

Contents

1 Segmentation of Liver Tissue in computed tomography	2
1.1 Introduction	2
1.2 Publicly available datasets & open challenges	3
2 Classical semantic segmentation methods	5
2.1 Image Intensity based methods	6
2.2 Region Growing	8
2.3 Geometric deformable models	10
2.4 Graph-theory based methods	14
2.5 Probabilistic Atlases	16
2.6 Statistical Shape Models	18
2.7 Hybrid methods	22
3 Deep Learning based methods	25
3.1 General Deep Imaging introduction	25
3.2 State-of-the-art DL implementations	26
3.2.1 FCN-based architectures	26
3.2.2 Cascaded architectures	28
3.2.3 Input type (2D, 2.5D, 3D, patches, multiphase)	29
3.2.4 Training strategies (pre-processing, pre-trained networks...)	30
3.2.5 Inference schemes (post-processing, ensemble learning)	31
3.2.6 Conclusion	32
4 Semantic segmentation applied to the study of HCC	33
4.1 Motivations	33
4.2 Datasets	33
4.3 Cascaded architecture	35
4.4 U-Net network	36
4.5 Multiphase information	38
4.6 Experiments	38
4.6.1 Data pre-processing	38
4.6.2 Training	40
4.6.3 Experiments and results	41
4.6.4 Conclusions	43

1 Segmentation of Liver Tissue in computed tomography

1.1 Introduction

As explained in the **Medical Context** section, the liver is an organ with a key function in the human body. Being a soft organ, it presents a large variation in shape among the population (inter-patient variation), and can also be subject to deformations due to respiratory and heart motions in case of multiphasic acquisition after injection of contrast agents (intra-patient variation).

On computed tomography images, the intensities of the liver itself and those of nearby organs (heart, spleen, stomach) are very similar. The images are often acquired after the injection of contrast agents to offer critical information (e.g. *wash-in/wash-out*). Nonetheless, the acquisition protocol may differ among the institutions (type of contrast agent, moments of injection, injection rate, ...) and the results often differ with the patient characteristics (e.g. blood flow). Moreover, the liver can easily be affected by diseases such as cirrhosis, and is often the site of either primary or secondary cancers with varying contrast levels (hyper-/hypo-intense tumors, peritumoral enhancement) and shapes (non-smooth margins, hypoattenuating halos, ...). All these reasons cause the automatic or semi-automatic segmentation of the liver and its structures to be challenging.

The automatic segmentation of the tumors brings by definition a volumetric information, which corresponds to a better predictor when compared to the diameter only, according to the *Response Evaluation Criteria in Solid Tumor* (RECIST) standard [1, 2]. A precise volumetric segmentation of both the liver and its tumors is a prerequisite in case of treatment planning (*TARE*, *TACE*, *percutaneous thermal ablation*, *radiotherapy surgical resection*, ...) [3–6]. Associating liver and tumors segmentation allows the computation of the *tumor burden*, which has an importance in case of surgical resection, or when estimating the efficiency of a given treatment [7–14].

Although not corresponding to a true gold standard [15] and often requiring several experts to segment the different structures for reducing the operator-dependant bias [16, 17], relying on human expert delineations currently corresponds to the most common way to assess segmentation methods [18].

In order to compare digitally obtained segmentations with the expert outputs, different metrics have been implemented over time (see “*Semantic Segmentation Metrics*” in Appendix). However, as explained above, manually segmenting the liver and its different structures is a tedious, time-consuming (up to 90 minutes per patient [19]) and operator-dependent task, as seen in the figure 1. Furthermore, the constitution of an open database needs to respect several ethical aspects, and the patient data privacy. This is the main reason why only a small number of samples have been publicly released over time.

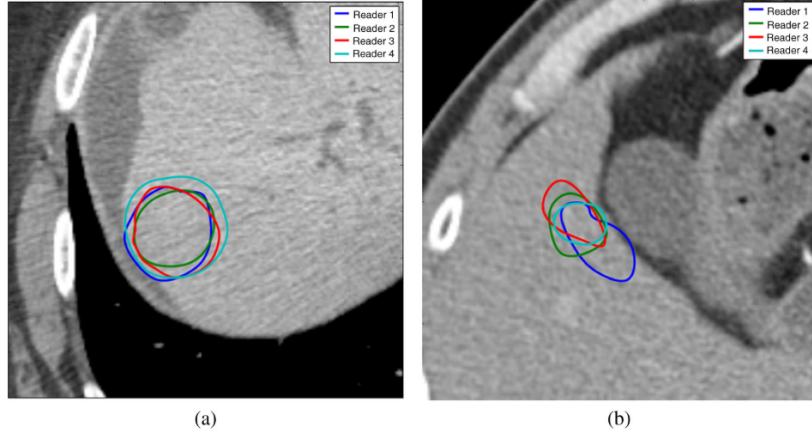


Fig. 2 Two samples of 14 slices selected for segmentation by all four readers. Each closed boundary represents a different reader's segmentation. (a) Relatively high overlap (54%) and (b) relatively low overlap (23%).

Figure 1: Interobserver variability as illustrated by ©Echegaray et al. [16]

1.2 Publicly available datasets & open challenges

We are first covering the different publicly available datasets as long as the different challenges launched to increase interest in liver semantic segmentation. We are then investigating the historical methods developed to perform the semantic segmentation of the liver and/or liver tumors, before presenting most recent advances brought by deep learning. The “classical” methods were often guided by manual interactions, and commonly implemented engineered features whereas deep learning allowed relevant features to be learned directly from the input data.

Up to today, and compared to other organs (brain, breast, lung) [20], only a small number of computed tomography images datasets have been publicly available for liver tissue segmentation purposes. A table presenting details of the publicly available datasets can be found in the table 1.

Table 1: Publicly available datasets

Name	#Cases	Liver GT	Tumor GT	Tumor Type	Other GT	#Experts	Volume size (pixels)	Voxel size (mm)	Challenge	Hidden Cases	Phases
3Dircad-01	20	true	true	Undefined	Blood vessels	1	axial: 512x512 z : 74 - 260	axial:0.56 - 0.87 z: 1 - 4	-	-	Undefined
3Dircad-02	2	true	true	FNH	-	1	axial: 512x512 z: 167 - 219	axial: 0.96 z: 1.8 - 2.4	-	-	1 AR 1 PV
Sliver07	30	true	true	Undefined	-	1	-	axial: 0.55 - 0.8 z: 1 - 3	MICCAI07	10	PV
MIDAS	4	false	true	Undefined	-	5	-	-	-	-	Undefined
LITS	131	true	true	Undefined	No	3	axial: 512x512 z: 42 - 1026	axial:0.56 - 1 z: 0.45 - 6	MICCAI17 ISBI 2017	70	Undefined
TCIA	97	false	false	HCC	No	-	axial:512x512 z: 26 - 192	axial:0.58 - 0.98 z: 0.8 - 5	-	-	mixed
CHAOS	40	true	false	Healthy Livers	No	3	axial: 512x512 z: 77 - 105	axial: 0.7 - 0.8 z: 3 - 3.2	ISBI 2019	-	PV
ImageClef Liver	50	true	true	(Bounding Box)	No	-	x: 190- 308 y: 213-238 z: 41 - 588	axial: 0.68 - 1.01 z: 0.40 - 2.5	ImageClef 2015	10	Undefined

Some of the available datasets have been launched by medical-related institutions. IRCAD for example, provided both 3DIrcadb sets [21], with the first containing 20 abdominal CT scans, where 15 of the patients suffered from liver tumor, and the second containing only 2 CT scans with patients suffering from FNH (*Focal Nodular Hyperplasia*). The *National library of Medicine* (NLM) offered the MIDAS dataset containing annotated pathological CT scans from 4 patients [22] whereas the cancer imaging archive (TCIA) proposed a set of 97 unlabeled volumes from patients suffering from hepatocellular carcinoma (HCC) [23].

Thanks to the different segmentation challenges launched throughout the time, some other sets have been provided and offered real benchmarks. The *Segmentation of the Liver* (Sliver'07) was organized in conjunction with MICCAI 2007 and offered 40 CT volumes [24], where 30 volumes were available during the challenge, and 10 hidden cases were used to evaluate the proposed methods. In 2015 the Image Clef challenge offered 50 CT scans with expert annotations [25], whereas 201 volumes have been provided by the organizers of both ISBI and MICCAI in 2017 for the *Liver Tumor Segmentation* (LITS) benchmark [18]. In 2019 the organizer of ISBI further granted access to 40 cases for the *Combined (CT-MR) Healthy Abdominal Organ Segmentation* (CHAOS) [26].

In these different datasets, the volumes were annotated by one or multiple experts, with predominantly pixel-wise delineations of the areas of interest (the ImageClef dataset is the only one with bounding box annotations for the lesions). Only one annotation volume per case is given which is designed to translate the consensus among the experts, however, the inter-expert variations can not be correctly caught by this format. Moreover, most of the available databases present heterogeneous data in terms of geometry (number of slices per patient, volume size, resolution), pathology (the number of lesions per volume in the LITS dataset goes from 1 to more than 70 as depicted in the figure 2) and regarding the acquisition protocols (e.g. the acquisition phase is undefined for the LITS data where arterial and portal venous volumes are mixed up).

The limited access to data and the heterogeneity present in the publicly available datasets render thus the liver tissue segmentation task very challenging.

2 Classical semantic segmentation methods

The first methods that have been used for the semantic segmentation of liver tissue relied mainly on traditional computer vision and machine learning methods. In most of the cases, prior anatomical knowledge about the intensity, the shape or the position of the liver is incorporated in the process.

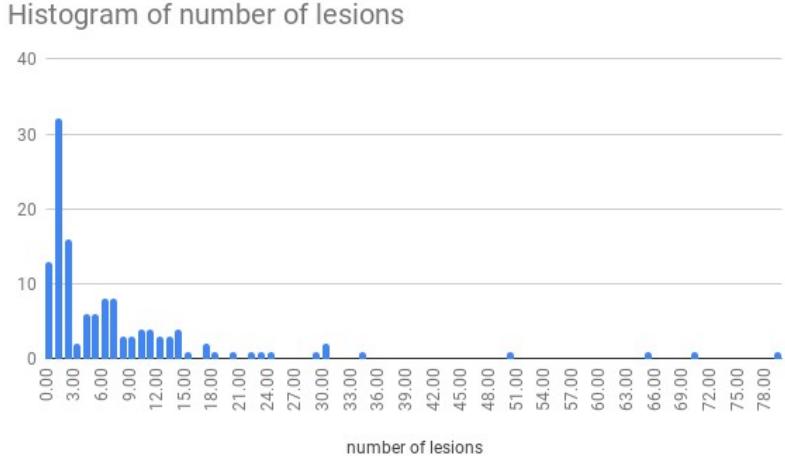


Figure 2: Histogram of the number of lesions per patient in the LITS dataset

The different methods are supported by several pre- and post-processing steps, and may require one or multiple interactions with experts. Among the most utilized methods, we will first present those based on the intensity of the voxels, before mentioning the ones requiring manual interactions such as region growing or graph-theory based methods. We will then introduce the geometric deformable models that can be exploited to add local constraints, before introducing methods requiring prior shape knowledge such as probabilistic atlases (PA) or statistical shape models (SSM). Finally, we will mention the different studies that are combining the aforementioned methods.

2.1 Image Intensity based methods

Lim et al. combined prior-knowledge such as the location of the liver and its intensity distribution with active contours to segment the liver [27–29]. The rough estimation is determined via thresholding or morphological operations in a multiscale fashion. K-means clustering is applied to remove unwanted tissues from the candidate area. Active contour, as described later, is then implemented to refine the liver boundary by first detecting the boundary between liver and adjacent organs. The segmentation pipeline is depicted in the figure 3.

Lee et al. [30], proposed a multimodal contextual neural network using pixel intensity and neighborhood dynamic to get a rough liver segmentation. The refinement step is performed using fuzzy rules based on prior knowledge such as location or textural properties. Liu et al. [31], developed a method to segment the liver by first detecting the potential contour implementing a Canny Edge detector algorithm, modified by the template liver position obtained via thresholding. A Gradient Vector Flow (GVF) field computation is used by the snake algorithm to obtain an accurate liver boundary detection. The

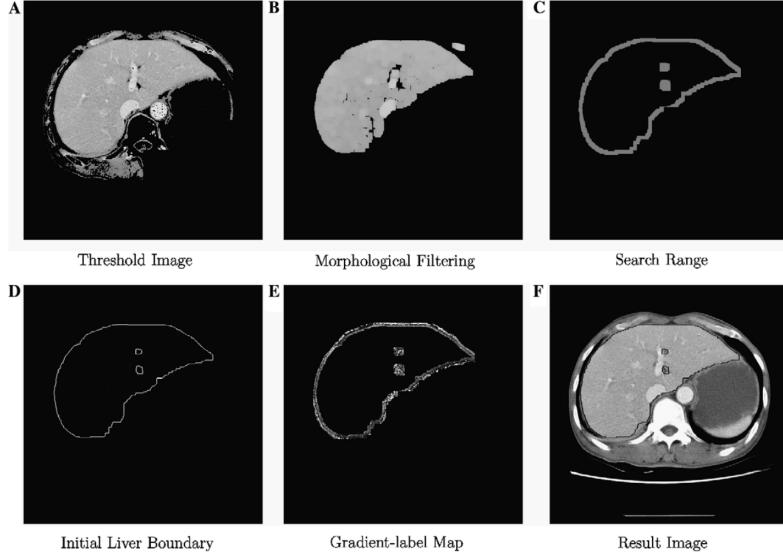


Fig. 14. Experiment of the patient 1.

Figure 3: Intensity based segmentation workflow used by ©Lim et al. [29]

final segmentation is obtained semi-automatically by asking the user to select the slices where the segmentation gave accurate results, and propagate this segmentation iteratively to adjacent slices. An example is depicted in the figure 4.

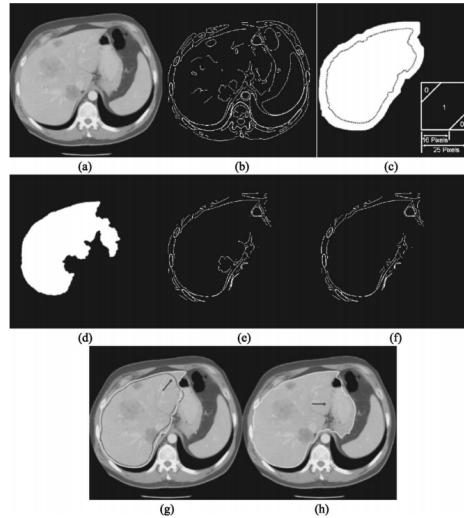


FIG. 6. Illustration of the segmentation of the liver volume. (a) Original image to be segmented. (b) Initial edge map. (c) Mask obtained by dilating the result from the segmented liver (black dotted contour) in the previous slice using the structure element shown at the right lower corner. (d) The liver template. (e) Edge map after modification by the mask and the liver template. Some edges of the lesion and fissure remain inside the liver. (f) Final edge map after modification by the concavity removal algorithm. There is no edge within the liver. (g) The original image overlapped with the final contour (white solid). It also shows the initial contour refined by the snake algorithm (arrow). (h) The GVF snake may leak into stomach if no mask is applied (arrow).

Figure 4: Liver segmentation as described by ©Liu et al. [31]

Seo and Park [32] combined thresholding techniques and morphological operations to obtain a final segmentation. Specific filters are firstly used to smooth the distribution and reduce the noise. Adaptive multi-modal threshold is then applied to find the intensity range of the contrast enhanced liver. The left partial histogram threshold (LPTH) algorithm is then implemented to detect the liver and remove neighboring organs, before morphological operations are performed to refine the candidate ROI. Kim et al. [33], obtained a rough segmentation of the liver by comparing the gray level dynamic of the test

patient with gray-level distributions obtained from liver ROIs in the training dataset. The ROI is refined using watershed transform followed by anatomical constraints such as the presence of smooth boundaries. Campadelli et al. [34], proposed a new fully automatic liver segmentation technique. They first constrained the research by considering only the area below the heart. Different organs are afterwards classified, in particular the liver which corresponds to the biggest connected component found between the different organs. The liver region is further refined using a region growing method to ensure 3D consistency. Foruzan et al. [35], computed the liver intensity range using expected-maximisation (EM) algorithm, and applied thresholding operation in a bounding box constrained by the position of the ribs and the heart as illustrated in figure 2.1. The final candidate region of interest is then refined by anatomical constraints such as liver anatomy and assumed intensity range.

Fig. 3 Typical ROI estimation based on rib bones in different slices. **a** #4, no. 55, **b** #19, no. 49

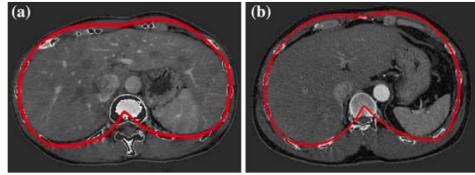


Figure 5: Estimation of the liver ROI based on the anatomical rib position as described by ©Foruzan et al. [35]

Other studies decided to first rely on the gray level distribution of the liver in the training images to get an initial segmentation during the inference, before applying a refinement step consisting on active contours, region growing, or morphological techniques with anatomical constraints.

2.2 Region Growing

Contrary to many other methods, region growing strategies often do not require prior knowledge to perform a segmentation. This set of methods is based on the placement of an initial seed point/region and evolves based on the image properties. As an example, Beck and Aurich [36], proposed to manually place seed points iteratively until the entire liver is segmented by 3-D filling using non-linear interpolation rules. Qi et al. [37], tackled the lesion segmentation problem by using a region growing strategy with manually placed seeds. The lesion is modeled as a bag of Gaussians. A new voxel is added to the region by comparing its intensity and neighborhood to each of the lesions Gaussian through the Bhattacharyya distance. Since manual placement of seeds still suffers from classical problems such as inter and intra-operator variability, some studies decided to automatically define seed points to start the algorithm with.

Rusko et al. decided to define the original seed region by considering a set of rules such as the intensity

range often occupied by the liver or its anatomical position relative to the heart [38, 39]. Susomboon et al. [40] decided to partition the abdominal region via intensity-based EM algorithm to find the distribution of soft tissues in the whole volume. The liver region used as seed was further isolated using quad-tree decomposition and textural features. Kumar et al. [41], proposed a region growing method to segment the liver where the initial seed point corresponds to the centroid of a rough liver region obtained based on several anatomical and intensity constraints. To ensure that the centroid does not lie on lesions or dark objects, its intensity must be in the liver likelihood range. The process is depicted in the figure 6.

