Medical Image Analysis 53 (2019) 104–110

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Automatic needle detection and real-time Bi-planar needle visualization during 3D ultrasound scanning of the liver

Muhammad Arif[a,b,*], Adriaan Moelker[b], Theo van Walsum[a,b]

[a] Department of Medical Informatics, Erasmus MC, University Medical Center Rotterdam, the Netherlands
[b] Department of Radiology & Nuclear Medicine, Erasmus MC, University Medical Center Rotterdam, the Netherlands

ARTICLE INFO

Article history:
Received 23 July 2018
Revised 8 January 2019
Accepted 1 February 2019
Available online 2 February 2019

Keywords:
Needle tracking
Ultrasound image guidance
Percutaneous interventions
Convolutional neural network
Automatic needle segmentation
Real-time

ABSTRACT

2D ultrasound (US) image guidance is used in minimally invasive procedures in the liver to visualize the target and the needle. Needle insertion using 2D ultrasound keeping the transducer position to view needle and reach target is challenging. Dedicated needle holders attached to the US transducer help to target in plane and at a specific angle. A drawback of this is that, the probe is fixed to the needle and cannot be rotated to assess the position of the needle in a perpendicular plane. In this study, we propose an automatic needle detection and tracking method using 3D US imaging to improve image guidance and visualization of the target in the liver with respect to the needle during these interventional procedures. The method utilizes a convolutional neural network for detection of the needle in 3D US images. In a subsequent step, the output of the convolutional neural network is used to detect needle candidates, which are fed into a final tracking step to determine the real needle position. The needle position is used to present two perpendicular cross-sectional planes of the 3D US image containing the needle in both directions. Performance of the method was evaluated in phantoms and *in-vivo* data by calculating the needle position distance and needle orientation angle between segmented needles and reference ground truth needles, which were manually annotated by an observer. The method successfully detects the needle position and orientation with mean errors of 1 mm and 2°, respectively. The proposed method yields a robust automatic needle detection and visualization at a frame rate of 3 Hz in 3D ultrasound imaging of the liver.

1. Introduction

Generally, minimally invasive techniques are used for diagnosis and treatment. Compared to open surgery, these procedures are aimed to reduce the risk of surrounding tissues injury and lead to less trauma and shorter recovery time (Slakey et al. 2013). Percutaneous biopsy in the liver is a minimally invasive procedure where a needle is used to remove tissue samples for histology-based diagnosis. Needle insertion is also performed in therapeutic procedures such as in prostate brachytherapy and in liver tumor thermal ablation (radiofrequency, microwave or cryo ablation). Image guidance is required for such procedures to visualize both, the target and the needle. UltraSound (US) image guidance is widely used in such minimally invasive procedures, as it is relatively safe, cheap and real-time imaging modality. Conventionally, radiologists use 2D US to visualize the needle during insertion in the liver and mainly use two techniques for needle insertion: guided and freehand (Chapman et al., 2006). In the guided needle insertion approach, a detachable needle guide is attached onto the 2D US probe to guide the needle; this only allows needle insertion at a few pre-defined planes and angles. In the freehand insertion technique, the needle is inserted without use of a guide. This provides greater flexibility in choosing the angle of insertion of the needle. However, the freehand approach is more challenging than the guided approach as it requires the interventionalist to continuously align the US image with the needle for visualization. 3D US has the potential to address some of the challenges associated with 2D US: It may allow easier free-hand needle insertion and provide better image guidance as the transducer is imaging a large region. One could even fix, e.g. by a probe holder device, the transducer such that the interventionalist can focus only on the needle insertion. In addition, a 3D volume allows visualization of the target in multiple directions enabling the interventionalist to hit the target in the center more easily. This is advantageous, especially in ablative therapies. However, the visualization of the needle trajectory in 3D US imaging is challenging (Arif et al., 2017), as direct volume rendering is not a good solution. If the needle position in the 3D

* Corresponding author at: Department of Medical Informatics, Erasmus MC, University Medical Center Rotterdam, Wytemaweg 80, Room Na 2506 Erasmus MC, 3015 CN Rotterdam, the Netherlands.
E-mail address: a.muhammad@erasmusmc.nl (M. Arif).

https://doi.org/10.1016/j.media.2019.02.002
1361-8415/

US image is known, one can generate multi-planar reconstructions (MPR) aligned with the needle. This makes it possible to generate views which are conventionally used in 2D US without being limited to certain introduction angles, and without having to manually align the US image with the needle. The purpose of our work is to develop and evaluate such a needle tracking approach for 3D US images.

