Automated method for detection and segmentation of liver metastatic lesions in follow-up CT examinations

Avi Ben-Cohen
Eyal Klang
Idit Diamant
Noa Rozendorn
Michal Marianne Amitai
Hayit Greenspan

Automated method for detection and segmentation of liver metastatic lesions in follow-up CT examinations

Avi Ben-Cohen,^{a,†} Eyal Klang,^{b,†} Idit Diamant,^a Noa Rozendorn,^b Michal Marianne Amitai,^{b,‡} and Hayit Greenspan^{a,*}

^aTel Aviv University, Faculty of Engineering, Department of Biomedical Engineering, Medical Image Processing Laboratory, Tel Aviv 69978, Israel

^bSheba Medical Center, Diagnostic Imaging Department, Abdominal Imaging Unit, Tel Hashomer 52621, Israel

Abstract. This paper presents a fully automated method for detection and segmentation of liver metastases in serial computed tomography (CT) examinations. Our method uses a given two-dimensional baseline segmentation mask for identifying the lesion location in the follow-up CT and locating surrounding tissues, using nonrigid image registration and template matching, in order to reduce the search area for segmentation. Adaptive region growing and mean-shift clustering are used to obtain the lesion segmentation. Our database contains 127 cases from the CT abdomen unit at Sheba Medical Center. Development of the methodology was conducted using 22 of the cases, and testing was conducted on the remaining 105 cases. Results show that 94 of the 105 lesions were detected, for an overall matching rate of 90% making the correct RECIST 1.1 assessment in 88% of the cases. The average Dice index was 0.83 ± 0.08 , the average sensitivity was 0.82 ± 0.13 , and the positive predictive value was 0.87 ± 0.11 . In 92% of the rated cases, the results were classified by the radiologists as acceptable or better. The segmentation performance, matching rate, and RECIST assessment results hence appear promising. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.2.3.034502](https://doi.org/10.1117/1.JMI.2.3.034502)]

Keywords: detection; segmentation; follow-up examination; liver metastasis; computed tomography sequential image analysis.

Paper 14139PRR received Oct. 15, 2014; accepted for publication Jul. 6, 2015; published online Aug. 19, 2015.

1 Introduction

The liver is one of the most common organs to develop metastases. Metastatic liver lesions can be derived from different primary cancer types, including melanoma, breast cancer, colon cancer, neuroendocrine, and pancreatic cancer. Most liver metastases are multiple and vary extensively in size. In some cases, a growing metastatic lesion compresses adjacent liver parenchyma, causing atrophy and forming a connective tissue rim. A common feature of large metastatic lesions is the appearance of necrosis in the center of the lesion.¹ These features, among others, make computerized liver lesions detection a difficult task.

Computed tomography (CT) is one of the most common modalities used for detection, diagnosis, and follow-up (FU) of liver lesions.² The images are acquired before and after intravenous injection of a contrast agent. Radiologists usually detect and diagnose liver lesions based on the different density of the lesions at different points in time of the scan with optimal detection on the portal phase (60 to 80 s post-injection) images. When the diagnosis indicates metastatic liver lesions, FU CT examinations are conducted every few months. The FU examinations are used to assess the success of a given therapy, where the main criterion for therapy evaluation is change in tumor size. This is done by comparing the diameter between corresponding lesions (RECIST criteria).³ Tracking lesions over time and measuring diameter size is a time-consuming task requiring the radiologist to compare two three-dimensional (3-D) CT scans, which may include multiple metastases. The difficulty of this task highlights

the need for computerized analysis to assist clinicians in the detection and evaluation of the size of liver metastases in FU examinations.

Several methods for detection of matching lesions have been introduced in the literature. To the best of our knowledge, only a few works have focused on liver metastases. Moltz et al.⁴ detected a matching lesion using gray value filtering, and detected circular structures with a Hough transform and template matching. Their data contained lung nodules, liver metastases, and lymph nodes. The detection rate for liver metastases was 84.1% for data that included mostly hypodense lesions. Another method for matching lesions was proposed by Charnoz et al.⁵ They computed a deformation field for the liver based on vessel tree matching and applied it to the lesions. This method was tested on a large synthetic database.

Liver lesion segmentation is a challenging research problem that has attracted a great deal of attention in recent years. The “3D Segmentation in the Clinic: A Grand Challenge I” at MICCAI 2008 workshop,⁶ provided a good overview of the methods developed up to 2008. The most accurate automated method at this workshop was presented by Shimizu et al.,⁷ who used the AdaBoost technique to separate liver tumors from normal liver tissue based on several local image features. Kuhnigk et al.⁸ proposed a method for lung nodules, combining a threshold-based approach with model-based morphological processing that was recently adapted to liver metastases by Moltz et al.⁹ and reported a volumetric overlap error of 31%. Schwier et al.¹⁰ suggested an object-based analysis for automatic detection and segmentation of hypodense liver lesions. In 89% of the cases, the segmentation delivered appropriate quality as evaluated by two radiologists. Militzer et al.¹¹ presented

*Address all correspondence to: Hayit Greenspan, E-mail: hayit@eng.tau.ac.il

[†]These authors equally contributed.

[‡]These authors equally contributed.

a system for automatic detection and segmentation of focal liver lesions in CT images. Their system utilized a probabilistic boosting tree to classify points in the liver as either lesion or parenchyma, providing both detection and segmentation of mainly hypodense lesions. They achieved a detection rate of 71% with 14.4 false positives (FPs) per patient. For the segmentation of correctly detected lesions, they achieved 0.95 sensitivity and 0.93 specificity.

Automatic methods for FU examinations are of great interest in clinical practice. Various works have stressed the need for an automatic lesion tracking tool that can identify the target lesions in the FU CT and evaluate the lesion size measurements. Moltz et al.¹² evaluated the clinical benefit of an automatic lesion tracking tool that identifies the target lesions in the FU CT and precomputes the lesion volumes with manual refinement steps in case of failure. They showed that their automatic lesion tracking tool can make interpretation of FU CT examinations quicker and provide results that are less reader-dependent. Their data included lung lesions, liver lesions, and lymph nodes. Keil et al.¹³ compared semi-automated software measurements with manual assessment of different types of liver metastases. They found a strong correlation between measurement techniques. Xu et al.¹⁴ described a method to automatically and reproducibly identify and segment abnormal lymph nodes in serial CT exams. They obtained a search area using the baseline image and then applied an adaptive region-growing algorithm.

