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

Haobo Chen1,2 · Yuqun Wang3 · Jie Shi1,2 · Jingyu Xiong1,2 · Jianwei Jiang3 · Wanying Chang3 · Man Chen3 · Qi Zhang1,2,4

Received: 16 July 2021 / Accepted: 9 November 2021 / Published online: 26 November 2021 © Taiwanese Society of Biomedical Engineering 2021

Abstract
Purpose The automated segmentation of lymph nodes (LNs) in ultrasound images is challenging, largely because of speckle noise and echogenic hila. This paper proposes a fully automatic and accurate method for LN segmentation in ultrasound that overcomes these issues.
Methods The proposed segmentation method integrates diffusion-based despeckling, U-Net convolutional neural networks and morphological operations. First, the speckle noise is suppressed and the lymph node edges are enhanced using Gabor-based anisotropic diffusion (GAD). Then, a modified U-Net model is used to segment the LNs excluding any echogenic hila. Finally, morphological operations are undertaken to segment the entire LNs by filling in any regions occupied by echogenic hila.
Results A total of 531 lymph nodes from 526 patients were segmented using the proposed method. Its segmentation performance was evaluated in terms of its accuracy, sensitivity, specificity, Jaccard similarity and Dice coefficient, for which it achieved values of 0.934, 0.939, 0.937, 0.763 and 0.865, respectively.
Conclusion The proposed method automatically and accurately segments LNs in ultrasound images, enhancing the prospects of being able to undertake artificial intelligence (AI)-based diagnosis of lymph node diseases.

Keywords Segmentation · U-Net · Lymph nodes · Ultrasound · Gabor-based anisotropic diffusion

1 Introduction
Lymph nodes (LNs) assist a body's immune system in building an immune response. LNs swell and develop lymphadenopathy in cases of invasion by cancer and immune disorders. Adequate assessment of LN status is accordingly crucial when diagnosing diseases and making treatment decisions. Ultrasound is generally the preferred method for the diagnosis of lymphadenopathy because it offers real-time imaging, is non-invasive, and is widely available and flexible. Quantitative assessment of lymphadenopathy using ultrasonography involves image segmentation to localize LN areas and find their borders. Currently, the segmentation of LNs in ultrasound images is generally performed manually by professional experts, such as experienced radiologists or ultrasonologists. However, this is very time-consuming, tedious and subjective. As a result, there is an urgent need for computer algorithms that can segment images
automatically, accurately and quickly, without any need for human intervention.

Medical image segmentation technology is already extremely helpful in quantifying tissue volume, assisting diagnosis, localizing pathologies, and enabling the study of anatomical structures [1, 2]. A variety of methods have been developed for lesion segmentation in ultrasound images, including textural-classifiers [3, 4], active contour models [5, 6] and graph theory-based approaches [7, 8]. Textural-classifiers distinguish lesions from adjacent soft tissue on the basis of intensity information for different image regions [3]. Active contour models, based on approaches such as snake methods and level set methods, use a dynamic curve to find lesion boundaries [5, 6]. Graph theory-based methods, such as graph cut and grab cut, use a min-cut to minimize the energy function, thereby revealing the foreground and background pixel sets [8]. These traditional segmentation methods are sensitive to edges, but cannot effectively distinguish between lesion and non-lesion regions. The presence of serious levels of speckle noise or complex lesions in ultrasound images can further undermine the capacity of the above methods to achieve ideal segmentation results.

Convolutional neural networks (CNNs) have recently become very popular in the fields of machine learning and computer vision [9, 10]. In addition to their notable success in the context of natural image computing, CNNs have also shown promise in a variety of medical image analysis tasks [11–13]. In relation to medical image segmentation, Ciresan et al. [14] have proposed a boundary prediction method for electron microscopy that uses a CNN as a pixel classifier. Avendi et al. [15] used a CNN for the automatic detection of left ventricles in cardiac magnetic resonance imaging. Cha et al. [16] have developed a CNN-based system combined with cascading level sets for bladder segmentation in CT urography. Nida et al. [17] have proposed a model for melanoma lesion detection that segments dermoscopic images by 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 the input for a CNN for classification. These patch-based methods are computationally intensive because patches often overlap, leading to global information loss because of their limited receptive fields.

