

Automated seeding for ultrasound skin lesion segmentation

Péter Marosán^{a,b,*}, Klára Szalai^c, Domonkos Csabai^{a,b}, Gergely Csány^{a,b}, András Horváth^a, Miklós Gyöngy^{a,b}

^a Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Práter u. 50/A, 1083, Budapest, Hungary

^b Dermus Kft., Kanizsai u. 2-10 C/11, Budapest, Hungary

^c Department of Dermatology, Venerology and Dermatocology, Semmelweis University, Mária u. 41, 1085 Budapest, Hungary

ARTICLE INFO

Keywords:

Skin ultrasound
Computer vision
Automated layer segmentation
Automated lesion localization
Seed selection

ABSTRACT

The segmentation of cancer-suspicious skin lesions using ultrasound may help their differential diagnosis and treatment planning. Active contour models (ACM) require an initial seed, which when manually chosen may cause variations in segmentation accuracy. Fully-automated skin segmentation typically employs layer-by-layer segmentation using a combination of methods; however, such segmentation has not yet been applied on cancerous lesions. In the current work, fully automated segmentation is achieved in two steps: an automated seeding (AS) step using a layer-by-layer method followed by a growing step using an ACM. The method was tested on images of nevi, melanomas, and basal cell carcinomas from two ultrasound imaging systems ($N = 60$), with all lesions being successfully located. For the seeding step, manual seeding (MS) was used as a reference. AS approached the accuracy of MS when the latter used an optimal bounding rectangle based on the ground truth (Sørensen–Dice coefficient (SDC) of 72.3 vs 74.6, respectively). The effect of varying the manual seed was also investigated; a 0.7 decrease in seed height and width caused a mean SDC of 54.6. The results show the robustness of automated seeding for skin lesion segmentation.

1. Introduction

1.1. Motivation

The last 30 years have shown a great increase in the incidence of skin diseases such as skin cancer: for instance, in the past 10 years, melanoma incidence in the US has increased by more than 54 percent [1–3]. Information about the geometric extent and shape of a skin lesion is necessary for surgery planning and also provides features for lesion classification [4–6]. For both these applications, it is advantageous to have a fully automatic, robust, and precise method for skin lesion segmentation of dermatological ultrasound images. In the following subsections, an overview is provided of relevant existing segmentation approaches, followed by the aims of the current work.

1.2. Overview of existing approaches

1.2.1. General ultrasound segmentation

Numerous segmentation methods have been applied to B-mode ultrasound images in the past decades [7,8]. In general, the classical segmentation approaches can be classified into active contour methods (ACM), clustering, and threshold-based methods [9]. In addition to

classical approaches, there is a great increase in popularity of using deep learning methods [10–13] for ultrasound segmentation. These techniques can be considered as an extension of classical clustering-based approaches; however, they require large training data sets in addition to validation and testing data sets [8]. Similarly, statistical methods making use of pixel distributions can be argued to be a generalization of threshold-based methods that imply an underlying distribution.

Following the above classification, the overview below of existing skin ultrasound image segmentation techniques groups them into ACM, layer-by-layer methods (generally employing a mixture of methods), and statistical methods. While ACMs require an initial seeding step, layer-by-layer methods are fully automatic but have not yet, up to now, been used to seed lesions. Lastly, statistical approaches are a fully automatic methods with great potential; however, their validation remains to be carried out.

1.2.2. Skin segmentation using active contour models (ACM)

A common approach for segmentation of skin lesions in ultrasound images is to employ a manual seeding step, where the lesion

* Corresponding author at: Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Práter u. 50/A, 1083, Budapest, Hungary.

E-mail addresses: marosan.peter@itk.ppke.hu (P. Marosán), klaraszalai@gmail.com (K. Szalai), domonkos.csabai@dermusvision.com (D. Csabai), csany.gergely@itk.ppke.hu (G. Csány), horvath.andras@itk.ppke.hu (A. Horváth), gyongy.miklos@itk.ppke.hu (M. Gyöngy).

<https://doi.org/10.1016/j.ultras.2020.106268>

Received 27 March 2020; Received in revised form 30 September 2020; Accepted 30 September 2020

Available online 5 October 2020

0041-624X/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

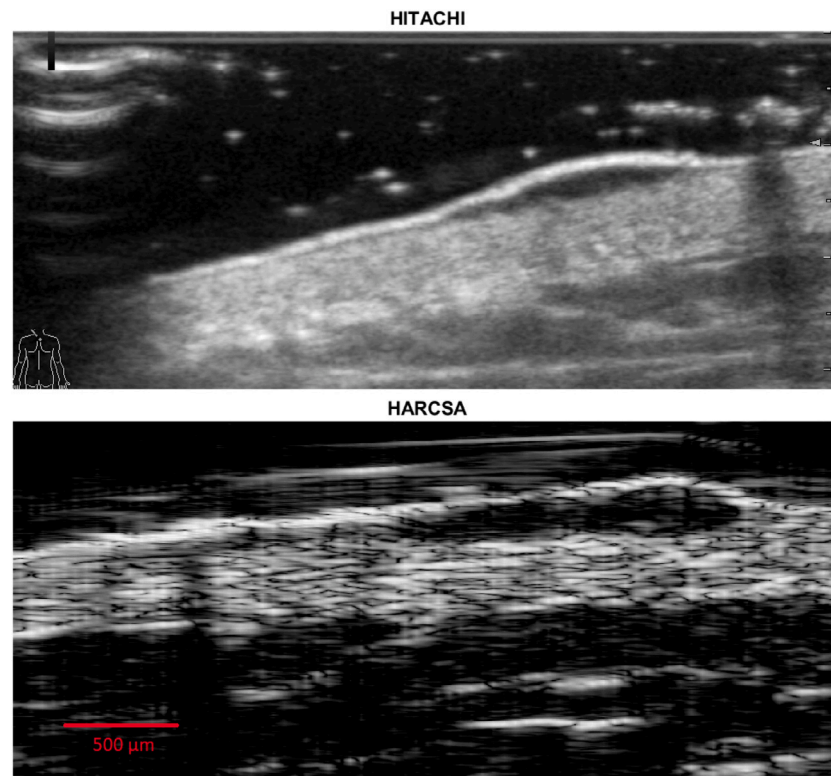


Fig. 1. Illustration of skin lesion ultrasound images of both examined datasets. While the top image was recorded by a general commercial device (from Hitachi), the bottom one was recorded using our own custom device (called HARCSA). The bottom, HARCSA image has been cropped before display, for simple comparison. It is noticeable that the records of the datasets are highly different with regard to resolution, brightness, contrast and speckle noise. The main regions, such as the above-skin region, epidermis, dermis, and the lesion are all easily identifiable qualitatively. Some specific artefacts such as the scatter noise, echogenic shadow, gaps in features, bubbles, and hair are also visible in the images, for which care needs to be taken during image processing.

is first selected using a point or rectangle, followed by automatic segmentation.

In semi-automatic lesion detection, a manually chosen inner rectangle or box provides a seed for the segmentation algorithm. However, variations in choosing the box may impact segmentation accuracy.

These methods have generally use level-set segmentation with special additional terms fitted for the current task [14–17].

1.2.3. Layer-by-layer segmentation using combination methods

Automated skin lesion localization can rely on the extraction of the main features of skin images, the segmentation of main skin layers. A large literature exists for detecting skin layers using a combination of methods [18–24]; however, none of the cited works use full skin layer detection for lesion detection. Recently, the work of Anxionnat [9] performed a fully automated segmentation for acne vulgaris. The algorithm is based on the assumption that an acne vulgaris lesion causes a curvature on the surface of the skin. For the present aim of detecting cancerous skin lesions, this assumption may not always work. Unfortunately, the detection accuracy of the lesion is not presented, merely the accuracy of the skin layer detection.

Statistical approaches

Pereyra et al. [25] presented an automated skin lesion segmentation technique, using a spatially coherent generalized Rayleigh mixture model. Its advantage lies in the fact that no initialization or supervised training is needed; however, a region of interest (ROI) does need to be provided, in this sense requiring manual intervention from the user. Moreover, no segmentation accuracy results are provided.

1.3. Aims of current work

The current work combines the advantages and results of previous approaches. The approach for detecting skin layers is similar to a combination of methods used in Section 1.2.3. This layer-by-layer segmentation provides an initial seeding step for lesion detection, which is followed by a simple Chan–Vese [26] ACM model. Although arguably, more complex ACMs have since been used (as in Section 1.2.2), the Chan–Vese model is a simple model requiring less parameters that acts as a proof of concept for the approach.