The current work focuses on metastatic lesions in the liver using serial CT examinations (portal phase). We present a fully automated algorithm for detection and tracking of lesions, along with segmentation of the lesions given a baseline segmentation mask. The underlying methodology is an adaptation of Xu et al.¹⁴ to liver lesions. Additional features were added to deal with the large variability of liver metastatic lesions in size, shape, and texture. The presented work is an extension to our earlier work in Ref. 15 with augmented methodology and substantially increased set of cases.

The contribution of this work includes the following:

- The method presented in this work utilizes information taken from the baseline CT as context to the FU CT.
- The algorithm was designed to imitate the radiologist's way of work in tracking and segmenting liver metastatic lesions.
- Results are shown on data taken from a real-world clinical setting with varying scanner properties and varying types of liver metastatic lesions.

The method proposed is presented in Sec. 2. Experimental results are shown in Sec. 3. The experiments include overall parameter optimization analysis as well as comparison across alternative algorithmic solutions to the task. In Sec. 4, a discussion of the results is presented followed by a conclusion of the work.

2 Methods

The algorithm inputs include a single cross-sectional slice of a baseline CT scan (portal phase) with identified and circumscribed single lesion marked by an expert radiologist as well as FU CT slices. In actual clinical practice, for therapy evaluation purposes, physicians compare the longest diameter (LD) between corresponding lesions between a baseline and a FU scan. A cross-sectional slice is found containing the lesion

where its LD can be measured. This slice, as selected by the expert, is used in this work as the manually segmented slice in the baseline CT.

The proposed analysis framework is composed of four main stages as can be seen in the block diagram of Fig. 1: (1) finding a slice containing the lesion in the FU scan; (2) constraining the search region in the FU scan based on the prior CT study; (3) segmenting the corresponding lesions; and (4) propagating the segmentation mask to 3-D segmentation. Specific features were added in certain difficult special cases. These cases include peripheral lesions (Sec. 2.1), low contrast or lesions with high variability (Sec. 2.3), and small lesions (Sec. 2.5). The special cases are automatically detected by the algorithm and are treated with special care. In the following, we provide additional detail for each stage.

2.1 Registration and Alignment: Finding a Slice Containing the Lesion

Given an input slice in the baseline CT, the first step of the algorithm determines a corresponding slice in the FU CT. In the search for the slice of interest, we used several key steps. These include automated delineation of the liver volume within the 3-D volume, registration of each slice within the extracted volume to the baseline CT slice, and finally localization of the lesion within the selected set of slices. Given a complete FU CT volume, the first step uses coarse volume segmentation, via clustering, in order to efficiently find a subset of slices that contained the liver organ. We apply mean-shift clustering¹⁶ to the midcoronal slice set of the FU CT. Mean shift is a robust unsupervised clustering method that does not require prior knowledge of the number of clusters. It classifies the pixels into a few clusters according to the discontinuity of the intensity distribution. Each pixel is, thus, replaced by the local density maximum. This procedure yields many homogeneous clusters. In practice, the mean-shift algorithm uses the following parameters: the spatial bandwidth (h_s), the range bandwidth (h_r), and the size in pixels of the smallest allowed region. The spatial bandwidth defines a spatial search window, and the range bandwidth controls the gray-level intensity resolution. Based on empirical studies, we selected the following set of parameters: $h_s = 6$ pixels (approximately 12 mm for coronal slices and 6 mm for axial slices); $h_r = 10$ HU (Hounsfield units); smallest allowed output

Fig. 1 Overview of the proposed algorithm.

region defined as 10 pixels (approximately 10 mm² for coronal slices and 3 mm² for axial slices). This set of parameters was used for all processing steps involving mean-shift throughout the work. Following the clustering phase, physiological-based rules that focus on the relative location, size, and expected range of HU values for the liver organ are used to select a relevant set of slices for further analysis. We search for the largest region in the right half of the mean-shift map that is above the lower quartile of the scan and in the range of 75 to 130 HU. The selected intensity range includes normal and abnormal liver parenchyma values in the portal phase [Figs. 2(a) and 2(b)]. Liver size depends on several factors such as age, sex, body size, and shape. Normal length values of the liver in the mid-sternal line vary from 4 to 8.5 cm depending on these factors.¹⁷ Typically, there is suspicion of an enlarged liver when metastatic lesions occur. Thus, if the extracted region height is less than 10.5 cm, the next largest region that meets the above criterion is considered as a potential liver region. In the current work, we used constant liver specification values for all patients. These values can be easily adjusted for different sex, height, and weight. Following the mean-shift-based clustering step, which generates a coarse segmentation of the liver in the z-axis, a subset of slices is extracted for the second analysis step. In this second step, all selected transverse plane slices are registered to the baseline slice. A two-dimensional (2-D) affine registration algorithm is initially applied, minimizing the mean square error (MSE) [Fig. 2(c)]. The slice with the minimum MSE along with 10 neighboring slices (5 above and 5 below) are selected and further processed. The affine transformation captures only the global motion of the liver. An additional transformation is required, which models the local deformation of the liver. The nature of the local deformation of the liver can significantly vary across patients as a result of breathing. Therefore, it is difficult to describe the local deformation via parameterized transformations. We next apply 2-D cubic B-spline registration to minimize the residual complexity (RC) as proposed by Myronenko and Song.¹⁸ This similarity measure favors a registration with minimum compression complexity of the difference (residual) image between the two registered images; empirical tests have shown that RC outperforms other similarity measures, including sum-of-squared-differences, mutual information, correlation ratio, and correlation coefficient (CC).¹⁸ The transformation was modeled using

Fig. 2 (a) The right half of the mid-coronal computed tomography (CT) slice; (b) the output of applying mean-shift clustering technique on (a); (c) illustration of a comparison between slices in the follow-up (FU) CT search area and the baseline CT slice containing the lesion; and (d) illustration of the template-matching stage using the baseline lesion mask as the template.

a free-form deformation transformation with three hierarchical levels of B-spline control points.¹⁹

Finally, in the third processing step, we perform a search for the lesion that is most similar to the baseline lesion within the set of registered slices. The slice in which this lesion was found is selected as the slice of interest for further processing and refined lesion segmentation. The bounding box of the lesion in the baseline input is used as a template. A template-matching procedure is conducted in a small search area on each registered slice in the FU CT, surrounding the baseline lesion mask at a distance of $d = 20$ mm from each edge of the bounding box [Fig. 2(d)]. The measure used for lesion matching combines the largest normalized cross-correlation coefficient found in each slice along with its distance from the center of the lesion in the baseline, using the following framework:

