Segmentation of Lymph Nodes in Ultrasound Images using U-Net Convolutional Neural Networks and Gabor-based Anisotropic Diffusion

Haobo Chen
Shanghai University
Yuqun Wang
Tongren Hospital Shanghai Jiaotong University School of Medicine
Jie Shi
Shanghai University
Jingyu Xiong
Shanghai University
Jianwei Jiang
Tongren Hospital Shanghai Jiaotong University School of Medicine
Wanying Chang
Tongren Hospital Shanghai Jiaotong University School of Medicine
Man Chen
Tongren Hospital Shanghai Jiaotong University School of Medicine
Qi Zhang (zhangq@t.shu.edu.cn)
Shanghai University https://orcid.org/0000-0003-4641-3858

Research Article

Keywords: segmentation, U-Net, lymph nodes, ultrasound, Gabor-based anisotropic diffusion

Posted Date: July 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-724851/v1

Version of Record: A version of this preprint was published at Journal of Medical and Biological Engineering on November 26th, 2021. See the published version at https://doi.org/10.1007/s40846-021-00670-8.

Segmentation of Lymph Nodes in Ultrasound Images using U-Net Convolutional Neural Networks and Gabor-based Anisotropic Diffusion

Haobo Chen[1,2,3,#] , Yuqun Wang[4,#], Jie Shi[1,2,3], Jingyu Xiong[1,2,3], Jianwei Jiang[4], Wanying Chang[4], Man Chen[4*], Qi Zhang[1,2,3*]

[1]Shanghai Institute for Advanced Communication and Data Science, The SMART (Smart Medicine and AI-based Radiology Technology) Lab, Shanghai University, Shanghai, China

[2]Institute of Biomedical Engineering, Shanghai University, Shanghai, China

[3]School of Communication and Information Engineering, Shanghai University, Shanghai, China

[4]Department of Ultrasound, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[#]These authors contributed equally and are co-first authors.

***Authors for Correspondence:**

Qi Zhang, PhD
Professor
Institute of Biomedical Engineering, Shanghai University, Shanghai, China
E-mail address: zhangq@t.shu.edu.cn
Tel: +86-21-66137256
Address: Room 803, Xiangying Building, Shanghai University, No. 333, Nanchen Rd, Shanghai 200444, China

Man Chen, MD, PhD
Professor
Department of Ultrasound, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
E-mail: maggiech1221@126.com
Address: 1111 Xianxia Rd, Shanghai 200050, China

Abstract

Objective Automated segmentation of lymph nodes (LNs) in ultrasound images is a challenging task mainly due to the presence of speckle noise and echogenic hila. In this paper, we propose a fully automatic and accurate method for LN segmentation in ultrasound.

Methods The proposed segmentation method integrates diffusion-based despeckling, U-Net convolutional neural networks and morphological operations. Firstly, we suppress speckle noise and enhance lymph node edges using the Gabor-based anisotropic diffusion (GAD). Secondly, a modified U-Net model is proposed to segment LNs excluding echogenic hila. Finally, morphological operations are adopted to segment entire LNs by filling the regions of echogenic hila.

Results A total of 531 lymph nodes from 526 patients were included to evaluate the proposed method. Quantitative metrics of segmentation performance, including the accuracy, sensitivity, specificity, Jaccard similarity and Dice coefficient, reached 0.934, 0.939, 0.937, 0.763 and 0.865, respectively.

Conclusion The proposed method automatically and accurately segments LNs in ultrasound, which may assist artificially intelligent diagnosis of lymph node diseases.

Keywords: segmentation, U-Net, lymph nodes, ultrasound, Gabor-based anisotropic diffusion

1 Introduction

Lymph nodes (LNs) assist the immune system in building an immune response, and LNs swell and develop lymphadenopathy in cases of invasion by cancer and immune disorders. Adequate assessment of LN status is crucial to diagnose diseases and make treatment decisions. Ultrasound is generally the preferred method for the diagnosis of lymphadenopathy due to its real-time imaging, non-invasiveness, vast availability, and flexibility. In order to quantitatively assess lymphadenopathy using ultrasonography, it requires image segmentation for localizing areas of LNs and finding their borders. However, segmentation of LNs in ultrasound images is generally performed manually by professional experts such as experienced radiologists or ultrasonologists, which is very time-consuming, tedious and subjective. Due to the slow process and tedious nature of the manual segmentation approaches, there is a critical demand for computer algorithms that segment images automatically, accurately and quickly without human interactions.

Recently, convolutional neural networks (CNNs) have become very popular in the field of machine learning and computer vision [1][2]. Besides the success in tasks of natural image computing, CNNs have also shown promising performance in a variety of medical image analysis tasks [3][4][5]. Specifically for medical image segmentation, Ciresan et al. [6] propose a boundary prediction method for electron microscopy by using a CNN as a pixel classifier. Avendi et al. [7] use a CNN for automatic detection of left ventricles from cardiac magnetic resonance imaging. Cha et al. [8] develop a CNN-based system combined with cascading level sets for bladder segmentation in CT urography. Nida et al. [9] propose a model for melanoma lesion detection and segmentation on dermoscopic images using a deep region-based CNN and fuzzy C-means clustering. These methods perform pixel-wise segmentation, in which the patches around each pixel are regarded as input of a CNN for classification. Obviously, the patch-based methods are computationally intensive due to overlapped patches and may lead to global information loss due to the limited receptive fields of patches.

