linear search starts detecting the liver from the first slice to the last slice with the confident score of 0.1 to ensure the detected liver range fully cover the actual liver range.

Based on the results from the experiment, we evaluate the proposed algorithm on 3D bounding box localization and compared its performance to that of several state-of-the-art 3D bounding box liver detectors (see Table 5). The algorithm successfully detects the liver in 223 out of 224 CT volumes. From the table, we can see that LRS achieves the lowest wall distance and centroid distance, yielding scores of 5.4 ± 4.4 mm and 9.7 ± 9.2 mm, respectively; meanwhile it can be seen that the performance of LRS obtained a 3D IoU score of 85.5 ± 9% which is comparable with that of the best reported score amongst the state-of-the-art algorithms (87%).

In addition, the evaluation shows that LRS can detect the 3D bounding box in 0.5 s on average with GPU GTX 1080 Ti and 0.3 s on average with GPU RTX 8000, which are faster than most of other detectors. Note that the total processing time includes both the data loading time from SSD storage to the memories and the inference time. Meanwhile, the computational burdens for Modified 3D faster R-CNN, YOLOv4 + linear search, and YOLOv4 + RLS are not much different and be very small compared to the resources of most of available modern PC nowadays.
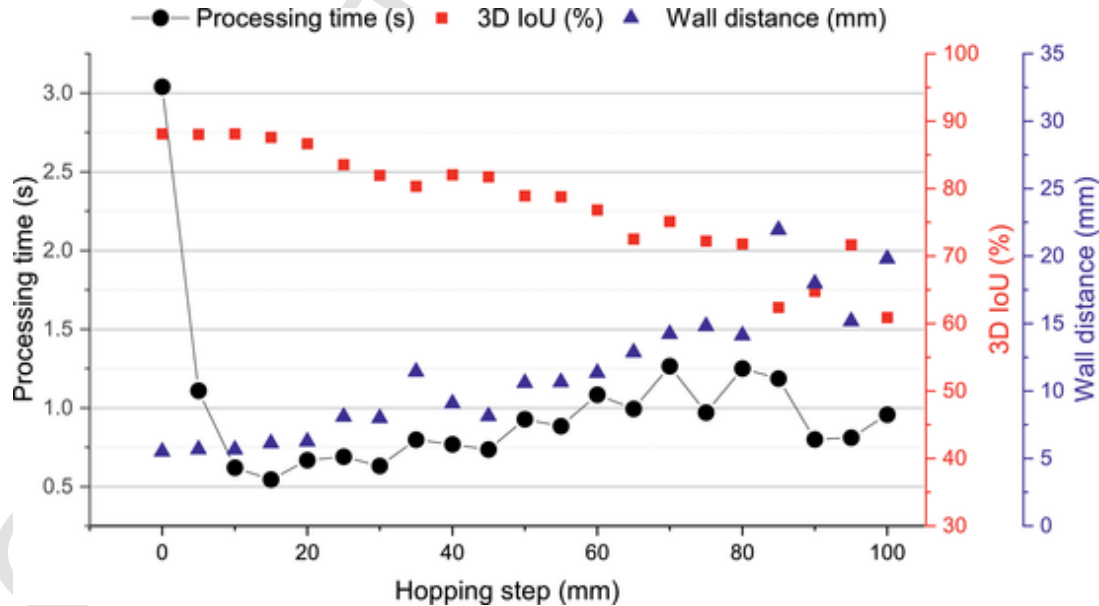

Fig. 6. The effect of varying the number of hopping step HS in LRS algorithm on the accuracy of the liver detection and processing time.

Table 3

Comparison of the speed and accuracy of well-known 2D object detectors on the test dataset (LiTS and EMC datasets). The experiment was carried out using GPU GTX 1080 Ti.

Method	Backbone	Inference time (s)	AP (%)	$AP@50$ (%)	$AP@75$ (%)	AP_S (%)	AP_M (%)	AP_L (%)
SSD (Liu et al., 2016)	VGG-16	0.5	64.4	86.3	73.1	0.3	50.7	69.7
Faster R-CNN (Ren et al., 2017)	ResNet-50	0.07	67.9	89.5	75.7	**28.9**	**64.3**	72.8
YOLOv4 (Bochkovskiy et al., 2020)	CSPDarknet-53	**0.02**	**71.1**	**93.1**	**80.1**	24.4	58.2	**75.9**

Table 4

Comparison of the precision and inference time of the state-of-the-art 2D liver detection methods on CT images. The **RC** of YOLOv4 is computed with the IoU at higher than 50% (Xia and Yin, 2019).

Method	Test dataset	RC (%)	Inference time (s)	GPU
Combined Deep Network (VGG16 + DenseNet) (Xia and Yin, 2019)	IRCAD dataset	91.3	~0.2	GTX 1080 Ti
YOLOv3-arch (Pang et al., 2019)	1000 2D CT images	96.9	~0.008	-
YOLOv4 (this study)	$LiTS$	96.2	~0.02	GTX 1080 Ti
	EMC_{diag}	90.7		
	EMC_{intra}	89.1		

Figure 7 illustrates an example of the liver bounding boxes determined by LRS (the red box), and the Modified 3D faster R-CNN (the blue box) w.r.t. the ground truth of the liver range (the green box). It can be seen that the liver range determined by LRS eliminates the problem of missing liver range at the bottom of the liver which can be seen with the Modified 3D faster R-CNN estimate.