Our proposed seeding method is compared to a manual seeding protocol by applying the same ACM on the two seeds. Both the segmentation accuracies of the automated and optimally chosen manual seeds are compared, and the sensitivity of the segmentation accuracy in manual seeding to deviations from the optimal.

2. Materials and methods

2.1. Overview

This section first describes the ultrasound images used in the current work. The section also details the two-step localization and segmentation framework, where the first step, the so-called lesion seeding is presented in the first place. Both the proposed, automated seeding (AS) approach and the reference manual seeding (MS) protocol are detailed. While the seeding step localizes and roughly delimits the lesion region inside the ultrasound image, a commonly used, automated growing step is also demonstrated briefly, which uses the outputs of the two seeding procedures to retrieve an accurate final region boundary. Finally, the current section describes the evaluation framework to compare the performance of the proposed method to the reference.

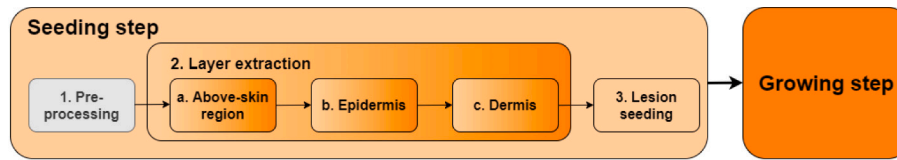


Fig. 2. Flow diagram of fully-automated skin lesion detection method. The zeroth step is pre-processing of the images which handles the differences between the images from distinct data sets. Lesion seeding is preceded by the examination of the skin structure by the extraction of all the main layers. While the main novelty of the current method is the robust localization (seeding step) of the skin lesions with the help of a two dimensional heat map, precise boundary delimitation of the lesions (growing step) also stands as a post-processing step.

2.2. Data acquisition

In order to demonstrate the robustness of the proposed algorithm, images were collected from two different sources. The first dataset (from now on referred to as HITACHI dataset) was recorded using a commercial high-frequency ultrasound imager (HI VISION Preirus with 5–18 MHz EUP-L75 transducer connected to Hitachi Preirus, Hitachi, Tokyo, Japan) transmitting in high frequency mode. The second dataset (from now on referred to as HARCSA dataset) was obtained using a custom system based on a manually scanning single element transducer (V317, Olympus, Tokyo, Japan; diameter: 6.3 mm; $f\#$: 3) driven at 8 MHz. For further details about the custom system, see Csányi et al. [27,28]. Both datasets were recorded during in vivo human skin experiments at the Department of Dermatology, Venereology and Dermatonecology, Semmelweis University, Budapest, Hungary. The experiments were carried out as part of an ethically approved study with informed consent from the participating patients for the anonymized use of the data in research and publication.

30 representative ultrasound images were selected from each dataset so that they contain an equal number ($N = 10$) of nevi, basal cell carcinomas (BCC) and melanomas. A training and a test set were constructed based on the collected 60 records. The training set contained 40 images, 6 or 7 instances from all lesion types and each data sets. The remaining 20 images were used for validating the proposed method. One typical skin lesion ultrasound image from each dataset can be seen in Fig. 1, which shows the layers of the skin, as well as a melanoma lesion.

2.3. Seeding step

Seeding is the first step of the two-step localization and segmentation framework. The aim of this task is to find an initial point or a set of points inside the lesion region, which can be the basis of a well-chosen automated segmentation algorithm to find the region borders.

While the proposed AS approach gives an automated and deterministic solution to extract the set of seed points, the reference MS method is highly dependent on the manual selection, which causes variance in the accuracy of boundary delimitation.

2.3.1. Proposed automatic seeding

In the following section, the AS approach is outlined briefly. The current subsection focuses on the main steps of the proposed algorithm, while the technical details are discussed later, in Appendix.

The algorithm consists of a pre-processing step, layer extraction steps, and lesion detection within the dermis, see Fig. 2. Each layer extraction task consists of two main steps. First, as an initializing step an intensity-based clustering or thresholding method combined with geometric prior information is performed to localize the region of interest (ROI) in the image. Next, an active contour-based refinement step and a post-processing method uses morphological operations to segment the exact borders. (Fig. 3)

The pre-processing step is responsible for making records more similar to each other in the aspect of histogram distribution, image resolution and noise level. The pre-processed images are ready for the extraction of the main regions, such as the above-skin region, the epidermis and the dermis. Extraction of the latter regions rely on the

previously delimited ones, and the above mentioned two-step segmentation framework is applied for all three cases. The first step stands for a robust, but not so accurate localization of the required region. The general pipeline of this step contains linear filtering, multilevel image-thresholding or clustering, and biological information extraction about the regions size, shape and echogenicity. The second step uses the rough output, morphological operators, and geodesic active contour models to refine the boundaries of the different layers of the skin.

The lesion localization algorithm is searching for the lesion at the epidermis–dermis interface. Clustering on the adaptive histogram equalized images provides the list of all potential lesion blobs. While most of the potential blobs can be pre-filtered based on their size and pixel intensity, post-processing steps are applied on the remaining ones, before doing the lesion selection. Selecting the lesion from all the potential blobs is based on a two dimensional heat map. The vertical term of the heat map takes into account the distance of the blob from the epidermis–dermis interface, while the horizontal term takes into consideration the skin border interface curvature, the size, and the echogenicity of the blobs (Fig. 4). The blob selected in this way will give the seeding mask of the lesion.

2.3.2. Reference manual seeding

Semi-automatic lesion segmentation approaches, which are already detailed in Introduction apply manual seeding step before the growing step. A few methods require a single seed point, others start with a rectangle (or any polygon) as an initial seeding blob. [15,16]

In the first set of results (Section 3 - Segmentation accuracy: Automated vs optimal manual seeding), the seed rectangle is chosen to be the largest area rectangle (LAR) inside the ground truth lesion mask, which is optimal in the sense that it covers the largest possible area within the lesion and is therefore expected to converge quickly and accurately to the true mask. The accuracy of this expectation can be checked in the Segmentation accuracy analysis of manual seeding selection section. Human experts commonly mark the seed blob to be completely inside the lesion area and be as large as possible.

The methodology and effect of deviating from this optimal seed is described in Section 4 - Segmentation accuracy analysis of manual seeding selection. Reference seed rectangle selection methods and further analysis for measuring the influence of seeding point for final segmentation results are detailed here.

2.4. Growing step

The validation of the proposed AS method was carried out by applying a common growing step on its output and on the reference MS method output masks. The chosen growing algorithm bases on the active contour model (ACM) designed by Chan & Vese [26]. This approach is commonly mentioned or used [14] as a reference for validating the performance of other works in the topic.

Selection of the initial lesion seed is followed by the use of the Chan–Vese ACM on the given seed blobs with the exact same parameters in case of both seeding approaches. Since the growing step following the AS and MS methods is the same, the influence of the initial localization algorithm can be assessed, making the performance of the AS and MS methods directly comparable.

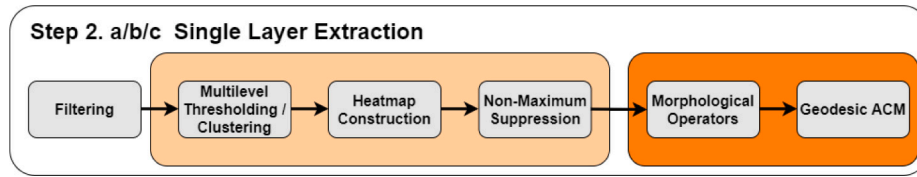


Fig. 3. Flow diagram for single layer extraction. Although every single region is extracted in a different way, a general process flow can be defined which describes all of them. After applying different filtering methods, a special extension of the Otsu-method is used to separate different regions and tissue types based on echogenicity. Subsequent steps extract all the potential blobs and selects one as an initial mask based on the constructed heat map. With the help of morphological operators and the use of an active contour model, the final mask is obtained.

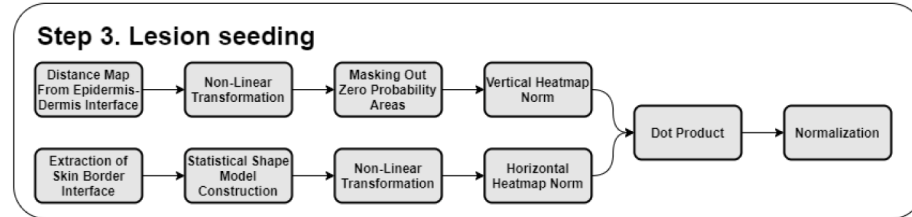


Fig. 4. Construction of lesion seeding heat map. After extracting all potential blobs, a two-component heat map is constructed based on geometrical and biological information. The base of the vertical one is the position of the lesion seed, which can be found around the interface of the epidermis and dermis layers. Horizontal heat map uses lesion property-based features among the curve of epidermis lane.