To solve these problems, Long et al. [10] propose the fully convolutional network (FCN) for semantic image segmentation. The FCN is an end-to-end network that can learn semantic information simply and efficiently from whole-image inputs. The DeepLab [11] and PSPNet [12] are FCN-based semantic segmentation methods that achieve state-of-the-art performance. SegNet [13] proposes an encoder-decoder segmentation model, in which the encoder is a 13-layer VGG16 network and the decoder up-samples feature maps with lower resolutions.

However, most of these FCN architectures are developed and tailored for natural image segmentation rather than medical image segmentation. Fortunately, there are emerging models proposed specifically for medical image segmentation. The U-Net is one of the most popular models for medical image segmentation, which yields a u-shaped network architecture [14]. Based on the U-Net, Yuan et al. [15] propose a fully automated method for skin lesion segmentation on dermoscopic images. Alom et al. [16] demonstrate the effectiveness of the U-Net model on the segmentation of various medical imagining modalities, including retina blood vessel segmentation on color retinal images, skin cancer lesion segmentation on dermoscopic images, and lung segmentation in CT images.

Two problems should be considered before the adoption of U-Net based frameworks to medical ultrasound, i.e., the inherent speckle noise and the presence of the echogenic hilum. On one hand, speckle degenerates the signal-to-noise ratio of ultrasound and disturbs ultrasound image segmentation. Thus, it is extremely difficult to accurately extract the edges of LNs from ultrasonic images. In view of speckle polluting medical ultrasound images, there is an urgent need for a denoising method to effectively suppress speckle noise. The classic anisotropic diffusion (AD) method, introduced by Perona et al., used a partial differential equation to gradually denoise an image via iterative diffusion [17]. Considering tissue edges in ultrasound exhibit obvious directionality while noise is randomly distributed, the directionality of edges may facilitate discrimination between edges from noise. The Gabor-based anisotropic diffusion (GAD) captures the edge directionality with the Gabor-based edge detector. The GAD not only suppresses speckle noise in ultrasound but also preserves and enhances tissue edges, structures, and details [18]. Thus, it has the potential to be explored for noise reduction and edge enhancement so as to more accurately segment LNs in ultrasound.

On the other hand, an echogenic hilum is a sonographic feature of most normal LNs, while metastatic, lymphomatous and tuberculous LNs may present with an echogenic hilum in their early stage of involvement [19][20][21]. In ultrasound, a hilum appears to be a depressed/concave area of the surface of an LN. The echogenicity of a hilum and its adjacent soft tissues is very similar in ultrasound and hence a hilum appears to be continuous with its adjacent soft tissues, which makes the detection of the border between a hilum and its adjacent tissues highly difficult and challenges automated segmentation of an entire LN, as seen in Fig. 1a. With this in mind, we design a multi-stage strategy to achieve better segmentation of LNs to cope with the problem of hila presence. In the first stage, we segment an LN excluding a hilum (if any) with a U-Net based model so as to detect the LN region with a concave representing a hilum, and in the second stage, we use morphological operations to refine the segmentation and obtain an entire LN by filling the concave at the hilum.

Inspired by the above observations, we propose a U-Net based framework integrated with the

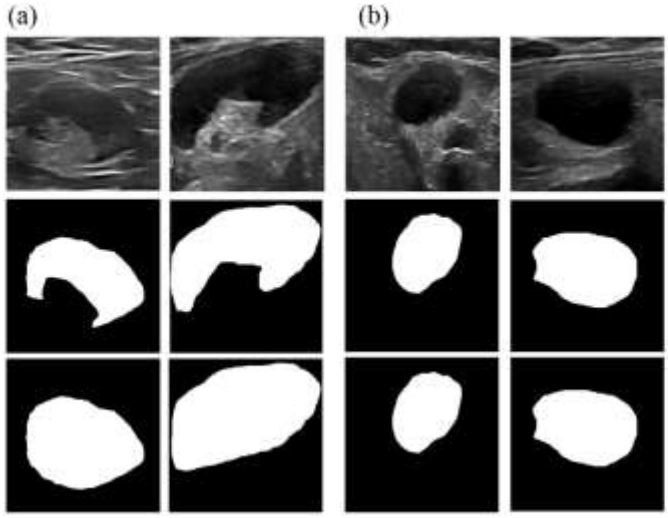

Fig. 1. Typical ultrasound images of lymph nodes (LNs). (a) LNs with hila, and (b) LNs without hila. The first row: ultrasound images. The second row: binary masks of LNs excluding hila. The third row: binary masks of LNs including hila.

GAD to reduce speckle noise and morphological operations to fill echogenic hila, which allows automatic segmentation of entire LNs in ultrasound images. This paper is organized as follows.

We introduce the details of the proposed U-Net based segmentation method in Section 2 and then report the experimental design and results in Section 3. Finally, we discuss our results, findings, and future work in Section 4 and conclude our study in Section 5.