To solve these problems, Long et al. [18] have suggested using fully convolutional networks (FCNs) for semantic image segmentation. An FCN is an end-to-end network that can learn semantic information simply and efficiently from whole-image inputs. DeepLab [19] and PSPNet [20] are both FCN-based semantic segmentation methods that have achieved state-of-the-art performance. SegNet [21] uses an encoder-decoder segmentation model, where the encoder is a 13-layer VGG16 network and the decoder up-samples feature maps with lower resolutions. However, most FCN architectures have been developed for natural image segmentation rather than medical image segmentation. Fortunately, a few of the most recent models have started to move in this direction. U-Net, for instance, has become a popular resource for this kind of segmentation. It yields a u-shaped network architecture [22], on the basis of which, Yuan et al. [23] have developed a fully automated method for skin lesion segmentation in dermoscopic images. Alom et al. [24] have also demonstrated the effectiveness of the U-Net model for several different medical imaging modalities, such as retina blood vessel segmentation in color retinal images, skin cancer lesion segmentation in dermoscopic images, and lung segmentation in CT images.

Two problems confront the adoption of U-Net based frameworks for medical ultrasound. One of these is the inherent presence of speckle noise in ultrasound images. The other is the presence of echogenic hila. Speckle noise degrades the signal-to-noise ratio and disrupts the ultrasound image segmentation process. This makes it extremely difficult to accurately extract the edges of LNs from ultrasonic images. The pervasiveness of speckle noise pollution in medical ultrasound images has resulted in an urgent need for a denoising method that can effectively suppress speckle noise. The classic anisotropic diffusion (AD) method for tackling this, first introduced by Perona et al. uses a partial differential equation to gradually denoise an image via iterative diffusion [25]. Tissue edges in ultrasound exhibit obvious directionality, while noise is randomly distributed. Thus, the directionality of edges can facilitate discrimination between the edges and noise. Gabor-based anisotropic diffusion (GAD) captures edge directionality with a Gabor-based edge detector. GAD not only suppresses speckle noise in ultrasound but also preserves and enhances tissue edges, structures and details [26]. It therefore has significant potential for noise reduction and edge enhancement when seeking to accomplish more accurate LN segmentation in ultrasound.

An echogenic hilum is a sonographic feature that is present in most normal LNs. However, metastatic, lymphomatous and tuberculous LNs can also present with an echogenic hilum in their early stage of development [27–29]. In ultrasound, a hilum appears as a depressed/concave area on the surface of an LN. The echogenicity of a hilum and the adjacent soft tissue is very similar in ultrasound, so a hilum usually appears to be continuous with the adjacent soft tissue, making detection of the border between them extremely challenging. This makes it very difficult to achieve the automated segmentation of an entire LN, as can be seen in Fig. 1a. With this in mind, we have designed a multi-stage strategy for LN segmentation that can cope with the presence of hila. In a first stage, we segment the LN excluding the hila (if any), by means of a U-Net-based model that can detect concave LN regions that are indicative of their presence. In a second stage, we
use morphological operations to refine the segmentation and obtain an entire LN image where the concave regions associated with hila have been filled in.

On the basis of the above design strategy, this paper proposes a U-Net-based framework integrated with the GAD to reduce speckle noise and the morphological operations to fill echogenic hila. This then allows for the automatic segmentation of entire LNs in ultrasound images.

The paper is organized as follows: The details of the proposed U-Net-based segmentation method are provided in Sect. 2. We then describe our experimental approach and report the experimental results in Sect. 3. The results and potential future work are discussed in Sect. 4 and we give our overall conclusions in Sect. 5.

2 Materials and Methods

2.1 Image Acquisition

This study drew upon ultrasound images of 531 LNs (231 with hila and 300 without hila) from 526 patients. The ultrasound examinations were performed by an experienced radiologist using the Mylab 90 system (Esaote, Genoa, Italy) with a 4–13 MHz probe (L523). All of the images had previously been manually segmented by the radiologist to identify the borders of the LNs and their echogenic hila (if any). Therefore, for each LN with a hilum, gold standard segmentation had been obtained for two regions: the LN region including the hilum; and the LN excluding the hilum, as shown in Fig. 1a. For each LN without a hilum, the gold standard for the two regions was exactly the same (Fig. 1b).