1. SET **maxCC** to zero and **minDist** to ∞ .
2. FOR all slices in the search region:
 - (a) Set **CC** and **dist** to be the maximum normalized cross-correlation coefficient value and its spatial distance from the center of the lesion in the baseline for the current slice, respectively.
 - (b) IF ($CC > 0.65$ AND $(\text{min Dist}-\text{dist}) > 10$) OR ($CC > \text{maxCC}$ AND $(\text{dist}-\text{min Dist}) < 10$) THEN
 - i. SET closest slice to be the current slice.
 - ii. SET **maxCC** to **CC**.
 - iii. SET **minDist** to **dist**.
3. END IF
4. END FOR

2.1.1 Handling the case of peripheral lesions

Peripheral lesions located adjacent to dark regions outside the liver boundary contain intensity values below -10 HU in the baseline bounding box. In such cases, the boundary between the liver and low-value pixels representing air or liquid dominates the difference between the lesion and the liver parenchyma; this causes the template matching to find the shape of the liver boundary instead of the lesion. For such cases, we therefore stop at the MSE comparison between the slices and do not proceed to the lesion-level processing.

2.2 Constraining the Search Region in the Follow-Up Scan

The underlying methodology to segment the lesion in a FU image is based on region growing. We augment the region-growing algorithm by utilizing the information provided by the baseline scan. First, seed pixels are taken from the baseline scan and found on the new scan using the alignment step in Sec. 2.1. Second, a set of masks, derived from the mean-shift regions in the baseline scan, is used to support the region-growing algorithm in the FU scan by constraining the search area. The main assumption is that the tissues surrounding a lesion should not substantially vary between each pair of CT studies. We identify surrounding tissues in the baseline scan and use them to constrain the search area in the FU scan.

Mean-shift clustering is used on the baseline CT image containing the lesion to identify specific regions of interest (ROI) surrounding the marked lesion. From the mean-shift segmentation map, we extract regions that are spatially close to the (expert-marked) lesion and have different intensity from the lesion (these may correspond to blood vessels, the liver boundary, or other tissues) as well as regions that are spatially close and have very similar intensity to the lesion intensity (such as other proximal lesions or surrounding tissues with very similar intensity). We term the first set of regions, R1, and the second set, R2, respectively.

The process of extracting the set of regions is as follows. We first compute the pixel mean intensity $I_{\text{lesion mean}}$ and standard deviation I_{std} of the expert-marked lesion. To find R1 regions, we search for regions with an absolute difference between their mean intensity and the lesion mean intensity ($I_{\text{lesion mean}}$) that is greater than: $w_a \cdot I_{\text{std}}$ ($w_a = 3$), and at a location fewer than d pixels ($d = 30$) from the lesion. To identify R2 regions, we search for similar-intensity regions that contain pixels with intensity values that are fewer than 10 HU away from $I_{\text{lesion mean}}$, and fewer than d pixels ($d = 30$) from the lesion in the baseline image but more than five pixels away (to ensure we do not select mean-shift regions from the originally marked lesion).

Three masks were defined to constrain the search area, as follows: Mask Ma removes regions R1 and R2 found in the baseline scan, from the FU slice of interest (using the alignment of Sec. 2.1). We apply dilation to the mask to prevent loss of areas that are part of the lesion. An intensity mask, Mb, removes pixels from the slice of interest in the FU scan, where the absolute difference between the pixels and $I_{\text{lesion mean}}$ is more than $w_b \cdot I_{\text{std}}$ ($w_b = 2$). We apply a median filter and opening to the mask to reduce noise. Mask Mc removes the ribs and muscles from the search area using the mean-shift segmentation

of the slice of interest in the FU scan. Generating Mc entails the following: all regions with intensity levels below 75 HU or above 130 HU are removed (a coarse thresholding that includes the normal intensities of internal organs). The mask contour is approximated by applying opening, dilation, and using a convex hull. The final mask that constrains the search area is a combination of the three masks $M_{\text{tot}} = M_a \cap M_b \cap M_c$.

Examples of the masks generated by the algorithm are shown in Fig. 3. Figure 3(a) shows a cropped baseline CT scan with a lesion outlined by a radiologist. Figure 3(b) shows the corresponding R1 regions (green) and R2 regions (red) for that example. Figure 3(c) shows the same lesion in the FU CT scan, as outlined by a radiologist. The corresponding extracted mask Mb is shown in Fig. 3(d). Figure 3(e) shows the entire FU slice of interest and Fig. 3(f) shows the thresholded result, which is used to create mask Mc in Fig. 3(g). Figure 3(h) shows the result of applying mask Mc on the FU slice of interest.

2.3 Segmentation of the Lesion in the Follow-Up Scan

Next, adaptive region growing is performed on the FU scan. The goal of this stage is to be able to find very subtle lesion changes, while constraining the algorithm from diverging too widely within the liver and its surroundings. Two key points determine the behavior of the region-growing algorithm: the selection of an initial seed point, and the set of constraints used to support the termination of the region-growing process. Specifics of the steps taken in the current work are detailed next.

Seed pixels in the new scan are found as follows: starting with the lesion outline in the baseline scan, n -fold erosion is applied to the outline and the results are projected to the FU scan.¹⁴ An optimal threshold value t in the region-growing

Fig. 3 (a) Baseline CT scan (cropped), lesion outlined by radiologist (zoom-in view); (b) the baseline CT scan from (a). In green, R1 regions that include ribs and background; in red, R2 regions that include the intercostal muscles; (c) FU CT scan (cropped), lesion outlined by radiologist for ground-truth purposes (zoom-in view); (d) the FU CT scan from (c). In blue, areas masked out by Mb (the FU intensity mask); (e) FU CT slice of interest; (f) thresholding the mean-shift clustering output of (e), taking regions in the range of 75–130 HU; (g) the ribs and muscles mask, Mc, after applying morphological operators and convex hull on (f); and (h) the result of applying Mc on (e).

algorithm was chosen to maximize the compactness of the boundary,¹⁴ defined as $C(t) = A/P^2$, where A is the area and P is the perimeter, with two main conditions:

1. The difference between the mean intensity of the lesion in the baseline and the region found is less than 25 HU.
2. The relative difference between the intensity STD of the lesion in the baseline and the region found is below 50%.