2 Materials and Methods

2.1 Image Acquisition

This study includes 531 LNs (231 with hila and 300 without hila) from 526 patients. Ultrasound examinations have been performed by an experienced ultrasonologist using the Mylab 90 system (Esaote, Genoa, Italy) with a 4–13 MHz probe (L523). All images have been manually segmented by the ultrasonologist to achieve the borders of LNs and their echogenic hila (if any). Therefore, for each LN with a hilum, the gold standard of segmentation has been obtained for two regions, namely the LN region including the hilum and that excluding the hilum, as shown in Fig. 1a. For each LN without a hilum, the gold standards for the two regions are exactly the same (Fig. 1b).

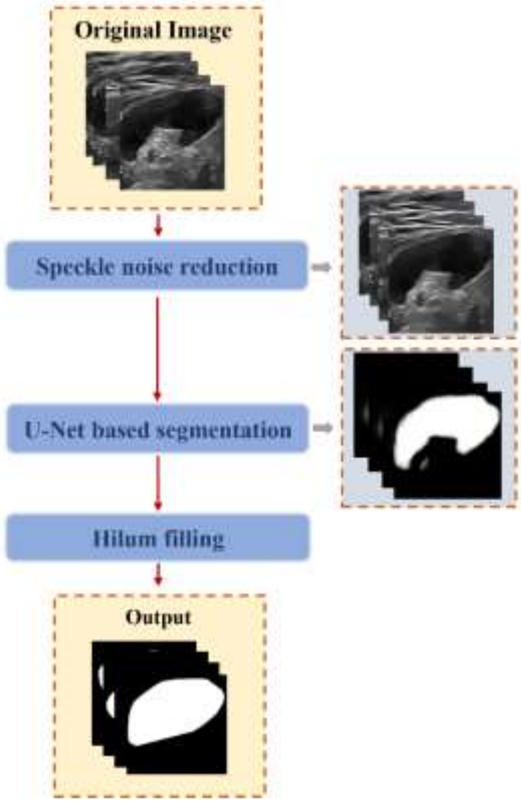

Fig. 2. The flowchart of the proposed automated segmentation method for LNs.

2.2 Overview of the Automatic Segmentation System

In this work, we present a method for LN segmentation in ultrasound images consisting of the following three steps, as illustrated in Fig. 2. Firstly, we adopt the GAD to reduce speckle noise in ultrasound and enhance lymph nodal edges. Secondly, a U-Net model is modified to be adapted for LN ultrasound images and it is trained on the gold standards of the LNs excluding hila. Finally, we fill the concaves at hila and segment the whole LNs through a set of morphological operations.

2.3 Gabor-based Anisotropic Diffusion for Speckle Noise Reduction

In this section, we introduce the GAD to suppress speckle and enhance nodal edges in the medical ultrasonography of LNs. The GAD is a speckle reduction method for denoising ultrasound images by employing a new edge detector based on the Gabor transform into the anisotropic diffusion [18]. If an input image is denoted as $I(x, y)$, its Gabor transform is the convolution of $I(x, y)$ with a family of Gabor kernels:

$$G_d(x,y) = I(x,y) * \mathrm{imag}\left[g_d(x,y)\right] \tag{1}$$

where * represents the convolution operator, imag[·] denotes the imaginary part, and $G_d(x, y)$ is the d-th convoluted image obtained by convolving the d-th Gabor kernel with the input image. Here, only the imaginary part of the Gabor kernel is utilized for convolution [18]. An edge detector based on the Gabor transform, called the Gabor-based edge detector, is hence given by:

$$G_{\mathrm{sd}}(x,y) = \sqrt{\frac{1}{D-1}\sum_{d=0}^{D\text{-}1}[G_d(x,y)]^2} \tag{2}$$

The partial differential equation of the GAD model is described as:

$$\begin{cases} \dfrac{\partial I(x,y)}{\partial t} = \mathrm{div}[c\left(G_{\mathrm{sd}}\right) \cdot \nabla I(x,y)] \\ I(x,y;t=0) = I_0(x,y) \end{cases} \tag{3}$$

where div is the divergence operator, $c(\cdot)$ is the diffusion coefficient, ∇ represents the gradient operator, t is the diffusion time, and I_0 is the initial image.

2.4 U-Net based Segmentation of Lymph Nodes excluding Hila

Due to the intensity of an echogenic hilum in ultrasound appears similar to that of its adjacent soft tissues, we do not intend to segment an entire LN directly but to first segment the LN excluding the hilum and then fill the concave echogenic hilum in the LN. Here in this section we propose a modified U-Net model for segmentation of LNs excluding hila in ultrasound.

2.4.1 U-Net Architecture

The U-Net architecture, which is an encoder-decoder, consists of an encoding path to capture image features and a symmetrical decoding path for precise localization [22]. As shown in Fig. 3, we have modified the original U-Net in the following ways to adapt it to our small ultrasound dataset: (a) setting 240×240 as the size of an input image and generating feature maps with different sizes to fit our ultrasound dataset; (b) using the convolution with zero padding to avoid cropping and to generate the output as the same size of the input; (c) using the deconvolutional layer with a kernel size of 3×3 and a stride of 2×2 instead of the deconvolutional layer with a kernel size of 2×2 so as to enlarge the receptive field of the kernel and obtain more useful information.