2.5. Evaluation metrics

Evaluation of the AS and MS methods-based final masks is carried out using ground truth lesion masks marked by a radiologist with more than 5 years of experience in the field of skin lesions for all the 60 images. The performance of the detection algorithms was evaluated using four conventional metrics, as listed below.

Sensitivity (S) calculates the true positive rate of detected mask pixels over the reference area, but does not punish false negative pixels.

$$S = \frac{|A \cap GT|}{|GT|}, \quad (1)$$

where A notes the binary mask of the detected area, and GT stands for the mask of the ground truth area.

Precision (P) calculates true positive rate over the automatically segmented area, but does not measure how effectively can the algorithm find the whole lesion.

$$P = \frac{|A \cap GT|}{|A|}, \quad (2)$$

Sørensen–Dice coefficient (D) measures sensitivity and precision at the same time, and it also handles the imperfection of both metrics.

$$D = \frac{2|A \cap GT|}{|A| + |GT|}, \quad (3)$$

In contrast to the above unitless measures, the Mean Absolute Distance (MAD) calculates the average distance between the reference and the segmented contour.

$$MAD = \sum_{x \in \delta A} \frac{d(x|GT)}{2ctr(\delta A)} + \sum_{x \in \delta GT} \frac{d(x|A)}{2ctr(\delta GT)} \quad (4)$$

where $d(x|GT)$ is the distance of point x to the ground truth lesion mask, and $ctr(\delta GT)$ is the number of pixels in the boundary δGT .

3. Segmentation accuracy: Automated vs optimal manual seeding

The current section is structured as follows: First, the layer-by-layer extraction of the proposed AS algorithm is presented and discussed. Next, the quantitative results of the AS method-based segmentation accuracy are compared to the masks, computed, based on the reference MS method. Finally, the computational times of the algorithm are discussed.

3.1. Validation of layer extraction

Using a representative image from both the HARCSA and HITACHI image sets (Section 2.2 - Data acquisition), the step-by-step functioning and qualitative performance of the fully automated segmentation procedure (Fig. 2) is presented. Fig. 5 shows the layer extraction of the above-skin region and epidermis; Fig. 6 presents the layer extraction of the dermis; while Fig. 7 illustrates the lesion detection step. A discussion of the observed performance follows each step.

In the detection of the above-skin region and epidermis (Fig. 5), the Multi-threshold Otsu method is shown to provide good discrimination between the skin and above skin region, allowing the binarization to separate the two regions well. Following appropriate blob extraction (Section 2.3.1 - Proposed automatic seeding), the skin region becomes even more separated from other regions. This is incorporated into a vertical heatmap that reflects the ever-decreasing probability of a region being below the skin. This allows the extraction of the above skin layer (red, Final masks, Fig. 5), followed by the extraction of the epidermis, discussed below.

The extraction of the epidermis is shown to work well, validating the approach of using prior information about the epidermis structure. The epidermis consists of anywhere between 50 to 100 cell layers with significantly high echogenicity, and having a specific thickness of $0.1 \text{ mm} \pm 20 \mu\text{m}$. Relying on this prior biological and visual information, pixels from the highest intensity bin, provided by the multi-threshold Otsu method, yield all the potential blobs, while the extracted skin border and epidermis thickness information could determine a reliable heat map for blob selection. The second, refinement step used a morphological closing operator with an arbitrary structuring element to fill all the gaps of the epidermis stripe. The above selected ACM was applied with low degree of smoothness and minor contour growing outwards. These parameters and the small number of iterations ensure that the algorithm closely follows the finer details of the boundaries.

Delimitation of the dermis layer is illustrated in Fig. 6. The initial mask was constructed based on visual information and biological knowledge of the properties of the examined layer. The parameters for the initial mask were set to always encompass the lesion even at the cost of overestimating the dermal region.

Before the final step of lesion detection, the previously discussed heat maps are combined with the above-mentioned blobs to detect the most likely lesion (Section 2.3.1 - Proposed automatic seeding).

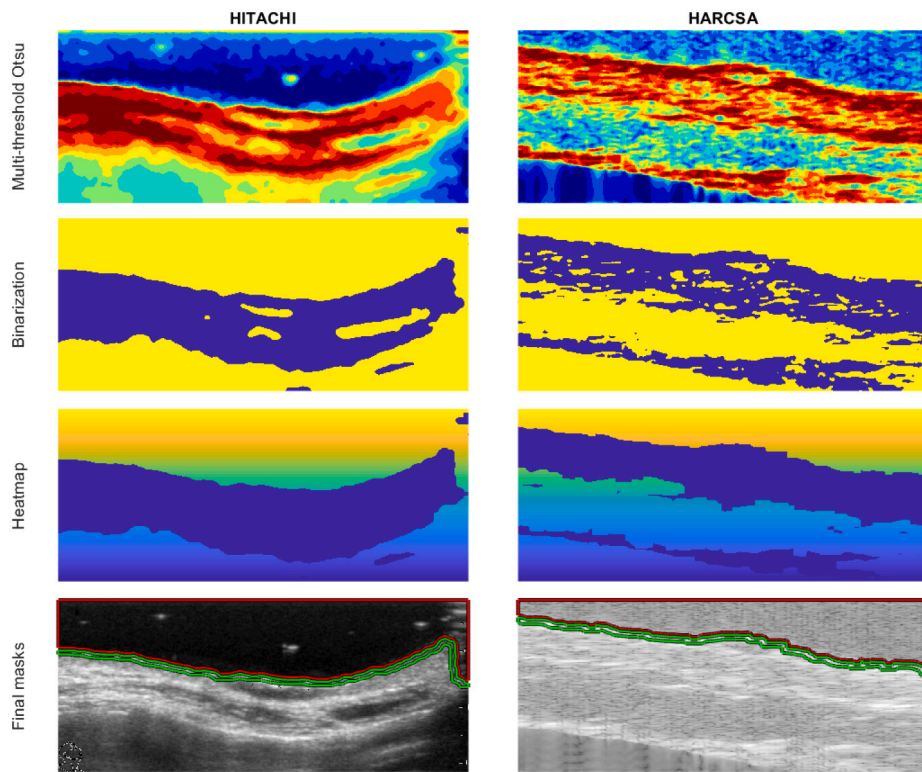


Fig. 5. Extraction of the region above the skin and the epidermis. Current method uses the echogenic contrast between the region above the skin and epidermis to extract their interface. Geometric structure of ultrasound skin records, biological knowledge about the width and components of the epidermis and other assumptions help to distinguish these regions. Images show results of multi-threshold Otsu-method, present binarization based on the described factors, illustrate heat map for detection of the region above the skin, and give the final masks for the region above the skin and the epidermis.

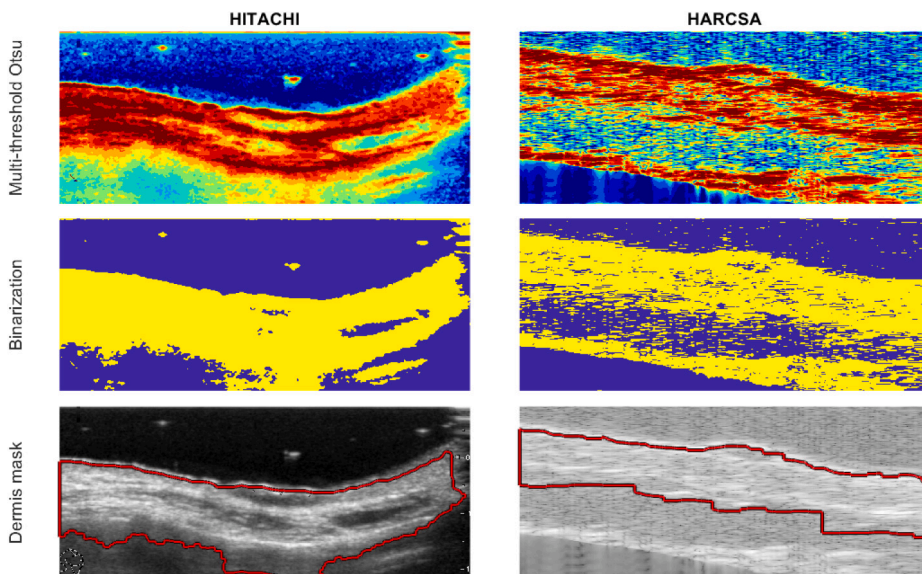


Fig. 6. Segmentation of the dermis layer. The dermis layer contains the skin lesion, making its robust segmentation crucial. After the extraction of the epidermis–dermis interface, the next step is to find the lower interface, which is the dermis–hypodermis interface. Use of biological, geometrical, structural, and echogenic knowledge of the layer results in a robust segmentation algorithm for emphasizing the dermis.