2.2 Overview of the Automatic Segmentation System

In this work, we present a method for the LN segmentation of ultrasound images that consists of three steps, as illustrated in Fig. 2. First, a GAD is used to reduce the speckle noise in the ultrasound and enhance the lymph nodal edges. Second, a modified U-Net model that has been specifically adapted for LN ultrasound images is trained on the gold standard segmentation of the LNs excluding hila. Third, we fill in the hila-related concave areas and segment each whole LN through a set of morphological operations.

2.3 Gabor-Based Anisotropic Diffusion for Speckle Noise Reduction

This section introduces how GAD is used to suppress speckle noise and enhance the nodal edges in the medical ultrasonography of LNs. GAD is a speckle reduction method that can be used to denoise ultrasound images by employing anisotropic diffusion based on the Gabor transform for the purposes of edge detection [18]. If an input image is denoted \( I(x, y) \), its Gabor transform is the convolution of \( I(x, y) \) with a family of Gabor kernels, i.e.:
where \( \ast \) represents the convolution operator; \( \text{imag}[\cdot] \) 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 [26]. An edgedetector based on the Gabor transform, called a Gabor-based edge detector, is hence given by:

\[
G_{sd}(x, y) = \frac{1}{D-1} \sum_{d=0}^{D-1} |G_d(x, y)|^2
\]  

(2)

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

\[
\begin{aligned}
\frac{\partial I(x, y)}{\partial t} &= \text{div}[c(G_{sd} \cdot \nabla I(x, y))] \\
I(x, y; t = 0) &= I_0(x, y)
\end{aligned}
\]

(3)

where, \( \text{div} \) is the divergence operator; \( c(\cdot) \) is the diffusion coefficient; \( \nabla \) 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

As the intensity of an echogenic hilum in an ultrasound image is similar to that of the adjacent soft tissue, we do not intend to segment an entire LN directly. We first segment the LN excluding the hilum and then fill in the concave region associated with the hilum. In this section, we describe the modified U-Net model we have designed for the segmentation of LNs excluding hila.
Decoder Each block in the decoding path starts with a deconvolutional layer with a kernel size of 3 × 3 and a stride of 2 × 2. This doubles the size of the feature maps but decreases their number by two. Thus, the size of the feature maps increases from 15 × 15 to 240 × 240 (see Table 1). After the deconvolutional layer, a skip connection is used to concatenate the feature maps from the encoding path and the feature maps from the deconvolution. Two convolutional layers are then used to reduce the number of feature maps. Finally, another convolutional layer with a kernel size of 1 × 1 is used to reduce the number of feature maps to two, thereby reflecting the probability of each pixel belonging to the foreground or the background. The final output is therefore a ‘probability’ map. Unlike from the original U-Net architecture, we use zero padding to maintain the size of the output feature maps for all the convolutional layers in both the encoding and decoding paths. Other details of the network are shown in Table 1.

2.4.2 U-Net Training

Loss Function As a part of the approach we are using here, the dice loss described in [31] is used as the network loss function. This can be considered a differentiable form of the original dice coefficient. The dice loss for N images is computed as follows:

\[
L(X, Y) = \frac{1}{N} \sum_{i=1}^{N} \frac{2|X \cap Y| + k}{|X| + |Y| + k}
\]  

(4)

where, X and Y denote the predicted segmentation and the ground truth (i.e., the gold standard segmented by the radiologist), 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 [32]. For this, we use an adaptive moment estimator (Adam) [33] to estimate the parameters. Originally, Adam utilized the first and second moments of each gradient to update and correct the moving average of the current gradient. The parameters in our Adam optimizer are set to a learning rate of 0.0001 and a maximum number of epochs of 100. All the weights are initialized by a normal distribution with a mean of 0 and a standard deviation of 0.01 and all the biases are initialized as 0.

Data Augmentation To improve the robustness of the proposed U-Net based model, we artificially produce more training data from the original data by undertaking a set of image transformations, as summarized in Table 2. These are:

- Geometric transformations, such as flip, shift and rotation, which can result in displacement fields in the images. Shear operations, meanwhile can slightly distort the global shape of the LNs in the horizontal direction.
- Intensity transformations, which randomly jitter the intensity of the images by using a Gaussian random factor. This includes transformations of brightness and contrast.
- Elastic transformations [34], which generate more training data by using arbitrary but reasonable shapes. This ensures sufficiently variable training data, as LNs have no definite shapes.