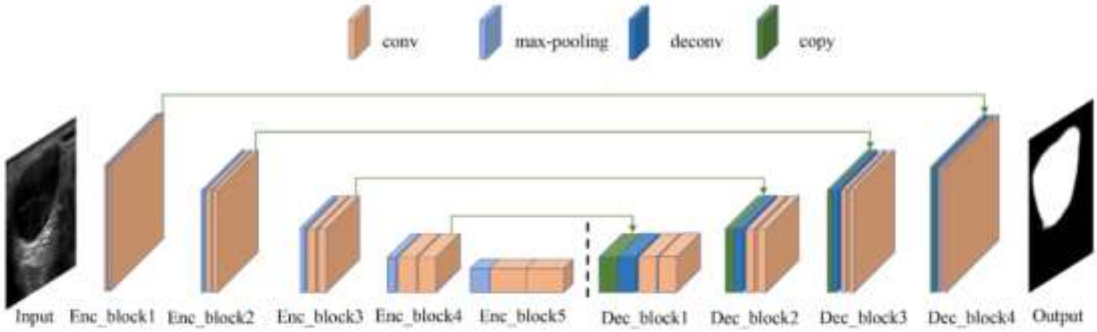

Fig. 3. The architecture of the modified U-Net for segmentation of LNs excluding hila.

Table 1. Architecture details of the proposed U-Net model

Encoder	Output size	Decoder	Output size
Input	240 × 240 × 1	Dec_block1	30 × 30 × 512
Enc_block1	240 × 240 × 64	Dec_block2	60× 60 × 256
Enc_block2	120 × 120 × 128	Dec_block3	120 × 120 × 128
Enc_block3	60× 60 ×256	Dec_block4	240 × 240 × 64
Enc_block4	30 × 30 × 512	*Output*	240 × 240 × 2
Enc_block5	15 × 15 × 1024		

Encoder: There are five convolutional blocks in the encoding path. Each block has two convolutional layers with a kernel size of 3×3. Through the path, the number of feature maps is increased from 1 to 1024, as shown in Table 1. At the end of each block except the last block, a max pooling layer with a stride of 2×2 is applied to down-sampling the size of the feature map by two. Hence, the size of feature maps decreases from 240×240 to 15×15 (Table 1).

Decoder: Each block of the decoding path starts with a deconvolutional layer with a kernel size of 3×3 and a stride of 2×2, which doubles the size of feature maps but decreases the number of feature maps by two. Therefore the size of feature maps increases from 15×15 to 240×240 (Table 1). Following the deconvolutional layer, a skip connection is used to concatenate the feature maps from the encoding path and the feature maps from deconvolution. Then two convolutional layers are used to reduce the number of feature maps.

Finally, an additional convolutional layer with a kernel size of 1×1 is used to reduce the number of feature maps to two that reflects the probabilities of each pixel belonging to the foreground and background, and thus the final output is called the probability map. Different from the original U-Net architecture, we use the zero padding to maintain the size of output feature maps at all the convolutional layers in both encoding and decoding paths. Other details of the network are tabulated in Table 1.

2.4.2 U-Net Training

Loss Function: In this work, the dice loss described in [23] is used as the loss function of the network which can be considered as a differentiable form of the original dice coefficient. The dice loss of *N* images is computed by:

$$L(\mathrm{X},\mathrm{Y}) = \frac{1}{N}\sum_{i=1}^{N}\frac{2|X \cap Y\ |+k}{|X|+|Y|+k} \tag{4}$$

where X and Y denote predicted segmentation and the ground truth (i.e., the gold standard segmented by the ultrasonologist), respectively, and $k\ \in (0,1)$ denotes the smoothing coefficient.

Adam Stochastic Optimization: Training deep neural networks requires stochastic gradient-based optimization to minimize the loss function with respect to its parameters [24]. We adopt the adaptive moment estimator (Adam) [25] to estimate the parameters. In general, Adam utilizes the first and second moments of gradients for updating and correcting the moving average of the current gradients. The parameters of our Adam optimizer are set as: the learning rate = 0.0001 and the maximum number of epochs = 100. All weights are initialized by a normal distribution with a mean of 0 and a standard deviation of 0.01, and all biases are initialized as 0.

Data Augmentation: In order to improve the robustness of the proposed U-Net based model, we artificially produce more training data from the original data with a set of image transformations summarized in Table 2.

- Geometric transformation such as flipping, shift and rotation can result in displacement fields to images. Shear operation can slightly distort the global shape of LNs in the horizontal direction.
- Intensity transformation randomly jitters the intensities of the images by a Gaussian random factor, including the transformation of brightness and contrast.
- Elastic transformation [26] generates more training data with arbitrary but reasonable shapes

to gain sufficient variable training data, as LNs have no definite shapes.

Table 2. Summary of the applied image transformations.