Some external tracking devices (Barratt et al., 2001; Wei et al., 2004; Fronheiser et al., 2007; De Lorenzo et al., 2011) have been proposed to track the needle in 3D US imaging to improve image guidance. However, these external guidance systems require additional equipment, time to calibrate and specific skills to operate such systems. Image based guidance techniques (Uhercik et al., 2013; Zhao et al., 2013; Pourtaherian et al., 2014) have also been developed. Pourtaherian et al. (2014) presented a 3D Gabor transformation method for needle detection. They selected needle candidate voxels using directionally-sensitive Gabor wavelet transforms and estimated the needle axis. The study was performed on a limited set (32 3D US images) of data, created by flipping in x and y directions eight different 3D US images acquired from a chicken breast phantom using an iU22 US system. Uhercik et al. (2013) used line filtering, voxel classification and model fitting techniques to detect a needle in 3D US images. Zhao, Cachard et al. (2013) used a combination of RANSAC and a Kalman filter to detect and track a needle in 3D US images. In both studies high quality simulated US data were used to evaluate the performance of the method. In clinical cases, 3D US images may not be of such high quality, which may lead to a decreased robustness and increased processing time for the methods. Image based guidance methods do not need any additional external devices, which is an advantage over tracking based methods and it is the same approach we will pursue in this study.

Convolutional neural networks (CNN) have recently shown to be very effective in segmentation of the organs in various medical imaging modalities (Ronneberger et al., 2015; Milletari et al., 2016) and also in US (Baka et al., 2017). Particularly the U-net and V-Net architecture and variants thereof have demonstrated superior performance compared to traditional, non-deep learning approaches in several medical imaging challenges. These biomedical image segmentation networks are composed of several convolutional layers and with data augmentation. After brief training, data can be generalized well and can predict output as binary segmented images.

In our approach, we have combined such a convolutional neural network with conventional image processing techniques for automatic detection and tracking of needles in 3D US images of the liver. The contributions of our work are: (1) a robust approach for automatic needle detection and tracking in subsequent images, (2) real-time visualization of the tracked needle using 2D cross-sectional planes of 3D US data and (3) performance evaluation of the proposed method in phantoms and partially in-vivo.

2. Method

Fig. 1 shows the block diagram of the proposed method of automatic needle detection in 3D US images. The method consisted of four main stages. At stage-1, a CNN model was used to get an initial needle segmentation from the 3D US image; at stage-2, needle candidates were extracted from this initial segmentation; at stage-3, needle selection was performed and at stage-4 the resulting needle was used to generate 2D cross-sections out of 3D US data containing the complete needle for visualization and image guidance. The CNN model, needle candidate extraction, needle detection and visualization step are described in Fig. 1.

2.1. CNN model

The needle is a straight linear structure and in order to accurately and robustly discriminate needle voxels from other similar echogenic structures such as bones and tissue boundaries in 3D US images, a CNN approach was used instead of traditional image based methods (Zhao et al., 2013; Pourtaherian et al., 2014). In literature, use of a compressed CNN (a network with fewer convolutional layers and fewer features) has shown advantages in prediction time (Guo et al., 2018). Therefore, an adapted compressed version of the V-net (Ronneberger et al., 2015) model was implemented to segment the needle in 3D US images. While in a V-Net, the network consists of one to three convolutions in each layer, the network utilized in this paper contains only a single 3D convolution with $3 \times 3 \times 3$ kernel size, followed by a Rectified Linear Unit (ReLU). At the 5th and 6th layer of the network, a kernel size filter of $3 \times 3 \times 1$ was used due to small image size along z-axis. Furthermore, the input of each stage was not added to output of last convolutional layer of that stage as described for the V-net to learn residual function. Batch normalization (BN), was added after each 3D convolution to improve convergence speed during training (Ioffe and Szegedy, 2015). In the up-sampling part, instead of using deconvolutions in every step, 3D up-sampling of the feature map was used. Additionally, a concatenation with the corresponding computed featured map from down-sampling part, and one 3D convolution with a $3 \times 3 \times 3$ kernel size, ReLU and BN were performed after up-sampling. At the final layer a 3D convolution with a kernel of $1 \times 1 \times 1$ size was used to predict the needle segmentation from the computed features. In each convolution appropriate padding was used. The complete network design is shown in Fig. 2.