This process is depicted in Fig. 7. In all, the whole detection process incorporates biological, geometrical, and visual information to yield robust detection. Fig. 8 illustrates the robustness of the constructed heat map — while a simple curvature term would fail in some cases where the curvature is not significant enough, the extended term set handles all cases.

3.2. Evaluation of lesion detection

Table 1 gives a summary of the performance of the proposed AS algorithm using the metrics listed in the Evaluation metrics section. Current table presents the metrics, evaluating only the seeding step, without applying any growing algorithm on it. Table 1 shows that even the worst Sørensen–Dice coefficient out of all lesions was above 45.0%;

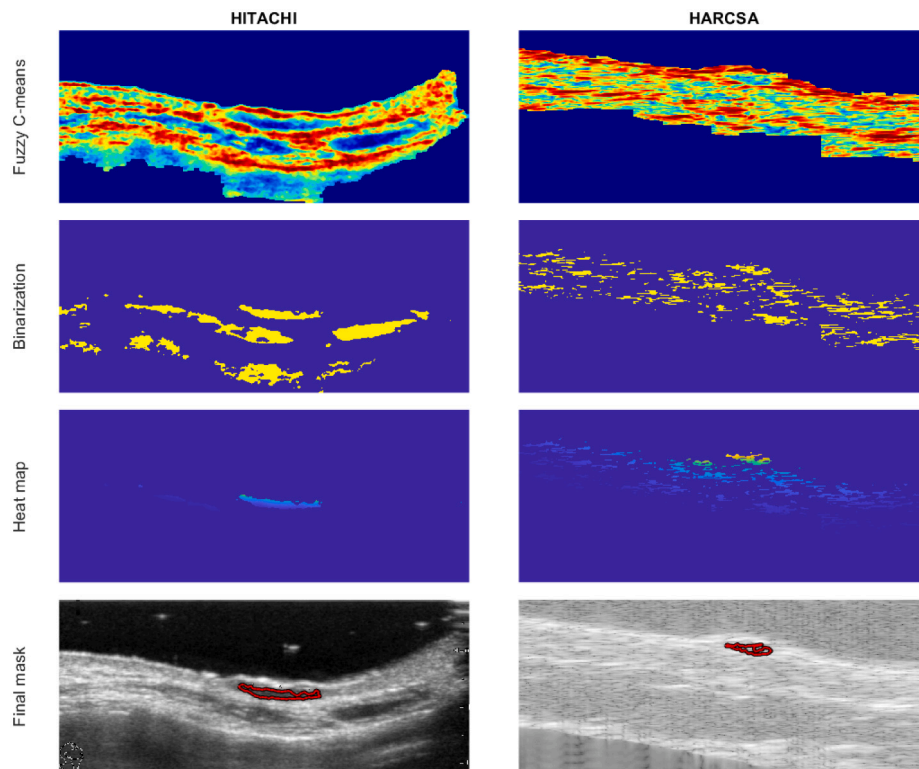


Fig. 7. Heat map construction for skin lesion detection. The united region of the epidermis and dermis can contain several possible lesion blobs. To find the true lesion in this set of blobs the above described two-component heat map is used. Current figure shows an example for the current process. The first row shows contrast enhancement and fuzzy C-means clustering method applied on the region of the epidermis and the dermis, the second one presents the dermis with potential blobs, the third one illustrates the horizontal and vertical heat maps together projected on blob set, while the last one notes the output masks of the united heat map-based lesion detection.

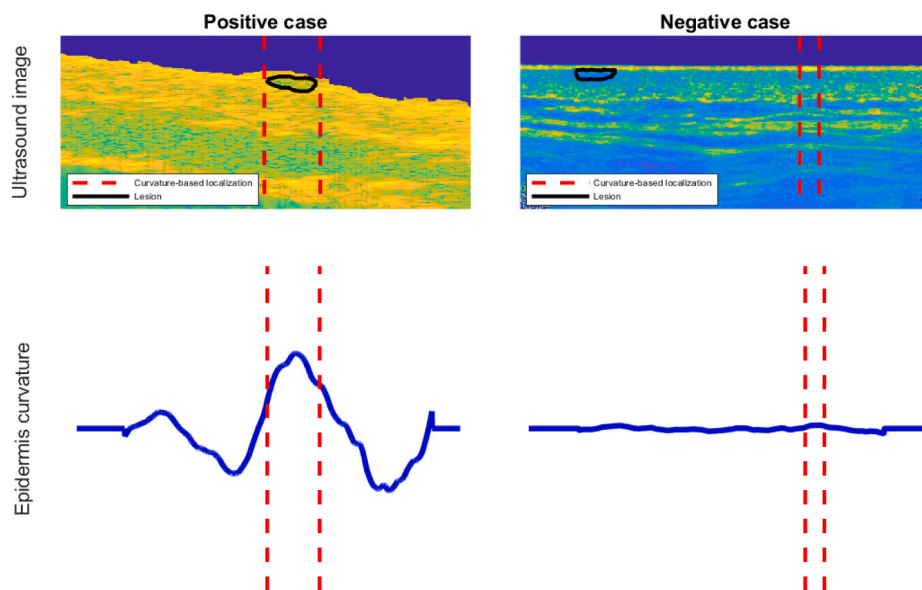


Fig. 8. Issue of epidermis curvature. Horizontal heat map counts with the curvature of the epidermis lane. There are several cases where this factor is a useful and effective term of the heat map, but automated detection algorithm cannot rely on only this. Images of the first column show an example where the curvature of epidermis can determine the horizontal position of the lesion, on the other hand second column images present a counterexample.

the mean Sørensen–Dice coefficient for HITACHI images was 59.1%, and for HARCSA it was 69.1%. The centre of mass for all ground truth masks were inside the detected lesion blobs for all records. These results demonstrate that lesion localization was successfully accomplished for all the test images thanks to the precise skin structure processing and lesion detection method.

The geodesic active contour model was used on both the outputs of the proposed AS method and the MS reference algorithm with the optimized parameters, detailed in [Appendix. Table 2](#) presents the accuracy values which show that the AS performance closely follows that of the MS method, the latter requiring extensive manual input. In terms of the Sørensen–Dice coefficient, a mean similarity of 72.28%

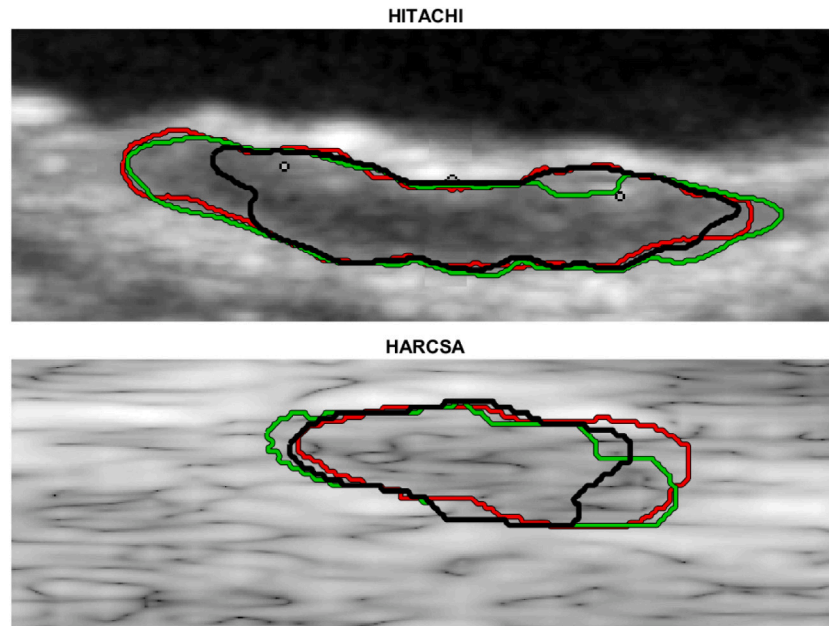


Fig. 9. Comparison between lesion masks of own fully automated seeding (AS) approach, manual seeding (MS) method and ground truth. The green line stands for the lesion border provided by proposed AS algorithm, the black line means the lesion detected by the MS method, while the red line is the ground truth manually marked by highly trained expert. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

Measurement of lesion blob detection performance. Table shows similarity of detected blob masks without any further boundary segmentation approach to the ground truth masks. While the output masks give a relatively high mean value for Sørensen–Dice coefficient, the minimum values also show that the own automated seeding (AS) approach can localize lesion blobs in the case of all the examined ultrasound images.