Fig. 5 Liver and lesion segmentation process. **a** Original image, **b** simplified image, **c** eroded image (1st erosion), **d** cross indicates the centroid (seed point) on the eroded image, **e** region grown liver, **f** segmented liver after postprocessing, **g** AFM clustered lesion, **h** segmented lesion, **i** segmented liver and lesion outlined in the original image

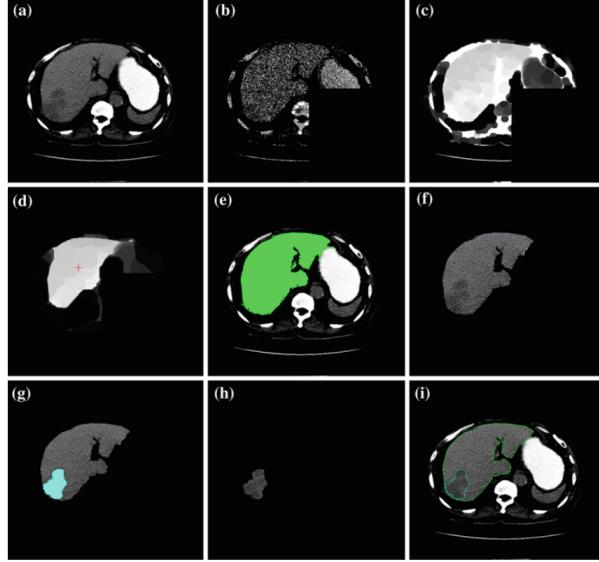


Figure 6: Liver segmentation pipeline after the placement of a seed at the estimated liver centroid position as detailed by ©Kumar et al. [41]

Geometrical rules or local intensity properties often guide the growth of the region. In Rusko et al. Susomboon et al., Kumar et al. the intensity in a given n-voxels-sized neighborhood is analyzed to decide whether or not a point is added to the region [38, 40, 41]. Phole and Toennies decided to iteratively test the homogeneity of the region after adding candidate points [42]. In Beck and Aurich, the filling step is controlled by a nonlinear coupling criterion, which computes the weighted intensity difference between a seed point and its neighborhood. Weights are controlled by a non-normalized Gaussian function applied on both the distance to the seed point and on intensity difference [36]. However region growing methods have the disadvantage to poorly handle pathological livers, and often fail due to leakage in presence of weak boundaries.

2.3 Geometric deformable models

Both geometric deformable models (aka active contours) and level set based methods rely on boundary tracking to perform the segmentation.

The basic idea of *active contours* (introduced in 1988) is to come from an initial contour and to deform it until it reaches the contour of the object. The deformation function reaches its minimum when the boundary of the object is found. The function used to deform the contour is an energy function that tries to control the smoothness of the curve and that is attracted by the object boundaries.

One of the main drawbacks of this method is that the final curve will always have the same topology as the initial one, meaning for example that multiple objects can't be detected. To overcome those limitations, new models, based on the theory of curve evolution and geometric flows, were designed. Caselles et al. proposed a formulation where finding the contour will be equivalent to finding a geodesic curve of minimal weighted length, thus the formulation of *geodesic active contours* [43].

In the active contours formulation, a parametric characterization of the boundary is utilized, whereas in the level set paradigm, they are embedded as a time-dependent partial differential equation (PDE). Osher and Fedkiw proposed a definition, where the boundary of an object can be regarded as a zero-level set of a time-dependent function which is evolving based on the speed at which the contour evolves along its normal direction [44]. The negative values correspond to voxels belonging to the liver regions, and positive values to the outside. This paradigm was first implemented by Pan and Dawant, who used the level set technique for the automatic segmentation of liver in CT scans [45]. They realized that the classical formulation can't accurately handle cases where weak boundaries are present in noisy or non-uniform images. Therefore, they decided to use an accumulative curvature-dependent speed function that would depend on the front past history. In order to be compliant with abdominal CT scans segmentation, a priori anatomical knowledge was added to the process by considering the distance from the liver to the skin and the ribs. Initial detection of the skin and the ribs is illustrated in the figure 7.

Wang et al. combined probabilistic atlas (PA) and level set to perform the segmentation of the liver [46]. Training volumes were pre-processed before the liver region was segmented, and stored with the denoised CT volumes. During the inference phase, the different training denoised CT volumes are used to compute the weighted PA for the current test patient. The liver ROI is then extracted using the weighted PA, and improved by estimating 5 different regions (heart, liver, right kidney, spleen and bone). Finally, a combined shape-intensity model is generated from the training samples, and used to refine the mask by level-set based segmentation.

Garamendi et al. proposed a method for segmenting the liver based on a geometric level set method [47].

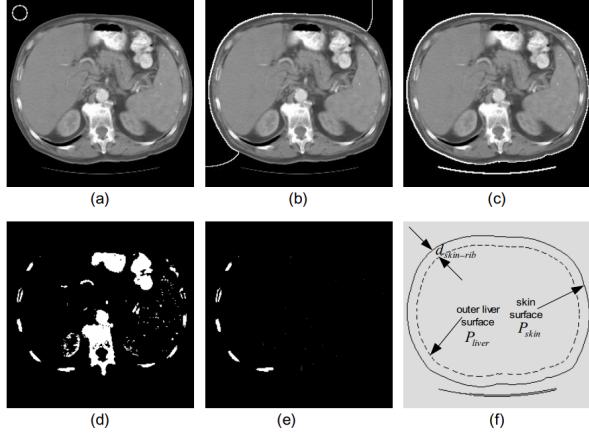


Figure 5. Localization of the skin surface in CT image (panels (a)-(c)) and extraction of the ribs (panels (d)-(e)).

Figure 7: Skin surface localization followed by the extraction of the ribs as described by ©Pan et al. [45]

After manually placing a circle in the liver region to segment, different candidates ROI were iteratively obtained by computed gray-level properties inside and outside the given region, and shifting the boundary by looking at local information around the border pixels. This procedure is repeated until the border stabilizes. Chi et al. proposed a segmentation technique based on active contour [48]. They placed the initial boundary by analyzing anatomical properties of the patient such as the position of the rib cage, the heart and the lungs. K-means clustering is added to reduce the variation present in the candidate surface by detecting blood vessels, remaining heart tissues, the liver, the kidney and the tumors. Finally, active contour is used based on GVF (*gradient vector flow*) combined with distance transform formulation which detects the liver boundary based on the boundary found in the previous slice. Furukawa et al. combined rough segmentation based on MAP estimation and level set method to achieve liver segmentation [49]. To perform the rough segmentation, the lungs are initially segmented and used to normalize the different images. The liver as long as 3 other components (right-kidney, heart and other tissues/organs or muscles) are segmented using MAP with a PA as reference. Active contours are used to refine the segmentation by re-using the definition provided by Catelles et al. and adding a term based on the distance to the human body contours.

Lee et al., computed a speed image through a fast marching algorithm, by looking at gradient magnitude in the original image [50]. A rough segmentation is retrieved by performing a 2.5D propagation after placing seed points on the top and the bottom of the liver. A level set method is implemented to refine the obtained area. The evolution of the level-set based contour is controlled by the initial speed image values and by the curvature. Vessels near the liver boundary can be added to the final map using a rolling ball algorithm. Wimmer et al. semi-automatically segmented the liver by extending 2D cross-sectional annotations, in 3D using RBF and applied a level set technique to refine the obtained

volume [51]. The initial area is obtained by manually performing 2 cross-sectional segmentations per axis, and extending them in 3D via RBF after sampling N points along the given contours, as depicted in the figure 8.

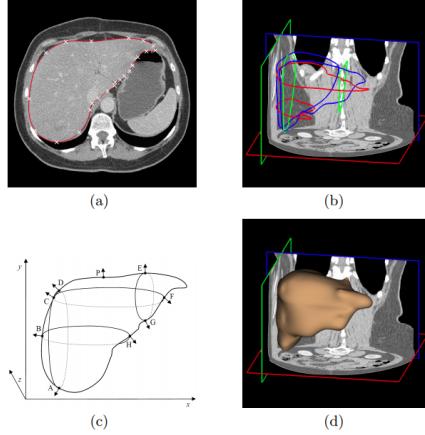


Fig. 2. (a) Manual delineation of a liver cross-section; spline control points are denoted with "x". (b) Six cross-sectional contours are defined, two for each orthogonal view orientation. (c) Visualization of calculated 3-D normals; at the points A – H, the planar normal of the contour defined in the x/y plane is extended to 3-D using the z -components of normals of intersecting contours defined in x/z and y/z planes, respectively; the z -component of the normal at P is determined by linear interpolation between the z -components at D and E according to the arc length distances from D to P and from P to E. (d) Rendering of the interpolated surface.

Figure 8: Semi automatic segmentation pipeline as proposed by ©Wimmer et al. [51]

The level set reconstruction surface was then used to refine the obtained shape. The function used takes into account the points initially present after manual segmentation. This method however suffers from errors when an initial surface includes high-contrast boundaries not belonging to the liver, because they will tend to attract the surface during the deformation process. Later, Wimmer et al. used an implicit active shape model to perform the segmentation of the liver [52]. To build the active shape model, one segmentation is used as reference, and the rest of the training segmentations are aligned to it. To convert the obtained masks into an implicit representation, a signed distance map is computed for each mask where grid points are assigned positive or negative Euclidean distance to the boundary. The shape model is expressed based on the principal component analysis of the signed distance maps. A level set function is formulated based on the given shape model. To build the boundary, all grid points in a narrow-band around the zero-level set are considered, and their boundary profile is computed. A probability is assigned to each point by considering the learned appearance model. Candidates points are projected back to the zero level set, and the points with the highest probability are kept. The obtained probability map is used in GVF (*Gradient Vector Flow*) to get the best contour. Massoptier and Casciaro used a patch-based intensity approach to distinguish hepatic tissues from other abdominal organs, and refine the obtained boundary by implementing active contour with GVF [53]. The liver segmentation was followed by a liver lesions segmentation via k-means clustering. Wimmer et

al. built later a shape model based on nonparametric density estimates [54]. The appearance model is obtained by computing profiles for points in a narrow band along the zero-level set boundary. To evaluate the boundary probability for the points present in the narrow-band, they used a kNN classifier. A region model built upon a cascade of classifiers is implemented to complete the boundary model, and prevent it to stop at local extremum. Finally, a shape model based on Parzen density estimation constrains the level set evolution. Song et al. pre-segmented the liver using an edge detector, and refined the segmentation with a curvature-based level set algorithm [55]. Three B-splines surface models are built knowing the position of the lung, and assuming that the bottom of the liver corresponds to the bottom of the left lobe. Those three surface models are implemented in a graph-based optimal surface fitting scheme to remove possible false positives and the remaining part is used as initialization for the level set evolution.

Suzuki et al. preprocessed the image via anisotropic diffusion filtering to reduce the noise on the CT images, and combined it with a scale-specific gradient magnitude filter to enhance the liver boundaries [56]. The obtained image is handled by experts to put seed points for the fast marching level-set algorithm. Next to the boundary estimation, the geodesic active contour level set segmentation refined the initial contour. Their technique uses the gradient-enhanced image as a speed function, so that the front expansion speed slows down in regions having high gradient and accelerates where the gradient is low. The different steps of their method is depicted in the figure 9.

Platero et al. first aligned the different training shapes, and captured their variability via principal components analysis when building the level set function [57]. The narrow-band technique is used to select boundary candidates whose profiles are computed. Their method, incorporating shape based priors as long as edge and region-based knowledge is then implemented. Jimenez et al. proposed an optimized level set where the parameters are defined at each stage by means of multi-curvature, and where the segmentation is iteratively corrected in a pyramidal process [58]. A fine details strategy tries to prevent problems that can come from the use of a narrow-band technique when searching for the boundary, by removing outlier pixels. An additional step is semi-automatically used to impose local curvature constraints.

Algorithms based on active contours were historically considered as one of the most utilized segmentation methods. Later after the introduction of active-contours and level set, a new technique called Graph Cut segmentation was introduced by Boykov et al. to propose an alternative to boundary-based approaches [59].

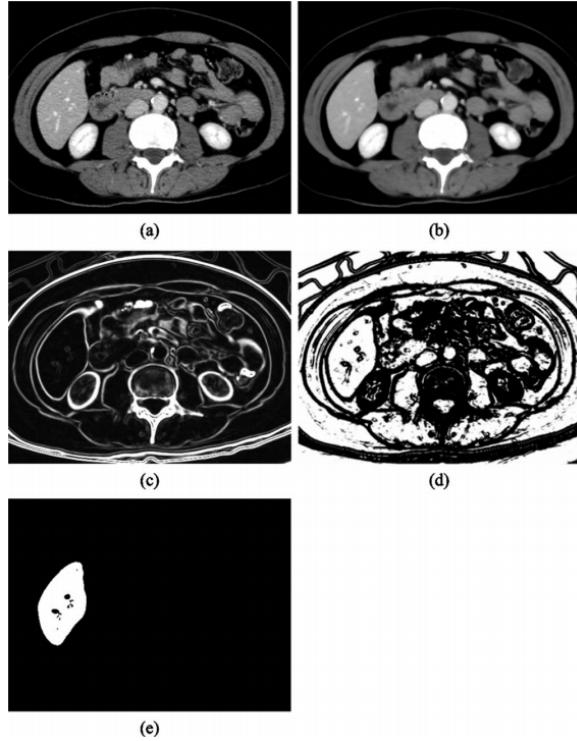


FIG. 3. Illustration of the resulting images at each step in our scheme. (a) Original CT image. (b) Anisotropic diffusion noise reduction. (c) Scale-specific gradient magnitude calculation. (d) Nonlinear grayscale conversion. (e) Geodesic active contour segmentation.

Figure 9: The different steps of the segmentation process developed by ©Suzuki et al. [56]

2.4 Graph-theory based methods

The random walker algorithm, which was introduced by Grady et al., and the Graph Cut (GC) method correspond to semi-automatic methods where the user provides seeds for both the background and the region to segment [60]. They both interpret pixels of the image as nodes on a graph where edges represent adjacency between pixels. The weights on the edges correspond to the similarity between adjacent pixels.

The GC will then try to find the minimum cost function between all possible cuts on the graph to separate the object from its background.

GC often suffers from the “small cut” problem where only the seeds are separated from the rest of the image, whereas the random walker method does not suffer from this problem since it is not aiming for the smallest boundary, as represented in the figure 10.

The random walker algorithm has been implemented by Maier et al. to proceed the segmentation of the liver [61]. Seed points were automatically detected in several regions in a slice-wise fashion. Based on the skin surface and some anatomical assumptions, the rib cage and liver seed points are generated.

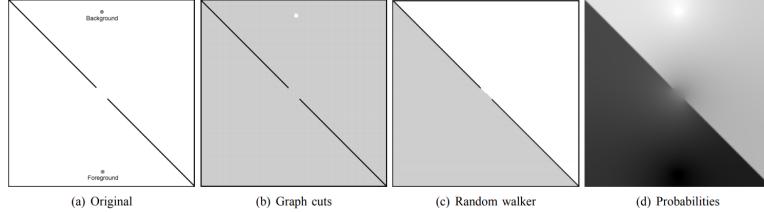


Fig. 5. Comparison of random walker algorithm to graph cuts for a weak boundary with small seeds. Note that a 4-connected graph was used in these experiments. a) Original (synthetic) image created with a diagonal black line with a section completely erased. b) Graph cuts solution — Since surface area of seeds is smaller than the weak boundary, the smallest cut minimally surrounds the seeds. c) Random walker solution. d) Probabilities associated with the random walker algorithm offer a notion of segmentation confidence at each pixel.

Figure 10: Examples of failure caused by the GC algorithm as depicted by ©Grady et al. [60]

Background seeds are obtained by detecting the air, the fat and the bones.

Dong et al. proposed a random walker implementation where adjacent slices are considered as prior knowledge when segmenting the current slice [62]. Two slices are chosen so that they present the largest cross sectional area of both the liver and the spleen. Seeds are manually placed on these slices and updated on adjacent slices via *Gaussian Mixtures Modeling* and intensity constraints through *Narrow Band Thresholding*, before the modified random walker segment the organs in a slice-wise fashion. Details of the execution can be found in the figure 11.

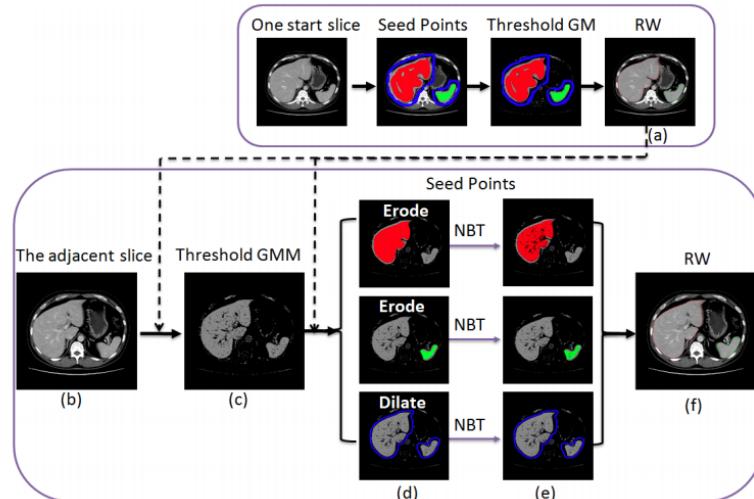


Fig. 3 Steps of the RWNBT method. (a) The segmented liver (red) and spleen (green) of the previous slice; (b) The current slice; (c) Candidate pixel by applying a threshold to the GMM; (d) The rough objects (red and green) and background (blue) seed points; (e) The fine seed points using the NBT method; (f) The segmentation results by RWNBT.

Figure 11: Segmentation pipeline as detailed by ©Dong et al. [62]

Beichel et al. used a graph-cut method to propose an initial liver region which can interactively be segmented by the users [63]. The cost of the graph cut segmentation is specified by taking region and boundary properties into account. The region term evaluates the gray-level similarity in the neighborhood of a given point and compares it to the intensity of the seed points. The boundary term

is computed from the local gradient information. *Massoptier and Casciaro 2007*, introduced a GC method initialized by an adaptive threshold. The thresholding was implemented by detecting the most likely liver intensity based on patch partition of the abdominal slices. The energy function took both a region and a boundary-based term into account. Shimizu et al. combined a statistical atlas-based approach and a graph cuts algorithm to segment the liver [64]. The graph supports an energy function including a local term, a boundary term and a shape-based term computed from the gradient of the initial region.

The key part of graph cut methods is to define a relevant energy function, which can sometimes be difficult.

Probabilistic atlases (PA) and Statistical Shape Models (SSM) both utilize prior information on the liver. The main goal of those techniques is to improve the segmentation by not relying only on the gray-level distribution since it was proven not to be sufficient, especially knowing the intensity similarity that the liver presents with close organs [65–68].