We also evaluate the performance of LRS for 3D liver detection w.r.t. various slice spacings. We choose 46 CT volumes with a spacing smaller than 1 mm from the test dataset and down sample the volumes along the Z-axis by factors from 2 to 5 times the original spacing. We analyze the liver detection accuracy compared to the ground truth. The experimental result is summarized in Table 6. The results show that when reducing the slice spacing, the accuracy of both $3D\ IoU$ and CD does not change statistically significantly, but it statistically significantly affects WD score of the 3D detected bounding box. The reason for the better detection with larger spacing is that for the small spacing, there are more small-liver-slices which have higher rate of miss detection and thus LRS may stop before reaching the extreme end of the liver slices, while this can be skipped with the volume with the larger spacings. Note that the down sampling is to simulate the volume with the spacing up to 5 mm, which is much smaller than the large hopping steps in the experiment in Section 3.4.1 (where the detection accuracy scores remarkably reduce by the effect of the large hopping steps).

Furthermore, we divide the test dataset into two groups, one with a slice spacing equal to or larger than 3 mm (143 volumes), and the other with a slice spacing smaller than 3 mm (70 volumes). We then analyze

the difference in the liver detection accuracy between the two groups. The experimental result is summarized in Table 7. It can be seen that the $3D\ IoU$ and CD scores of the two groups are not statistically significantly different, but the WD scores of the two groups are statistically significantly different ($p = 0.04$). This result is consistent to the result from the previous experiment.

3.4.4. Assessment of liver range detection accuracy and excess radiation dose reduction

Table 8 shows the liver range detection accuracy of LRS. The evaluation results are listed for each phase of the test dataset, because our aim is to provide the second attempt for the liver range detection of the subsequent image in the event that the first attempt of liver detection on the scout volume is not successful. The results show that LRS successfully detects the liver range on 223 out of 224 CT volumes, with the failed one being in the arterial phase. In addition, the liver range error at the lower side is an average of 7.8 mm and an average of 2.1 mm for the upper side. In general, MLL and MLU chosen as 7.8 mm and 2.1 mm, respectively, can compensate for the missing liver range detection by LRS.

Table 9 summarizes the liver excess range for radiographers and the automatic scan range generation on the EMC_{intra} subset and the Hanoi's dataset. The excess ranges were computed on the arterial phase images and the delayed images. For 53 scanning sessions from 46 patients (10 sessions for the EMC_{intra} subset, 29 sessions for the $H108_{diag}$ subset and 14 sessions for the $H108_{post}$ subset), the automatic method failed for one session only in the $H108_{post}$ subset with unsuccessful estimation of the liver in an arterial phase scan. With the EMC_{intra} subset, where the radiographer aimed to scan only the liver during the intervention, the excess scan range of the radiographer and the proposed method are 44.9 mm and 46.2 mm on average, respectively, and there is no statistically significant difference in the performance of the radiographer versus the automatic method. For the $H108$ dataset, the automatic method achieved the similar performance to that of the EMC_{intra} subset while the original scans performed by the radiographers contains much larger excess scan range (194–292 mm on average).

3.4.5. Assessment of excess radiation reduction

Furthermore, based on the CT dose index ($CTDI_{vol}$) reported in the DICOM tag of the $H108_{diag}$ subset, we calculated the Dose Length Prod-

Table 5

Comparison of the state-of-the art 3D bounding box liver localization on CT volume. The experiment is on LRS without adding MLU and MLL values. The upper part of the table (above the thick line) is the list of the reported results from the original papers. The results in the lower part are experimentally carried out in this study using the test dataset and the same hardware.