Similarity index	HARCSA	HITACHI	ALL
Sensitivity	64.7% (35.9%–91.9)	47.2% (39.6%–76.3%)	56.0% (29.9%–91.9%)
Precision	85.9% (60.3%–100%)	91.3% (67.5%–99.6%)	88.6% (60.3%–100%)
Sørensen–Dice	69.1% (51.9%–88.6%)	59.1% (45.6%–73.2%)	64.1 (45.6%–88.6%)%
<i>MAD</i> [μm]	326 (1829–77)	336 (978–29)	331 (1829–29)

Table 2

Comparison of own automated seeding-based (AS) approach with the manual seeding-based (MS) reference method using the LAR (largest area rectangle) blobs as initial masks. Sørensen–Dice coefficients show that the own AS-based approach can closely reach the performance level of the MS-based reference algorithm, when the seeding rectangle is selected to be optimal.

Dataset	AS-based	MS-based (LAR)
HARCSA	71.82%	73.73%
HITACHI	72.74%	75.51%
ALL	72.28%	74.62%

is achieved for the AS method, compared to 74.62% for the reference in optimal case. This is an acceptable result, considering the fully automated nature of the former. Furthermore this value was achieved, when the reference method used the largest area rectangle (LAR) inside the ground truth lesion mask as a seeding blob.

Fig. 9 illustrates the final lesion masks gained by the proposed AS-based approach, the MS-based reference method and the ground truth. Finally, images of Fig. 10 give a clear and comprehensive view of the detection and segmentation of all the main regions of a typical skin ultrasound image.

3.3. Running time

The algorithm was originally implemented in MATLAB (MathWorks, Inc. Natick, Massachusetts, USA), with additional MEX files, built from C++ source files, regarding to reduce the running time of the proposed approach.

It was also an aim to reduce the computational cost as much as possible. The running time for an average image from the HARCSA dataset (size: 3000 $\times$ 1000 pixels) was about 15 s, and for an image from the HITACHI set of records (size: 900 $\times$ 400 pixels) was less than 10 s (Intel Core i7-7500U CPU @ 2.70 GHz, 16 GB RAM). Table 3 details the computational cost of the proposed and reference algorithm. While thresholding methods, filtering related algorithms and morphological operations run fast, geodesic active contour model related operations require the significant part of the running time. Note that if only detection is required without accurate segmentation, the running time goes down to a mere 4–7 s. Measurements on both data sets showed, that it takes an average of 3 s for an expert to find the lesion and draw a seed rectangle into it.

4. Segmentation accuracy analysis of manual seeding selection

To analyse the effect of deviating from an optimal rectangle seed (Section 2.3.2), a number of variations in the seed are introduced, as described below.

4.1. Half-Area Rectangle (HAR) seeds

HAR uses a sliding window inside the previously chosen LAR, where the width and height of the seeding rectangles are 0.7 times the width and height of the LAR. These HARs have approximately half the area of the LAR. As before, the growing step is applied to each seed and compared with the true lesion segmentation. The mean segmentation accuracy – in terms of the Sørensen–Dice coefficient – was then calculated over all windows.

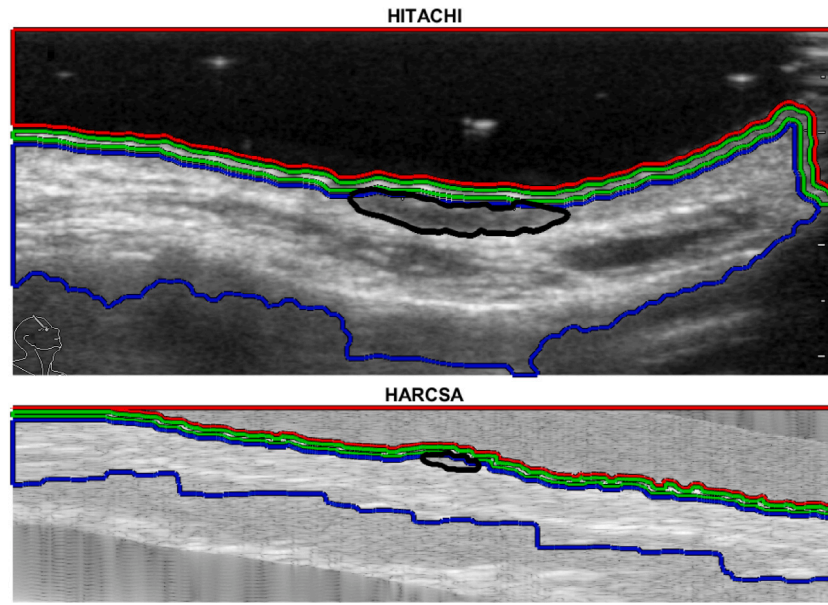


Fig. 10. All the detected and segmented main regions of the skin ultrasound images. The red line is the border of the above-region skin, the green line stands for the epidermis, the blue line marks the dermis, while the black line means the detected lesion. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3

Mean computation cost of main parts of the proposed fully automatic AS-based and reference semi-automatic, MS-based methods. It is noticeable, that automated detection of the lesions is not as time consuming as the following active contour model-based segmentation, which is common in both algorithms. It is also important to mention that the automated detection algorithm requires almost as small amount of time as the manual initialization of the reference algorithm.

	Proposed approach			Reference method		
	Auto seeding	Growing	Sum	Manual seeding	Growing	Sum
HARCSA	7 s	20 s	27 s	3 s	21 s	24 s
HITACHI	4 s	12 s	16 s	3 s	12 s	14 s

While the mean AS-based segmentation accuracy was above 72% for both data sets, the mean accuracy for HAR was only around 54.60% (50.9% for HARCSA and 58.28% in case of HITACHI images), with a standard deviation of 11.80%. These values show that the proposed automated approach can beat the sub-optimal manual seeding. It should be noted that such sub-optimal seeding is deemed realistic in the case of a manual operator, as operators may wish to avoid having the rectangle touch the lesion outline in case it accidentally crosses the outline.

4.2. Point-like Area Rectangle (PAR) seeding

While HAR simulates the conservative choice of a manual user, PAR seeds show the effect of starting the seed from a point-like area, shedding light on the relative differences between starting the ACM from different locations inside the lesion.

In case of PARs, the smallest interpretable (25×25 pixel) seeding rectangle was selected and slid along the bounding box of the lesion ground truth mask in a pixel-wise way.

Fig. 11 illustrates the input and output of the PAR technique. While there is an area on the bottom image with higher Sørensen–Dice coefficients, the corresponding area on the upper image does not contain any feature that is visible by naked eyes and can indicate this phenomenon. This suggests that there is no immediately clear reason for the high variation in segmentation accuracy results with the changing location of the seed, and therefore no relevant guidance that can be given to the operator choosing the manual seed as to which areas to include or avoid.

4.3. Fixed Upper-Left Corner (FULC) seeds

In addition to the above investigations, it would be useful to examine how changing the area of the seed affects results, while keeping a corner of the seed fixed. To this end, the FULC technique set the upper-left corner of the LAR as the upper-left (UL) corner of the FULC seeds, and varying the location of the opposite (lower-right) corner across every pixel in the LAR.

Fig. 12. illustrates the FULC technique. Generally, there is a trend for the Sørensen–Dice coefficient to grow as the seeding rectangle area grows. However, the growth is not entirely monotonic, let alone linear. As an extreme example, even a change of one pixel in the location of the lower-right corner may severely impact segmentation accuracy (Figs. 13 and 14). Similarly to the previous subsection, there is nothing immediately apparent on the B-mode image that would warrant such a change in accuracy.

5. Discussion and conclusions

A fully-automated approach was presented to localize lesions in skin ultrasound images. The proposed method detects lesions of two datasets recorded by entirely different type of devices and different recording parameter sets. Segmentation accuracies (as defined by the Sørensen–Dice coefficient) arising from automated seeding (AS) closely approaches the performance of manual seeding (MS) (72.3% vs 74.6%, correspondingly) when the MS is optimally chosen. Moreover, it has been shown that seed selection within the lesion can greatly influence segmentation accuracy, with values as low as 54.6% possible; therefore, the results from AS the automated seeding algorithm are deemed especially robust.