2.5 Probabilistic Atlases

To construct a probabilistic atlas, images from the training set are registered to a reference image, whereas manual delineations are warped onto the template image and averaged to get a liver appartenance probability for each voxel of the space. The generated atlas is then incorporated in the segmentation process. The studies differ one another on the way they construct the atlas, and on how it is incorporated in the segmentation process. Park et al. used the MIAMI method (*Mutual Information for Automatic Multimodality Image Fusion*) which combines a TPS-based (*Thin Plate Splines*) registration method with MI (*Mutual Information*) as a similarity metric [66]. The TPS is a registration method that uses control points as constraints to the interpolating function. They used 36 control points that were placed in different organs by experts, and they performed the registration separately for each organ to not be biased by the presence of more control points on the liver surface. The different points can be seen in the figure 12.

Zhou et al. extracted the so-called “anatomical structure” of each patient, which is composed of the bone structure (determined using gray-level thresholding) and the diaphragm (based on the shape of the air within the lungs) [65]. A matrix transformation of the TPS is calculated between each one of the training cases and a predefined standard anatomical structure [69]. The atlas is then computed by combining the obtained positions using a voting strategy. When performing the segmentation, the same strategy is performed: the current patient’s anatomical structure is estimated, and deformed to

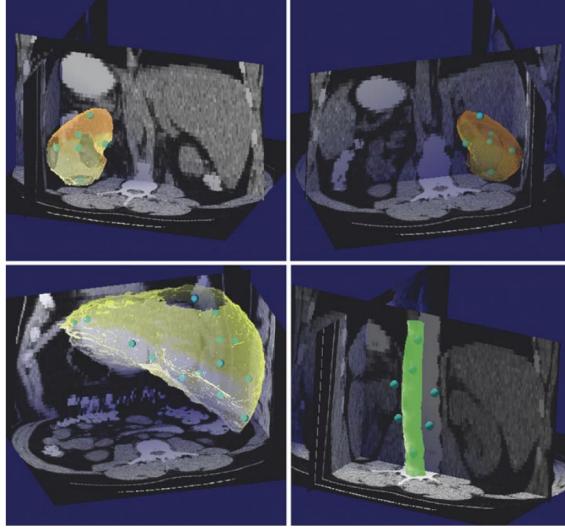


Fig. 1. Distribution of control points for organs of interests. From top left, left kidney model and associated six control points; (top right) right kidney model and associated six control points; (bottom left) liver model and associated 17 control points; and (bottom right) spinal cord model and associated seven control points in the reference volume.

Figure 12: Example of control points distribution used in the MIAMI method as described by ©Park et al. [66]

the standard one. The gray-level distribution is also used to compute the final probability as well as the spatial location as depicted in the figure 13.

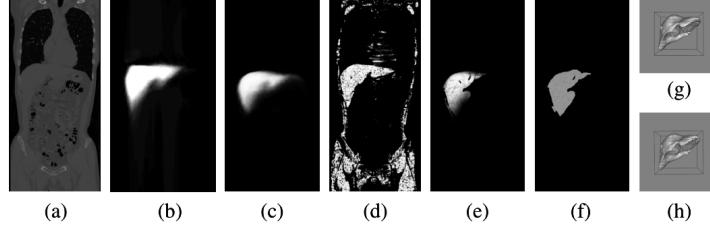


Fig. 4. An example of liver segmentation results. (a) original CT image (1 coronal slice), (b) liver atlas, (c) $P(B)$: liver probability on spatial location, (d) $P(CIB)$: liver probability on density, (e) $P(A)$: liver probability, (f) liver segmentation result (1 coronal slice), (g) ground truth of liver (3-D view), (h) liver segmentation result (3-D view).

Figure 13: Example of liver segmentation with the different slices results as described by Zhou et al. [65]

Slagmolen et al. resampled all training images to a selected one using affine registration with MI as the similarity metric, followed by a non-rigid registration guided by a surface distance metric [67]. A mean morphology is defined by averaging, for each patient in the training set, its deformation field to every other image. Each patient is then deformed using its corresponding mean deformation field. And finally the obtained images are averaged on their intensities to obtain the atlas.

Van Rikxoort et al. decided to first find the vertical range that most likely contains the liver by looking at HU values covering the lungs and extract slices above them [68]. They trained a kNN classifier with both gray level features (obtained from both the gradient computation and the Gaussian derivatives at

various scales), positional features through the coordinates in the liver space, and an additional feature obtained by computing the proportion of the liver above, behind and to the left of a given voxel. No concrete atlas was created here, but the current test volume is registered to each training sample in a *non-rigid multi-atlas segmentation*, before applying the classification to refine the obtained region. Li et al. constructed a liver and a rib-cage probabilistic atlas . Both atlases were built iteratively to reduce the dependence to the reference patient [70]. Linguraru et al. created the atlas by normalizing the organ coordinates of each patient relative to the xiphoid, and used those coordinates in the size-preserving affine registration process, by randomly selecting a patient as reference [71].

In those different studies, the segmentation method applied is often related to the way the atlas is created. Therefore Park et al. decided to complete the segmentation step by incorporating the atlas information in a bayesian framework which combines MAP and EM, with gray level distribution as basis [66]. The segmentation is then refined using a MRF (*Markov Random Field*) regularization step. In the other studies, non-rigid or elastic registration is performed from the atlas to the target case [67, 70, 71], after constraining the input, using either manually defined ROI around the liver [67], or analyzing the gaussian intensity distribution to remove irrelevant parts [70].

2.6 Statistical Shape Models

Statistical shape models on the other hand try to extract features of the model instead of building an average map.

They were first described by Cootes et al., who introduced the concept of Active Shape Models (ASM) which are built from a set of segmented training images and consist of 2 parts, a geometrical model, and a local appearance model [72]. The geometrical model describes the shape and is represented by a PDM (*Point Distribution Model*), a dense collection of landmark points on the surface of the object. Given the location of the landmarks for each training case, a principal component model can be built in order to approximate all valid shapes. The local appearance model describing the boundary, is used as an additional one to detect the modeled shape in the new image. It is based on the local gray value appearance around the boundary. A PCA or a kNN model can be used to determine the possible profiles at each landmark. The model is then guided by internal forces that should keep the shape of the deformable model similar to the one of the underlying SSM, and external forces that drive the deformable surface towards the best fit to the data.

Montagnat and Delingette followed this scheme by deforming the model with global constraints (shape constraints and transformation constraints) to reduce the degrees of freedom, and registered the

obtained model on the volumetric data including an additional external force based on local gradient computation [73]. Lamecker et al. proposed a model that combined global constraints that minimised the geometric distortion for shape correspondence, and local constraints by considering the intensity profiles for the boundary search [74]. The evolution of the boundary search for one example can be seen in the figure 14.

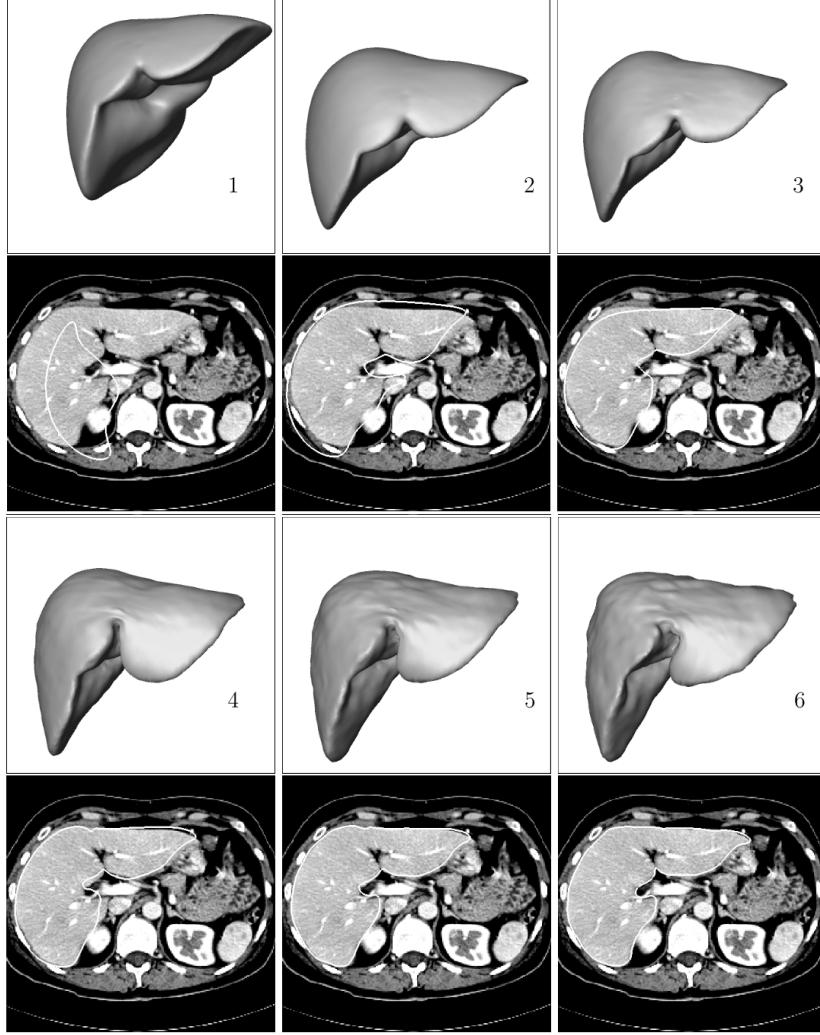


Figure 9: Steps of the segmentation process for a data set contained in the statistical liver model. At each of the six steps the model at its current position and shape is shown as a 3D surface. Below each surface one axial slice with the intersection line of the surface is displayed. Step 1: initial positioning of the mean model in the CT data such that the upper borders of the bounding boxes match. Step 2: result of the initial optimization of the position parameters. Step 3-5: results of the combined position and shape optimization with increasing number of shape modes (3, 9, 21). Step 6: final result (42 shape modes).

17

Figure 14: Evolution of boundary search in one example as described in ©Lamecker et al. [74]

Heimann et al. used an evolutionary algorithm to create a population of possible shape configurations and evaluated all the population members using a local appearance model based on a kNN-classifier utilizing the profiles [75]. Saddi et al. constrained the boundary of the liver to fit the global-shape learned from the training samples, and refined it locally using a template-matching algorithm [76]. For both problems, they used the intensity distribution inside and outside the liver to find the best region

to segment. Zhang et al. performed the detection of the liver using a 3-D generalized Hough transform so that each vertex on the surface of the average shape model is stored in a table, by considering its coordinates relative to the liver centroid [77]. The most probable liver centroid is computed for a given test image. A search is then performed to determine the points belonging to the boundary knowing the intensity distribution inside the liver and the properties of the gradient along the boundary. An optimal surface detection algorithm is finally applied to refine the obtained boundary by using a graph-search strategy. The different steps of the algorithm can be visualized in the figure 15.

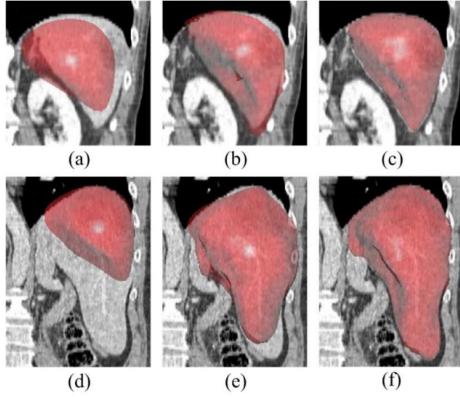


Fig. 5. State of the surface mesh in a coronal slice from back after each segmentation step. The top row shows a relatively easy case (testing data 5) while the bottom row shows a relatively difficult case (testing data 4). (a) and (d) After localization by 3-D GHT. (b) and (e) After model subspace initialization. (c) and (f) Final result after optimal surface detection.

Figure 15: Evolution of the surface mesh after the segmentation steps described by ©Zhang et al. [77]

Ling et al. went further by proposing a hierarchical segmentation of the liver in a coarse-to-fine fashion [78]. For the training process, a hierarchical shape model is obtained by downsampling the resolution of an initially obtained dense mesh. Each layer of the pyramid then contained the mean shape as long as the different modes that capture shape variations for the given resolution. During the inference, the pose estimation is performed using a MSL-based approach (*Marginal Space Learning*) to reduce the dimensionality of the research [79]. The model is then upsampled to the finest resolution using local boundary refinement. At the finest resolution, a patch-based approach is implemented to refine the obtained boundary. Seghers et al. used multiple local shape models instead of a global one [80]. A grid search is used for the landmarks candidates detection, then, an intensity model selects the best landmarks in a given test image by analyzing the intensity profiles. This search is performed in a multi-resolution manner by changing the size of the grid. A shape model was built at each edge of the geometrical model, where an edge is connecting two landmarks. Erdt et al. tried to determine the regions where high deformations are expected by looking at the curvature [81]. They first built the model based on the data present in the training set, and then constrained it locally using the curvature and the gradient at each landmark point. Examples of how the incorporation of local curvature can

improve the segmentation results are depicted in the figure 16.

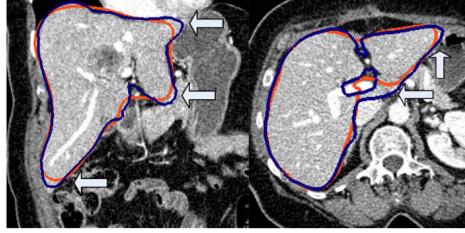


Figure 3. Using local constraints, our multi-tiered model (dark outline) leaves the learned space (bright outline) without leaking into other neighboring structures.

Figure 16: Example of how the incorporation of the curvature improved the quality of the segmentation as reported by ©Erdt et al. [81]

A common drawback of SSMs is their lack of flexibility. To combat this limitation, some studies combined the SSM with a free form deformation step to achieve better performances. For example, Kainmüller et al. built a SSM based on a PCA analysis of the given shapes present in the training dataset, and to add flexibility to the model, a free-form deformation step based on the computation of 3-vector fields was conducted [82]. During the deformation, they considered the general intensity distribution of the liver, the local displacements and they also tried to preserve the shape features of the initial model.

Okada et al. combined both PA and SSM in their study [83]. The PA was constructed from the different training cases by first registering the volumes to a standardized patient space using non-rigid registration of the abdominal cavity. The registration target was defined by experts describing the standard position and shape of the liver. The atlas was then computed by averaging the binary registered liver masks. The multilevel shape model (ML-SSM) is obtained by dividing the liver shape recursively into patches, and performing a PCA for each one of them. An example of construction of ML-SSM can be seen in the figure 17.

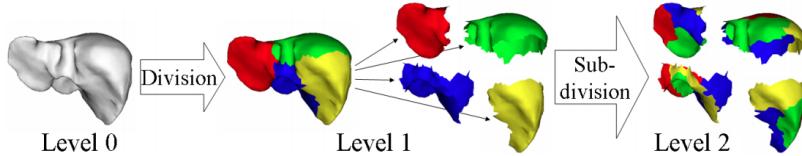


Fig. 1. Hierarchical division of the liver shape for multi-level statistical shape model

Figure 17: Division of the liver as described by ©Okada et al. [83]

For each test case, the atlas was combined with the gray-level distribution of the patient to obtain an initial liver region. A shape model was then built from this area and the ML-SSM is employed to segment

the different patches, by adding an adhesive constraint to eliminate inconsistencies between adjacent patches. They later improved the efficiency of their method by constructing the PA hierarchically, where the different structures are constrained by the shape of the organ in the next highest hierarchy level. 3 levels were handled, the abdominal cavity, the liver, and finally both the vena cava & the gallbladder. They extended the ML-SSM to an organ-based relationship and defined the MO-SSM (Multi-Organ SSM) [83]. The MO-SSM is constructed by combining ML-SSM for the different organs of interest. As a result they improved the accuracy in the regions concerned by the MO-SSM.

Apart from the lack of flexibility, SSMs suffer from a low number of training samples, and also hardly deal with low contrast between the liver and its surrounding or within the liver directly [74, 76]. Examples of those limitations are depicted in the figure 18.

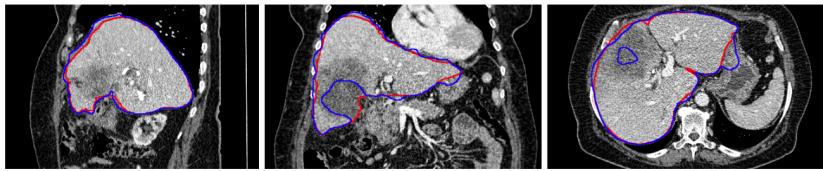


Figure 18: Example of errors obtained during the segmentation process as exposed by ©Saddi et al. [76]

2.7 Hybrid methods

Some studies decided to combine the aforementioned techniques with other machine learning methods such as standard neural networks or SVM (*Support Vector Machine*) to segment the liver.

Tsai and Tanashi used a neural network based on regional histogram to get a rough liver segmentation [84]. Several post-processing steps were added, such as Laplacian filtering to obtain the boundary followed by a smoothing using B-spline functions. Gao et al. first located the liver by extracting peaks on gray-level histograms, and applied domain-based knowledge to remove irrelevant tissues [85]. The liver region is refined based on adjacent slices properties, before a boundary refinement is performed via a deformable contour technique. Schimdt et al. incorporated a set of intensity distributions, neighboring relationship between organs and geometrical constraints to segment different parts of the abdomen [86]. Their study was based on several anatomical assumptions, but fails when the liver presents a slightly non-standardized aspect, especially when large tumors are present. Freiman et al. applied a multi-resolution, multi-class smoothed Bayesian classification followed by morphological adjustments and an active contour refinement for the segmentation of the liver [87]. The classification was performed by iteratively looking at liver and potential tumor gray-level distributions. MAP rule is applied for the two classes. The liver region is adjusted using morphological operations, before an

active contour refinement step based on gray-level intensity gives the final segmentation. Freiman et al. further proposed a method for the semi-automatic segmentation of liver tumors, where a SVM classification is employed to separate healthy and tumorous tissues [88]. A 3D energy function is then applied using affinity constraints to get the final VOI. Florin et al. proposed a shape model that will describe the shape variations only using a small number of key slices [89]. This model, called Sparse Information Model, consists of key slices that are selected to be sufficient so when combined with an interpolation function, the 3D volume can be reconstructed. Those key slices are further used for the segmentation but the result is highly sensitive to initialisation.