Transformations	Values
Flipping horizontally and vertically	50% probability on both directions
Shift horizontally and vertically	10% on both directions
Shear	20% on the horizontal direction
Rotation	[−40°, 40°]
Brightness and contrast	[0.7, 1.3]
Elastic transformation	$\alpha = 720$, $\sigma = 24^{*}$

* α and σ control the degree of the elastic transformation

2.4 Hilum Filling with Morphological Operations

The echogenic hilum is a sonographic feature for some (231/531) of the LNs in our dataset. Thus, following the U-Net based segmentation of LNs excluding hila, the morphological operations are performed on the detected LNs for filling the concaves at the echogenic hila and thus segmenting the entire LNs [27].

In the procedure of hilum filling, the probability maps derived from the U-Net model are first thresholded to get the binary maps of the LNs excluding hila. The opening operation is then applied to the binary maps to remove isolated debris wrongly detected as LNs by the U-Net. Afterwards, the closing operation is employed to fill the small gaps in the LNs. Finally, the hilum appears to be a concave in a binary LN map, and the concave region is filled to obtain a complete LN by using the convex hull operation.

Thresholding: With regard to the probability map, we chose a threshold of 0.5. All pixels below the threshold are set to zero while the pixels above it are set to one.

Opening and closing: The opening operation is the dilation of the erosion of a binary image while the closing operation is the erosion of the dilation of the image. The former removes small objects from the foreground (the white pixels) and places them into the background, while the latter removes small holes in the foreground and changes small islands of background into the foreground.

Convex hull computing: Computing the convex hull means that a non-ambiguous and efficient representation of the required convex shape is constructed. The concave region in a U-Net-detected LN, which represents the hilum of the LN, can be filled by using convex hull computing.

3 Experiments and Results

3.1 Experimental Settings

The proposed LN segmentation method was implemented with Python3.6 based on Keras package. The experiments were conducted on an Ubuntu 16.04 desktop with 2 CPUs (Intel Xeon), 2 GPUs (NVIDIA GTX 1080ti 11Gb), and 256Gb RAM.

As introduced in Section 2.1, we got a total number of 531 LNs extracted from 526 patients. We randomly separated the LNs into three parts: 390 for training, 51 for validation and 90 for independent test. We normalized the dataset to the standard Gaussian distribution to reduce the internal covariate shift. We resized the images to 240 × 240 so as to be suitable for the U-Net based model. To augment the dataset, we performed the data augmentation techniques (Section 2.4.2) on training and validation datasets and thus the image sample number was augmented to 20 times. Finally, we obtained a training dataset of 7800 images and a validation dataset of 1020 images.

3.2 Quantitative Evaluation

The segmentation performance in the test set was measured by the accuracy (ACC), sensitivity (SEN) and specificity (SPC) of classifying pixels into positives (inside a LN) or negatives (outside a LN):

$$\mathrm{ACC} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

$$\mathrm{SEN} = \frac{TP}{TP+FN} \tag{6}$$

$$\mathrm{SPC} = \frac{TN}{TN+FP} \tag{7}$$

where TP, TN, FP, and FN denoted the numbers of true positives, true negatives, false positives, and false negatives, respectively. These evaluation metrics evaluated comprehensively the segmentation performance from different aspects. All of them were values between 0 and 1.

We used Dice coefficient (DC) and Jaccard similarity (JS) to further measure the performance of LN segmentation. The DC is expressed as:

$$\mathrm{DC} = \frac{2TP}{2TP+FP+FN} \tag{8}$$

The JS is represented by using

$$\mathrm{JS} = \frac{TP}{TP+FP+FN} \tag{9}$$

Due to the non-normal distribution of the segmentation indices, the medians and interquartile ranges (IQRs) of the indices were calculated. The Wilcoxon signed-ranks test was adopted to compare the segmentation indices of the original ultrasound images and those of the GAD filtered images. Statistical significance was set at 0.05.

3.3 Results

We first visualized the results at each step in the segmentation process, using a typical LN in Fig. 4 as an example. Fig. 4(a)(b) illustrate the original LN image and the boundary on it. The image denoised by the GAD is shown in Fig. 4(c) and the result of the U-Net is shown in Fig. 4(d). Fig. 4(e) shows the results after the thresholding operation. Fig. 4(f) shows the results after the operations of opening, closing and convex hull. In the following sub-sections, we illustrate in detail the effectiveness of the GAD despeckling, U-Net segmentation, and morphological operations.

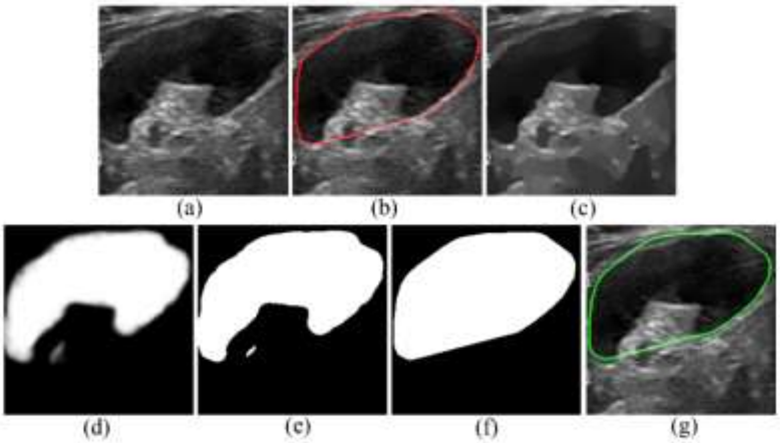

Fig. 4. The results at each step in the segmentation process. (a) The LN in the original ultrasound image. (b) The manual annotation in the original ultrasound image is marked with a red contour. (c) The image filtered with the GAD. (d) The segmentation result of the probability map. (e) The region of interest depicted as a binary mask. (f) Images after the morphological operations of the opening, closing and convex hull computing. (g) The result of morphological operations is marked in the original ultrasound image with a green contour.