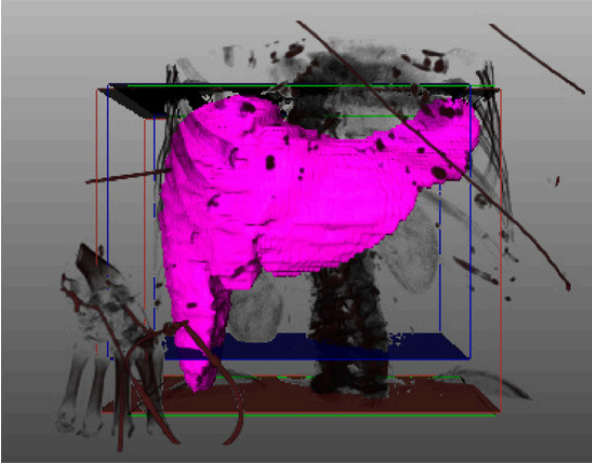| Method	Dataset	WD (mm)	CD (mm)	3D IoU (%)	Average processing time (s)	Hardware	Max VRAM usages (MB)	Max RAM usages (MB)
RRFs (Criminisi et al., 2013)	318 CT scans for training; 82 CT scans for testing	15.7 ± 14.5	-	-	4	CPU	-	-
Cascade of RRFs (Gauriau et al., 2015)	50 CT scans for training; 80 CT scans for testing	10.7 ± 4	-		3.2	CPU	-	-
BoBNet (de Vos et al., 2017)	200 low dose chest CT; 100 cardiac CT angiography; 100 CT abdomen	8.9 ± 15	16.9 ± 11.5	-	6.4	Tesla K40	-	-
Multi-label ConvNet (Humpire-Mamani et al., 2018)	1884 CT scans (60% training; 20% validation; 20% testing)	5.8 ± 12.7	-	-	4	-	-	-
3D triple-branch FCN Xu et al. (2019a)	LiTS dataset 131 for training and validation, 70 for testing	-	-	**87**	3.7	GTX 1080Ti	-	-
CycleGAN + Yolov3 (Hammami et al., 2020)	CT images from 50 patients	6.9 ± 3.4	-	-	8	Tesla V100	-	-
Modified 3D faster R-CNN Xu et al. (2019b)	From this study	9.15 ± 4.1	14.9 ± 11.8	75.9 ± 9.7	1.4 ± 0.2	GTX 1080 Ti	**1228**	3758
					0.9 ± 0.2	RTX 8000		
YOLOv4 + Linear search	From this study	21.6 ± 28	39.9 ± 45.2	63.4 ± 26.7	4.2 ± 4.6	GTX 1080 Ti	1389	3132
					2.9 ± 3.2	RTX 8000		
YOLOv4 + LRS (proposed)	From this study	**5.4** ±4.4	**9.7** ±9.2	85.5 ± 9	0.5 ± 0.2	GTX 1080 Ti	1389	**3087**
					0.3 ±0.1	RTX 8000		

Fig. 7. An example of 3D bounding box liver detections on an intra-interventional CT image of the liver: the green lines illustrate the ground truth of the upper and the lower liver bounding box. The dark blue planes denote the upper and lower extents of the liver bounding box detection by Modified 3D faster R-CNN (Xu et al., 2019b), while the dark red planes show the upper and the lower liver bounding box determined by LRS. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 6

Evaluation on the performance of LRS on CT volume w.r.t. various slice spacings (without adding MLU and MLL). The numbers in the parentheses are the p-values (T-test) compared to the original detections. The experiment was carried out using GPU GTX 1080 Ti.

	Orginal (Spacing)	2×Spacing	3×Spacing	4×Spacing	5×Spacing
Average processing time (s)	0.7 ± 0.2	0.6 ± 0.2	0.5 ± 0.1	0.5 ± 0.1	0.5 ± 0.1
3D IoU (%)	86.3 ± 10.3	86.3 ± 10.5 (0.89)	86.1 ± 10.7 (0.81)	86.7 ± 10.6 (0.58)	86.4 ± 9.2 (0.94)
CD (mm)	8.1 ± 7.7	8.5 ± 8.8 (0.73)	8.9 ± 9.2 (0.36)	8.6 ± 9.0 (0.65)	7.9 ± 6.0 (0.77)
WD (mm)	6.8 ± 5.6	5.7 ± 4.5 (0.05)	5.3 ± 4.2 (0.01)	5.0 ± 4.1 (<0.01)	4.6 ± 3.0 (<0.001)

Table 7

Analysis on the effect of different slice spacing on the accuracy of the 3D liver detection using LRS. The experiment was carried out using GPU GTX 1080 Ti.

Spacing (mm)	Average processing time (s)	3D IoU %	CD (mm)	WD (mm)
≤ 3	0.7 ± 0.2	86.6 ± 8.3	7.9 ± 6.1	6.4 ± 4.7
≥ 3	0.4 ± 0.2	85.8 ± 8.1	9.3 ± 10.0	5.0 ± 2.7
p-**value**	-		0.51	0.13	0.04

uct (DLP) and effective radiation dose via the methods recommended by AAPM (McCollough et al., 2008). The results show that the effective dose resulting from the original scans performed by the radiographer is 17.9 ± 7.4 mSv on average, while that of the range-reduced scans performed by the proposed method is 15.4 ± 6.3 mSv. This result is statistically significantly different (t-test, $p < 0.0001$). The absolute reduction in effective dose is 2.6 ± 1.3 mSv on average while the relative effective dose reduction is 14.5 ± 4.1% on average.

3.4.6. Radiologist assessment

In this experiment, we aim to verify whether the scan range reduction affects the clinical decision, i.e., the status of patients, the lesion characteristics and the treatment strategy for the diagnostic scans and the treatment outcome for the post treatment scans. Three radiologists

Table 8

Qualitative evaluation results of LRS (without adding MLU and MLL values) for liver range detection on the $LiTS$ and EMC datasets.