Schenk et al. combined the “livewire” algorithm, which is a semi-automatic 2D segmentation method, with a shape-base model to approximate the contour between the manually defined slices [90]. The interpolation used to combine the user-defined segmentations is an object-based interpolation considering the distance transforms. Maklad et al. proposed the segmentation of the liver by analyzing the blood vessels structure [91]. Abdominal blood vessels are first enhanced via bias field correction as depicted in the figure 19, and then extracted through thresholding and region growing techniques. Those vessels are further classified into hepatic and non-hepatic ones by applying distance transforms. The boundary between the liver and its surrounding organs is then constructed equidistantly from the hepatic vessels and the non-hepatic ones. A post-processing step is finally applied to refine the obtained area, by filling holes and classifying boundary tumors as belonging or not to the liver based on their relation to hepatic blood vessels.

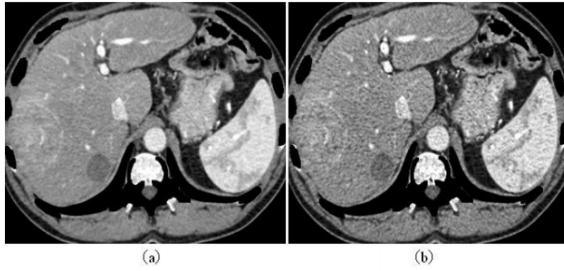


Figure 1. Blood vessel enhancement, (a) CT dataset before enhancement, (b) after enhancement

Figure 19: Results of the blood vessels enhancement initially performed by ©Maklad et al. [91]

Chartrand et al. started from a set of user-defined contours, and obtained an initial shape using variational interpolation [92]. Boundary points were then extracted using a template matching method. Intensity profiles were computed in a narrow band around the previously established boundary, before a non-rigid registration scheme based on Laplacian mesh optimization deformed it to the real liver boundaries. Goryawala et al. first extracted 5 different regions on the volumes using a k-means algorithm (the liver, surrounding tissues, peripheral muscles, rib/spinal cord and the air) where clusters

are initialized by user-defined seed points on the central liver slice [93]. An intensity based region growing algorithm is then applied using two more seed points on the top and bottom of the liver, and refined using a 3-axis growing strategy. The slice presenting the largest cross-sectional liver part is automatically identified before being proposed to the user to get a precise boundary which will then be copied in all slices containing liver voxels. This boundary is finally refined using a localized region growing method. One advantage of the presented method is that it seems to be weakly influenced by the user-defined initializations. Li et al. computed the total variation and the L1 norm to initialize the liver shape, before applying a level set method guided by local and global energies, and refining the obtained region using gray-level co-occurrence matrices [94].

Mostafa et al. proposed an artificial bee colony clustering technique to segment the liver, which is then refined using morphological operations [95]. The final step consists of a region growing method to enhance the segmentation obtained previously. Shi et al. started with a blood vessel shape initialization, and deformed the liver shape using a region-specific deformable framework [96]. The blood-vessel shape allows the algorithm to get a more accurate initialization based on the patient specificities. Al-Shaikhli et al. proposed a level set formulation guided by a combined region-based and voxel-wise cost function [97]. For the global image term, textural features (GLCM) as long as intensity-based and volume properties were computed. Initialization is performed based on prior knowledge about liver geometry and intensity distribution. The local information is represented by the shape prior obtained from a hierarchical patch-based division of the liver. The newly constructed level-set formulation is then iteratively calculated to get the liver boundary. Xu et al. performed a multi-organ segmentation via context learning followed by SIMPLE atlas selection (*Selective and Iterative Method for Performance Level Estimation*) [98]. Context-Learning was done via intensity distribution analysis with a GMM (*Gaussian Mixture Model*). The different organs were modelised by several atlases using the SIMPLE method. During inference the different organs segmentations were fused using joint level fusion before a final abdominal segmentation can be obtained.

Wang et al. learned the liver shape model from a set of training samples, by implementing a Sparse Shape Composition (SSC) [99]. When segmenting a new case, they first initialized the liver boundary that is then used to construct a rough polygonal mesh representation of the liver. The mesh is finally refined via homotopy-based optimization using the SSC as reference. Huang et al. first localized the liver using a trained AdaBoost classifier where image features (intensity based and contextual information) were treated as input [100]. The liver was then registered using the SSM incorporating all training shapes in combination with a kNN classification. A Free-form deformation is finally applied to refine the segmentation using the SSM to compute forces applied to the mesh. The entire pipeline is

depicted in the figure 20.

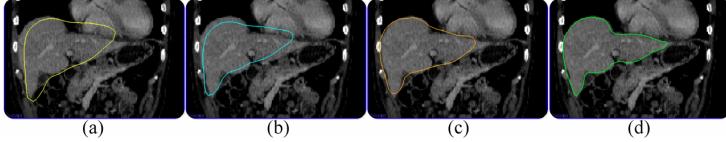


Figure 1: The four steps of liver segmentation framework: (a) liver model location; (b) registration with liver distance map; (c) shape fitting under appearance guidance; (d) free-form deformation.

Figure 20: Liver segmentation pipeline as described by ©Huang et al. [100]

Anter et al. first detected the liver boundary using fuzzy c-means clustering based on the gray level local distribution, then refined the liver segmentation map via connected component analysis [101].

It has been proven that methods relying only on intensity are not robust enough to produce acceptable results especially when dealing with some primary tumors that present high textural heterogeneity. Prior knowledge and engineered features were added over time to increase the performances, but remained sensitive to the size of the dataset used in the training process, and the amount of interaction to incorporate in the workflow. Historically, semi-automatic methods were used for liver tissue segmentation, while they now tend to be replaced by automatic techniques, offering reproducibility and fewer interactions with experts. However, newer methods still suffered when dealing with pathological livers, often presenting irregular shape or intensity patterns. Recently, deep learning has changed the way of comprehending different computer vision related problems, especially in the medical imaging field, where state-of-the-art methods rely now on these techniques.

3 Deep Learning based methods

3.1 General Deep Imaging introduction

The main breakthrough brought by DL was its ability to detect morphological properties in images only by using the pixel intensities as input, whereas traditional machine learning methods often required sophisticated hand-crafted features to achieve descent results [102, 103]. Deep learning networks achieved state-of-the-art results in many medical-related applications among which classification, localisation, detection, registration and segmentation [104]. Even with a small number of training cases, those performances were realized thanks to architecture choice, data augmentation or transfer learning [105, 106]. As exposed previously, the automatic segmentation of the tumors brings a volumetric information, that is more powerful than the diameter only, and that allows to compute the tumor burden, which has an importance when estimating the efficiency of a given treatment [9–14].

In this regard, several studies have been conducted to perform automatic liver and liver lesions segmentation.

3.2 State-of-the-art DL implementations

The different DL studies launched for the liver tissue semantic segmentation purpose share some common properties that will be exposed hereafter. Details regarding the reviewed studies can be found in the [Appendix](#).

3.2.1 FCN-based architectures

The use of fully convolutional networks (FCN) was quickly democratized, and classical imaging-related architectures such as VGG-16 or AlexNet were transformed to FCNs by replacing dense layers with convolution layers [107, 108]. An illustration of the resulting architecture is given in the figure 21.

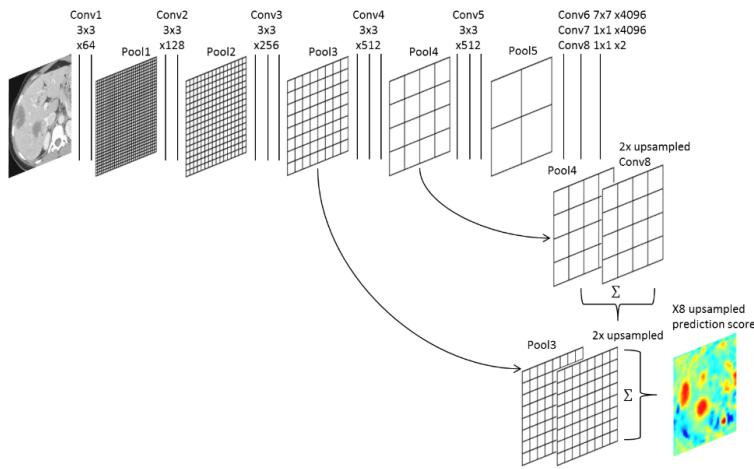


Figure 21: Example of a FCN architecture as described by ©Bellver et al. [108]

The FCN allows an architecture to predict as output an object having the same size as the input. Those architectures were perfectly suitable for the required segmentation tasks.

The most recent studies showed a predominance for U-Net like networks [109, 110]. The U-Net was introduced by Ronneberger et al. initially for the segmentation of cells in microscopic images [111]. An overview of its architecture is given in the figure 22. It consists of 2 parts, the first one where the information of the image is compressed, going from the extraction of low-level features, to the extraction of more semantic-related features, and the second part where the information is resampled back to the original image resolution. The main advantages of this architecture reside in the fact that it allows the network having both the input and the output sharing the same resolution, and also the

so-called *skip-connections* that were introduced to reinject the features learned from the contraction part, in the decoding part.

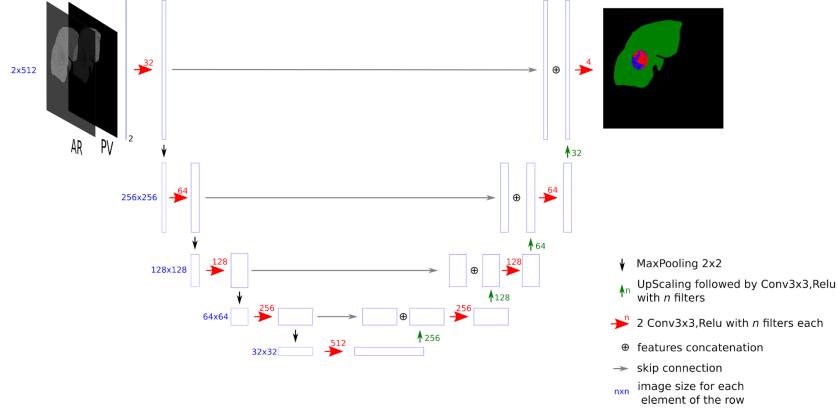


Figure 22: U-Net architecture as originally implemented and described by ©Ronneberger et al. [111]

Two other famous architectures named ResNet and DenseNet were implemented in some other studies, either alone or in combination with the U-Net [112–116].

The residual network (*ResNet*) developed by He et al. in 2015 introduced the concept of residual connections, that help deep networks combat the gradient vanishing problem [117]. At each stage of the network, identity connections, as illustrated in the figure 23, allow the information to be passed from one block to the following, so that early features are not lost.

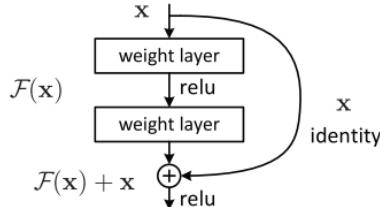


Figure 23: Residual connection as described by ©He et al. [117]

Densely connected networks (*DenseNet*) introduced by Huang et al. in 2017 share the same motivation of allowing early layers features being kept all throughout the network [118]. Contrary to the ResNet, the DenseNet requires less parameters and enables faster computation since it uses concatenation whereas ResNet added features map that needed to be kept [118]. Example of DenseNet architecture built for a classification task is depicted in the figure 24.

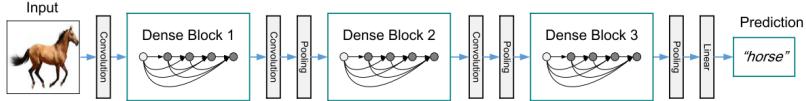


Figure 2: A deep DenseNet with three dense blocks. The layers between two adjacent blocks are referred to as transition layers and change feature-map sizes via convolution and pooling.

Figure 24: DenseNet architecture as initially described by ©Huang et al. [118]

3.2.2 Cascaded architectures

To segment the liver and the lesions it might contain, the vast majority of the studies used a cascaded architecture where a coarse liver segmentation is performed, before either being refined [110] or directly used to segment the lesions [107, 112, 115, 116, 119]. An example of cascaded architectures is given in the figure below.

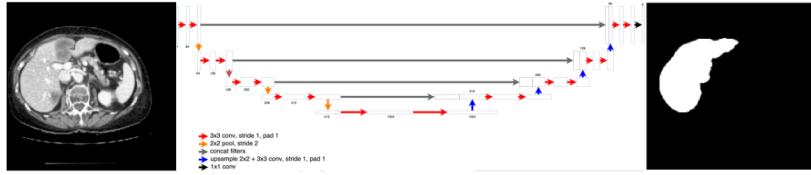


Figure 5: Step 1 of Cascaded FCN: The first U-Net learns to segment livers from a CT slice.

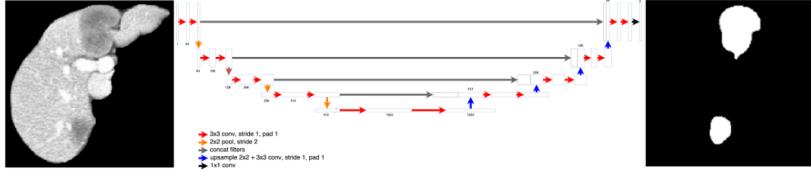


Figure 6: Step 2 of Cascaded FCN: The second U-Net learns to segment lesions from a liver segmentation mask segmented in step 1 of the cascade

Figure 25: Cascaded liver and lesion segmentation architecture as initially implemented by ©Christ et al. [119]

Interestingly, some studies decided to extract features before combining them to get a final segmentation map. For example, Bi et al. proceed first to an extraction of both liver and lesion features, before combining them with the original image to perform a pixel-wise categorical classification, as illustrated in the figure 26 [114].

Whereas Vorontsov et al. separated the extraction of liver features and lesion features, as detailed in the figure 27 [109]. They added a classifier at each step which used features concatenated from 3 adjacent slices to get a segmentation of the middle slice.

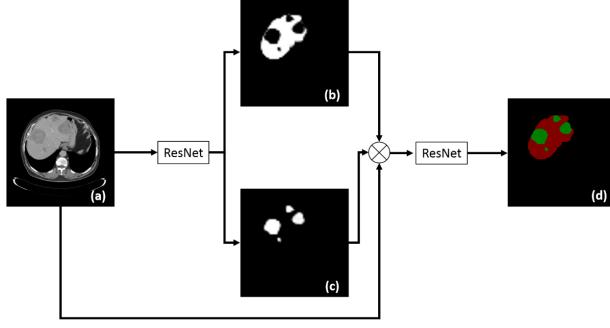


Fig. 2. Overview of the proposed cascaded ResNet architecture, where (a) is the input CT image; (b) and (c) are the produced probability map of the liver and the liver lesions, respectively; and (c) is the final predication from the cascaded ResNet for the liver and the liver lesions.

Figure 26: Custom cascaded architecture implemented by ©Bi et al. [114] where the liver and lesions features are first extracted before being combined to produce the final segmentation map

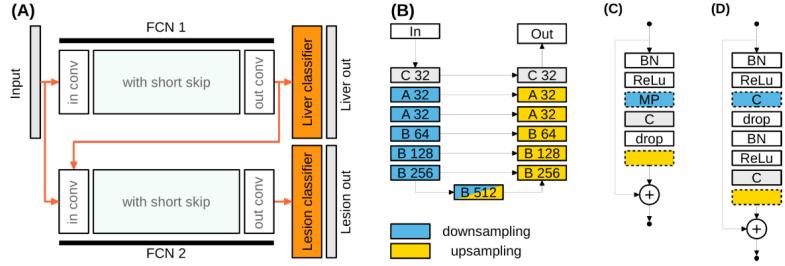


Fig. 1. (A) Two FCNs, FCN 1 and 2, each take a 2D axial slice as input. FCN 1 produces a segmentation mask for the liver; FCN 2 for lesions. The latent representation produced by FCN 1 is passed as an additional input to FCN 2. (B) FCN structure with the number of convolution filters noted in each block. Blocks coloured blue perform downsampling while those coloured yellow perform upsampling. “C” denotes a 3×3 pixel convolution layer; “A” and “B” denote blocks A and B, shown in (C) and (D), respectively. “BN”, “ReLU”, and “MP”, denote batch normalization, rectified linear units, and max pooling, respectively. Blocks with dashed lines are used in only the upsampling or the downsampling path, as denoted by colour.

Figure 27: Architecture implemented by ©Vorontsov et al. [109] where the liver and lesions features are computed separately before being combined using a classifier

3.2.3 Input type (2D, 2.5D, 3D, patches, multiphase)

The first deep-learning related study that performed the semantic segmentation of the liver or its lesions, used patches as input. Li et al. trained a classification model where patches were considered as positive if at least 50% of its pixels are tumoral [120]. Details of their network are given in the figure 28.

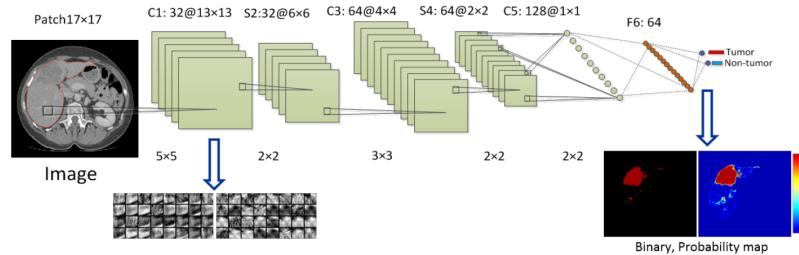


Figure 28: Patch-wise architecture as detailed and described by ©Li et al. [120]

Later, Frid-Adar et al. also implemented a patch-based segmentation, but by separating *non-lesion*

patches on whether they are located in the *liver interior* or in the *liver boundary* [121]. The classification was performed through a multi-scale approach, followed by a CNN-based FP reduction step.

However, the majority of the studies implemented a slice-by-slice segmentation [113–115], whereas some others chose to introduce volumetric consistency by using 3 or sometimes 5 adjacent slices as input in a 2.5D-manner [108, 110, 112].

Worth noting that only a few studies incorporated some real 3D layers in their architecture, such as Li et al. who computed 3D inter-slice features in 12-slices blocks all along the different CT scans [116]. Rafiei et al. decided to incorporate 3D layers only in the encoding part of their *3D-2D-FCN* model [122]. The connection between the encoding part and the decoding part was done by custom *skip connections* to concatenate the middle slice of the 3D volume with its corresponding 2D features map in the decoding part. Dou et al. were the first to build an entire 3D network with large kernels, and had intermediate supervised layers to combat the gradient vanishing problem [123]. Details of their network are given in the figure 29.