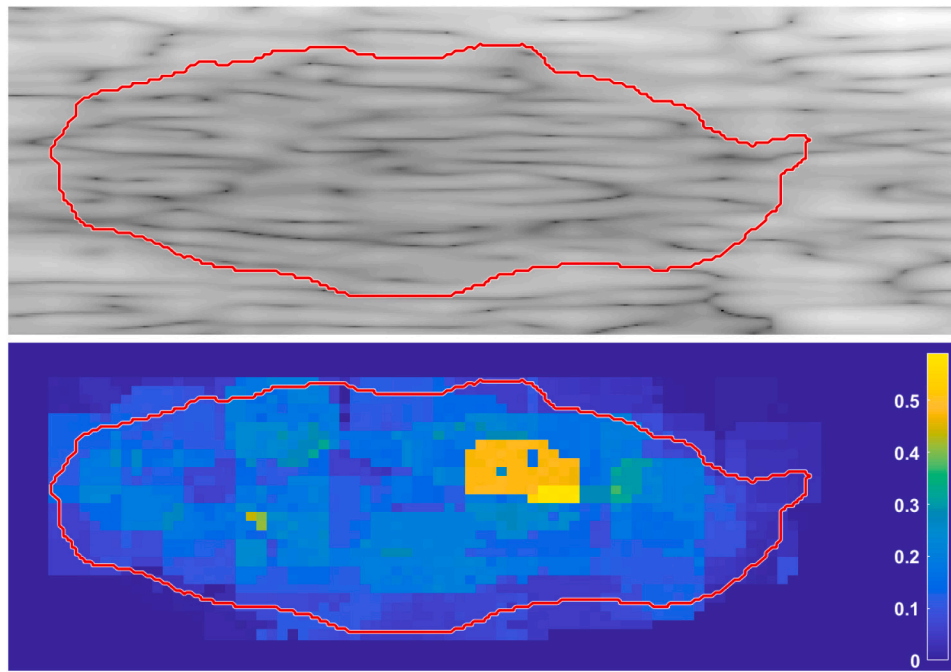


Fig. 11. The influence of the manual seeding rectangle selection was measured by the Point-like Area Rectangles (PAR) technique. Current technique used a sliding window with a fixed size of 25×25 for examine how are the final segmentation results depend on the selection of the seeding rectangle. The upper image represents the input ultrasound record of the lesion and its surroundings while the bottom one is a heat map, which displays the Sørensen-Dice coefficients calculated by final lesion masks, based on the seeding rectangle around the corresponding pixel. The ground truth boundary of the examined lesion is marked by a red line in both images. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

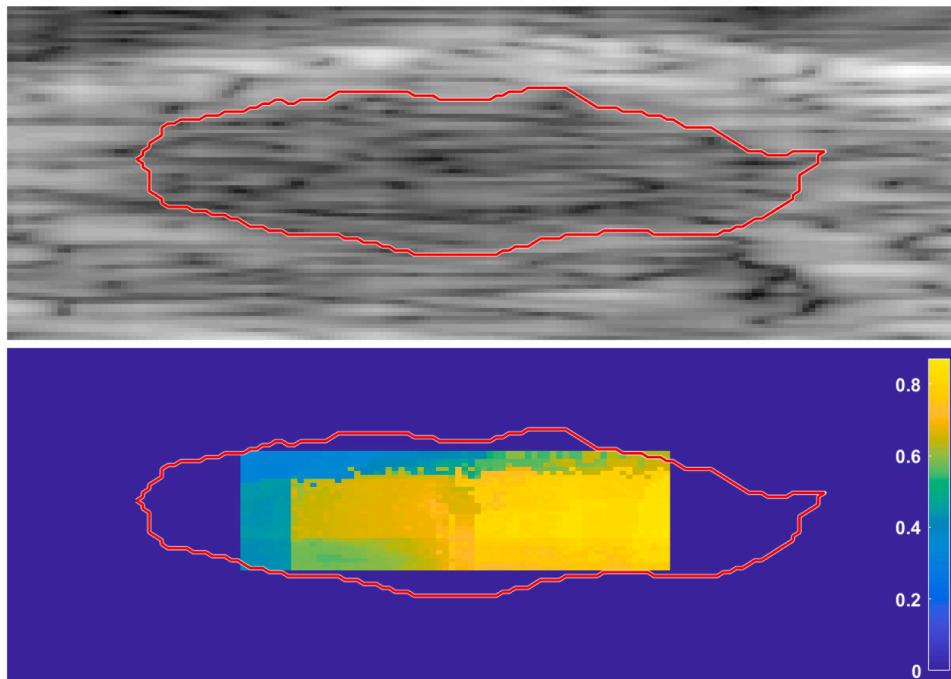


Fig. 12. Altering-Sized Rectangles with Fixed Upper-Left Corner (FULC) technique examines the sensitivity of the final lesion mask in aspect of the selection of the seed. Upper image shows the copped original ultrasound image. The bottom one represents the Sørensen-Dice coefficient-based heat map of the appropriate region. All pixel values are Sørensen-Dice values calculated by comparing the ground truth makes and the final segmentation mask, which was constructed based on a seeding rectangle with the corresponding pixel as a lower-right corner point. The lesion boundary is marked on both images by red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The authors note that the automatic detection algorithm can be combined with any semi-automatic lesion segmentation algorithm that requires an initial seed. Therefore, future work should apply more complex ACMs that have been shown to achieve better results on skin

lesion segmentation, such as in the work of Sciola et al. [14–16] or Nguyen et al. [17].

While the current work places emphasis on the detection of lesions at risk of being cancerous, other skin lesions such as scars, burn injuries

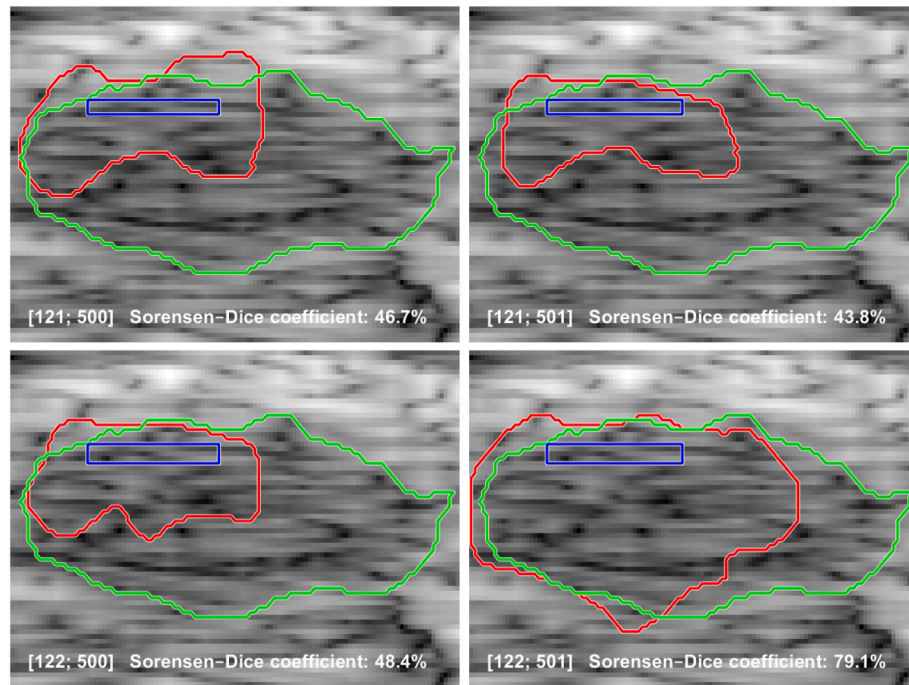


Fig. 13. FULC technique presents that even one pixel difference in the position of the lower-right corner of the seeding rectangle can cause significant changes in the output segmentation mask. On the four images, the lower-right corner points of the seeding (blue) rectangles are neighbouring ones, with coordinates [121, 500], [121, 501], [122, 500], [122, 501]. Red lines mark the output masks of the active contour model, based on the appropriate seeding rectangles, while green lines represent the ground truth boundary. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