Dataset	Phase	Acc	Lower side		Upper side	
			MLR (mm)	ER (mm)	MLR (mm)	ER (mm)
$LiTS$	**Portal venous**	99.0%	3.8 ± 4.6	1.8 ± 1.5	1.8 ± 1.5	4.1 ± 4.1
EMC_{diag}	**Non-contrast**	94.7%	8.8 ± 12.6	-	3 ± 0	10.8 ± 12.6
	Arterial	93.9%	10.6 ± 14.6	15.5 ± 8.6	2 ± 1.2	4.6 ± 2.9
	Portal venous	97.4%	7.9 ± 8.8	10.6 ± 11.2	2.8 ± 1.9	7.2 ± 4.8
	Delay	96.4%	10.3 ± 13.9	-	-	4.3 ± 2.2
EMC_{intra}	**Non-contrast**	94.1%	11.4 ± 8.8	9.8 ± 1.1	-	11.7 ± 12.7
	Arterial	91.4%	11.8 ± 13	3 ± 0	2.1 ± 1.3	4.5 ± 3
	Portal venous	95.1%	11.3 ± 13.1	19.5 ± 18.1	1.4 ± 1.4	5 ± 1.7
	Delay	97.9%	2.4 ± 1	5.7 ± 4.7	-	5.7 ± 4.7
All		97.3%	7.8 ± 10.5	15.1 ± 14.9	2.1 ± 1.4	6.5 ± 7.1

Table 9

Comparison of the mean excess range obtained by the radiographers and by the proposed approach.

Dataset	Radiographer's excess range (mm)	Proposed method's excess range (mm)	p-value	Absolute reduction (mm)	Relative reduction (RR)
EMC_{intra}	44.9 ± 14.7	46.2 ± 18.1	~0.81	-	-
$H108_{diag}$	194 ± 82.5	46 ± 25.8	<0.0001	152 ± 83	70.92 ± 21.2%
$H108_{post}$	292 ± 25	42 ± 13.5	<0.0001	250 ± 25	85.21 ± 4.3%

(with three, four and eleven years of experience) from two hospitals independently read both the original and range-reduced CT volumes from the Hanoi's dataset. The radiologists independently concluded that there is no difference in their clinical decisions when using either the original or the range-reduced datasets.

4. Discussion

This study addresses a method that automatically delineates the liver scan range in CT imaging. We evaluate three well-known fast CNN-based detectors for 2D liver detection and evaluate the proposed LRS for liver bounding box detection and liver range detection in a 3D volume using multiphase CT scans from several hospitals. We also compare the performance of the proposed detection method to state-of-the-art liver detectors. Subsequently, we estimate the maximum range of the live motion in subsequent scans using the Gaussian models. To this end, we calculate and compare the excess scan ranges performed by the proposed method and the radiographers, and assess the potential excess effective radiation dose reduction using the proposed method.

The experiments and comparison of the performance of the three well-known object detectors, SSD, Faster R-CNN and YOLOv4 on 2D liver detection (see Table 3), showed that YOLOv4 performed slightly better with an $AP@50$ score of 93% and an inference time of 0.02 s with GPU GTX 1080 Ti. In comparison to the state-of-the-art 2D liver detections (see Table 4), YOLOv4 achieved comparable results with a RC score of 96.2% on the $LiTS$ dataset and around 90% on the EMC dataset. These results are in line with the results reported in several studies using YOLO-based approaches for object detection. The low inference time is an essential factor that enables the detection the liver range sufficiently fast for use in clinical practice.

For the 3D liver bounding box localization (see Table 5), LRS performs the best with a processing time of less than 1 s on average, and outperforms most of other state-of-the-art liver detection methods which require several seconds on average for a 3D liver detection. This processing time is relatively small compared to the time between two CT scans (which is at least 15 s). Note that in multiphase CT imaging of the liver, timing is also critical due to the short period of contrast enhancement in the liver. Consequently, a long processing time may affect the clinical process. The key point of the fast performance of LRS is that it ignores a large amount of the unnecessary slices while other 3D approaches need to load the whole volume to the memory and the GPUs before the inference can be run. However, with the aid of modern GPUs, we think that Modified 3D faster R-CNN is also sufficiently fast for the multiphase CT scanning of the liver applications. In addition, LRS also achieved the smallest WD and CD scores (5.4 ± 4.4 mm and 9.7 ± 9.2 mm on average, respectively), while the $3D\,IoU$ score (85.5%) is comparable to the best state-of-the-art of 3D liver bounding box localization (87%). The marginally lower score for $3D\,IoU$ may be explained by the hopping detection, which may skip the slices of the largest liver size. From Table 5 we also see that YOLOv4 + linear search achieved quite poor performance compared to the other methods. The false positive detections due to the confidence score of 0.1 for the whole volume search causes a larger range of the detected liver than the ground truth range. LRS, which first detects the *anchor* and then extends the liver range, does not have this issue.

Most of the object detectors have an issue in detecting small objects, which may result in missing the upper and lower extent of the liver. In our liver range detection experiment (see Section 3.4.4), the extent of missing liver range for LRS at the lower and upper end are 7.8 mm and 2.1 mm, respectively. This asymmetry can be explained by the fact that the lower liver has a narrow shape while the upper liver has a round shape, resulting in the difference in liver detection performance. We suppose the values also depend on the quality of the datasets and the slice spacing. Nevertheless, we suggest that further studies need to adjust the values for better accuracy; we expect that the difference may not be more than one or two slices of the scan.

Liver motion is one of the major factors that may affect the scan range for the subsequent scans in multiphase CT imaging of the liver. Anatomically, liver motion is primarily driven by respiratory motion. We did not try to estimate respiratory state from one single CT image. As reported in study by Demircioğlu et al. (2021), where scan range of the lung was investigated, a margin of 2 cm was added to the lower lung to deal with the lung motion between the scans without a clear reason. In this study, we estimated the liver motions from the available data, and determined maximum displacements of the upper and the lower liver margin of 15.9 mm and 15.3 mm, respectively. This result is quite similar to numbers previously reported by Demircioğlu et al. (2021) and Tsai et al. (2018). Note that, whereas lung motion drives the liver motion, the liver displacement may differ from the exact lung motion. However, to guarantee the extreme ends of the liver in the subsequent scans are not missed when the patient may breathe more deeply, a larger margin may be implemented.