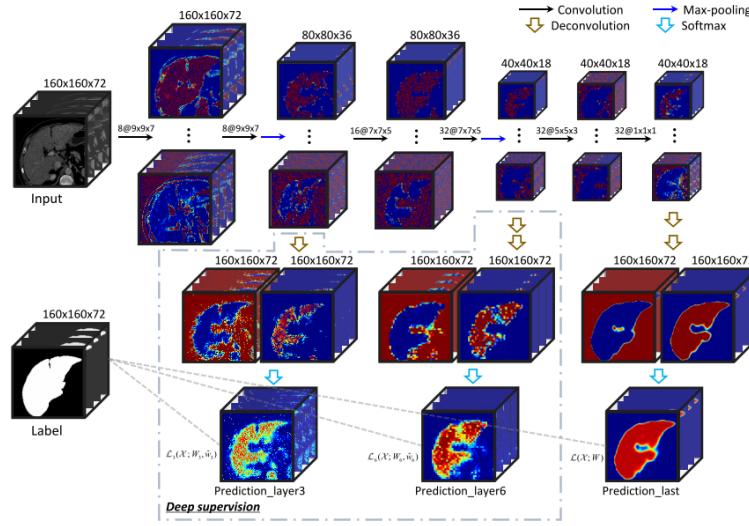


Figure 29: Full 3D architecture as detailed and implemented by ©Dou et al. [123]

Concerning the incorporation of temporal information, Sun et al. were the first to utilize multiphase information during the DL-driven automatic segmentation of liver lesions [124]. One FCN network per channel was trained to extract features before they were merged through a features fusion layer.

3.2.4 Training strategies (pre-processing, pre-trained networks...)

Regarding the training strategies, no real consensus can be established.

Concerning the pre-processing steps, the majority of the studies decided to perform the liver coarse segmentation on a down-sampled representation of the images to limit the computational cost, before

using the original resolution to perform the lesion segmentation in order to not lose any details [109, 110, 112, 115, 116].

Data normalization was also commonly applied after clipping the HU intensities to a given range, but no consensus can be deduced concerning the clipping range. Interestingly, Kaluva et al. decided to apply 3 different windowing ranges to the original image when carrying out the lesion segmentation.

Almost all the reviewed studies performed data augmentation to artificially increase the size of the dataset, by most of the time employing classical geometrical transformations such as rotations, flips, ships, scalings or elastic deformations [107–110, 112, 114, 116, 119, 121, 122].

Worth also noting that no consensus can also be deduced regarding the training from scratch or the application of fine-tuning after pre-training the architecture on another dataset such as ImageNet [108, 114, 119]. The same differences are remarkable for the choice of slices to keep during the training. Yuan et al. and Kaluva et al. decided to keep only liver slices plus a certain margin to train the liver segmentation network, whereas only slices presenting one or multiple lesions were used to train the lesion segmentation network [110, 115]. Bi et al. selected half of the slices to present both the liver and the lesions, whereas none of them were present in the other half [114]. The type of slices chosen is directly linked to the loss function used to train the network, where some studies implemented a weighted version of the cross entropy (categorical or binary) [107, 108, 112, 119], and some others decided to directly associate the target metric with the loss function through the Dice or Jaccard distance [109, 110, 113]

3.2.5 Inference schemes (post-processing, ensemble learning)

Finally, the inference schemes present also some differences among the different proposed methods.

The vast majority of the studies applied a post-processing step to refine the obtained liver segmentation by extracting the largest connected component [108, 110, 112, 115, 116]. They also often took advantage of the cascade paradigm where lesions activations outside of the predicted liver can be removed [109, 110, 116]. In other studies, the reduction of the FP (*False Positives*) was done through an additional classifier filtering the different lesions candidates. Chelbus et al. for example implemented an object based random forest classifier [113], Bellver et al. proposed a patch-based lesion detector as depicted in the figure 30 [108], whereas Frid-Adar et al. integrated an additional CNN designed to detect FP [121].

Some other studies decided to add a more sophisticated step by implementing Conditional Random Fields (CRF) [119, 122, 123], however, these models tend to be difficult to train, and may also increase the inference time.

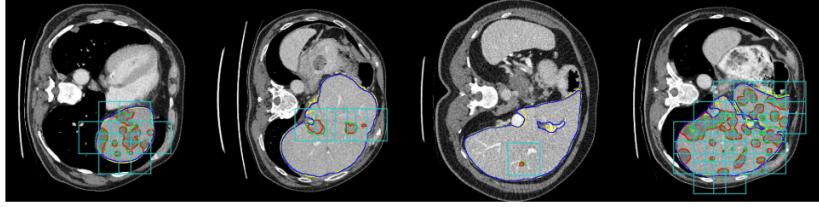


Figure 2: Results of the lesion detection network. Blue and red lines indicate the liver and lesion ground truth, respectively. Yellow and green lines are the segmentation results for liver and lesion. The light blue bounding boxes are the windows detected as having a lesion. All positive pixels at the output of the segmentation network will be removed if they disagree with the results of lesion detection.

Figure 30: Example of lesion segmentation results, as reported by ©Bellver et al. [108]

The combination of different networks in a so-called *ensemble-learning* way was also commonly realized by the different studies. Yuan et al. combined the different networks obtained through a cross-validation process to give the final segmentation [110]. Whereas Chlebus et al. trained one network per axis [113]. Bi et al. decided to train different networks following a multiscale-strategy [114], and Vorontsov et al. trained networks on differently oriented images [109].

3.2.6 Conclusion

As exposed previously, no real consensus can be found to design a generic DL-related liver tissue semantic segmentation study. However, the state-of-the-art studies share some common properties such as their cascaded architecture, the use of both pre- and post-processing or the use of FCN-based networks as basis. The different teams that participated in either the MICCAI17 or ISBI17 LITS challenge were reviewed by the organizers, who came to the same conclusions [18]. They also realized that the best tumor-segmentation results were obtained for large lesions, and within a specific lesion-liver HU difference interval (a difference between -10 and -60). They also noticed that top-ranked methods used some 3D approaches in their architecture, showing perspectives to capture the whole volume context in the future. They are planning to relaunch the same type of competition in the future, by particularly providing multiple ground truths and potentially splitting the lesion segmentation task into large and small lesions, since current methods still struggle in segmenting the small lesions.

More recent DL-related studies have been using publicly available databases such as LITS as a way to compare themselves to state-of-the-art methods. As a matter of facts, the key common elements that have been proposed previously served as a basis for new studies. Jin et al. for example proposed a network that integrated both U-Net and attention residual mechanism to proceed the segmentation of both the liver and the lesions [125]. The residual attention mechanism has been introduced in 2017 by Wang et al. to perform image classification, with the idea that attention mechanism can help the

network focusing on specific parts of the image [126]. The study from Jin et al. was the first to use the attention mechanism for semantic segmentation purpose. On the hidden test set of LITS, they outperformed a lot of 2D-based methods, but were still far from the top-ranked teams [125].

However, new paradigms such as the self-attention mechanism, in combination with state-of-the-art 2D and 3D architectures are certainly an avenue for the improvement of the automatic liver and tumors segmentation tasks [127].

4 Semantic segmentation applied to the study of HCC

As explained previously, state-of-the-art deep learning techniques allow modern hardware to perform automatic segmentation of both the liver and its structures with a precision close to the one obtained by the experts themselves.

4.1 Motivations

We implemented different architectures to validate the hypothesis that current deep networks can perform automatic delineation of both the liver and its inner structures. We also proved that incorporating the temporal information through the use of contrast-enhanced images can improve the performances of the network.

The features computed by the deep networks will further be used in a radiomics purpose, as explained previously.

As described previously, no real consensus was made regarding the design of the DL-related semantic segmentation studies, so we decided to launch several experiments on selected datasets to determine the impact of the training strategies on the performances of the networks. We will first describe the different datasets used and the implemented networks, before presenting the experiments. After selecting an optimal set of hyperparameters, we evaluated the advantages brought by a cascaded architecture, and the way to implement it, before describing the way to incorporate the multiphase information in our architecture. The results and the conclusions of this work were presented in the literature [128].

4.2 Datasets

The different datasets used in our research work are detailed in the table 2.

Table 2: Datasets used in our experiments

Db Name	Volume dimension	Axial voxel size (mm)	Slice thickness (mm)	Available contrast-enhanced phases	Contains Tumor	Tumor type	Liver GT	Tumor GT	Necrosis GT	#Experts
TheraHCC-dB	104 2D slices	0.66-0.97	0.7-1.25	NECT, AR & PV	Yes	HCC	true	true	true	4
LITS-dB	131 3D vol.	0.55-1	0.45-6	Single phase images but mixed (AR & PV)	Yes	Mixed	true	true	false	3
3Drcad-dB	15 3D vol.	0.57-0.87	1.25-4	Single phase images but mixed (AR & PV)	Yes	Mixed	true	true	false	-
TCA-dB	18 3D vol.	0.62-0.90	2.5-7	AR&PV	Yes	HCC	false	true	true	1
G-dB	79 3D vol.	0.58-0.98	0.8-5	AR&PV	Yes	HCC	false	true	false	1

As we can see in the table, and except for **LITS-dB** and **3DIrcad-dB** (**3DIrcad-dB** being a subset of the **LITS-dB** [18]), they are all different in their construction. Differences can be found on the annotated areas (chosen sparse slices or entire 3D volumes) and on the annotated tissues that can be found in the dataset (some of them only contain ground truth annotations for the liver and the tumors it might contain such as **LITS-dB**, whereas some others contain only ground truth annotation for the tumor such as **TCIA-dB**). Another crucial difference concerns the contrast enhanced phases available in each dataset (**LITS-dB** contains only single phase images whereas **TCIA-dB**, **TheraHCC-dB** and **G-dB** present multiphasic images).

To prove the ability of the deep learning to perform the automatic segmentation of both the liver and its internal tissues, such as the parenchyma and both the active and the necrotic part of the tumor, we first performed our experiments on **TheraHCC-dB** since this database was previously used for the same task [129], because it presents multiphase images and finally because this is the only available dataset with complete ground truth for both liver parenchyma, active and necrotic part of the tumors.

TheraHCC-dB is composed of images from seven patients, all suffering from *HCC* and who underwent CECT (Contrast-Enhanced Computed Tomography) examinations, resulting in a total of 13 CT sequences.

More details about the standard CECT examination protocol can be found in the **Medical Context section**. In our case, images were acquired at 4 different moments: one before the injection of the contrast medium (NECT: Non Enhanced CT), and the 2 others after the injection to reflect both the arterial (AR) (~20-25s after injection) and the portal venous (PV) phases (~60-70s after injection).

Eight regularly sampled slices across each one the 13 sequences were segmented by 4 experts, resulting in 104 labeled slices.

The segmentation maps obtained from each one the 4 experts were fused using the STAPLE algorithm to reach a consensus map [130].

4.3 Cascaded architecture

As exposed previously, no real consensus was made regarding the way to implement a semantic segmentation network dedicated to the delineation of both the liver and its internal structures. However, we noticed that several studies implemented a sequential pipeline that can be modeled as cascaded architecture. The cascade consists in several networks trained to perform a specific simpler task, and that are sequentially connected, to provide a final annotation map. In our case, the goal is the segmentation of the liver and its internal structures with the differentiation between parenchyma, and

both the active and the necrotic parts of the hepatic tumor.

Our cascade will consequently be composed of 3 networks, as depicted in the figure 31. The first one will be specialized to segment the liver in abdominal axial slices, the second will delineate the contours of the tumor in the predicted liver ROI, whereas the final one will differentiate between active and necrotic parts within the obtained tumor ROI [128]. Each one of the given networks will share the same *U-Net* architecture.

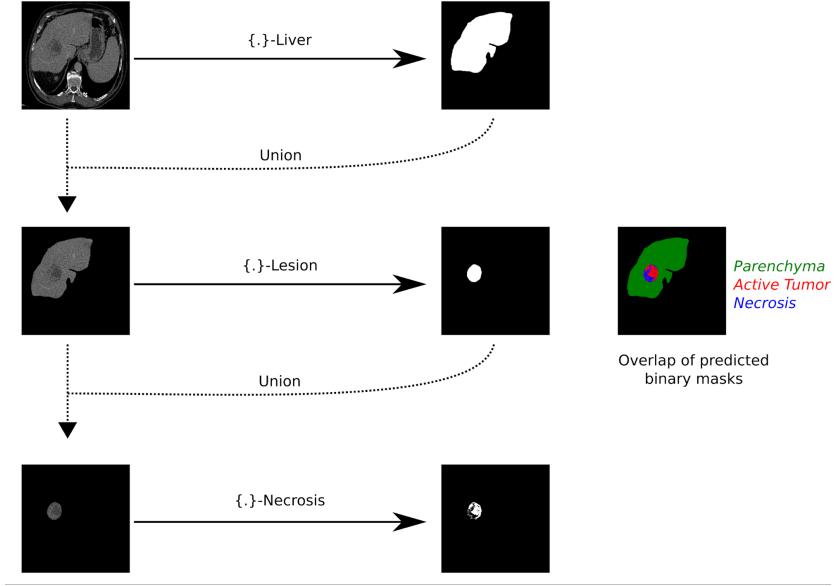


Figure 31: Cascaded network: The first network takes as input a CT image and segments the liver. The resulting segmentation map is used to remove non-liver pixels in the input data of the second network which performs the segmentation of lesions. The last network segments the necrosis within the lesions. The three binary masks are combined in the final segmentation map.

4.4 U-Net network

As a basis architecture, we have decided to implement *U-Net* like networks because it has been previously used for the semantic segmentation task, and has proven to give good results even with a small number of training samples.

The original *U-Net* architecture was developed by Ronneberger et al. and initially designed for the delineation of cells in microscopic images [111]. As detailed previously, the network can be divided into two subparts, a contraction one where the information contained within the images is compressed, through the extraction of high to low level features, and a decoding part where the compressed information is used to reconstruct a high resolution segmentation. The architecture was initially composed of 19 convolutional layers, with a rectified linear unit function as activation. The input image

had an initial size of 512×512 pixels, and at each stage of the encoding part, 2 convolutional layers are stacked with an increased number of filters. The initial pair of convolution layers used 64 filters each, whereas the final stage of the encoding stage used 1024 filters. After each pair of convolutional layers, a max pooling layer is applied to halve the spatial dimension of the features maps. As a result, a $30 \times 30 \times 1024$ features map is produced at the bottleneck of the network. This representation is then reformatted in the decoding part to obtain a segmentation map with a size equivalent to the one of the input image. In order to increase the spatial dimension of the features maps, Up-Convolutional layers are implemented. It is worth noting that the number of filters used in each of the pair of convolutional layers is decreased from the bottleneck to the final layer of the network. The last convolutional layers will map the obtained features to the final number of dimensions of the segmentation maps, which will correspond to the number of classes to predict. The final layer implements a softmax function, to simulate a prediction of appartenence to each one of the output class.

In comparison to the original architecture, we have implemented zero-padding convolutions to preserve the image size and obtain a segmentation map with the same spatial dimension as the input image. We conserved the same settings as in the original architecture concerning the number of filters to use at each stage, starting with a pair of convolutions of 64 filters each, and reaching a $32 \times 32 \times 1024$ features map in the bottleneck part of the network.

The same naming-system as in our study will be used here [128]. Single-phase elementary networks will be referred to by both the input phase and the segmentation target, as an example, PV-Lesion will refer to the network responsible for the segmentation of the lesion, with PV phase images as input. The complete *U-Net* architecture for this specific elementary network is depicted in the figure 32.

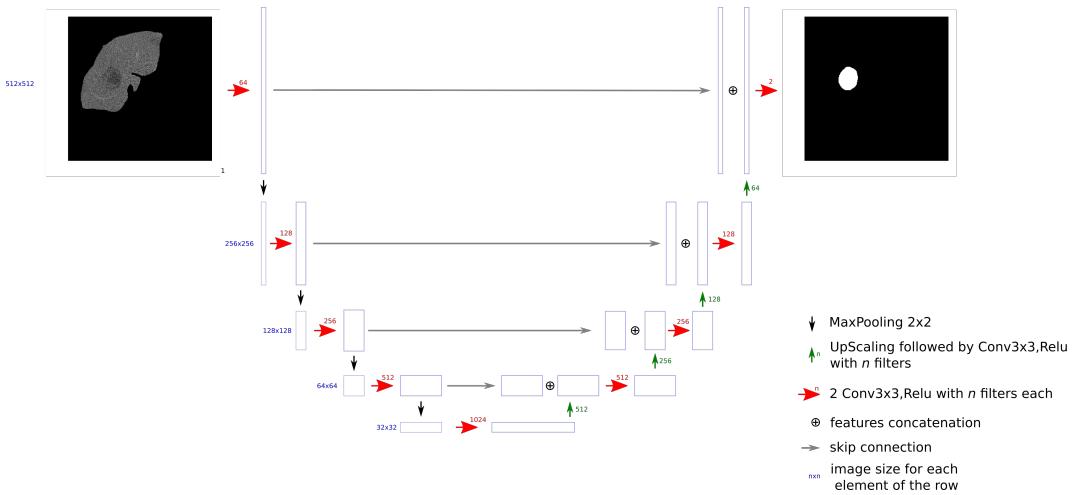


Figure 32: PV-Lesion network used to segment lesions within the liver with a PV image as input

In order to evaluate the improvements brought by the cascaded architecture, we also trained the original

U-Net architecture to perform simultaneously the whole internal tissues segmentation task. The same naming system as previously will be used where “Full” corresponds to the simultaneous segmentation task, thus, AR-Full will refer to the network dedicated to the segmentation of both the parenchyma, the active and the necrotic part of the lesions simultaneously, with AR images as input. An illustration of the network is given in the figure 33.

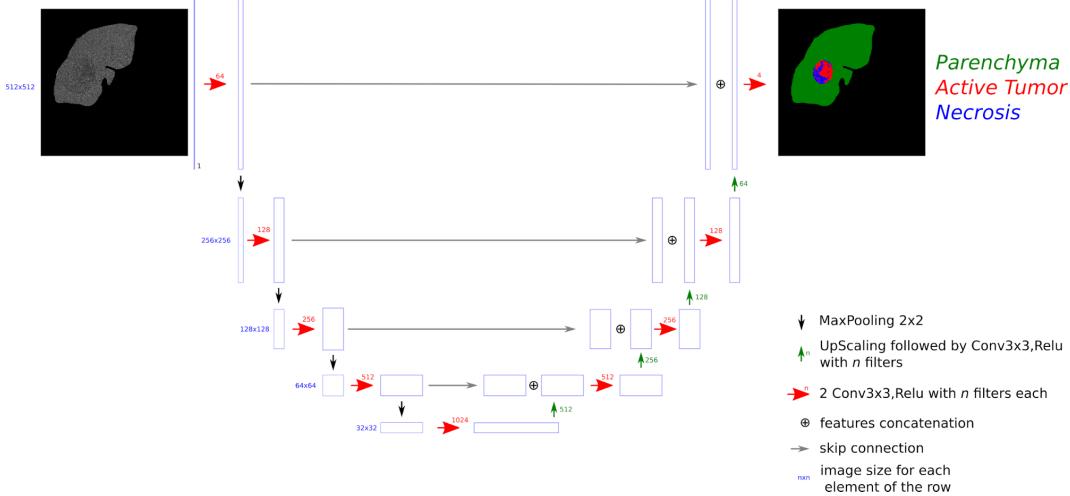


Figure 33: AR-Full refers to the network trained with AR images as input (values outside the liver are masked), and that outputs a label map, with parenchyma, active and necrotic parts annotated

4.5 Multiphase information

Only the AR and PV phases were considered in the multiphase networks because NECT phase images do not provide enough inter-tissue contrast.