The two conditions above were added to make sure that the segmentation process is done on the corresponding lesion by maintaining similar HU to the lesion in the baseline scan. The final segmentation is defined as the convex hull of the region, found by the region-growing procedure followed by opening and closing operations to obtain smooth boundaries.

2.3.1 Algorithm refinement for low contrast and highly variable lesions

Cases of low contrast can be automatically detected when no optimization is found for the threshold, t , or $90 \text{ HU} < I_{\text{mean}} < 110 \text{ HU}$ (the lesion has low contrast as compared to normal parenchyma values). In such cases, a refinement of the selected seed points is done as follows:

1. Thresholding using the mean intensity of the parenchyma P , estimated as the histogram peak value (in the range of 70 to 130 HU), and I_{mean} . For hyperdense lesions with I_{mean} higher than P , pixels whose intensity level is lower than $P + 15 \text{ HU}$ or higher than 200 HU are removed. For hypodense lesions with I_{mean} lower than P , pixels whose intensity level is higher than $P - 15 \text{ HU}$ or lower than $I_{\text{mean}} - 35 \text{ HU}$ are removed.
2. Applying template matching using the lesion in the baseline, finding the object with max cross correlation.
3. Applying n -fold erosion to obtain new seed points for the adaptive region growing.

Another refinement of the result is needed for lesions that have high variability expressed by a large set of regions (e.g., 12) within the lesion area in the baseline scan (following the mean shift). In this case, we use the mean-shift segmentation of the FU scan. Regions that are partially outside the boundary, as found in the final region-growing result, are added with their external part. Regions that overlap with the masks found in Sec. 2.2 are not added. Figure 4 presents an example of the different stages presented in our algorithm. In Figs. 4(a) and 4(b), a baseline CT scan and a FU CT scan are presented, respectively, with the lesion outlined by the radiologist. Figure 4(c) shows a zoom-in view of Fig. 4(b). Figure 4(d) shows an example of the baseline mask M_a (white), where the areas in black are masked out by M_a . Similarly, Fig. 4(e) shows the intensity mask M_b on the FU CT scan. Figure 4(f) shows mask M_c , the ribs and muscles removal mask, and 4(g) shows the final search area created by applying the masks on the FU slice of interest. Finally, Fig. 4(h) shows the output result: FU CT scan with lesion automatically outlined by applying adaptive region growing and mean-shift refinement.

2.4 Propagating in Three Dimensions

The final result of the previous step is a 2-D segmentation mask of the lesion within the most similar slice as found by our method. Two main considerations prompted us to extend this result to 3-D. The first is the goal of segmenting out lesions in 3-D, along with their 3-D properties (such as volume). The second consideration relates to algorithm evaluation needs. Human expert segmentation in the FU CT volume is conducted on a slice selected by the expert. The automated algorithm slice selection often fails to return the result in the exact same slice, which makes it imperative to propagate the extracted 2-D segmentation mask in the z -direction until we have both manual and automated segmentation results on one slice, after which we can use various contour matching evaluation measures for the evaluation.

The propagation process is based on similar steps as in the main segmentation algorithm. We start with the selected slice from Sec. 2.1 as the baseline slice with the lesion segmentation mask found by the proposed automated algorithm. Each neighboring slice is then treated as a FU slice for which we need to find the corresponding lesion mask. To support the process, mean-shift clustering is performed on each of the neighboring slices. The area within a new slice, in which the lesion is searched, is constrained by *a priori* context masks similar to the procedure discussed in Sec. 2.2. We propagate the known mask to the neighboring slice and rank each region in the mean-shift map, which falls in the propagated mask area, based on relative size and overlap with the propagated mask. Regions that have low likelihood of being part of the lesion are filtered out. We assume that if a region is part of the lesion it will have a large overlap area with the propagated mask. The final overall area should not be much larger than the area of the propagated mask. The relative weight of the size and overlap characteristics in the mean-shift region selection process is manually adjusted using sigmoid functions.¹⁰ The output of these functions is combined by addition. The initial segmentation includes all regions that fall in the propagated mask; regions with low ranking are removed one at a time until similar compactness (± 0.02), mean intensity ($\pm 25 \text{ HU}$), and STD ratio (< 0.5) as the lesion in the neighboring slice is achieved. The final segmentation of each slice is followed by convex hull, opening, and closing. If the neighboring (FU) segmentation mask is found to be 30% smaller than the propagated mask, we assume that the next segmentation mask will show a similar trend. In this case we add a condition that limits the next segmentation mask area to be smaller. This can be problematic if the lesion is not convex, so we also include a safety range of 50 pixels: if a decreasing trend is detected, we use the following condition: $\text{Area}_{\text{next}} \leq \text{Area}_{\text{previous}} + 50 \text{ pix}$.

2.5 Final Segmentation Refinement

The output of the previous step is a 3-D segmentation of the FU scan. To improve the segmentation quality, we added a case-specific refinement step. For each slice containing a segmentation mask whose LD is less than 27 mm, we use active contours (without edges)²⁰ employing the segmentation mask for initialization. The influence of the convex hull and the morphological operators on smaller lesions is greater, so the use of the active contour refinement step minimizes the FPs and can also add true positives in the case of expansion of the contour. We also use active contours to refine the segmentation of lesions whose

Fig. 4 (a) Baseline CT scan, lesion outlined by radiologist; (b) FU (FU) CT scan, lesion outlined by radiologist (for GT purposes); (c) FU CT scan (cropped), lesion outlined by radiologist (zoom-in view); (d) baseline mask Ma (white). In black are the areas masked out by Ma; (e) intensity mask Mb on FU CT scan; (f) Mc, the ribs and muscles removal mask; (g) final search area: created by applying (d), (e) and (f) on (b); and (h) output result: FU CT scan (cropped) with lesion automatically outlined by applying adaptive region growing and mean-shift refinement on (g) (zoom-in view).

intensity mean and STD under the segmentation mask satisfy $I_{\text{mean}} > 80 \text{ HU}$ and $I_{\text{std}} > 19 \text{ HU}$. These lesions have an intensity range that overlaps with the liver parenchyma; this refinement step helps distinguish between them using the segmentation mask. For large lesion segmentation masks, whose LD is larger than 50 mm, we use the geodesic active contour method²¹ biased toward shrinking inward, assuming that the initial segmentation boundary is outside the lesion. For lesions whose LD is less than 27 mm, another boundary refinement

step is applied. For each point in the boundary of the segmentation mask, we open a 3×3 ROI, checking which pixel intensities satisfy

$$|\text{ROI}_{\text{pix}} - I_{\text{mean}}| \leq w \cdot I_{\text{std}} \quad \text{and} \\ |\text{ROI}_{\text{pix}} - \text{ROI}_{\text{center}}| \leq \text{threshold},$$

where ROI_{pix} is a pixel in the ROI and outside the lesion boundary, $\text{ROI}_{\text{center}}$ is the center pixel of the ROI which is a part of

the lesion boundary, and I_{mean} and I_{std} are the mean intensity and STD of the lesion in the segmentation mask, respectively. Pixels satisfying the following condition are added to the segmentation mask by local convex hull in each ROI. The final result is smoothed by opening and closing. We used $w = 2$ and threshold = 20 HU in our method.