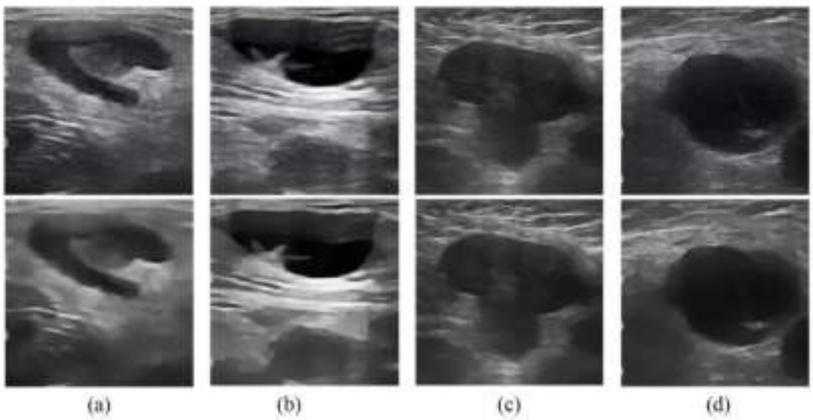

Fig. 5. The results of the GAD denoising (top: input images; bottom: output images).

3.3.1 Results of GAD Denoising

As shown in Fig. 5, speckle noise contaminates ultrasound images, especially in the areas surrounding LNs. From the filtered images, we can see that the GAD simultaneously reduced speckle noise obviously and enhanced the LN edges effectively, which would facilitate more accurate image segmentation for LNs.

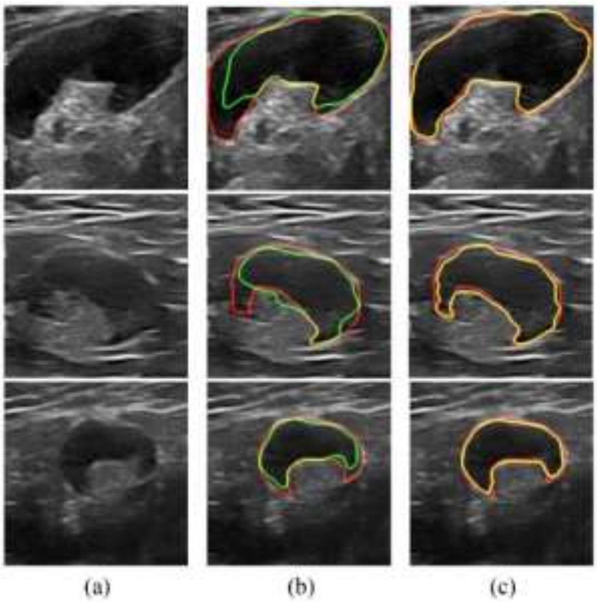

Fig. 6. The segmentation results for LNs excluding hila. (a) The original ultrasound images. (b) The results on the original images. (c) The results on the GAD filtered images. The red lines denote the contours from the manual segmentation, and the green and yellow lines represent the contours automatically segmented from the original images and the GAD filtered images, respectively.

Table 3. The segmentation results for LNs excluding hila on the original ultrasound images and the GAD filtered images.

		ACC	SEN	SPC	JS	DC
Median	Original	0.938	0.899	0.952	0.745	0.854
	GAD	0.939	0.879	0.967	0.763	0.866
IQR	Original	0.066	0.165	0.078	0.293	0.205
	GAD	0.055	0.172	0.070	0.262	0.177
p-values		0.046	0.009	<0.001	0.001	0.002

3.3.2 Results of Segmentation Excluding Lymph Hilum

Automated segmentation results were compared with the corresponding ground truth by the ultrasonologist. In order to observe the visual similarity of the shapes of the detected LNs, the

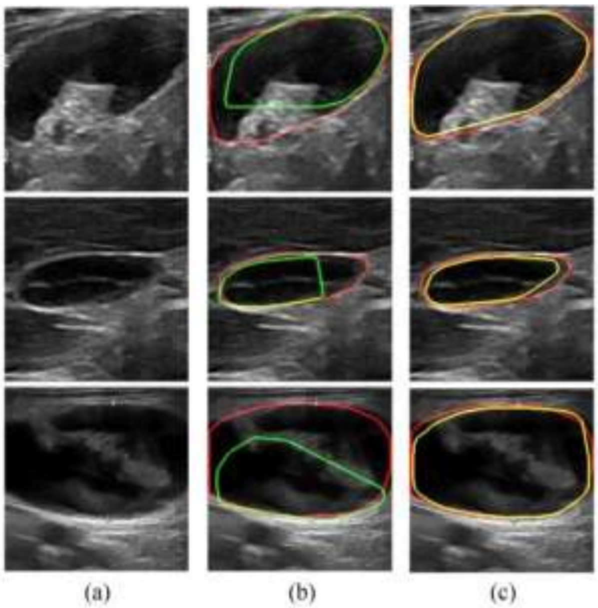

Fig. 7. Final segmentation results by filling the concaves at hila with morphological operations. (a) The original ultrasound images. (b) The results of the original images. (c) The results of the GAD filtered images. The red lines denote the contours from the manual segmentation, and the green and yellow lines represent the contours automatically segmented from the original images and the GAD filtered images, respectively.