In order to incorporate the multiphase information in our pipeline, we investigated 2 different strategies. The first one, referred to as DMP (Dimensional MultiPhase), consists in concatenating both the AR and PV images as input to the network (see figure 34). The second one referred to as MPF (MultiPhase Fusion), consists in performing both the encoding and the decoding separately for each phase, before merging the output maps (simple addition on the obtained features maps), as depicted in the figure 35.

4.6 Experiments

4.6.1 Data pre-processing

The first task to implement in this study was the inter-phase registration so that environmental effects, such as respiratory motions, will not affect the performances of the networks, and to ensure that a

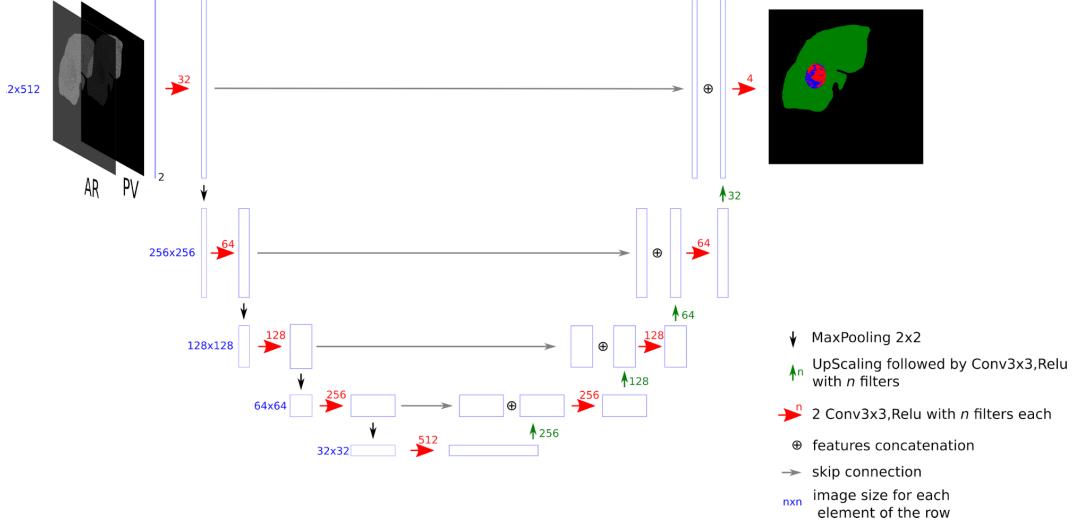


Figure 34: DMP-Full network that combines the AR and the PV images as an input to segment the parenchyma and both the active and the necrotic parts of the lesions. Here, the two channels are considered as features for the first layer

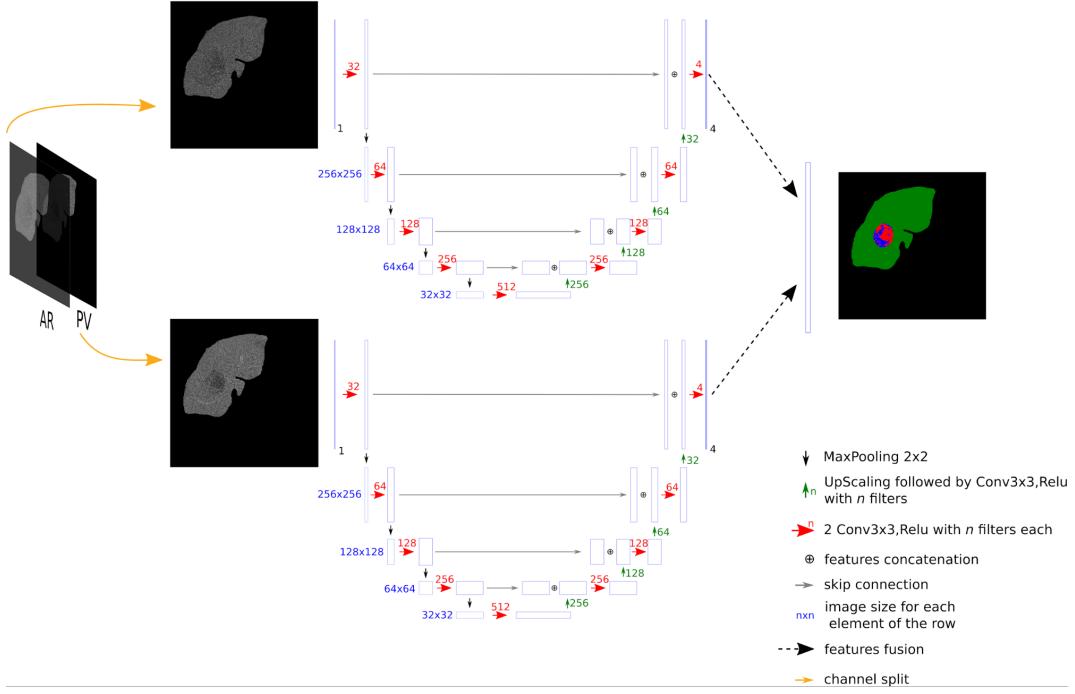


Figure 35: MPF-Full network: initially, AR and PV images are processed separately. The resulting maps are merged (by simple addition) at the end

given voxel is at the exact same position for the different CECT volumes of a patient. The registration was performed using a diffeomorphic deformable registration algorithm, where the PV images were used as reference, since they contained the original expert annotations [107, 119, 129, 131].

Another bias that can affect the deep semantic segmentation networks training is the heterogeneous

image sizes and voxel resolutions present within the training images. To avoid this bias, it has been decided to scale them so that they all have a 512x512 axial size and an isotropic voxel resolution of 0.97mm².

The data normalization is another aspect that needs to be considered before feeding the images in the deep network. In order to reduce the effect of extreme values from regions present in the tomographic images (such as the bones or the air), and to enhance the intensity of the liver voxels, we first clipped the *HU* values to be in the range [-100, 400], corresponding to the most commonly observed liver intensities range. The retained intensities were finally mapped to the interval [0, 1].

4.6.2 Training

In order to validate our hypotheses, we have decided to first run experiments on the **3Dircad-dB** to set the most crucial hyperparameters such as the learning rate, the decay, the depth of the network or the type and the amount of data augmentation. The selection process is detailed in the **Appendix**, and here is the list of the chosen hyperparameters:

- Lr: 1e-4
- Decay: 1e-4
- Number of epochs: 20
- Optimizer: Adam
- Number of filters at bottleneck: 1024
- Input image size: 512
- Data augmentation
 - Rotations in the interval [0, 40]
 - Translations with shift in the interval [-0.1, 0.1]
 - Horizontal and vertical flips
- Augmentation factor: 20
- Dropout: 0.2

When transferring those settings to the **TheraHCC-dB**, we removed the dropout and implemented both translation and the addition of gaussian noise in the training since it slightly improved the performances.

In order to remove any bias, the same set of hyperparameters has been used for both the $\{\cdot\}$ -Liver, $\{\cdot\}$ -Lesion, $\{\cdot\}$ -Necrosis and $\{\cdot\}$ -Full networks when training the **TheraHCC-dB**, regardless of the type of input (single phase or multiphase).

4.6.3 Experiments and results

Mean DSCs were computed in a slice-wise fashion for each target class. The different methods were statistically compared using the Wilcoxon signed paired rank tests, since slice-wise DSCs did not follow a normal distribution. We first compared $\{\text{NECT, AR, PV, DMP, MPF}\}$ -Liver networks to evaluate which phase allows better liver segmentation. We then trained $\{\text{NECT, AR, PV, DMP, MCF}\}$ -Lesion and $\{\text{NECT, AR, PV, DMP, MCF}\}$ -Necrosis networks separately on the **TheraHCC-dB** by masking all values outside the liver using ground truth annotations in order to assess whether multiphase information is really useful for the segmentation. The results are given in table 3.

Multiphase performed significantly better than single-phase for segmenting the liver (DMP vs PV, $P = 0.001$; DMP vs AR, $P = 0.005$, DMP vs NECT, $P < 0.001$) and the active part of the lesions (DMP vs PV, $P < 0.001$; DMP vs AR, $P = 0.003$; DMP vs NECT, $P < 0.001$). When comparing single-phase alone, PV achieved significantly better DSCs than AR or NECT for all the segmentation tasks except for the liver segmentation. When comparing multiphase methods, DMP carries out significantly better than MPF for the segmentation of the liver (DMP vs MPF, $P = 0.004$), the parenchyma (DMP vs MPF, $P < 0.001$) and the active part of the lesions (DMP vs MPF, $P = 0.005$).

Since both DMP-Lesion and DMP-Necrosis led to the best results, we combined them in a cascade as explained before, and compared it to both $\{\text{NECT, AR, PV, DMP, MPF}\}$ -Full networks. We evaluated them in terms of liver tissue classification performance on images where the values outside the ground truth liver area were masked. The mean DSCs are reported in table 4. Examples of segmentation results are given in figure 36.

The results highlighted that the cascaded version performed significantly better than $\{\cdot\}$ -Full networks for segmenting the active part of the lesion (Cascaded DMP vs PV-Full, $P = 0.001$). The resulting segmentation maps allow us to estimate the necrosis rate (which will be the ratio between the necrotic and the tumor areas). This metric is commonly used for diagnosis and prognosis of the treatment outcome. Our workflow provided estimates of this valuable biomarker with a mean error rate of 13.0 %, which is accurate enough for clinical application.

Table 3: Segmentation results using single-phase vs multiphase methods on TheraHCC-dB.

Input	Target	Network				
		NECT	AR	PV	DMP	MPF
Raw CT	Liver	81.1 ± 27.7	89.5 ± 13.2	88.7 ± 11.4	89.9 ± 15.6	88.2 ± 16.0
True liver mask	Parenchyma	86.5 ± 13.7	82.5 ± 18.6	88.7 ± 15.4	90.5 ± 13.2	86.9 ± 17.8
True liver mask	Lesion	77.4 ± 24.1	77.4 ± 20.2	87.8 ± 9.7	88.5 ± 11.7	86.6 ± 10.3
True lesion mask	Necrosis	67.5 ± 15.5	69.7 ± 16.3	77.8 ± 12.4	78.5 ± 13.3	78.8 ± 11.7
True lesion mask	Active Tumor	65.6 ± 20.4	63.9 ± 22.6	71.6 ± 20.7	75.5 ± 17.4	73.2 ± 18.6

Table 4: Segmentation results using $\{\cdot\}$ -Full vs cascaded architectures on Db .

Target	Network					
	NECT-Full	AR-Full	PV-Full	DMP-Full	MPF-Full	Cascaded DMP
Parenchyma	85.3 ± 14.9	84.1 ± 17.9	87.0 ± 19.0	82.9 ± 19.7	87.9 ± 15.9	90.5 ± 13.2
Necrosis	62.6 ± 18.6	63.5 ± 21.5	75.7 ± 14.4	73.7 ± 14.1	75.6 ± 13.4	75.8 ± 15.1
Active Tumor	42.2 ± 24.0	43.2 ± 26.1	53.5 ± 24.2	51.3 ± 25.6	52.0 ± 23.3	59.6 ± 22.5

When compared with another study that have been using a different dataset composed of MR images, we achieved slightly better segmentation results [132]. We evaluated our method on the same database used in [129], where a manual expert interaction was required for the segmentation phase, which is not the case in the present deep learning approach. To allow fair comparison, the evaluation was conducted on the areas where all the experts reached an agreement as in [129]. The mean patient-wise segmentation DSCs are depicted in table 5. Our method enabled a better segmentation of the lesions and both necrotic and active parts. Therefore, we were able to predict the patient-wise necrosis rate with a slightly better precision.

We finally combined the DMP-Liver, the DMP-Lesion and DMP-Necrosis networks in a cascaded

Table 5: Average patient-wise segmentation DSCs (on full agreement expert area) with the semi-interactive approach of [129] and the cascaded DMP method

	Semi-interactive method [129]	ours
Parenchyma	93.7 ± 3.4	92.2 ± 4.7
Lesion	90.7 ± 6	91.8 ± 4
Necrosis	83.0 ± 12.9	83.6 ± 11.7
Active Tumor	75.2 ± 10.9	82.0 ± 6.4
Necrosis rate error	7.84 ± 4.4	7.10 ± 1.4

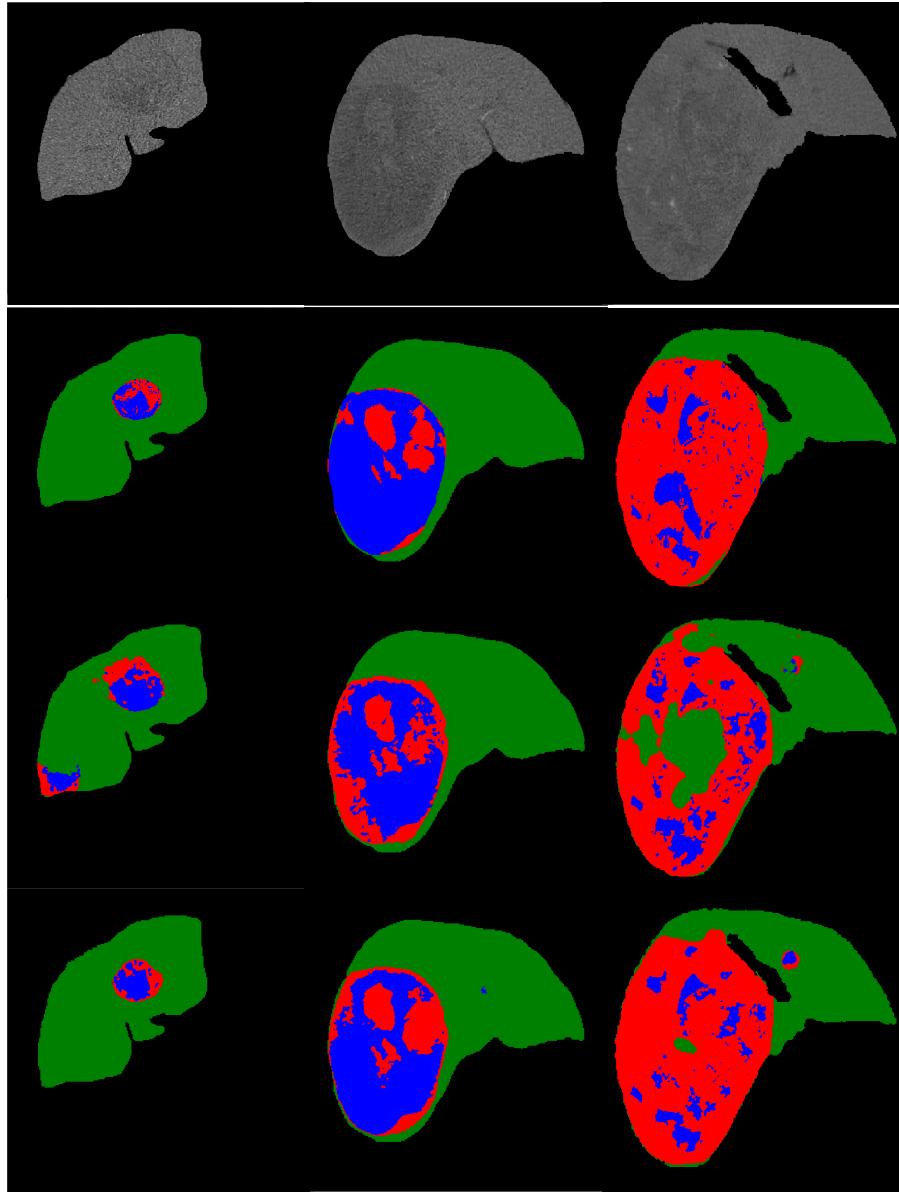


Figure 36: From top to bottom : Raw images with HU values inside the liver, Ground truth, DMP-Full segmentation, Cascaded DMP segmentation

fashion, as described in the figure 31. This allowed us to perform a fully automatic segmentation of the raw (unmasked) CT images. We reached average slice-wise DSCs of 78.3 ± 22.1 for the segmentation of the parenchyma, 50.6 ± 24.6 for the segmentation of the active part, and 68.1 ± 23.2 for the necrotic part of the lesions. We also provided a necrosis rate per patient with a mean error of 15.9% when compared to expert necrosis rate estimation. Examples of fully automatic segmentation results are given in figure 37.

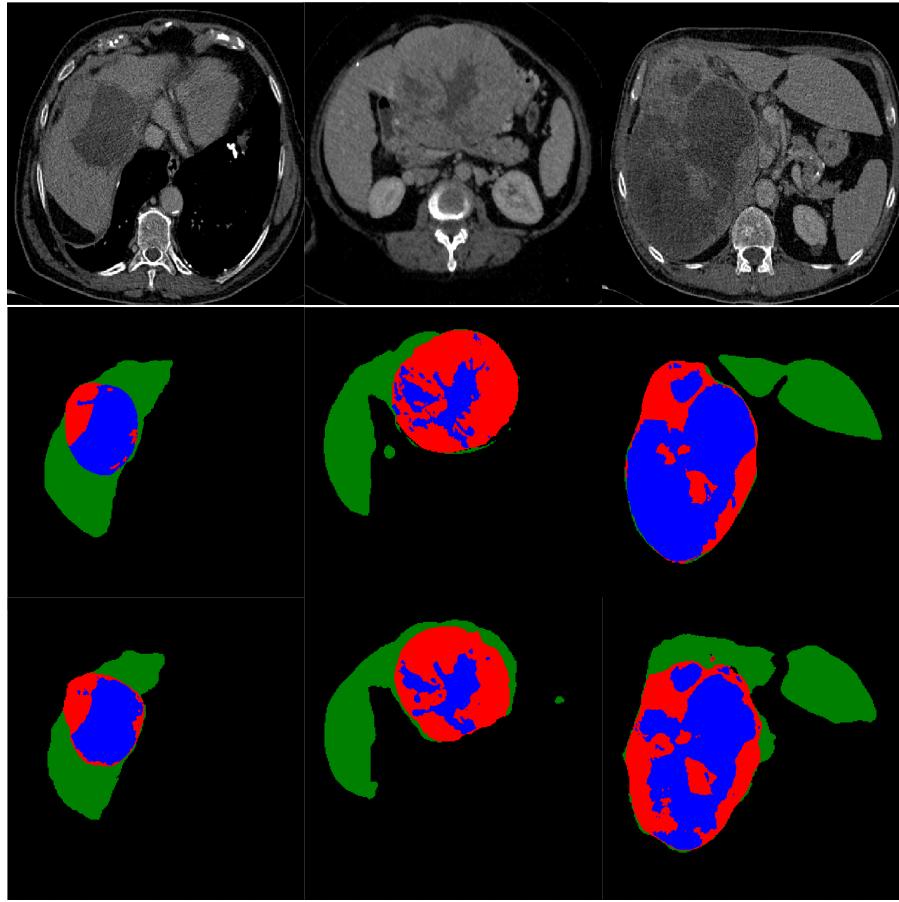


Figure 37: From top to bottom : Raw images, ground truth and results of the fully automatic segmentation of liver tissue

4.6.4 Conclusions

When comparing results obtained by the specialized networks, we proved that the addition of the multiphase information provided better segmentation results than when only single phase images are used. Statistically significant improvement was obtained for the segmentation of both the liver and the active part of the tumors.