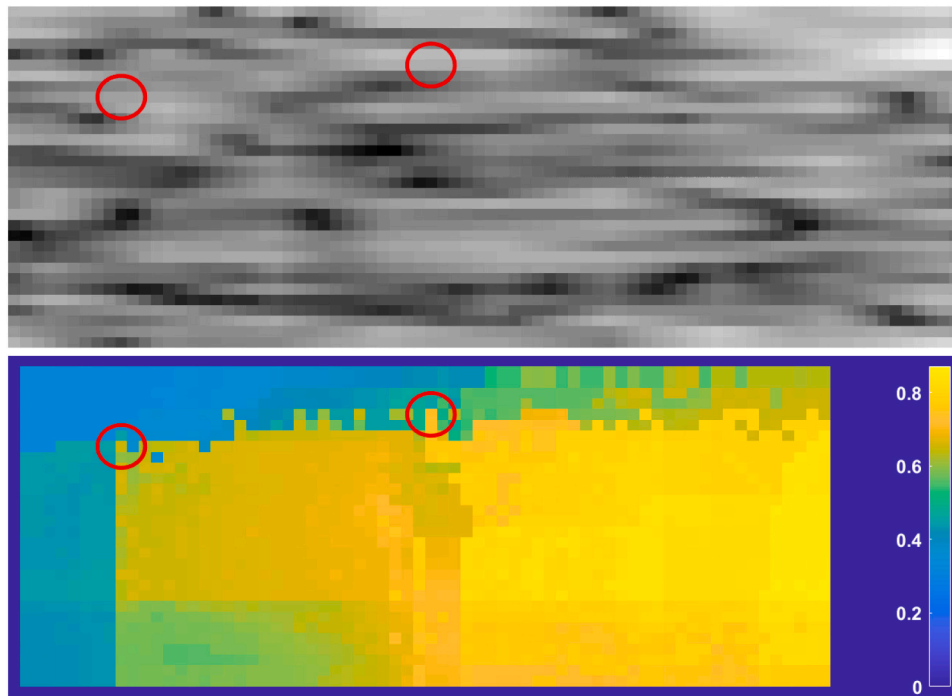


Fig. 14. The bottom image was made by the FULC technique based on the upper one in a pixel wise way. Red circles are located around neighbouring pixels where the output value of the Sørensen-Dice coefficient jumps in a sharp way, but no certain (visible by naked eyes) features belong to them on the ultrasound images.

Table 4

Parameters for single layer extraction. Although the extraction of the main skin layers uses the same process flow, the input parameters, heat maps of the single methods differ from each other. (In case of the vertical heat map construction parameters, δ_{skin} notes the above skin–epidermis interface, δ_{epi} stands for the epidermis–dermis interface and $dist$ gives the Euclidean distance from these lines.) Average skin layer thickness values are converted from μm to pixel number, based on image resolution.

Processing step	Parameter	Region above skin	Epidermis	Dermis
Pre-processing	Smoothing filter	Median	–	Median
	Iteration number	10	–	5
	Kernel size	[7,7]	–	[3,3]
Initialization	Multilevel Otsu bin	1st	5th	3 middle bin
	Avg skinlayer thickness	–	$thick_{epi} = 100 \mu\text{m}$	$thick_{dermis} = 2 \text{ mm}$
Vertical heat map	Equation	$I_{height} - row(x),$ $x \in I$	0, $if x \in Mask_{aboveskin}$	0, $if x \in Mask_{aboveskin}$ $\cup Mask_{epidermis}$
			$I_{height} - dist(x, \delta_{skin}),$ $if dist(x, \delta_{skin})$ $\leq thickness_{epi}$ $I_{height} -$ $exp(dist(x, \delta_{skin})),$ $if dist(x, \delta_{skin})$ $> thickness_{epi}$	$I_{height} - dist(x, \delta_{epi}),$ $if dist(x, \delta_{epi})$ $\leq thickness_{dermis}$ $I_{height} -$ $exp(dist(x, \delta_{epi})),$ $if dist(x, \delta_{epi})$ $> thickness_{dermis}$
Morphological operators	Type & num. of it	closing (10 it.)	dilation (5 it.)	closing (10 it.)
	Structural element	'rectangle', [3,3]	'rectangle', [2,4]	'rectangle', [3,3]
Active Contour Model	Model	GAC	GAC	GAC
	Smooth factor	0.7	2	0.7
	Contraction bias	–0.1	–0.1	–0.1
	Max num of iterations	200	100	200

Table 5

Parameters for lesion detection and boundary delimitation. The above mentioned training set of ultrasound images was used for optimization.

Processing step	Parameter	Value
Initialization	Fuzzy C-mean clustering	2 lowest value clusters from 5
Vertical heat map weights	$\alpha(curvature)$	0.4
	$\beta(blobsize)$	0.3
	$\gamma(blobecho)$	0.3
Active contour model	Model	Chan–Vese
	Smooth factor	0.7
	Contraction bias	–0.29
	Max. num. of it.	500

may also benefit from robust layer extraction methods, so hopefully this approach can be used in other fields also.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Janka Hatvani and Ákos Makra for providing feedback on the manuscript; Eszter Katalin Bakacs for giving useful advice in the topic of ultrasound related noise elimination; and Krisztián Füzési for designing and producing the casing for the HARCSEA device used for clinical data acquisition.

Appendix. Proposed seeding step appendix

Detailed description of the automated seeding step is presented here with all the applied parameters.

A.1. Pre-processing

The aim of this work was to develop an algorithm to localize lesions in general from different datasets using only a single predefined parameter set. To this end, number of image pre-processing steps were required.

Contrast enhancement was achieved by applying a global histogram equalization operator (T_{histeq}) and images were re-scaled ($T_{rescaled}$) to share the same resolution grid. The global mean of the gradient magnitude was calculated to measure the noise level (σ) of the records. A two-dimensional Gaussian smoothing kernel was used to achieve a common target noise level σ target for all images ($T_{denoise}$).

The composition of the aforementioned transformations was used during layer extraction and boundary delimitation:

$$I_{proc} = T_{rescale} \circ T_{histeq} \circ T_{denoise}(I), \quad (5)$$

where I is the input image and I_{proc} is the pre-processed output image.

A.2. Single layer extraction

For each three layers – the above-skin region, the epidermis, and the dermis – the segmentation was carried out in two steps. The same techniques were used in all three cases, but the parameters were different for each (Table 4).

A general pipeline illustrating the process of the interface extraction for a single region can be seen in Fig. 3. First, repetitive use of the median filter with an image resolution dependent kernel size made the images more homogeneous. Next, a multilevel image-thresholding version of the Otsu method was applied on the smoothed images. Based on empirical observation of the intensity histograms, 5 threshold levels were calculated and used for binarization of the regions with different echogenicity properties. After applying the thresholding on the images, a layer-specific vertical heat map helped to identify the initial mask of the single regions. For each successive layer, heat maps

were constructed, using the previously detected regions above them. Table 4 contains the equations of the heat maps constructed based on geometrical and biological knowledge.

Shadowing in the dermis causes these regions to be not detected as part of the dermis; one of the tasks of the second (refinement) step is to fill these undetected regions using morphological operators, such as a flood-fill operation. Final boundary delimitation was achieved by using further morphological operators and a well-parameterized geodesic active contour model [29] on the extracted interfaces.

A.3. Lesion detection and segmentation

Lesions occur at the epidermis–dermis interface, so the joint mask of these regions was selected as a ROI. An adaptive histogram equalizer was used on the ROI to amplify the contrast between hypoechogenic and hyperechogenic areas. Afterwards, a fuzzy C-means clustering method [30] was used to get all the potential blobs which can be classified as the lesion in demand. While all blobs below a given size (0.1mm^2) and above a certain intensity value were filtered out, the remaining ones were made more compact by filling the holes inside them using a flood-fill operation on the background pixels.

A two-dimensional heat map was constructed for lesion detection (Fig. 4). The vertical component applies biological knowledge about the lesion location, such as the earliest stage skin lesions can be found at the interface of the epidermis and dermis or in the upper part of the dermis as a rule. Applying non-linear transformations and masking out the zero probability areas gave the vertical heat map Eq. (2).

$$H_{\text{vertical}} = \begin{cases} \exp\left(\frac{\max(\text{dist}(x, \delta_{\text{epidermis}}) - \text{dist}(x, \delta_{\text{epidermis}}) * 10)}{\text{numel}(I)}\right), & \text{if } x \in \text{Mask}_{\text{dermis}} \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where $\delta_{\text{epidermis}}$ is the epidermis–dermis interface, dist gives the Euclidean distance from this line, and $\text{numel}(I)$ is the total number of pixels in the image. The pixel weights are zero outside of the dermis region.