3 Experiments and Results

3.1 Database

The data used in the current work include CT scans from Sheba Medical Center, conducted in the period from 2009 to 2014. Different CT scanners were used with 0.7090–1.1719 mm pixel spacing, 1.25–5 mm slice thickness, 120 KVP, and different convolution kernels (GE—standard/CHST, Siemens—B31s, Phillips—B/C). The scans were selected and marked by a radiologist. They include 79 patients with overall 127 lesion FU cases. Twenty-seven patients had more than one lesion marked; five patients had more than one FU scan conducted at intervals of several months. The data were divided into two sets as follows: development set with 22 cases of 12 patients and an evaluation set with 105 cases of 67 patients. The data include various liver metastatic lesions derived from different tissues showing melanoma, breast cancer, colon cancer, neuroendocrine, and pancreatic cancer.

3.2 Algorithm Performance Evaluation

We computed the number of lesions that were correctly identified in the FU: 94 out of the 105 lesions were detected for an overall matching rate of 90%.

We evaluated the algorithm's segmentation performance using the Dice index, and calculated the sensitivity and positive predictive value (PPV) according to the following equations:

$$\text{Dice} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FN} + \text{FP}}, \quad (1)$$

$$\text{Sens} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (2)$$

$$\text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (3)$$

The results are shown in Tables 1 and 2. The tables present a comparison of the automated algorithm results with those of the human expert "reference standard" [see Fig. 4(c) for example]. We obtained an average Dice index of 0.83 ± 0.08 , an average

Table 1 Performance of the proposed automated algorithm as compared to the "reference standard" for the development set.

	Dice	Sensitivity	PPV
Mean	0.86	0.86	0.88
SD	0.06	0.11	0.08
Min	0.72	0.59	0.70
Max	0.96	0.99	0.99

Table 2 Performance of the proposed automated algorithm as compared to the "reference standard" for the evaluation set.

	Dice	Sensitivity	PPV
Mean	0.83	0.82	0.87
SD	0.08	0.13	0.11
Min	0.65	0.48	0.54
Max	0.96	0.99	1.00

sensitivity of 0.82 ± 0.13 , and an average PPV of 0.87 ± 0.11 for the evaluation set (Table 2).

The mean LD of all lesions on the baseline CT was 29.0 ± 15.3 mm and 29.6 ± 14.7 mm for lesions on the FU CT. When using the RECIST 1.1 criteria and comparing the sum of LDs per patient for each FU CT, 58 out of 66 cases (88%), resulted in the same RECIST diagnosis as the "reference standard." In three cases, which did not result in the same RECIST diagnosis and were rated by radiologists (they were a part of the data presented in Table 3), two were rated as good or very good and one as acceptable.

To evaluate the algorithm's performance for lesions of different sizes, we divided the lesions into three groups: small lesions with LD less than 2 cm, large lesions with LD more than 5 cm, and medium-sized lesions with LD in the range of 2–5 cm. As mentioned above, there are several classes of lesions that are more problematic for detection and segmentation. These include heterogeneous lesions and lesions with low contrast. We evaluate the results in the different groups next. Table 4 presents the algorithm's performance for 13 small lesions, 2 large lesions, 22 medium lesions, and 9 heterogeneous and low-contrast lesions that were classified by a radiologist for a set of 37 test cases.

Table 3 Segmentation ratings by expert radiologists for 37 test cases.

	–	–	0	+	++
Ratings	2	1	4	15	15
Percentage	5.4%	2.7%	10.8%	40.5%	40.5%

Table 4 Performance of the proposed automated algorithm as compared to the "reference standard" on small lesions, large lesions, medium lesions, and heterogeneous lesions or lesions with low contrast.

	Number of cases	Dice	Sensitivity	PPV
Small	13	0.79 ± 0.07	0.79 ± 0.15	0.84 ± 0.13
Large	2	0.89 ± 0.002	0.92 ± 0.001	0.90 ± 0.02
Medium	22	0.84 ± 0.07	0.83 ± 0.09	0.87 ± 0.1
Heterogeneous/low contrast	9	0.80 ± 0.08	0.82 ± 0.11	0.79 ± 0.16
Total	37	0.83 ± 0.07	0.82 ± 0.12	0.86 ± 0.11

Fig. 5 Bland-Altman plot showing the mean difference in longest diameter (LD) from baseline to manual and automated segmentation (solid middle line) and 95% limits of agreement (red lines) for 37 cases.

Next, we evaluated the system results by the human experts for the same set of 37 test cases. A two-reader consensus was provided for each segmentation result. A five-point scale was used: very bad, bad, acceptable, good, or very good. The results are summarized in Table 3. In 92% of the cases, the results were classified as acceptable or better (0, +, ++).

Agreement between our segmentation and the manual “reference standard” was assessed by concordance correlation²² and the Bland-Altman analysis²³ of the LD for 37 test cases. The concordance correlation between the “reference standard” measure of the LD and the automated segmentation measure was 0.98. The mean difference between the LD measures was -0.6 mm with 95% limits of agreement of 5.6 mm (Fig. 5).

Examples of the algorithm results rated as “very good” are shown in Fig. 6. A peripheral hypodense lesion is shown in Fig. 6(a). In this case, it is difficult to distinguish between

the metastatic lesion and adjacent muscle that has almost the same density. In Fig. 6(b), the lesion has hypodense and hyperdense parts. This makes it difficult for the adaptive region-growing algorithm to find the optimal threshold. In the case shown in Fig. 6(c), the lesion has a similar intensity as the parenchyma and is very small, which makes it difficult to distinguish between them. An example of an algorithm result that was rated as “very bad” is shown in Fig. 7.