We also investigated the performances of the single phase networks, and noticed that the PV phase was the one providing the best results, where significant improvement was obtained for all the segmentation tasks, except for the liver segmentation.

Regarding the internal liver tissues segmentation, the elementary networks providing the best results were the DMP-Lesion and DMP-Necrosis ones. When combined in a cascaded way, they performed better than each one of the $\{\cdot\}$ -Full architecture, with a statistically significant improvement for the segmentation of the active part of the lesions, meaning that combining several specialized networks together is more efficient than training a single network addressing all tasks simultaneously.

With the same experimental conditions, our solution provided a better accuracy for the segmentation

of the lesions, and for both their active and the necrotic parts, than the one obtained using a semi-automatic technique with expert interaction [128, 129]. We also obtained equivalent results than those from a similar study considering MR images as input [133].

We concluded that the combination of multiphase registered images used in a cascaded architecture allows us to automatically perform the semantic segmentation of liver tissues.

References

- [1] E A Eisenhauer, P Therasse, J Bogaerts, L H Schwartz, D Sargent, R Ford, J Dancey, S Arbuck, S Gwyther, M Mooney, L Rubinstein, L Shankar, and L Dodd. New response evaluation criteria in solid tumours : Revised RECIST guideline (version 1 . 1). *European Journal of Cancer*, 45(2):228–247, 2008.
- [2] Xiao Dan Ye, Zuguo Yuen, Jian Zhang, and Zheng Yuan. Radiological biomarkers for assessing response to locoregional therapies in hepatocellular carcinoma: From morphological to functional imaging (Review). *Oncology Reports*, 37(3):1337–1346, 2017.
- [3] A. Al-Nahhas, T. Szyszko, P. Tait, O. Damrah, and R. Canelo. Selective Internal Radiation therapy. In *Liver and Biliary Tract Surgery*, volume 10, pages 409–418. Springer Vienna, Vienna, nov 2014.
- [4] R. Yamada, M. Sato, M. Kawabata, H. Nakatsuka, K. Nakamura, and S. Takashima. Hepatic artery embolization in 120 patients with unresectable hepatoma. *Radiology*, 148(2):397–401, 1983.
- [5] Kathy S. Albain, R. Suzanne Swann, Valerie W. Rusch, Andrew T. Turrissi, Frances A. Shepherd, Colum Smith, Yuhchyau Chen, Robert B. Livingston, Richard H. Feins, David R. Gandara, Willard A. Fry, Gail Darling, David H. Johnson, Mark R. Green, Robert C. Miller, Joanne Ley, Willliam T. Sause, and James D. Cox. Radiotherapy plus chemotherapy with or without surgical resection for stage III non-small-cell lung cancer: a phase III randomised controlled trial. *The Lancet*, 374(9687):379–386, 2009.
- [6] S Rossi, M Di Stasi, E. Buscarini, P Quarrelli, F Garbagnati, L Squassante, C T Paties, D E Silverman, and L Buscarini. Percutaneous RF interstitial thermal ablation in the treatment of hepatic cancer. *American Journal of Roentgenology*, 167(3):673–759, sep 1996.
- [7] Bernard Nordlinger, Marguerite Guiguet, Jean Christophe Vaillant, Pierre Balladur, Karim Boudjema, Philippe Bachellier, and Daniel Jaeck. Surgical resection of colorectal carcinoma metastases to the liver: A prognostic scoring system to improve case selection, based on 1568 patients. *Cancer*, 77(7):1254–1262, 1996.
- [8] S. Jagannath, W. S. Velasquez, S. L. Tucker, L. M. Fuller, P. W. McLaughlin, J. T. Manning, L. B. North, and F. C. Cabanillas. Tumor burden assessment and its implication for a prognostic model in advanced diffuse large-cell lymphoma. *Journal of Clinical Oncology*, 4(6):859–865, 1986.
- [9] Paolo G. Gobbi, Chiara Broglia, Giuseppe Di Giulio, Monica Mantelli, Paola Anselmo, Francesco Merli, Pier L. Zinzani, Gabriele Rossi, Vincenzo Callea, Emilio Iannitto, Marco Paulli, Lorena

- Garioni, and Edoardo Ascari. The clinical value of tumor burden at diagnosis in Hodgkin lymphoma. *Cancer*, 101(8):1824–1834, 2004.
- [10] Hans Christian Bauknecht, Valentina C. Romano, Patrik Rogalla, Randolph Klingebiel, Claudia Wolf, Lars Bornemann, Bernd Hamm, and Patrick A. Hein. Intra-and interobserver variability of linear and volumetric measurements of brain metastases using contrast-enhanced magnetic resonance imaging. *Investigative Radiology*, 45(1):49–56, 2010.
- [11] Lars Bornemann, Volker Dicken, Jan Martin Kuhnigk, Dag Wormanns, Hoen Oh Shin, Hans Christian Bauknecht, Volker Diehl, Michael Fabel, Stefan Meier, Oliver Kress, Stefan Krass, and Heinz Otto Peitgen. OncoTREAT: A software assistant for cancer therapy monitoring. *International Journal of Computer Assisted Radiology and Surgery*, 1(5):231–242, 2007.
- [12] Claus Peter Heußel, S. Meier, S. Wittelsberger, H. Götte, P. Mildenberger, and H. U. Kauczor. Quantitative CT-verlaufskontrolle von lebermalignomen nach RECIST und WHO im vergleich zur volumetrie. *RoFo Fortschritte auf dem Gebiet der Rontgenstrahlen und der Bildgebenden Verfahren*, 179(9):958–964, 2007.
- [13] Jan Martin Kuhnigk, Volker Dicken, Lars Bornemann, Annemarie Bakai, Dag Wormanns, Stefan Krass, and Heinz Otto Peitgen. Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans. *IEEE Transactions on Medical Imaging*, 25(4):417–434, 2006.
- [14] Michael Puesken, Boris Buerke, Joachim Gerss, Barbara Frisch, Florian Beyer, Matthias Weckesser, Harald Seifarth, Walter Heindel, and Johannes Wessling. Prediction of lymph node manifestations in malignant lymphoma: Significant role of volumetric compared with established metric lymph node analysis in multislice computed tomography. *Journal of Computer Assisted Tomography*, 34(4):564–569, 2010.
- [15] Tobias Heimann, B. van Ginneken, M.A. Styner, Yulia Arzhaeva, Volker Aurich, Christian Bauer, Andreas Beck, Christoph Becker, Reinhard Beichel, et al. Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets. *IEEE Transactions on Medical Imaging*, 28(8):1251–1265, aug 2009.
- [16] Sebastian Echegaray, Olivier Gevaert, Rajesh Shah, Aya Kamaya, John Louie, Nishita Kothary, and Sandy Napel. Core samples for radiomics features that are insensitive to tumor segmentation: method and pilot study using CT images of hepatocellular carcinoma. *Journal of Medical Imaging*, 2(4):041011, 2015.

- [17] Jan Hendrik Moltz, Lars Bornemann, Jan Martin Kuhnigk, Volker Dicken, Elena Peitgen, Stephan Meier, Hendrik Bolte, Michael Fabel, Hans Christian Bauknecht, Markus Hittinger, Andreas Kießling, Michael Pusken, and Heinz Otto Peitgen. Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans. *IEEE Journal on Selected Topics in Signal Processing*, 3(1):122–134, 2009.
- [18] Patrick Bilic, Patrick Ferdinand Christ, Eugene Vorontsov, Grzegorz Chlebus, Hao Chen, Qi Dou, Chi-Wing Fu, Xiao Han, Pheng-Ann Heng, Jürgen Hesser, Samuel Kadoury, Tomasz Konopczynski, Miao Le, Chunming Li, Xiaomeng Li, Jana Lipkovà, John Lowengrub, et al. The Liver Tumor Segmentation Benchmark (LiTS). pages 1–43, 2019.
- [19] Akshat Gotra, Lojan Sivakumaran, Gabriel Chartrand, Kim Nhien Vu, Franck Vandenbroucke-Menu, Claude Kauffmann, Samuel Kadoury, Benoît Gallix, Jacques A. de Guise, and An Tang. Liver segmentation: indications, techniques and future directions. *Insights into Imaging*, 8(4):377–392, 2017.
- [20] Grand challenge. <https://grand-challenge.org>. Accessed: 2020-10-20.
- [21] 3dircadb. <https://www.ircad.fr/research/3d-ircadb-01/>. Accessed: 2020-10-20.
- [22] Midas. <http://insight-journal.org/midas/collection/view/38>. Accessed: 2020-10-20.
- [23] Kenneth Clark, Bruce Vendt, Kirk Smith, John Freymann, Justin Kirby, Paul Koppel, Stephen Moore, Stanley Phillips, David Maffitt, Michael Pringle, Lawrence Tarbox, and Fred Prior. The cancer imaging archive (tcia): Maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6):1045–1057, Dec 2013.
- [24] Bram van Ginneken, Tobias Heimann, and Martin Styner. 3D segmentation in the clinic: A grand challenge. *International Conference on Medical Image Computing and Computer Assisted Intervention*, 10:7–15, 2007.
- [25] Imageclef. <https://www.imageclef.org/2015/liver#access>. Accessed: 2020-10-20.
- [26] Chaos. https://chaos.grand-challenge.org/Combined_Healthy_Abdominal_Organ_Segmentation/. Accessed: 2020-10-20.
- [27] Seong-Jae Lim, Yong-Yeon Jeong, Chil-Woo Lee, and Yo-Sung Ho. Automatic segmentation of the liver in CT images using the watershed algorithm based on morphological filtering. *Medical Imaging 2004: Image Processing*, 5370(December 2015):1658, 2004.
- [28] Seong-Jae Lim, Yong-Yeon Jeong, and Yo-Sung Ho. Segmentation of the Liver Using the Deformable Contour Method on CT Images. In Yo-Sung Ho and Hyoung Joong Kim, editors,

Advances in Multimedia Information Processing - PCM 2005, volume 3767 of *Lecture Notes in Computer Science*, pages 570–581. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

- [29] Seong Jae Lim, Yong Yeon Jeong, and Yo Sung Ho. Automatic liver segmentation for volume measurement in CT Images. *Journal of Visual Communication and Image Representation*, 17(4):860–875, 2006.
- [30] Chien Cheng Lee, Pau Choo Chung, and Hong Ming Tsai. Identifying Multiple Abdominal Organs from CT Image Series Using a Multimodule Contextual Neural Network and Spatial Fuzzy Rules. *IEEE Transactions on Information Technology in Biomedicine*, 7(3):208–217, 2003.
- [31] Fan Liu, Binsheng Zhao, Peter K. Kijewski, Liang Wang, and Lawrence H. Schwartz. Liver segmentation for CT images using GVF snake. *Medical Physics*, 32(12):3699–3706, 2005.
- [32] Kyung Sik Seo and Jong An Park. Improved automatic liver segmentation of a contrast enhanced CT image. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3767 LNCS:899–909, 2005.
- [33] Pil Un Kim, Yun-Jung Lee, Youngjin Jung, Jin Ho Cho, and Myoung Nam Kim. Liver extraction in the abdominal CT image by watershed segmentation algorithm. *IFMBE Proceedings*, 14(1):2563–2566, 2007.
- [34] Paola Campadelli, Elena Casiraghi, and Andrea Esposito. Liver segmentation from computed tomography scans: A survey and a new algorithm. *Artificial Intelligence in Medicine*, 45(2-3):185–196, 2009.
- [35] Amir H. Foruzan, Reza Aghaeizadeh Zoroofi, Masatoshi Hori, and Yoshinobu Sato. Liver segmentation by intensity analysis and anatomical information in multi-slice CT images. *International Journal of Computer Assisted Radiology and Surgery*, 4(3):287–297, 2009.
- [36] Andreas Beck and Volker Aurich. Hepatux - a semiautomatic liver segmentation system. *3D Segmentation in The Clinic: A Grand Challenge*, pages 225–233, 2007.
- [37] Yingyi Qi, Wei Xiong, Wee-Kheng Leow, Qi Tian, Jiayin Zhou, Jiang Liu, Thazin Han, Sudhakar K Venkatesh, and Shih-chang Wang. Semi-automatic segmentation of liver tumors from CT scans using Bayesian rule-based 3D region growing. *MICCAI Workshop*, 41(43):1–10, 2008.
- [38] L Rusko and G Bekes. Fully automatic liver segmentation for contrast-enhanced CT images. *International Conference on Medical Image Computing and Computer-Assisted Intervention. Segmentation in the Clinic: a grand challenge, 2007*, pages 143–150, 2007.

- [39] László Ruskó, György Bekes, and Márta Fidrich. Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images. *Medical Image Analysis*, 13(6):871–882, 2009.
- [40] Ruchaneewan Susomboon. A hybrid approach for liver segmentation. *Proc. MICCAI Workshop 3-D Segmentat. Clinic: A Grand, Challenge,,* i:151–160, 2007.
- [41] S. S. Kumar, R. S. Moni, and J. Rajeesh. Automatic liver and lesion segmentation: A primary step in diagnosis of liver diseases. *Signal, Image and Video Processing*, 7(1):163–172, 2013.
- [42] Regina Pohle and Klaus D. Toennies. Segmentation of medical images using adaptive region growing. In Milan Sonka and Kenneth M. Hanson, editors, *Medical Imaging 2001: Image Processing*, volume 4322, pages 1337–1346, jul 2001.
- [43] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic Active Contours. In *International Journal of Computer Vision*, volume 22, pages 61–79. 1997.
- [44] Stanley Osher and Ronald Fedkiw. *Signed Distance Functions*, volume 153. 2003.
- [45] Shiyang Pan and Benoit M. Dawant. Automatic 3d segmentation of the liver from abdominal ct images: a level-set approach. *Medical Imaging 2001: Image Processing*, 4322:128–138, 2001.
- [46] Jinke Wang, Yuanzhi Cheng, Changyong Guo, Yadong Wang, and Shinichi Tamura. Shape–intensity prior level set combining probabilistic atlas and probability map constrains for automatic liver segmentation from abdominal CT images. *International Journal of Computer Assisted Radiology and Surgery*, 11(5):817–826, 2016.
- [47] J. F. Garamendi, N. Malpica, J. Martel, and E. Schiavi. Automatic segmentation of the liver in CT using level sets without edges. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4477 LNCS(PART 1):161–168, 2007.
- [48] Ying Chi, Peter M M Cashman, Fernando Bello, and Richard I Kitney. A Discussion on the Evaluation of A New Automatic Liver Volume Segmentation Method for Specified CT Image Datasets. *Heart*, pages 167–175, 2007.
- [49] D Furukawa, A Shimizu, and H Kobatake. Automatic Liver Segmentation Method based on Maximum A Posterior Probability Estimation and Level Set Method, T. *3D Segmentation in The Clinic: A Grand Challenge*, pages 117–124, 2007.
- [50] Jeongjin Lee, Namkug Kim, Ho Lee, Joon Beom Seo, Hyung Jin Won, Yong Moon Shin, and Yeong Gil Shin. Efficient Liver Segmentation exploiting Level-Set Speed Images with 2 .5D Shape Propagation. *3D Segmentation in the Clinic: A Grand Challenge*, pages 189–196, 2007.

- [51] Andreas Wimmer, Grzegorz Soza, and Joachim Hornegger. Two-stage semi-automatic organ segmentation framework using radial basis functions and level sets. *3D segmentation in the clinic: a grand challenge, LNCS, Springer Berlin, Heidelberg*, 1(c):207–214, 2007.
- [52] Andreas Wimmer, Joachim Hornegger, and Grzegorz Soza. Implicit active shape model employing boundary classifier. *Proceedings - International Conference on Pattern Recognition*, (June 2014), 2008.
- [53] Laurent Massoptier and Sergio Casciaro. A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from CT scans. *European Radiology*, 18(8):1658–1665, 2008.
- [54] Andreas Wimmer, Grzegorz Soza, and Joachim Hornegger. Organ Segmentation. *Image (Rochester, N.Y.)*, pages 26–33, 2009.
- [55] Yi Song, Andy J. Bulpitt, and Ken W. Brodlie. Liver Segmentation Using Automatically Defined Patient Specific B-Spline Surface Models. In *Professional differences in interprofessional working*, volume 22, pages 43–50. 2009.
- [56] Kenji Suzuki, Ryan Kohlbrenner, Mark L. Epstein, Ademola M. Obajuluwa, Jianwu Xu, and Masatoshi Hori. Computer-aided measurement of liver volumes in CT by means of geodesic active contour segmentation coupled with level-set algorithms. *Medical Physics*, 37(5):2159–2166, 2010.
- [57] Carlos Platero, María Carmen Tobar, Javier Sanguino, José Manuel Poncela, and Olga Velasco. Level set segmentation with shape and appearance models using affine moment descriptors. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6669 LNCS:109–116, 2011.
- [58] Daniel Jimenez-Carretero, Laura Fernandez-De-Manuel, Javier Pascau, Jose M. Tellado, Enrique Ramon, Manuel Desco, Andres Santos, and Maria J. Ledesma-Carbayo. Optimal multiresolution 3D level-set method for liver segmentation incorporating local curvature constraints. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 3419–3422, 2011.
- [59] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast Approximate Energy Minimization via Graph Cuts 1 Energy minimization in early vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.
- [60] Leo Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.