The horizontal component of the heat map uses the combination of several features. In a few special cases, the upper part of the skin has a concave protrusion at the lesion position, which can be detected by computer vision. However, according to radiologists, this curvature is not a reliable feature generally (illustrated in Fig. 8), so two further terms were also introduced to make horizontal lesion detection more robust. Both factors are dependent of a few attributes of the potential lesion blobs and causes probability peaks on the horizontal axis. First, the calculated size of peaks is directly proportional to the size of the blobs and second, it is inversely proportional to the mean echogenicity of the blob. Weights of the terms are optimized based on the training set of ultrasound images. The horizontal heat map can be calculated as follows:

$$H_{\text{horizontal}} = \alpha \text{term}_{\text{curvature}} + \beta \text{term}_{\text{blob_size}} + \gamma \text{term}_{\text{blob_echo}} \quad (7)$$

where $\text{term}_{\text{curvature}}$ takes epidermis curvature into account, $\text{term}_{\text{blob_size}}$ examines the size of potential lesion blobs and $\text{term}_{\text{blob_echo}}$ rewards low-echogenicity blobs. α , β and γ are the weights of the used terms.

A normalized joint heat map models all the biological, geometrical and visual knowledge by computing the elementwise product of the horizontal and vertical components.

$$H = \text{norm}(H_{\text{horizontal}}) \cdot \text{norm}(H_{\text{vertical}}) \quad (8)$$

All extracted blobs were labelled with the sum of the corresponding heat map pixels and the highest labelled one was selected to be the initial mask of the lesion. The detected blob mask represents only an inner part of the lesion itself because original blobs are chosen to get low pixel intensity values which, usually occur only at the less echogenic part of lesions. To get the final lesion mask, further

processing steps were required, such as a Gaussian image smoothing and use of an active contour model. Various level-set function based active contour models (ACM) are presented in the literature for semi-automatic skin lesion border segmentation; however, Chan–Vese model [26], an ACM without edges is the most frequently mentioned reference method for the proposed task. Parameters of the model, such as smooth factor and contraction bias, were calculated using a two-step scaling optimization process for reducing computational cost. Running the ACM with given parameters starts on the down-scaled version of the images and proceeds on the original resolution images to get the final lesion masks. Table 5 contains all the optimized parameters used during lesion detection and segmentation step.

References

- [1] American Cancer Society: Cancer Facts and Figures 2019, American Cancer Society, Atlanta, GA, 2019.
- [2] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2019, CA Cancer J. Clin. 69 (1) (2019) 7–34.
- [3] A.M. Glazer, R.R. Winkelmann, A.S. Farberg, D.S. Rigel, Analysis of trends in US melanoma incidence and mortality, JAMA Dermatol. 153 (2) (2017) 225–226.
- [4] M. Samimi, A. Perrinaud, M. Naouri, A. Maruani, E. Perrodeau, L. Vaillant, L. Machet, High-resolution ultrasonography assists the differential diagnosis of blue naevi and cutaneous metastases of melanoma, Br. J. Dermatol. 163 (3) (2010) 550–556.
- [5] D. Csabai, K. Szalai, M. Gyöngy, Automated classification of common skin lesions using bioinspired features, in: 2016 IEEE International Ultrasonics Symposium (IUS), IEEE, 2016, pp. 1–4.
- [6] K. Andrėkutė, G. Linkevičiūtė, R. Raišutis, S. Valiukevičienė, J. Makštieņė, Automatic differential diagnosis of melanocytic skin tumors using ultrasound data, Ultrasound Med. Biol. 42 (12) (2016) 2834–2843.
- [7] J.A. Noble, D. Boukerroui, Ultrasound image segmentation: a survey, IEEE Trans. Med. Imaging 25 (8) (2006) 987–1010.
- [8] K.M. Meiburger, U.R. Acharya, F. Molinari, Automated localization and segmentation techniques for B-mode ultrasound images: A review, Comput. Biol. Med. 92 (2018) 210–235.
- [9] A. Anxionnat, Segmentation of high frequency 3D ultrasound images for skin disease characterization, 2017.
- [10] M.H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, A.K. Davison, R. Martí, Automated breast ultrasound lesions detection using convolutional neural networks, IEEE J. Biomed. Health Inform. 22 (4) (2017) 1218–1226.
- [11] M.H. Yap, M. Goyal, F.M. Osman, R. Martí, E. Denton, A. Juetter, R. Zwiggelaar, Breast ultrasound lesions recognition: end-to-end deep learning approaches, J. Med. Imaging 6 (1) (2018) 011007.
- [12] Z. Zhuang, N. Li, A.N. Joseph Raj, V.G. Mahesh, S. Qiu, An RDAU-NET model for lesion segmentation in breast ultrasound images, PLoS One 14 (8) (2019) e0221535.
- [13] R.J. van Sloun, R.R. Wildeboer, C.K. Mannaerts, A.W. Postema, M. Gayet, H.P. Beerlage, G. Salomon, H. Wijkstra, M. Misch, Deep learning for real-time, automatic, and scanner-adapted prostate (zone) segmentation of transrectal ultrasound, for example, magnetic resonance imaging–transrectal ultrasound fusion prostate biopsy, Eur. Urol. Focus (2019).
- [14] B. Sciolia, P. Delachartre, L. Cowell, T. Dambry, B. Guibert, Multigrid level-set segmentation of high-frequency 3D ultrasound images using the hellingier distance, in: 2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA), IEEE, 2015, pp. 165–169.
- [15] B. Sciolia, L. Cowell, T. Dambry, B. Guibert, P. Delachartre, Segmentation of skin tumors in high-frequency 3-D ultrasound images, Ultrasound Med. Biol. 43 (1) (2017) 227–238.
- [16] B. Sciolia, P. Delachartre, L. Cowell, T. Dambry, B. Guibert, Improved boundary segmentation of skin lesions in high-frequency 3D ultrasound, Comput. Biol. Med. 87 (2017) 302–310.
- [17] K.L. Nguyen, P. Delachartre, M. Berthier, Multi-grid phase field skin tumor segmentation in 3D ultrasound images, IEEE Trans. Image Process. 28 (8) (2019) 3678–3687.
- [18] J.-M. Lagarde, J. George, R. Soulié, D. Black, Automatic measurement of dermal thickness from B-scan ultrasound images using active contours, Skin Res. Technol. 11 (2) (2005) 79–90.
- [19] F. Lacarrubba, A. Tedeschi, B. Nardone, G. Micali, Mesotherapy for skin rejuvenation: assessment of the subepidermal low-echogenic band by ultrasound evaluation with cross-sectional B-mode scanning, Dermatol. Therapy 21 (2008) S1–S5.
- [20] Y. Gao, A. Tannenbaum, H. Chen, M. Torres, E. Yoshida, X. Yang, Y. Wang, W. Curran, T. Liu, Automated skin segmentation in ultrasonic evaluation of skin toxicity in breast cancer radiotherapy, Ultrasound Med. Biol. 39 (11) (2013) 2166–2175.

- [21] Z.M. Wawrzyniak, M. Szyszka, Layer measurement in high frequency ultrasonography images for skin, in: *Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments 2015*, vol. 9662, International Society for Optics and Photonics, 2015, p. 96621T.
- [22] I. Bryjova, J. Kubicek, K. Molnarova, L. Peter, M. Penhaker, K. Kuca, Multiregional segmentation modeling in medical ultrasonography: extraction, modeling and quantification of skin layers and hypertrophic scars, in: *International Conference on Computational Collective Intelligence*, Springer, 2017, pp. 182–192.
- [23] B. Sciolla, J. Le Digabel, G. Josse, T. Dambry, B. Guibert, P. Delachartre, Joint segmentation and characterization of the dermis in 50 MHz ultrasound 2D and 3D images of the skin, *Comput. Biol. Med.* 103 (2018) 277–286.
- [24] J. Czajkowska, S. Korzekwa, E. Pietka, Computer aided diagnosis of atopic dermatitis, *Comput. Med. Imaging Graph.* 79 (2020) 101676.
- [25] M. Pereyra, N. Dobigeon, H. Batatia, J.-Y. Tourneret, Segmentation of skin lesions in 2-D and 3-D ultrasound images using a spatially coherent generalized Rayleigh mixture model, *IEEE Trans. Med. Imaging* 31 (8) (2012) 1509–1520.
- [26] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Trans. Image Process.* 10 (2) (2001) 266–277.
- [27] G. Csány, K. Szalai, K. Füzesi, M. Gyöngy, A low-cost portable ultrasound system for skin diagnosis, in: *Proceedings of Meetings on Acoustics 6ICU*, Vol. 32, No. 1, ASA, 2017, 020002.
- [28] G. Csány, K. Szalai, M. Gyöngy, A real-time data-based scan conversion method for single element ultrasound transducers, *Ultrasonics* 93 (2019) 26–36.
- [29] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, *Int. J. Comput. Vision* 22 (1) (1997) 61–79.
- [30] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, A.K. Nandi, Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering, *IEEE Trans. Fuzzy Syst.* 26 (5) (2018) 3027–3041.