The evaluation of the excess scan range by both radiographers and the automatic scan range generation in Section 3.4.3 indicated that when the radiologist aims to scan only the liver range, there is no statistically significant difference in the performance of the algorithm and the radiographer. In contrast, the experiment also showed that, for the Hanoi's dataset, when the radiographer did not explicitly aim to scan only the liver range in the subsequent scans, the scan range in practice is much larger than needed, and larger than that would have been the result of using an automatically generated scan range. Additionally, the reduction in the effective radiation dose in a scan session is 14.5% or 2.56 mSv on average, equivalent to approximate $1/8000$ chance of development of fatal cancers based on FDA's reported statistic (FDA, 2018), which is a statistically significant reduction ($p < 0.0001$).

Reducing the scan range automatically should not have an impact on subsequent image interpretation by clinicians. The result from a retrospective comparison of the full images and the automatically reduced scan range images by three radiologists showed that there was no difference in the diagnostic, treatment strategies and post treatment evaluation of the patients with the SIRT liver cancer treatment.

Our study still has some limitations. First, our study proposed an automatic scan range generation for multiphase CT liver imaging and reported the effective radiation dose reduction based on image analysis without investigating the actual performance of the CT scan. A further study may embed the method to the CT system for another practical evaluation. Second, we did not manage to improve the CNN's ability to detect small liver sections accurately. Instead, we added small margins to compensate for potentially missed liver sections. Further studies may develop a specific CNN for small liver detections and obtain better performance in these sections. Otherwise, we suggest that when the detection is not sufficiently accurate, the radiographer can just perform the scan range estimation manually. Finally, in the experimental section (Section 3.2), we determined the parameters of LRS, *anchor* threshold (AT) and searching threshold (ST), as 0.95 and 0.1, respectively. Yet, these values depend on the performance of the core CNN detector and quality of the training dataset, thus other applications may require different parameter settings.

5. Conclusions

In this study, we have proposed and evaluated an novel method for automatically generating scan ranges of the liver in multiphase CT images, and showed that machine-learning techniques are very effective in determining the scan range in multiphase CT scans in clinical practice. The method is based on a CNN model (YOLOv4) for detecting the liver in 2D slices, and LRS for fast liver range detection in a scout volume. Additionally, we estimate the liver motion range in the scan range generation using Gaussian models. Experiments on datasets from several hospitals showed that the liver detection can be performed within one second and the accuracy is comparable to the best state-of-the-art CNN detectors of the liver in CT images. The evaluation on the potential to reduce the scan range showed that the machine-generated scan range is not statistically different to the scan range manually obtained by the radiographer when the imaging focuses on the liver, while the method can significantly reduce the effective radiation dose in each session of scans by 14.5% on average, when the imaging does not only focus on the liver. Additionally, three medical experts concluded that the range reduction in both the arterial and delayed phase does not affect the clinical decisions.

Ethical statement

The local medical research ethics committees decided that the Medical Research Involving Human Subjects Act does not apply to this study. The *LiTS* and *Mayo* datasets are public datasets for research purpose.

CRediT authorship contribution statement

Manh Ha Luu: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Theo van Walsum:** Conceptualization, Methodology, Validation, Formal analysis, Writing – original draft, Writing – review & editing, Supervision. **Hong Son Mai:** Validation, Investigation, Writing – original draft. **Daniel Franklin:** Validation, Formal analysis, Writing – original draft, Writing – review & editing. **Thi Thu Thao Nguyen:** Validation, Formal analysis, Writing – original draft. **Thi My Le:** Validation, Formal

analysis. **Adriaan Moelker:** Conceptualization, Validation, Investigation, Resources, Writing – original draft, Supervision. **Van Khang Le:** Conceptualization, Validation. **Dang Luu Vu:** Conceptualization, Writing – original draft, Supervision. **Ngoc Ha Le:** Conceptualization, Supervision. **Quoc Long Tran:** Validation, Formal analysis, Writing – review & editing. **Duc Trinh Chu:** Validation, Supervision, Funding acquisition. **Nguyen Linh Trung:** Validation, Writing – original draft, Supervision, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgments

References

Bilic, P., Christ, P. F., Vorontsov, E., Chlebus, G., Chen, H., Dou, Q., Fu, C.-W., Han, X., Heng, P.-A., Hesser, J., Kadoury, S., Konopczynski, T., Le, M., Li, C., Li, X., Lipkovà, J., Lowengrub, J., Meine, H., Moltz, J. H., Pal, C., Piraud, M., Qi, X., Qi, J., Rempfler, M., Roth, K., Schenk, A., Sekuboyina, A., Vorontsov, E., Zhou, P., Hülsemeyer, C., Beetz, M., Ettlinger, F., Gruen, F., Kaissis, G., Lohöfer, F., Braren, R., Holch, J., Hofmann, F., Sommer, W., Heinemann, V., Jacobs, C., Mamani, G. E. H., van Ginneken, B., Chartrand, G., Tang, A., Drozdzal, M., Ben-Cohen, A., Klang, E., Amitai, M. M., Konen, E., Greenspan, H., Moreau, J., Hostettler, A., Soler, L., Vivanti, R., Szeskin, A., Lev-Cohain, N., Sosna, J., Joskowicz, L., Menze, B. H., 2019. The liver tumor segmentation benchmark (LiTS). arXiv:1901.04056 [cs]

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y. M., 2020. YOLOv4: optimal speed and accuracy of object detection. arXiv:2004.10934 [cs, eess].