2.2. Needle candidate extraction

The CNN model generates an image where voxels that belong to a needle having high values. The purpose of next step was to extract needle candidates from the CNN output. Therefore, in this step, we used a connected component labeling to obtain all connected segments $S_i\,(1 < i \leq N_{seg})$, where N_{seg} is number of segments. All segments with size (number of voxels, v_n) smaller than a threshold T_{seg} were assumed to be false detections and ignored in further processing. Subsequently, Principle Component Analysis (PCA) was applied to determine the direction vector $\vec{u}_i$ (eigenvector for largest eigenvalue) of the remaining segments, and the center of mass M_i of each segment was determined.

$$\vec{u}_i = PCA\ (N_i) \tag{1}$$

$$M_i = \frac{1}{N_i} \sum_{i=1}^{n} v_n \tag{2}$$

In a last step, we created a straight needle model by grouping the segments belonging to one straight line. To this end, segments were first sorted out on the basis of size. Starting from the longest segment S_l, the shortest distance between the line through the longest segment and center of mass of each remaining segment S_j was calculated and all remaining segments with a distance smaller than d belonged to one group.

$$Group(S_l) = \left\{ S_j | distance\ \left(S_l, S_j\right) < d \right\} \tag{3}$$

where,

$$distance\ \left(S_l, S_j\right) = \left| (M_l - M_j) - \left[\left(\left(M_l - M_j\right).\ \widehat{u_l} \right) \widehat{u_l} \right] \right| \tag{4}$$

with $\widehat{u_l}$ the normalized eigenvector belonging to the largest eigenvalue of S_l. This process was iterated until all segments were assigned a group, giving groups G_k, $(k = 1, 2, 3, \ldots k)$. For each group a needle model was obtained by fitting a line through all

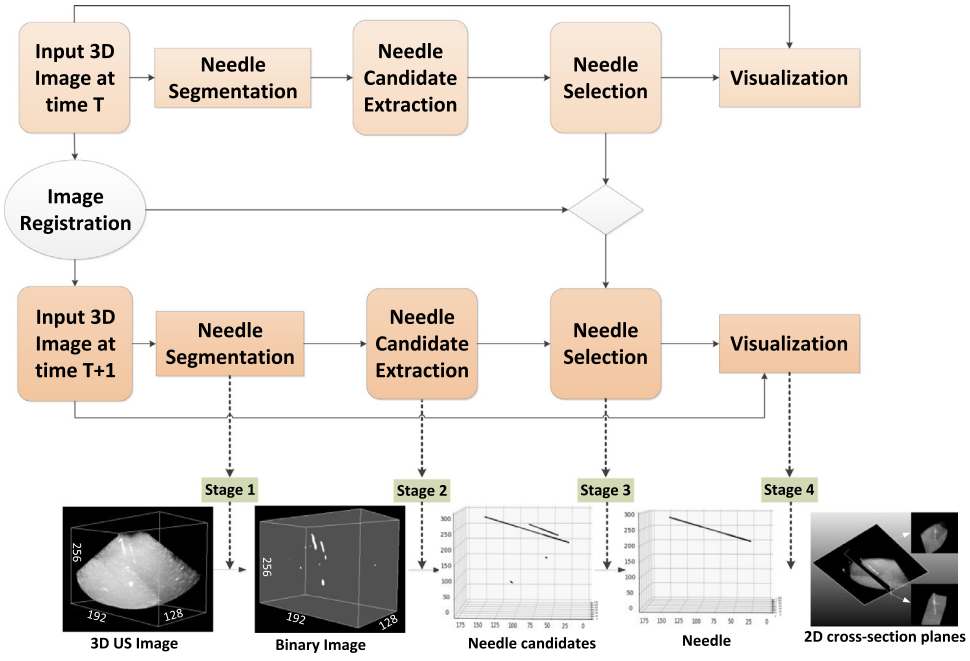

Fig. 1. Block diagram of automatic needle detection in 3D US images method.