2.5 Hilum Filling Via Morphological Operations

An echogenic hilum was a sonographic feature in 231 out of the 531 LNs in our dataset. Thus, after the U-Net based segmentation of the LNs, where hila were excluded, morphological operations were performed on the detected LNs to fill in the concave areas associated with the hila, thus making it possible to segment the entire set of LNs [35].

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

Thresholding We chose a threshold of 0.5 in relation to the probability map. All pixels below the threshold were set to zero, while the pixels above were set to one.

Opening and Closing The opening operation involves dilation of the erosion of the binary image, while the closing operation is erosion of the dilation of the image. The former removes small objects from the foreground (the white pixels) and places them in the background, while the latter removes

| Table 2 Summary of the applied image transformations |
|---------------------------------------------------|
| Transformations                            | Values                                |
| Flipping horizontally and vertically       | 50% probability in both directions     |
| Shifting horizontally and vertically       | 10% in both directions                |
| Shear                                     | 20% in the horizontal direction       |
| Rotation                                  | [− 40°, 40°]                           |
| Brightness and contrast                    | [0.7, 1.3]                            |
| Elastic transformation                     | \( \alpha = 720, \sigma = 24^a \)    |

\( a \) and \( \sigma \) control the degree of the elastic transformation.
small holes in the foreground and moves small islands of background to the foreground.

Convex Hull Computation Computing a convex hull involves constructing a non-ambiguous and efficient representation of the required convex shape. The concave region in a U-Net-detected LN, which represents the hilum, can be filled by adopting this approach.

3 Experiments and Results

3.1 Experimental Design and Settings

We compared the modified U-Net with two other traditional segmentation methods, namely reaction diffusion (RD) level set [6] and grab cut [8], which are representative of deformable models and graph theory-based models, respectively. We present below the segmentation results for the RD level set, the grab cut and our modified U-Net for LNs excluding hila and including hila, and for both the original ultrasound images and the GAD filtered images.

Taking the 531 LNs extracted from the 526 patients, we randomly split the LNs three parts: 390 for training; 51 for validation; and 90 for independent testing.

The experiments were conducted on an Ubuntu 16.04 desktop with 2 CPUs (Intel Xeon), 2 GPUs (NVIDIA GTX 1080ti 11 Gb), and 256 Gb RAM.

3.2 Quantitative Evaluation

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

\[
\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}
\]

\[
\text{SEN} = \frac{TP}{TP + FN} \tag{6}
\]

\[
\text{SPC} = \frac{TN}{TN + FP} \tag{7}
\]

where, TP, TN, FP, and FN denote the number of true positives, true negatives, false positives, and false negatives, respectively. These evaluation metrics comprehensively cover the segmentation performance from different perspectives. All of the values fall between 0 and 1.

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

\[
\text{DC} = \frac{2TP}{2TP + FP + FN} \tag{8}
\]

The JS is obtained using:

\[
\text{JS} = \frac{TP}{TP + FP + FN} \tag{9}
\]

Due to the non-normal distribution of the segmentation indices, their medians and interquartile ranges (IQRs) were also 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 will first look at the results for each step in the segmentation process, using the typical LN in Fig. 4 as an example.
Figure 4a and b illustrate the original LN image and its boundary. The GAD denoised image is shown in Fig. 4c and the U-Net result is shown in Fig. 4d. Figure 4e shows the results after the thresholding operation. Figure 4f shows the results after the opening, closing and convex hull operations. In the following sub-sections, we will look in detail at the effectiveness of the GAD despeckling, U-Net segmentation, and morphological operations.

3.3.1 GAD Denoising Results

As shown in Fig. 5, speckle noise contaminates ultrasound images, especially in the areas surrounding LNs. From the filtered images, we can see that applying GAD noticeably reduced the speckle noise, but also substantially enhanced the LN edges, thereby facilitating more accurate image segmentation.

3.3.2 Segmentation Results Excluding the Lymph Hilum

The automated segmentation results were compared with the corresponding ground truth by the radiologist. To observe the visual similarity of the shapes of the detected LNs, the contours of the LNs in the manual and automatic segmentation were extracted and marked with different colors. It can be seen from Fig. 6 that the modified U-Net with the GAD filtered images achieved the best segmentation effect. The modified U-Net located the position of the LNs precisely and segmented the contour of the LNs effectively, showing good consistency with the ground truth. The segmentation performance of the RD level set and grab cut were not satisfactory. The segmentation contour delivered by grab cut was rough and the lesion and non-lesion regions were not separated effectively. The RD level set was weak when it came to locating edges, especially in the case of images with strong speckle noise. On the whole, the segmentation effect for the GAD filtered images was better than it was for the original ultrasound images. As noted above, GAD can reduce speckle noise and enhance the LN edges, thus facilitating more accurate image segmentation.