Table 4. Final segmentation results (for LNs including hila) on the original ultrasound images and the GAD filtered images.

		ACC	SEN	SPC	JS	DC
Median	Original	0.929	0.949	0.925	0.736	0.848
	GAD	0.934	0.939	0.937	0.763	0.865
IQR	Original	0.075	0.086	0.095	0.341	0.242
	GAD	0.063	0.103	0.084	0.247	0.165
p-values		0.108	0.040	<0.001	0.025	0.041

contours of the LNs in the manual and automatic segmentations were extracted and marked with different colors, as shown in Fig. 6. We can see, in terms of contours, the results of the automatic segmentation from the GAD filtered images showed good consistency with the ground truth.

As shown in Table 3, the ACC, SEN, SPC, JS and DC of the GAD filtered images were 0.939

0.879, 0.967, 0.763 and 0.866 respectively. We also compared the ACC, SEN, SPC, JS and DC of the GAD filtered images to those of the original images. There were statistically significant differences between the ACC, SEN, SPC, JS and DC of the GAD filtered images and those of the original images (all $p<0.05$). From Table 3, it can be demonstrated that in terms of all indices, our final model, which was trained on the GAD filtered images, was better than the traditional U-Net model. For example, the DC and JS of the filtered GAD images (0.866 and 0.763) were significantly higher than those of the original images (0.854 and 0.745) by 0.012 and 0.018, respectively.

3.3.3 Results of final segmentation

Fig. 7 shows the final segmentation results after filling concaves at hila by using morphological operations. It can be seen that the segmentation performance on the GAD filtered images outperformed that on the original images.

The ACC, SEN, SPC, DC and JS in the GAD filtered images and in the original ultrasound images are listed in Table 4. The ACC, SEN, SPC, JS and DC of the GAD filtered images were 0.934 0.939, 0.937, 0.763 and 0.865 respectively. The GAD filtered images achieved higher DC and JS than the original ultrasound images. The statistical test further confirmed that the segmentation on the GAD filtered images was statistically more accurate than that on the original ultrasound images ($p<0.05$ for all indices except ACC; Table 4).

4 Discussion

We propose a novel framework for LN segmentation in ultrasound based on the U-Net model, collaborated with the GAD filtering and morphological operations. Firstly, we denoise the ultrasound image with the GAD to suppress speckle noise and enhance edges of LNs. Secondly, we propose a modified U-Net model to segment LNs excluding hila. Finally, the morphological operations are performed to fill the concaves at hila and thus achieve the final segmentation of the entire LNs. The experimental results have indicated that the proposed framework has the capacity to segment LNs automatically and accurately in ultrasound images.

We propose a two-stage segmentation framework for LNs, in which we first segment the LNs excluding LN hila and then fill the concaves at hila. To the best of our knowledge, this is the first time that the “excluding-then-filling hila” scheme is investigated in the segmentation of LNs in ultrasound images. This scheme accomplishes the segmentation of those LNs with echogenic hila, while the existence of hila challenges the direct segmentation of entire LNs because the echogenicity of hila and their adjacent soft tissues are too similar to distinguish.

The GAD is introduced to suppress speckle in ultrasound images which is beneficial for the segmentation of LNs. It employs a new edge detector based on the convolution of an input image with the Gabor kernels. Therefore, a good GAD filter can be seen as a well-trained convolutional layer that depends on the prior of speckle noise in ultrasound images. In a future study, we will develop an end-to-end convolutional network with the capability of denoising at its low layers and segmentation at its high layers.

Different numbers of GAD iterations lead to different filtered images. These multiple images filtered by the GAD can be used to generate a set of multi-channel images by concatenating them with the original image. These multi-channel images may supplement each other. If they are used as the input of the U-Net in a future study, the segmentation performance may be further improved.

The limited number of images is one of the main challenges in applying deep learning to medical image analysis. Our approach to addressing the lack of samples is generating samples artificially via data augmentation to expand the database. Three types of augmentation methods are performed to generate abundant samples including the geometric transformation, intensity transformation and elastic transformation, which create kinds of shapes and intensities of training samples to ensure the robustness to datasets.

Although our method has achieved promising segmentation performance, there are some drawbacks and future directions. First, in addition to the U-Net model, the denoising with the GAD and the morphological operations consume extra labors. Thus, an end-to-end model that fuses the three steps would be expected to be developed in the future. Second, other modifications to the U-Net model may be made to further improve its segmentation performance, such as combining multiple segmentation maps [28] and extending the U-Net with residual blocks [29]. Finally, the data augmentation method is used to address the problem of a limited number of images. Alternatively, transfer learning, which uses deep models trained on natural images and transfer them to medical images, has been proven to be highly effective in several applications and maybe useful for the segmentation for LNs in ultrasound images [30].