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 68 (6), 394–424. https://doi.org/10.3322/caac.21492.

Brenner, D.J., Hall, E.J., 2007. Computed tomography – an increasing source of radiation exposure. N. Engl. J. Med. 357 (22), 2277–2284. https://doi.org/10.1056/NEJMra072149.

Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., Siddiqui, K., 2013. Regression forests for efficient anatomy detection and localization in computed tomography scans. Med. Image Anal. 17 (8), 1293–1303. https://doi.org/10.1016/j.media.2013.01.001.

Demircioğlu, A., Kim, M.-S., Stein, M.C., Guberina, N., Umutlu, L., Nassenstein, K., 2021. Automatic scan range delimitation in chest CT using deep learning. Radiol. Artif. Intell. 3 (3), e200211. https://doi.org/10.1148/ryai.2021200211.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. ImageNet: a large-scale hierarchical image database, 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255. https://doi.org/10.1109/CVPR.2009.5206848. ISSN: 1063-6919

Devapalasundaram, A., Lau, K.K., Paul, E., Mak, L., Ardley, N., Buchan, K., Kuganesan, A., 2016. Liver detection algorithm: its efficacy for CT noise reduction in the liver. Abdom. Radiol. 41 (3), 493–499. https://doi.org/10.1007/s00261-015-0617-3.

de Vos, B.D., Wolterink, J.M., de Jong, P.A., Leiner, T., Viergever, M.A., Išgum, I., 2017. ConvNet-based localization of anatomical structures in 3-D medical images. IEEE Trans. Med. Imaging 36 (7), 1470–1481. https://doi.org/10.1109/TMI.2017.2673121.

de Vos, B.D., Wolterink, J.M., Jong, P.A.d., Viergever, M.A., Išgum, I., 2016. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks, Medical Imaging 2016: Image Processing. International Society for Optics and Photonics, p. 97841Y. https://doi.org/10.1117/12.2216971.

FDA, 2018. What are the Radiation Risks from CT? The United States Food and Drug Administration. https://www.fda.gov/radiation-emitting-products/medical-x-ray-imaging/what-are-radiation-risks-ct

Gauriau, R., Cuingnet, R., Lesage, D., Bloch, I., 2015. Multi-organ localization with cascaded global-to-local regression and shape prior. Med. Image Anal. 23 (1), 70–83. https://doi.org/10.1016/j.media.2015.04.007.

Goldman, A.R., Maldjian, P.D., 2013. Reducing radiation dose in body CT: apractical approach to optimizing CT protocols. Am. J. Roentgenol. 200 (4), 748–754. https://doi.org/10.2214/AJR.12.10330.

Goodman, G., Kuhns, L., Novak, W., Onitsuka, H., 1979. Liver localization by percussion for CT scanning. Am. J. Roentgenol. 132 (3). https://doi.org/10.2214/ajr.132.3.480. 480–480

Hammami, M., Friboulet, D., Kechichian, R., 2020. Cycle GAN-based data augmentation for multi-organ detection in CT images via Yolo, 2020 IEEE International Conference on Image Processing (ICIP). pp. 390–393. https://doi.org/10.1109/ICIP40778.2020.9191127. ISSN: 2381-8549

He, K., Zhang, X., Ren, S., Sun, J., 2014. Spatial pyramid pooling in deep convolutional networks for visual recognition. arXiv:1406.4729 [cs] vol. 8691, 346–361. 10.1007/978-3-319-10578-9_23

Hoang, H.S., Phuong Pham, C., Franklin, D., van Walsum, T., Ha Luu, M., 2019. An evaluation of CNN-based liver segmentation methods using multi-types of CT abdominal images from multiple medical centers, 2019 19th International Symposium on Communications and Information Technologies (ISCIT). pp. 20–25. https://doi.org/10.1109/ISCIT.2019.8905166. ISSN: 2643-6175

Hosang, J., Benenson, R., Dollár, P., Schiele, B., 2016. What makes for effective detection proposals? IEEE Trans. Pattern Anal. Mach. Intell. 38 (4), 814–830. https://doi.org/10.1109/TPAMI.2015.2465908.

Humpire-Mamani, G.E., Setio, A.A.A., Ginneken, B.v., Jacobs, C., 2017. Organ detection in thorax abdomen CT using multi-label convolutional neural networks, Medical Imaging 2017: Computer-Aided Diagnosis. International Society for Optics and Photonics, p. 1013416.