It can be seen from Table 3 that all three segmentation methods performed better on the GAD filtered images than on the original ultrasound images. The RD level set, however, did not effectively segment the LNs and it consistently produced the lowest index values. Grab cut achieved the best sensitivity but the other indices were relatively low. The modified U-Net achieved the best LN segmentation result and the ACC, SEN, SPC, JS and DC values were 0.939, 0.879, 0.967, 0.763 and 0.866 respectively. Compared with the indices for the original ultrasound images, the SPC (p < 0.001), JS (p = 0.001), DC (p = 0.002) and ACC (p = 0.046) when using the modified U-Net on the GAD filtered images improved significantly, though the SEN significantly decreased (p = 0.009).

3.3.3 Final Segmentation Results

Figure 7 shows the final segmentation results after filling the concave hilum regions using the morphological operations detailed in Sect. 2.4. It can be seen that the segmentation performance using the modified U-Net on the GAD filtered images outperformed the other methods for both the original images and the GAD filtered images.

The ACC, SEN, SPC, DC and JS results for the RD level set, grab cut and our modified U-Net for LNs including hilum for both the GAD filtered images and the original ultrasound images are listed in Table 4. The ACC, SEN, SPC, JS and DC using the modified U-Net on the GAD filtered images were 0.934, 0.939, 0.937, 0.763 and 0.865, respectively. The modified U-Net generally achieved higher indices than the RD level set and grab cut. Compared with the indices for the original ultrasound images, the SPC (p < 0.001), JS
Segmentation of Lymph Nodes in Ultrasound Images Using U-Net Convolutional Neural Networks and Gabor-Based Anisotropic Diffusion

(p = 0.025) and DC (p = 0.041) for the modified U-Net with the GAD filtered images improved significantly, though the ACC increase was not significant (p = 0.108) and the SEN result significantly decreased (p = 0.040).

4 Discussion

Standard ultrasound imaging analysis includes lesion segmentation, feature engineering and diagnostic analysis. The lesion segmentation performance directly affects the subsequent analysis. Segmentation errors can result in feature deviation and established diagnostic models may be not satisfactory. The lesion segmentation of ultrasound images is extremely important and is one of the keys for precise computer-aided diagnosis.

We have proposed a novel framework for LN segmentation in ultrasound images based on the U-Net model, but also incorporating GAD filtering and morphological operations. First, the ultrasound image is denoised using GAD to suppress the speckle noise and enhance the LN edges. Then, the modified U-Net model is used to segment the LNs, excluding any hila. Finally, morphological operations are performed to fill in the concave regions associated with the hila and thus achieve the final segmentation of each entire LN.

We have proposed a two-stage segmentation framework for LNs, where we first segment the LNs excluding LN hila, then fill in the hila-related concave regions. To the best of our knowledge, this is the first time that an “excluding-then-filling hila” scheme has been explored for the segmentation of LNs in ultrasound images. This scheme is able to segment
LNs with echogenic hilar areas. This has previously been challenging when undertaking the direct segmentation of entire LNs because of the echogenicity of the hilar regions and the similarity of the adjacent soft tissue.

Introducing GAD as a way of suppressing the speckle noise in ultrasound images further facilitates LN segmentation. It employs a new edge detector based on the convolution of an input image with Gabor kernels. A good GAD filter can be seen as a well-trained convolutional layer that depends on the prior speckle noise in ultrasound images. In the future, we aim to develop an end-to-end convolutional network that can carry out denoising in its lower layers and segmentation in its higher layers.

When our proposed approach was compared with traditional segmentation methods, i.e., the RD level set and grab cut, the results indicated that the proposed framework for LN segmentation has a greater capacity to segment LNs automatically and accurately in ultrasound images. Traditional methods are based on pixel intensity information and prior knowledge, but lack image semantics. The proposed framework uses deep learning to acquire deep semantic information about ultrasound images, so that it can distinguish between lesion and non-lesion regions.

A limited number of images is one of the main challenges to applying deep learning to medical image analysis. We have sought to address this lack of samples by generating samples artificially via data augmentation, thus expanding the database. Three types of augmentation methods can be used to generate a much greater number of samples: geometric transformation; intensity transformation; and elastic deformation.