3.3 Comparison to Possible Algorithm Variations

This section assesses three algorithm steps and evaluates their contribution to the results, as compared to possible variations.

3.3.1 Comparison to Xu et al.¹⁴

The current work was motivated by Xu et al.,¹⁴ in which abnormal lymph nodes in serial CT exams are segmented using context from the baseline image, followed by an adaptive region-growing algorithm. We next compare two algorithm steps of the presented methodology to the methodology described in Xu et al.¹⁴ For the next set of experiments, we used a randomly selected set of 59 test cases (out of the 127 lesion FU cases). One component is the alignment step, or registration. In Sec. 2.1, we started the matching process with a 2-D affine registration step to find the initial similar slice in the FU scan (as based on the marked slice in the baseline scan). It is worth inquiring why we did not use 3-D registration in this alignment step. The use of 3-D rigid registration as an initial step is part of the alignment method of Xu et al.¹⁴ In our first experiment, we assess the relative advantages of 2-D versus 3-D as follows: in the process of FU slice selection (Sec. 2.1), we use 3-D rigid registration instead of the 2-D affine registration. We used the Image registration toolkit (IRTK) and used the same parameters as in Ref. 14. Eventually, the parameters that gave the best results were manually established. These consisted of a hierarchical coarse-to-fine approach with control point spacings of 10, 5, and 2.5 mm. At each stage, the output transformation of the previous step was used as the starting point. The stopping condition for the optimization was either no further improvement in the sum of squared differences or reaching a maximum number of 10 iterations at each level. Table 5 summarizes the results of the comparison. It can be seen that in our experimentations, using 3-D rigid registration instead of the 2-D affine registration resulted in half of the detected lesions and reduced the matching percentage by half.

Fig. 6 Sample results of the segmentation algorithm: (a) peripheral hypodense lesion, (b) heterogeneous lesion, and (c) small and weakly hypodense lesion.

Fig. 7 Example of a poor segmentation result: heterogeneous lesion with calcifications. (a) Algorithm result and (b) outlined by radiologist.

Table 5 Matching rate results with three-dimensional (3-D) rigid registration (RR) and two-dimensional (2-D) affine registration.

	Number of cases	Detected cases	Matching rate [%]
3-D RR	59	28	47.5
2-D affine	59	55	93.2

The second experiment includes a comparison to the segmentation steps presented in Xu et al.¹⁴ We implemented a similar segmentation framework to the one described in their work, which includes four main steps: (1) identify tissues surrounding the lesion in a cross section of the baseline scan; (2) register the FU scan to the cross section in the baseline scan; (3) recognize tissue regions surrounding the lesion in the FU scan; and (4) segment the lesion by region growing and clustering. The segmentation framework used in this experiment is similar to ours except for three main components we added to their work. In comparing the methods, we note the following: (1) Mask Ma and mask Mb (Sec. 2.2) have similar properties to the masks presented in Xu et al.¹⁴ Mask Mc in our presented methodology was added to deal with peripheral lesions in order to mask out muscle tissue. (2) Two intensity-based conditions were added as part of the region-growing process (Sec. 2.3). The updates to the region-growing algorithm were made to ensure similar HU characteristics to the baseline lesion. (3) Algorithm refinements were added to support handling of special cases, such as peripheral lesions (Sec. 2.1), low contrast or lesions with high variability (Sec. 2.3), and small lesions (Sec. 2.5).

Table 6 summarizes the results of the comparison. Our proposed segmentation algorithm provided better results in terms of Dice index and PPV. The segmentation sensitivity results were higher using the Xu-based segmentation framework but resulted in a much lower PPV. Segmentation results of the two approaches for two cases are shown in Fig. 8. The top lesion [Figs. 8(a)–8(c)] is a peripheral lesion located adjacent to a

Table 6 Performance of the proposed automated segmentation algorithm and Xu et al. algorithm as compared to the “reference standard.”

	Dice	Sensitivity	PPV
Xu-based algorithm	0.76 ± 0.13	0.89 ± 0.14	0.71 ± 0.21
Proposed algorithm	0.84 ± 0.07	0.83 ± 0.11	0.87 ± 0.11

muscle with similar intensity. In this case, mask Mc masked out the muscle and improved the segmentation result. The bottom lesion [Figs. 8(d)–8(f)] shows the need for intensity constraints using the lesion from the baseline as reference for stopping the region-growing process in cases where strong differences in gray level values exist between the lesions.

3.3.2 Importance of the masks

In a third experiment, we evaluate the importance of the masks presented in Sec. 2.2. These masks are used to provide region constraints in the search for lesions. Table 7 shows the Dice index, sensitivity, and PPV with and without the masks. Note the increase in performance of around 5% (in PPV and Dice) with the use of the masks.

3.4 Parameter Optimization

The parameters used in the mean-shift algorithm were the spatial bandwidth (h_s) = 6 pixels (defines a spatial search window), the range bandwidth (h_r) = 10 HU (controls the gray-level intensity resolution), and the smallest allowed output region of 10 pixels to prevent oversegmentation. The additional parameters used in the method were intensity factors (w_a, w_b) for masks Ma and Mb and parameters related to the maximal intensity differences between the lesion in the baseline and the region found using adaptive region growing. The following optimization steps were applied on a smaller dataset consisting of 12 patients and 22 test cases. To find the optimized parameters

Fig. 8 Segmentation results for two lesions: (a, d) proposed algorithm result; (b, e) Xu et al. algorithm result; and (c, f) Lesion outline by radiologist.

Table 7 Performance of the proposed automated algorithm as compared to the “reference standard” with and without the masks used to constrain the search region for segmentation.

	Dice	Sensitivity	PPV
Without masks	0.78 ± 0.17	0.80 ± 0.17	0.82 ± 0.19
With masks	0.84 ± 0.07	0.83 ± 0.11	0.87 ± 0.11

Fig. 9 Mean Dice index as a function of (a) FU-intensity factor – w_b and (b) baseline intensity factor – w_a .

values, we varied h_r and h_s and calculated the performance for each value using the mean Dice index. Similarly, we varied the intensity factors (w_a, w_b) and calculated the mean Dice index (Fig. 9). Figure 9(a) shows that the mean Dice was relatively constant for $w_b \geq 2$. In a similar way, Fig. 9(b) shows the mean Dice was relatively constant for $w_a \geq 2.5$. Figure 10 indicates that the mean Dice index was relatively constant with h_r . In Fig. 11, a similar trend is shown when changing h_s . Hence, the algorithm is not sensitive to these parameters within the given test range. Based on these optimizations, we used $w_a = 3$, $w_b = 2$, $h_r = 10$ HU, and $h_s = 6$ in our experiments. The heuristics for the adaptive region growing and the 3-D propagation were manually established on the small dataset.