5 Conclusions

In this study, we present an automatic segmentation method for LN ultrasound images based on the U-Net model and GAD. Firstly, the original ultrasound images are despeckled by using the GAD filter. Secondly, three transformation methods are performed for ultrasound dataset augmentation, and a modified U-Net model is proposed to segment LNs excluding hila. Finally, the morphological operations are employed to complete the segmentation of LNs including hila. The segmentation accuracy, sensitivity, specificity, Jaccard similarity and Dice coefficient reach 0.934, 0.939, 0.937, 0.763 and 0.865, respectively, which indicates that the proposed method has the capacity to effectively segment LNs in ultrasound images and may potentially facilitate artificially intelligent diagnosis of LN diseases in future studies.

Acknowledgments

The work was funded by the National Natural Science Foundation of China (Nos. 62071285 and 61911530249).

Conflict of interest

None declared.

References

[1] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Int. Conf. Learn. Represent.*, vol. 14, no. 4, pp. 1–14, Aug. 2015.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[3] V. Gulshan *et al.*, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *JAMA*, vol. 316, no. 22, p. 2402, Dec.

2016.

[4] M. A. Al-antari, M. A. Al-masni, M. Choi, S. Han, and T. Kim, “A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification,” *Int. J. Med. Inform.*, vol. 117, no. May, pp. 44–54, Sep. 2018.

[5] G. Dimauro *et al.*, “Nasal cytology with deep learning techniques,” *Int. J. Med. Inform.*, vol. 122, no. October 2018, pp. 13–19, Feb. 2019.

[6] and J. S. D. Ciresan, A. Giusti, L. Gambardella, “Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images,” in *Advances in neural information processing systems*, 2012, pp. 2852--2860.

[7] M. R. Avendi, A. Kheradvar, and H. Jafarkhani, “A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI,” *Med. Image Anal.*, vol. 30, pp. 108–119, May 2016.

[8] K. H. Cha, L. M. Hadjiiski, R. K. Samala, H.-P. Chan, R. H. Cohan, and E. M. Caoili, “Comparison of bladder segmentation using deep-learning convolutional neural network with and without level sets,” 2016, vol. 43, no. 4, p. 978512.

[9] N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, “Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering,” *Int. J. Med. Inform.*, vol. 124, pp. 37–48, Apr. 2019.

[10] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, Apr. 2017.

[11] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” pp. 1–14, 2016.

[12] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid Scene Parsing Network,” 2016.

[13] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” pp. 1–14, 2015.

[14] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical

Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.

[15] Y. Yuan, M. Chao, and Y. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," *IEEE Trans. Med. Imaging*, vol. 36, no. 9, pp. 1876–1886, Sep. 2017.

[16] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, "Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation," *Desalination*, vol. 227, no. 1–3, pp. 327–333, Feb. 2018.

[17] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[18] Q. Zhang, H. Han, C. Ji, J. Yu, Y. Wang, and W. Wang, "Gabor-based anisotropic diffusion for speckle noise reduction in medical ultrasonography," vol. 31, no. 6, pp. 1273–1283, 2014.

[19] P. W. S. Rosário *et al.*, "Ultrasonographic Differentiation Between Metastatic and Benign Lymph Nodes in Patients With Papillary Thyroid Carcinoma," *J. Ultrasound Med.*, vol. 24, no. 10, pp. 1385–1389, Oct. 2005.

[20] R. M. Evans, A. Ahuja, and C. Metreweli, "The linear echogenic hilus in cervical lymphadenopathy — A sign of benignity or malignancy?," *Clin. Radiol.*, vol. 47, no. 4, pp. 262–264, Apr. 1993.

[21] M. Ying, A. Ahuja, F. Brook, and C. Metreweli, "Vascularity and Grey-Scale Sonographic Features of Normal Cervical Lymph Nodes: Variations with Nodal Size," *Clin. Radiol.*, vol. 56, no. 5, pp. 416–419, May 2001.

[22] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks," 2017, pp. 506–517.

[23] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.

[24] L. Bottou, "Stochastic Gradient Descent Tricks," 2012, pp. 421–436.

[25] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *AIP Conf. Proc.*, vol.

1631, pp. 58–62, Dec. 2014.

[26] P. Y. Simard, D. Steinkraus, and J. C. Platt, “Best practices for convolutional neural networks applied to visual document analysis,” in *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.*, 2003, vol. 1, pp. 958–963.

[27] E. Supriyanto and N. Zulkifli, “Abnormal tissue detection of breast ultrasound image using combination of morphological technique,” *Proc. 15th …*, pp. 234–239, 2011.

[28] B. Kayalibay, G. Jensen, and P. van der Smagt, “CNN-based Segmentation of Medical Imaging Data,” *Bioelectrochemistry*, vol. 75, no. 2, pp. 130–135, Jan. 2017.

[29] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, “The Importance of Skip Connections in Biomedical Image Segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10008 LNCS, pp. 179–187, Aug. 2016.

[30] N. Tajbakhsh *et al.*, “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016.