- [61] F Maier, A Wimmer, G Soza, J N Kaftan, D Fritz, and R Dillmann. Automatic Liver Segmentation using the Random Walker Algorithm. *Proceedings des Workshops vom 6. bis 8. April 2008 in Berlin Bildverarbeitung für die Medizin 2008*, pages 56–61, 2008.
- [62] Chunhua Dong, Yen Wei Chen, Lanfen Lin, Hongjie Hu, Chongwu Jin, Huajun Yu, Xian Hua Han, and Tomoko Tateyama. Simultaneous segmentation of multiple organs using random walks. *Journal of Information Processing*, 24(2):320–329, 2016.
- [63] Reinhard Beichel, Thomas Pock, Christian Janko, Roman B. Zotter, Bernhard Reitinger, Alexander Bornik, Kalman Palagyi, Erich Sorantin, Georg Werkgartner, Horst Bischof, and Milan Sonka. Liver segment approximation in CT data for surgical resection planning. 5370(316):1435, 2004.
- [64] Akinobu Shimizu, Keita Nakagomi, Takuya Narihira, Hidefumi Kobatake, Shigeru Nawano, Kenji Shinozaki, Koich Ishizu, and Kaori Togashi. Automated Segmentation of 3D CT Images Based on Statistical Atlas and Graph Cuts. In Bjoern Menze, Georg Langs, Zhuowen Tu, and Antonio Criminisi, editors, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 6533 of *Lecture Notes in Computer Science*, pages 214–223. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [65] Xiangrong Zhou, Teruhiko Kitagawa, Takeshi Hara, Hiroshi Fujita, Xuejun Zhang, Ryujiro Yokoyama, Hiroshi Kondo, Masayuki Kanematsu, and Hiroaki Hoshi. Constructing a probabilistic model for automated liver region segmentation using non-contrast X-ray torso CT images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4191 LNCS:856–863, 2006.
- [66] Hyunjin Park, Peyton H Bland, and Charles R Meyer. Construction of an abdominal Atlas. 22(4):483–492, 2003.
- [67] Pieter Slagmolen, An Elen, Dieter Seghers, Dirk Loeckx, Frederik Maes, and Karin Haustermans. Atlas based liver segmentation using nonrigid registration with a B-spline transformation model. *Proceedings of MICCAI workshop on 3D segmentation in the clinic: a grand challenge*, pages 197–206, 2007.
- [68] Eva Van Rikxoort, Yulia Arzhava, and Bram Van Ginneken. Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. *Proceedings of the MICCAI ...*, pages 101–108, 2007.
- [69] X. Zhou, T. Kitagawa, K. Okuo, T. Hara, H. Fujita, R. Yokoyama, M. Kanematsu, and H. Hoshi. Construction of a probabilistic atlas for automated liver segmentation in non-contrast torso CT images. *International Congress Series*, 1281:1169–1174, 2005.

- [70] Chang Yang Li, Xiuying Wang, Stefan Eberl, Michael Fulham, Yong Yin, and Dagan Feng. Fully automated liver segmentation for low- and high-contrast ct volumes based on probabilistic atlases. *Proceedings - International Conference on Image Processing, ICIP*, pages 1733–1736, 2010.
- [71] Marius George Linguraru, Zhixi Li, Furhawn Shah, See Chin, and Ronald M. Summers. Automated liver segmentation using a normalized probabilistic atlas. *Medical Imaging 2009: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 7262(February 2009):72622R, 2009.
- [72] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models - their training and application, 1995.
- [73] J. Montagnat and H. Delingette. Volumetric medical images segmentation using shape constrained deformable models. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 1205, pages 13–22. 1997.
- [74] H Lamecker, T Lange, and M Seebass. Segmentation of the liver using a 3D statistical shape model. *Konrad-Zuse-Zentrum für Informationstechnik Berlin*, 09(April):27, 2004.
- [75] Tobias Heimann, Sascha Münzing, Hans Peter Meinzer, and Ivo Wolf. A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4584 LNCS:1–12, 2007.
- [76] Kinda Anna Saddi and Christophe Chefd. Global-to-Local Shape Matching for Liver Segmentation in CT Imaging. *MICCAI (October 2007)*, pages 207–214, 2007.
- [77] Xing Zhang, Jie Tian, Kexin Deng, Yongfang Wu, and Xiuli Li. Automatic liver segmentation using a statistical shape model with optimal surface detection. *IEEE Transactions on Biomedical Engineering*, 57(10 PART 2):2622–2626, 2010.
- [78] Haibin Ling, S. Kevin Zhou, Yefeng Zheng, Bogdan Georgescu, Michael Suehling, and Dorin Comaniciu. Hierarchical, learning-based automatic liver segmentation. *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2008.
- [79] Yefeng Zheng, Adrian Barbu, Bogdan Georgescu, Michael Scheuering, and Dorin Comaniciu. Fast automatic heart chamber segmentation from 3D CT data using marginal space learning and steerable features. *Proceedings of the IEEE International Conference on Computer Vision*, pages 1–8, 2007.

- [80] Dieter Seghers, Pieter Slagmolen, Yves Lambelin, Jeroen Hermans, Dirk Loeckx, Frederik Maes, and Paul Suetens. Landmark based liver segmentation using local shape and local intensity models. *Proc. Workshop of the 10th Int. Conf. on MICCAI, Workshop on 3D Segmentation in the Clinic: A Grand Challenge*, (May 2014):135–142, 2007.
- [81] Marius Erdt, Sebastian Steger, Matthias Kirschner, and Stefan Wesarg. Fast automatic liver segmentation combining learned shape priors with observed shape deviation. *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, pages 249–254, 2010.
- [82] Dagmar Kainmueller, Thomas Lange, and Hans Lamecker. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. *Proc. MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge*, pages 109–116, 2007.
- [83] Toshiyuki Okada, Ryuji Shimada, Masatoshi Hori, Masahiko Nakamoto, Yen Wei Chen, Hironobu Nakamura, and Yoshinobu Sato. Automated Segmentation of the Liver from 3D CT Images Using Probabilistic Atlas and Multilevel Statistical Shape Model. *Academic Radiology*, 15(11):1390–1403, 2008.
- [84] Du Yih Tsai and Nobutaka Tanahashi. Neural-network-based boundary detection of liver structure in CT images for 3-D visualization. *IEEE International Conference on Neural Networks - Conference Proceedings*, 6:3484–3489, 1994.
- [85] Luomin Gao, David G. Heath, Brian S. Kuszyk, and Elliot K. Fishman. Automatic liver segmentation technique for three-dimensional visualization of CT data. *Radiology*, 201(2):359–364, 1996.
- [86] Günter Schmidt, Maria Athelogou, Ralf Schönmeye, Rene Korn, and Gerd Binnig. Cognition network technology for a fully automated 3d segmentation of liver. *MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge*, pages 125–133, 2007.
- [87] Moti Freiman, Ofer Eliassaf, Yoav Taieb, Leo Joskowicz, and Jacob Sosna. A Bayesian approach for liver analysis: Algorithm and validation study. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5241 LNCS(PART 1):85–92, 2008.
- [88] Moti Freiman, Ofir Cooper, Dani Lischinski, and Leo Joskowicz. Liver tumors segmentation from CTA images using voxels classification and affinity constraint propagation. *International Journal of Computer Assisted Radiology and Surgery*, 6(2):247–255, 2011.

- [89] Charles Florin, Nikos Paragios, Gareth Funka-Lea, and James Williams. Liver segmentation using sparse 3D prior models with optimal data support. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4584 LNCS(June 2014):38–49, 2007.
- [90] Andrea Schenk, Guido Prause, and Heinz Otto Peitgen. Efficient semiautomatic segmentation of 3D objects in medical images. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1935:186–195, 2000.
- [91] Ahmed S. Maklad, Mikio Matsuhiro, Hidenobu Suzuki, Yoshiaki Kawata, Noboru Niki, Noriyuki Moriyama, Toru Utsunomiya, and Mitsuo Shimada. Blood vessel-based liver segmentation through the portal phase of a CT dataset. *Medical Imaging 2013: Computer-Aided Diagnosis*, 8670(December 2013):86700X, 2013.
- [92] G. Chartrand, T. Cresson, R. Chav, A. Gotra, A. Tang, and J. DeGuise. SEMI-automated liver CT segmentation using Laplacian meshes. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 641–644. IEEE, apr 2014.
- [93] Mohammed Goryawala, Seza Gulec, Ruchir Bhatt, Anthony J. McGoron, and Malek Adjouadi. A low-interaction automatic 3D liver segmentation method using computed tomography for selective internal radiation therapy. *BioMed Research International*, 2014(1), 2014.
- [94] Dengwang Li, Li Liu, Jinhu Chen, Hongsheng Li, and Yong Yin. A multistep liver segmentation strategy by combining level set based method with texture analysis for CT images. *IEEE International Conference on Orange Technologies, ICOT 2014*, pages 109–112, 2014.
- [95] Abdalla Mostafa, Ahmed Fouad, Mohamed Abd Elfattah, Aboul Ella Hassanien, Hesham Hefny, Shao Ying Zhu, and Gerald Schaefer. CT liver segmentation using artificial bee colony optimisation. *Procedia Computer Science*, 60(1):1622–1630, 2015.
- [96] Changfa Shi, Yuanzhi Cheng, Fei Liu, Yadong Wang, Jing Bai, and Shinichi Tamura. A hierarchical local region-based sparse shape composition for liver segmentation in CT scans. *Pattern Recognition*, 50:88–106, 2016.
- [97] Saif Dawood Salman Al-Shaikhli, Michael Ying Yang, and Bodo Rosenhahn. Automatic 3D Liver Segmentation Using Sparse Representation of Global and Local Image Information via Level Set Formulation. pages 1–10, 2015.
- [98] Zhoubing Xu, Ryan P. Burke, Christopher P. Lee, Rebeccah B. Baucom, Benjamin K. Poulose, Richard G. Abramson, and Bennett A. Landman. Efficient multi-atlas abdominal segmentation

- on clinically acquired CT with SIMPLE context learning. *Medical Image Analysis*, 24(1):18–27, aug 2015.
- [99] Guotai Wang, Shaoting Zhang, Hongzhi Xie, Dimitris N. Metaxas, and Lixu Gu. A homotopy-based sparse representation for fast and accurate shape prior modeling in liver surgical planning. *Medical Image Analysis*, 19(1):176–186, 2015.
- [100] Cheng Huang, Xuhui Li, and Fucang Jia. Automatic liver segmentation using multiple prior knowledge models and free-form deformation. *CEUR Workshop Proceedings*, 1194:22–24, 2014.
- [101] Ahmed M. Anter, Aboul Ella Hassanien, Mohamed A. Abu ElSoud, and Mohamed F. Tolba. Neutrosophic Sets and Fuzzy C-Means Clustering for Improving CT Liver Image Segmentation. In Pavel Kömer, Ajith Abraham, and Václav Snášel, editors, *Advances in Intelligent Systems and Computing*, volume 303 of *Advances in Intelligent Systems and Computing*, pages 193–203. Springer International Publishing, Cham, 2014.
- [102] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42(1995):60–88, 2017.
- [103] Kenji Suzuki. Overview of deep learning in medical imaging. *Radiological Physics and Technology*, 10(3):257–273, 2017.
- [104] Justin Ker, Lipo Wang, Jai Rao, and Tchoyoson Lim. Deep Learning Applications in Medical Image Analysis. *IEEE Access*, 6:9375–9379, 2017.
- [105] Bo Hao Zheng, Long Zi Liu, Zhi Zhi Zhang, Jie Yi Shi, Liang Qing Dong, Ling Yu Tian, Zhen Bin Ding, Yuan Ji, Sheng Xiang Rao, Jian Zhou, Jia Fan, Xiao Ying Wang, and Qiang Gao. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary HCC patients. *BMC cancer*, 18(1):1148, 2018.
- [106] Zilong Hu, Jinshan Tang, Ziming Wang, Kai Zhang, Lin Zhang, and Qingling Sun. Deep learning for image-based cancer detection and diagnosis - A survey. *Pattern Recognition*, 83:134–149, 2018.
- [107] Avi Ben-Cohen, Idit Diamant, Eyal Klang, and Michal Amitai. Deep Learning and Data Labeling for Medical Applications. 10008:77–85, 2016.
- [108] Miriam Bellver, Kevis-Kokitsi Maninis, Jordi Pont-Tuset, Xavier Giro-i Nieto, Jordi Torres, and Luc Van Gool. Detection-aided liver lesion segmentation using deep learning. (Nips), 2017.

- [109] Eugene Vorontsov, An Tang, Chris Pal, and Samuel Kadoury. Liver lesion segmentation informed by joint liver segmentation. *Proceedings - International Symposium on Biomedical Imaging*, 2018-April:1332–1335, 2018.
- [110] Yading Yuan. Hierarchical Convolutional-Deconvolutional Neural Networks for Automatic Liver and Tumor Segmentation. *i*:3–6, 2017.
- [111] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. pages 1–8, 2015.
- [112] Xiao Han. Automatic Liver Lesion Segmentation Using A Deep Convolutional Neural Network Method. 2017.
- [113] Grzegorz Chlebus, Andrea Schenk, Jan Hendrik Moltz, Bram van Ginneken, Horst Karl Hahn, and Hans Meine. Automatic liver tumor segmentation in CT with fully convolutional neural networks and object-based postprocessing. *Scientific Reports*, 8(1):1–7, 2018.
- [114] Lei Bi, Jinman Kim, Ashnil Kumar, and Dagan Feng. Automatic Liver Lesion Detection using Cascaded Deep Residual Networks. 2017.
- [115] Krishna Chaitanya Kaluva, Mahendra Khened, Avinash Kori, and Ganapathy Krishnamurthi. 2D-Densely Connected Convolution Neural Networks for automatic Liver and Tumor Segmentation. *X(X)*:1–4, 2018.
- [116] Xiaomeng Li, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng Ann Heng. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes. *IEEE Transactions on Medical Imaging*, (1):1–13, sep 2017.
- [117] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. *Enzyme and Microbial Technology*, 19(2):107–117, dec 2015.
- [118] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-Janua:2261–2269, 2017.
- [119] Patrick Ferdinand Christ, Florian Ettlinger, Felix Grün, Mohamed Ezzeldin A. Elshaera, Jana Lipkova, Sebastian Schlecht, Freba Ahmaddy, Sunil Tatavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Felix Hofmann, Melvin D Anastasi, Seyed-Ahmad Ahmadi, Georgios Kaassis, Julian Holch, Wieland Sommer, Rickmer Braren, Volker Heinemann, and Bjoern Menze. Automatic Liver and Tumor Segmentation of CT and MRI Volumes using Cascaded Fully Convolutional Neural Networks. pages 1–20, 2017.

- [120] Wen Li, Fucang Jia, and Qingmao Hu. Automatic Segmentation of Liver Tumor in CT Images with Deep Convolutional Neural Networks. *Journal of Computer and Communications*, 03(11):146–151, 2015.
- [121] Maayan Frid-adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. *Patch-Based Techniques in Medical Imaging*, volume 10530 of *Lecture Notes in Computer Science*. Springer International Publishing, Cham, 2017.
- [122] Shima Rafiei, Ebrahim Nasr-Esfahani, Kayvan Najarian, Nader Karimi, Shadrokh Samavi, and S. M.Reza Soroushmehr. Liver Segmentation in CT Images Using Three Dimensional to Two Dimensional Fully Convolutional Network. *Proceedings - International Conference on Image Processing, ICIP*, pages 2067–2071, 2018.
- [123] Qi Dou, Hao Chen, Yueming Jin, Lequan Yu, Jing Qin, and Pheng-Ann Heng. 3D Deeply Supervised Network for Automatic Liver Segmentation from CT Volumes. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 9901 LNCS, pages 149–157. jul 2016.
- [124] Changjian Sun, Shuxu Guo, Huimao Zhang, Jing Li, Meimei Chen, Shuzhi Ma, Lanyi Jin, Xiaoming Liu, Xueyan Li, and Xiaohua Qian. Automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs. *Artificial Intelligence in Medicine*, 83:58–66, 2017.
- [125] Qiangguo Jin, Zhaopeng Meng, Changming Sun, Leyi Wei, and Ran Su. RA-UNet: A hybrid deep attention-aware network to extract liver and tumor in CT scans. (October):1–13, 2018.
- [126] Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang. Residual attention network for image classification. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, 2017-January(1):6450–6458, 2017.
- [127] Xiao Chen, Lanfen Lin, Dong Liang, Hongjie Hu, Qiaowei Zhang, and Yutaro Iwamoto. A DUAL-ATTENTION DILATED RESIDUAL NETWORK FOR LIVER LESION CLASSIFICATION AND LOCALIZATION ON CT IMAGES College of Computer Science and Technology , Zhejiang University , China Department of Radiology , Sir Run Run Shaw Hospital , China College of Informa. *2019 IEEE International Conference on Image Processing (ICIP)*, pages 235–239, 2019.
- [128] Farid Ouhmich, Vincent Agnus, Vincent Noblet, Fabrice Heitz, and Patrick Pessaux. Liver tissue segmentation in multiphase CT scans using cascaded convolutional neural networks. *International Journal of Computer Assisted Radiology and Surgery*, 2019.

- [129] Pierre-Henri Conze, Vincent Noblet, François Rousseau, Fabrice Heitz, Vito de Blasi, Riccardo Memeo, and Patrick Pessaux. Scale-adaptive supervoxel-based random forests for liver tumor segmentation in dynamic contrast-enhanced CT scans. *International Journal of Computer Assisted Radiology and Surgery*, 12(2):223–233, 2017.
- [130] Simon K. Warfield, Kelly H. Zou, and William M. Wells. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging*, 23(7):903–921, 2004.
- [131] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41, feb 2008.
- [132] Fan Zhang, Junlin Yang, Nariman Nezami, Fabian Laage-gaupp, Julius Chapiro, Ming De Lin, and James Duncan. Liver Tissue Classification Using an Auto-context-based Deep Neural Network with a Multi-phase Training Framework. In Wenjia Bai, Gerard Sanroma, Guorong Wu, Brent C Munsell, Yiqiang Zhan, and Pierrick Coupé, editors, *Patch-Based Techniques in Medical Imaging*, pages 59–66, Cham, 2018. Springer International Publishing.
- [133] Fan Zhang, Junlin Yang, Nariman Nezami, Fabian Laage-Gaupp, Julius Chapiro, Ming De Lin, and James Duncan. Liver tissue classification using an auto-context-based deep neural network with a multi-phase training framework. In Wenjia Bai, Gerard Sanroma, Guorong Wu, Brent C Munsell, Yiqiang Zhan, and Pierrick Coupé, editors, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11075 LNCS, pages 59–66, Cham, 2018. Springer International Publishing.