Our method uses nonrigid 2-D cubic B-spline registration for the final alignment. We compared this to the affine registration using a marker-based approach on a small dataset by marking five points inside the vertebra, three points inside the aorta, and 20 points inside and outside the liver adjacent to its boundary. We also measured the center of mass of the lesion in the baseline and its distance from the center of mass of the lesion in the FU after registration, using the expert's marks. For the affine registration, we computed an average of 1.7 errors in the location of the markers per case, with a mean distance of 7.8 mm from the center of mass of the lesion in the FU. The nonrigid registration had an average of 0.3 errors and a mean distance of 5.8 mm.

Therefore, the nonrigid registration was more accurate than the affine registration and more suitable for the final alignment.

4 Discussion

4.1 Matching Rate

The main difficulties involved in identifying the correct lesion are considered to be the presence of multiple metastases, the difference in contrast agents, and a considerable large change in the position of the patient between scans. The presence of multiple metastases may cause the algorithm to detect and segment a similar nearby lesion instead of the correct lesion. The difference in contrast agents and change in position between scans can cause the registration to fail and lead to erroneous slice detection. Overcoming all the difficulties listed above, our results show that the proposed method detected the correct lesion in the FU CT scan in 94 out of the 105 test cases, for a high matching rate of 90% using only a single baseline CT slice with an outlined lesion. Using the baseline slice instead of the entire 3-D CT examination¹⁴ provides a prior regarding important locations; i.e., it yields slices that are similar to the slice in the baseline rather than having to find the best alignment between two 3-D datasets with no prior. Our algorithm imitates the radiologists way of work by first searching for the corresponding slice in the FU scan, which should have similar

Fig. 10 Mean Dice index as a function of range bandwidth.

Fig. 11 Mean Dice index as a function of spatial bandwidth.

anatomical appearance, then searching for the lesion and finally measuring its size.

4.2 Segmentation

Table 2 shows the segmentation results. The Dice index for the 94 segmented lesions was 0.83 ± 0.08 . Table 4 shows that the segmentation measures for the small lesions (LD under 2 cm) were not significantly lower, with an average Dice index of 0.79 ± 0.07 . This makes sense since we had fewer pixels to take into account in the segmentation measure calculation. For large lesions (LD over 5 cm), the results were higher with an average Dice index of 0.89 ± 0.002 , which is still not significantly different than the measure for the entire lesion set. The slight increase is again due to the overall pixel amount present. For the heterogeneous and low contrast lesions classified by the radiologist, the average Dice index was 0.80 ± 0.08 , which is close to the average Dice index for the entire dataset. In Table 3, the expert's rating results show that in 92% of the cases the results were classified as acceptable or better and in 81% the results were classified as good or very good. Comparing the LD manually measured by the radiologist and automatically by our method yielded a concordance correlation coefficient of 0.98. Testing the RECIST 1.1 criteria using our method compared to "reference standard" indicated 88% success. In those cases, in which a discrepancy was found between the experts rating and the automated method (with the method generating an erroneous diagnosis according to RECIST 1.1 criteria), the experts still rated the automatic segmentation as acceptable or better. This implies that these cases are borderline cases.

4.3 Related Prior Art

The presented algorithm is an extended version of our previous work.¹⁵ In Ref. 15, the objective was to find the corresponding lesion in one slice and to provide a segmentation mask in that slice, ignoring the entire CT volume. In the current version, we added a 3-D propagation step and an improved alignment step to deal with the entire CT volume. Moreover, we extended our dataset from 27 cases to 127 cases, with 22 cases in the development set and 105 cases in the evaluation set.

The contribution of this method over the prior art includes the following:

- Most works presented in the literature include semi-automatic segmentation^{6,9,13} of lesions. In our method, the lesion is detected and automatically segmented in the FU CT.
- Other works may be automated but are shown with human refinements along the way (when needed).¹² The results shown in our work are presented with no human intervention.
- Our method is lesion specific. Given a specific lesion of interest and its segmentation mask in the baseline examination, the goal in the current work is to detect and segment the same lesion in a FU examination. This scenario is different from many other automated detection and segmentation methods,^{6,11} which are designed to work on all liver lesions.
- The method presented by Xu et al.,¹⁴ which is designed to work on lymph nodes, is augmented here to work on liver metastatic lesions. The main components that were

modified are the detection framework, the masking of irrelevant tissues, the adaptive region-growing process and specific features we added to deal with the large variability of liver metastatic lesions in size, shape, and texture, as well as change in lesion location due to liver movements between scans.

4.4 Limitations

Our algorithm relies on the assumption that tissues surrounding the lesion do not substantially vary between the baseline CT and the FU CT. Thus, we can use the information given in the baseline CT to locate and segment the lesion in the FU CT. This assumption can be violated if a patient has surgery between scans, if there is substantial change in the size of the lesion that causes the liver to deform, or any other condition where the surrounding structures change. In such cases, the registration may fail to find the correct slice and will produce an incorrect segmentation. Our algorithm was developed to segment lesions in 2-D. The segmentation mask is propagated to 3-D and we believe it is applicable to 3-D volumes, though validation requires further studies. Another limitation is that our algorithm cannot detect large morphological changes in the lesions such as merging, splitting, and disappearing. We used manual circumscription of the lesions in the baseline images, but our method is applicable to any type of initial segmentation in the baseline CT, such as automated segmentation. Our algorithm was developed to work on CT scans in the portal phase and was not tested on multiphase CT scans.

5 Conclusion

Our objective was to create an automated methodology for detection and segmentation of liver metastatic lesions in sequential CT scans, given a baseline segmentation mask. The initial results presented here look promising. We achieved high sensitivity and high Dice segmentation results as well as a high matching rate on a set of 105 lesion FU cases (67 patients). Using the RECIST diagnosis criteria, we obtained a success rate of 88%. Additionally, expert ratings of the automated system segmentation results were in strong agreement.

A key contribution of this work is the fact that the data were taken from a real-world clinical setting, with varying scanner properties and varying types of liver metastatic lesions including homogeneous, inhomogeneous, hypodense, and hyperdense lesions. We are currently expanding the method to 3-D segmentation and continuously increasing the dataset from the CT abdomen unit at Sheba Medical Center. It is our belief that working on real clinical setting data for a robust system has great potential value from a clinical perspective. The methodology presented comprised detection and tracking components that can be generalized to additional pathologies in the clinic.