**Table 4** Final segmentation results using the RD level set, grab cut and our modified U-Net for LNs including hilar regions for both the original ultrasound images and the GAD filtered images

|                | ACC       | SEN       | SPC       | JS        | DC        |
|----------------|-----------|-----------|-----------|-----------|-----------|
| **Original**   |           |           |           |           |           |
| RD level set   | 0.866 (0.205) | 0.900 (0.096) | 0.821 (0.363) | 0.676 (0.276) | 0.807 (0.208) |
| Grab cut       | 0.897 (0.040) | 0.943 (0.015) | 0.850 (0.092) | 0.713 (0.172) | 0.832 (0.121) |
| Modified U-Net | 0.929 (0.075) | 0.949 (0.086) | 0.925 (0.095) | 0.736 (0.341) | 0.848 (0.242) |
| **GAD**        |           |           |           |           |           |
| RD level set   | 0.905 (0.134) | 0.921 (0.150) | 0.919 (0.315) | 0.698 (0.209) | 0.822 (0.148) |
| Grab cut       | 0.898 (0.049) | 0.950 (0.060) | 0.879 (0.100) | 0.720 (0.223) | 0.837 (0.156) |
| Modified U-Net | 0.934 (0.063) | 0.939 (0.103) | 0.937 (0.084) | 0.763 (0.247) | 0.865 (0.165) |

The denoted indices are the median (interquartile range).
transformation. These create different kinds of shapes and intensities in the training samples, thus improving the robustness of the dataset.

Although our method has achieved promising segmentation performance, there are some issues that will require further attention. First, adding GAD denoising and morphological operations to the U-Net model increases the computational labor. This is a key impetus to the development of an end-to-end model that fuses these three steps, as mentioned above. Other modifications may also be needed to improve the segmentation performance of the U-Net model, such as combining multiple segmentation maps [36] and extending the U-Net model with residual blocks [37]. Finally, while data augmentation is currently being used to address the problem of a limited number of images, an alternative approach is transfer learning. This uses deep models trained on natural images and transfers the learning to medical images. It has proven to be highly effective in several applications and it is worth exploring whether it can assist with the segmentation of LNs in ultrasound images [38].

5 Conclusion

In this study, we have presented an automatic segmentation method for LN ultrasound images based on a modified U-Net model and GAD. The original ultrasound images are first despeckled by a GAD filter. Three transformation methods are then performed to augment the ultrasound dataset and the modified U-Net model segments the LNs, excluding any hila. After this, morphological operations are employed to complete the segmentation of the LNs, including the hila regions. During experiments, the segmentation accuracy, sensitivity, specificity, Jaccard similarity and Dice coefficient reached 0.934, 0.939, 0.937, 0.763 and 0.865, respectively. This indicates that the proposed method has the capacity to effectively segment LNs in ultrasound images and that it may have the potential to facilitate AI-based diagnosis of LN diseases in the future.

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

Declarations

Conflict of interest None declared.

References

1. Evans, K. D., Sammet, S., Ramos, Y., & Knopp, M. V. (2008). Image segmentation for evaluating axillary lymph nodes. Journal of Diagnostic Medical Sonography, 24(6), 329–336.

2. Zhang, Y., Ying, M. T., Yang, L., Ahuja, A. T., & Chen, D. Z. (2016, December). Coarse-to-fine stacked fully convolutional nets for lymph node segmentation in ultrasound images. In 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 443-448). IEEE.

3. Pham, D. L., Xu, C., & Prince, J. L. (2000). A survey of current methods in medical image segmentation. Annual Review of Biomedical Engineering, 2, 315–337.

4. Sarle, W. S. (2012). Algorithms for Clustering Data. Technometrics, 32(2), 227–229.

5. Zhang, K., Zhang, L., Song, H., et al. (2013). Reinitialization-free level set evolution via reaction diffusion. IEEE Transactions on Image Processing, 22(1), 258–271.

6. Zhang, Q., Song, S., Xiao, Y., et al. (2019). Dual-mode artificially-intelligent diagnosis of breast tumours in shear-wave elastography and B-mode ultrasound using deep polynomial networks. Medical Engineering & Physics, 64, 1–6.

7. McNerney, T., & Terzopoulos, D. (1996). Deformable models in medical image analysis: A survey. Medical Image Analysis, 1(2), 91–108.