Humpire-Mamani, G.E., Setio, A.A.A., Ginneken, B.v., Jacobs, C., 2018. Efficient organ localization using multi-label convolutional neural networks in thorax-abdomen CT scans. Phys. Med. Biol. 63 (8), 85003. https://doi.org/10.1088/1361-6560/aab4b3.

Hussain, M.A., Amir-Khalili, A., Hamarneh, G., Abugharbieh, R., 2017. Segmentation-free kidney localization and volume estimation using aggregated orthogonal decision CNNs. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2017. Springer International Publishing, Cham, pp. 612–620. https://doi.org/10.1007/978-3-319-66179-7_70.

Johnson, P.T., Mahesh, M., Fishman, E.K., 2015. Image wisely and choosing wisely: importance of adult body CT protocol design for patient safety, exam quality, and diagnostic efficacy. J. Am. College Radiol. 12 (11), 1185–1190. https://doi.org/10.1016/j.jacr.2015.02.021.

Lin, E.C., 2010. Radiation risk from medical imaging. Mayo Clin. Proc. 85 (12), 1142–1146. https://doi.org/10.4065/mcp.2010.0260.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: common objects in context. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision - ECCV 2014. Springer International Publishing, Cham, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48.

Lip-Pauwels, W., Dijkshoorn, M. L., Booij, R., Willemssen, F., Krestin, G. P., 2012. Optimizing 3-phase liver CT in hepatocellular carcinoma: achieve better lesion depiction and reduce radiation dose and contrast material.Publisher: European Congress of Radiology - ECR 2012, https://epos.myesr.org/poster/esr/ecr2012/C-0092.

Liu, S., Qi, L., Qin, H., Shi, J., Jia, J., 2018. Path aggregation network for instance segmentation, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8759–8768. https://doi.org/10.1109/CVPR.2018.00913. ISSN: 2575-7075

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A. C., 2016. SSD: Single shot multibox detector. arXiv:1512.02325 [cs] vol. 9905, 21–37. 10.1007/978-3-319-46448-0_2

Luu, H.M., Moelker, A., Klein, S., Niessen, W., van Walsum, T., 2018. Quantification of nonrigid liver deformation in radiofrequency ablation interventions using image registration. Phys. Med. Biol. 63 (17), 175005. https://doi.org/10.1088/1361-6560/aad706.

Luu, H.M., Walsum, T.v., Franklin, D., Pham, P.C., Vu, L.D., Moelker, A., Staring, M., VanHoang, X., Niessen, W., Trung, N.L., 2021. Efficiently compressing 3D medical images for teleinterventions via CNNs and anisotropic diffusion. Med. Phys. 48 (6), 2877–2890. https://doi.org/10.1002/mp.14814.

Mai, H.S., Ha, L.N., Bang, M.H., Bae, S., Giang, D.T., Thinh, N.T., Paeng, J.C., 2021. Diagnostic and prognostic value of 99m Tc-MAA SPECT/CT for treatment planning of 90 y-resin microsphere radioembolization for hepatocellular carcinoma: comparison with planar image. Sci. Rep. 11 (1), 3207. https://doi.org/10.1038/s41598-021-82887-w.

McCollough, C., Cody, D., Edyvean, S., Geise, R., Gould, B., Keat, N., Huda, W., Judy, P., Kalender, W., McNitt-Gray, M., Morin, R., Payne, T., Stern, S., Rothenberg, L., 2008. The Measurement, Reporting, and Management of Radiation Dose in CT. Technical Report. AAPM. https://doi.org/10.37206/97.

McCollough, C., Leng, S., 2020. Use of artificial intelligence in computed tomography dose optimisation. Ann. ICRP 49 (1_suppl), 113–125.

McCollough, C.H., Bartley, A.C., Carter, R.E., Chen, B., Drees, T.A., Edwards, P., Holmes, D.R., Huang, A.E., Khan, F., Leng, S., McMillan, K.L., Michalak, G.J., Nunez, K.M., Yu, L., Fletcher, J.G., 2017. Low-dose CT for the detection and classification of metastatic liver lesions: results of the 2016 low dose CT grand challenge. Med. Phys. 44 (10), e339–e352. https://doi.org/10.1002/mp.12345.

McGlynn, K.A., London, W.T., 2011. The global epidemiology of hepatocellular carcinoma: present and future. Clin. Liver Dis. 15 (2), 223–243. https://doi.org/10.1016/j.cld.2011.03.006.

Navarro, F., Sekuboyina, A., Waldmannstetter, D., Peeken, J.C., Combs, S.E., Menze, B.H., 2020. Deep reinforcement learning for organ localization in CT, Medical Imaging with Deep Learning. PMLR, pp. 544–554. ISSN: 2640-3498 http://proceedings.mlr.press/v121/navarro20a.html

Pang, S., Ding, T., Qiao, S., Meng, F., Wang, S., Li, P., Wang, X., 2019. A novel YOLOv3-arch model for identifying cholelithiasis and classifying gallstones on CT images. PLoS ONE 14 (6), e0217647. https://doi.org/10.1371/journal.pone.0217647. Publisher: Public Library of Science

Raman, S.P., Mahesh, M., Blasko, R.V., Fishman, E.K., 2013. CT Scan parameters and radiation dose: practical advice for radiologists. J. Am. Coll. Radi. 10 (11), 840–846. https://doi.org/10.1016/j.jacr.2013.05.032. Publisher: Elsevier

Redmon, J., Farhadi, A., 2018. YOLOv3: an incremental improvement. arXiv:1804.02767 [cs].

Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 39 (6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

Samarakoon, P.N., Promayon, E., Fouard, C., 2017. Light random regression forests for automatic multi-organ localization in CT images, 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). pp. 371–374. https://doi.org/10.1109/ISBI.2017.7950540. ISSN: 1945-8452

Shao, Y.-H., Tsai, K., Kim, S., Wu, Y.-J., Demissie, K., 2020. Exposure to tomographic scans and cancer risks. JNCI Cancer Spect. 4 (1). https://doi.org/10.1093/jncics/pkz072.

Jimenez-del Toro, O., Müller, H., Krenn, M., Gruenberg, K., Taha, A.A., Winterstein, M., Eggel, I., Foncubierta-Rodríguez, A., Goksel, O., Jakab, A., Kontokotsios, G., Langs, G., Menze, B.H., Salas Fernandez, T., Schaer, R., Walleyo, A., Weber, M.-A., Dicente Cid, Y., Gass, T., Heinrich, M., Jia, F., Kahl, F., Kechichian, R., Mai, D., Spanier, A.B., Vincent, G., Wang, C., Wyeth, D., Hanbury, A., 2016. Cloud-based evaluation of anatomical structure segmentation and landmark detection algorithms: VISCERAL anatomy benchmarks. IEEE Trans. Med. Imaging 35 (11), 2459–2475. https://doi.org/10.1109/TMI.2016.2578680.

Tsai, Y.-L., Wu, C.-J., Shaw, S., Yu, P.-C., Nien, H.-H., Lui, L.T., 2018. Quantitative analysis of respiration-induced motion of each liver segment with helical computed tomography and 4-dimensional computed tomography. Radiat. Oncol. 13 (1), 59. https://doi.org/10.1186/s13014-018-1007-0.

Xia, K., Yin, H., 2019. Liver detection algorithm based on an improved deep network combined with edge perception. IEEE Access 7, 175135–175142. https://doi.org/10.1109/ACCESS.2019.2953517.

Xu, X., Zhou, F., Liu, B., Bai, X., 2019. Multiple organ localization in CT image using triple-branch fully convolutional networks. IEEE Access 7, 98083–98093. https://doi.org/10.1109/ACCESS.2019.2930417.

Xu, X., Zhou, F., Liu, B., Fu, D., Bai, X., 2019. Efficient multiple organ localization in CT image using 3D region proposal network. IEEE Trans. Med. Imaging 38 (8),

1885–1898. https://doi.org/10.1109/TMI.2019.2894854.

Zanca, F., Demeter, M., Oyen, R., Bosmans, H., 2012. Excess radiation and organ dose in chest and abdominal CT due to CT acquisition beyond expected anatomical boundaries. Eur. Radiol. 22 (4), 779–788. https://doi.org/10.1007/s00330-011-2332-y.

Zhan, Y., Zhou, X.S., Peng, Z., Krishnan, A., 2008. Active scheduling of organ detection and segmentation in whole-body medical images. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008. Springer, Berlin, Heidelberg, pp. 313–321. https://doi.org/10.1007/978-3-540-85988-8_38.

Zhang, W., Mantlic, F., Zhou, S.K., 2010. Automatic landmark detection and scan range delimitation for topogram images using hierarchical network, Medical Imaging 2010: Image Processing, vol. 7623. International Society for Optics and Photonics, p. 762311.

Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D., 2007. Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features, 2007 IEEE 11th International Conference on Computer Vision. pp. 1–8. https://doi.org/10.1109/ICCV.2007.4408925. ISSN: 2380-7504

Zheng, Y., Georgescu, B., Comaniciu, D., 2009. Marginal space learning for efficient detection of 2D/3D anatomical structures in medical images. In: Prince, J.L., Pham, D.L., Myers, K.J. (Eds.), Information Processing in Medical Imaging. Springer, Berlin, Heidelberg, pp. 411–422. https://doi.org/10.1007/978-3-642-02498-6_34.

Zhou, X., Kojima, T., Wang, S., Zhou, X., Hara, T., Nozaki, T., Matsusako, M., Fujita, H., 2019. Automatic anatomy partitioning of the torso region on CT images by using a deep convolutional network with majority voting, Medical Imaging 2019: Computer-Aided Diagnosis. International Society for Optics and Photonics, p. 109500Z.

Zhou, X., Wang, S., Chen, H., Hara, T., Yokoyama, R., Kanematsu, M., Fujita, H., 2012. Automatic localization of solid organs on 3D CT images by a collaborative majority voting decision based on ensemble learning. Comput. Med. Imaging Graphics 36 (4), 304–313. https://doi.org/10.1016/j.compmedimag.2011.12.004.

Zinsser, D., Maurer, M., Do, P.-L., Weiß, J., Notohamiprodjo, M., Bamberg, F., Othman, A.E., 2019. Reduced scan range abdominopelvic CT in patients with suspected acute appendicitis - impact on diagnostic accuracy and effective radiation dose. BMC Med. Imaging 19 (1), 4. https://doi.org/10.1186/s12880-019-0304-x.