References

1. G. T. Sica, H. Ji, and P. R. Ros, "CT and MR imaging of hepatic metastases," *Am. J. Roentgenol.* **174**(3), 691–698 (2000).
2. K. D. Hopper, K. Singapuri, and A. Finkel, "Body CT and oncologic imaging 1," *Radiology* **215**(1), 27–40 (2000).
3. E. Eisenhauer et al., "New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)," *Eur. J. Cancer* **45**(2), 228–247 (2009).
4. J. H. Moltz, M. Schwier, and H.-O. Peitgen, "A general framework for automatic detection of matching lesions in follow-up CT," in *IEEE Int.*

- Symp. on Biomedical Imaging: From Nano to Macro, 2009. ISBI2009*, pp. 843–846, IEEE (2009).
5. A. Charnoz et al., “Liver registration for the follow-up of hepatic tumors,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, pp. 155–162, Springer (2005).
 6. X. Deng and G. Du, “Editorial: 3D segmentation in the clinic: a grand challenge II—liver tumor segmentation,” in *Medical Imaging Computing Computer Assisted Intervention Workshop* (2008).
 7. A. Shimizu et al., “Ensemble segmentation using AdaBoost with application to liver lesion extraction from a CT volume,” in *Proc. Medical Imaging Computing Computer Assisted Intervention Workshop on 3D Segmentation in the Clinic: A Grand Challenge II*, New York (2008).
 8. J.-M. Kuhnigk et al., “Morphological segmentation and partial volume analysis for volumetry of solid pulmonary lesions in thoracic CT scans,” *IEEE Trans. Med. Imaging* **25**(4), 417–434 (2006).
 9. J. H. Moltz et al., “Advanced segmentation techniques for lung nodules, liver metastases, and enlarged lymph nodes in CT scans,” *IEEE J. Sel. Top. Signal Process.* **3**(1), 122–134 (2009).
 10. M. Schwier, J. H. Moltz, and H.-O. Peitgen, “Object-based analysis of CT images for automatic detection and segmentation of hypodense liver lesions,” *Int. J. Comput. Assist. Radiol. Surg.* **6**(6), 737–747 (2011).
 11. A. Militzer et al., “Automatic detection and segmentation of focal liver lesions in contrast enhanced CT images,” in *20th Int. Conf. on Pattern Recognition (ICPR), 2010*, pp. 2524–2527, IEEE (2010).
 12. J. H. Moltz et al., “Workflow-centred evaluation of an automatic lesion tracking software for chemotherapy monitoring by CT,” *Eur. Radiol.* **22**(12), 2759–2767 (2012).
 13. S. Keil et al., “Semi-automated measurement of hyperdense, hypodense and heterogeneous hepatic metastasis on standard MDCT slices. Comparison of semi-automated and manual measurement of RECIST and WHO criteria,” *Eur. Radiol.* **18**(11), 2456–2465 (2008).
 14. J. Xu et al., “Automated temporal tracking and segmentation of lymphoma on serial CT examinations,” *Med. Phys.* **38**(11), 5879–5886 (2011).
 15. A. B. Cohen et al., “Automatic detection and segmentation of liver metastatic lesions on serial ct examinations,” *Proc. SPIE* **9035**, 903519, (2014).
 16. Y. Cheng, “Mean shift, mode seeking, and clustering,” *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(8), 790–799 (1995).
 17. D. C. Wolf, “Evaluation of the size, shape, and consistency of the liver,” *Clinical Methods: The History, Physical, and Laboratory Examinations*, 3rd ed., H. K. Walker, W. D. Hall, and J. W. Hurst, Eds., Butterworths, Boston (1990).
 18. A. Myronenko and X. Song, “Intensity-based image registration by minimizing residual complexity,” *IEEE Trans. Med. Imaging* **29**(11), 1882–1891 (2010).
 19. D. Rueckert et al., “Nonrigid registration using free-form deformations: application to breast MR images,” *IEEE Trans. Med. Imaging* **18**(8), 712–721 (1999).
 20. T. F. Chan and L. A. Vese, “Active contours without edges,” *IEEE Trans. Image Process.* **10**(2), 266–277 (2001).
 21. V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *Int. J. Comput. Vis.* **22**(1), 61–79 (1997).
 22. L. I. Lin, “A concordance correlation coefficient to evaluate reproducibility,” *Biometrics* **45**(1), 255–268 (1989).
 23. J. M. Bland and D. Altman, “Statistical methods for assessing agreement between two methods of clinical measurement,” *Lancet* **327**(8476), 307–310 (1986).

Avi Ben-Cohen is a PhD candidate at Tel Aviv University. He received his MSc degree in biomedical engineering in the field of medical image processing. His current research interests include medical image processing, liver lesions analysis, computer-aided diagnosis, and machine learning.

Eyal Klang is a senior radiologist at Chaim Sheba Medical Center. He works in the body imaging unit of the Radiology Department. His current research interests include computer-aided diagnosis of liver lesions and oncological imaging.

Idit Diamant is a PhD candidate at Tel Aviv University. She received her MSc degree in biomedical engineering in the field of musculoskeletal biomechanics. Her current research interests include medical image processing, liver lesions analysis, computer-aided diagnosis, and machine learning.

Noa Rozendorn: Biography is not available.

Michal Marianne Amitai was trained at the Department of Diagnostic Imaging at Chaim Sheba Medical Center, Tel Hashomer, Israel, 1989 to 1994. She has been a senior staff radiologist since 1994. She received a research and clinical fellowship in cross-sectional abdominal imaging at the University Health Network—University of Toronto Canada, 1999 to 2001. She has been the director of the CT service since 2010 and also author and coauthor of 40 peer-reviewed publications. She is a lecturer at Tel Aviv University.

Hayit Greenspan is a tenured associate professor at the Biomedical Engineering Department, Tel-Aviv University. Her research areas include medical image processing, computer-aided diagnosis, deep learning, and image retrieval. She received BS and MS degrees in EE from the Technion and a PhD in EE from CALTECH with a postdoc at U.C. Berkeley. In 1997, she joined TAU and founded the medical image computing lab. She has many publications and is the inventor of several patents. She is an associate editor for the *IEEE Transactions on Medical Imaging*.