8. Kuo, J., Mamou, J., Wang, Y., et al. (2017). Segmentation of 3-d high-frequency ultrasound images of human lymph nodes using graph cut with energy functional adapted to local intensity distribution. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 64(10), 1514–1525.

9. Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations, 1(4), 1–14.

10. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90.

11. Gulshan, V., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA, 316(22), 2402.

12. Al-antari, M. A., Al-masni, M. A., Choi, M., Han, S., & Kim, T. (2018). A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. International Journal of Medical Informatics, 117(May), 44–54.

13. Dimauro, G., et al. (2019). Nasal cytology with deep learning techniques. International Journal of Medical Informatics, 122, 13–19.

14. Cinizzola, D., Giusti, A., Gambardella, L., & Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. Advances in Neural Information Processing Systems, 25, 2843–2851.

15. Avendi, M. R., Kheradvar, A., & Jafarkhani, H. (2016). A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Medical Image Analysis, 30, 108–119.

16. Cha, K. H., Hadjiiski, L. M., Samala, R. K., Chan, H.-P., Cohan, R. H., & Caoili, E. M. (2016). Comparison of bladder segmentation using deep-learning convolutional neural network with and without level sets. Medical Imaging: Computer-aided Diagnosis, 43(4), 97512.

17. Nida, N., Irtaza, A., Javed, A., Yousaf, M. H., & Mahmood, M. T. (2019). Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. International Journal of Medical Informatics, 124, 37–48.

18. Shellhammer, E., Long, J., & Darrell, T. (2017). Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(4), 640–651.

19. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848.

20. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881–2890).

21. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder–decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 1–14.

22. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention, 9351, 234–241.

23. Yuan, Y., Chao, M., & Lo, Y. (2017). Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. IEEE Transactions on Medical Imaging, 36(9), 1876–1886.

24. Alom, M. Z., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net). IEEE National Aerospace & Electronics Conference, 227(1–3), 327–333.

25. Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7), 629–639.

26. Zhang, Q., Han, H., Ji, C., Yu, J., Wang, Y., & Wang, W. (2014). Gabor-based anisotropic diffusion for speckle noise reduction in medical ultrasonography. Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 31(6), 1273–1283.

27. Rosário, P. W. S., et al. (2005). Ultrasonographic differentiation between metastatic and benign lymph nodes in patients with papillary thyroid carcinoma. Journal of Ultrasound in Medicine, 24(10), 1385–1389.

28. Evans, R. M., Ahuja, A., & Metreweli, C. (1993). The linear echogenic hilus in cervical lymphadenopathy—A sign of benignity or malignancy? Clinical Radiology, 47(4), 262–264.

29. Ying, M., Ahuja, A., Brook, F., & Metreweli, C. (2001). Vascularity and grey-scale sonographic features of normal cervical lymph nodes: Variations with nodal size. Clinical Radiology, 56(5), 416–419.

30. Dong, H., Yang, G., Liu, F., Mo, Y., & Guo, Y. (2017). Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks. Med Image Underst Anal Conf, 723, 506–517.

31. Milletari, F., Navab, N., & Ahmadi, S. A. (2016, October). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In 2016 fourth international conference on 3D vision (3DV) (pp. 565–571). IEEE.

32. Bottou, L. (2012). Stochastic gradient descent tricks. Neural Networks: Tricks of the Trade, 7700, 421–436.

33. Kingma, D. P. & Ba, L. (2014). Adam: A method for stochastic optimization. http://arxiv.org/abs/1412.6980.

34. Simard, P. Y., Steinkraus, D., & Platt, J. C. (2003, August). Best practices for convolutional neural networks applied to visual document analysis. In Int Conf Doc Anal Rec, (vol. 3, pp. 958–963).

35. Supriyanto, E., & Zulkifli, N. (2011). Abnormal tissue detection of breast ultrasound image using combination of morphological technique. In WSEAS CSCC Multiconference, (pp. 234–239).

36. Kayalibay, B., Jensen, G., & van der Smagt, P. (2017). CNN-based segmentation of medical imaging data. Bioelectrochemistry, 75(2), 130–135.

37. Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., & Pal, C. (2016). The importance of skip connections in biomedical image segmentation. Lecture Notes in Computer Science, 10008, 179–187.

38. Tajbakhsh, N., et al. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5), 1299–1312.