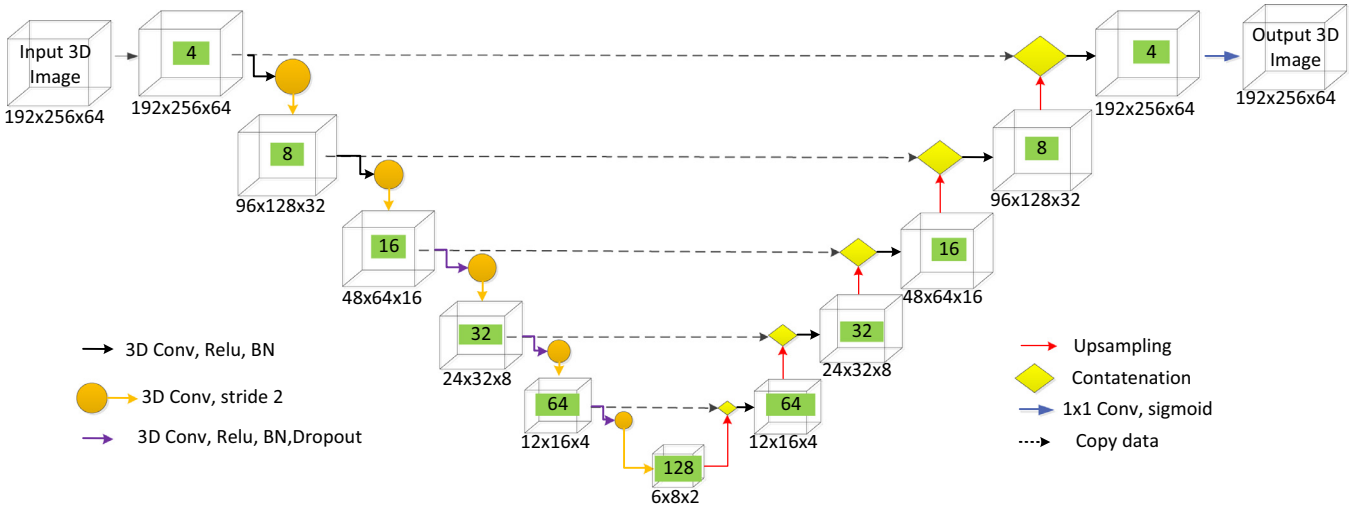

Fig. 2. A schematic representation of our convolutional neural network (CNN) model in the V-net architecture. The number of channels is mentioned in each 3D volume and corresponds to a multi-channel feature map.

voxels of all segments in group k. Distal and proximal end points (D_k, P_k) were determined from the needle voxels that projected onto the extreme ends of these lines.

2.3. Needle tracking

The needle candidate extraction results in N_{needle} candidates. Purpose of the needle tracking step was to select the real needle from these candidates, using information from the previous time step. To that end, a needle tracking step was used, where the needle extraction result was combined with result from the previous time step. Assuming that the needle only moved forward or backward in the image, a needle selection algorithm was used to calculate distance between the current position (N_t) and the previous position (N_{t-1}) of the needle:

$$Dist\_needle = D(N_{t-1}, N_t) \Big/ \frac{(\vec{n}_{t-1}.\vec{n}_t)}{(|\vec{n}_{t-1}|.|\vec{n}_t|)} \quad (5)$$

Here $D(N_{t-1}, N_t)$ is Euclidean distance, calculated as the average point-line distance between two end points of the initial and new positions of the needle. In Eq. (5), the denominator part is used to calculate geometric (orientation) distance between two needle positions. For parallel needles geometric distance approaches to 1 and the distance between the needles at two temporal positions depends on the Euclidean distance only. For diverging needles, the Euclidean distance is scaled with a value that becomes larger as a function of the angle between the needles. At the end, needle with the shortest distance was selected.

The assumption of linear needle motion was invalid in case of transducer motion (translation or rotation) and it may have affected accuracy of needle selection. To compensate motion effect relative transformation between current image and previous image was determined using a fast and robust rigid image registration for 3D US images (Banerjeeet al., 2015) and was used to transformed the needle position of the current image to the previous image before calculating the shortest distance.

2.4. Needle visualization

The needle model that resulted from previous steps, allowed an effective visualization of the 3D US imaging data. Inspired by conventional visualization using 2D US, two perpendicular cross-sectional planes (parallel and perpendicular to transducer) of the 3D US image containing full-length needle were displayed using position of the tracked needle in 3D US image. Tracked needle position was shown by two planes in three dimensions with respect to target and provided better view of collateral structures such as vessels. Two plane visualization of the needle and anatomical structures is comparatively easy to interpret for the radiologist and could improve image guidance (Chapman et al., 2006).

3. Experiments and results

We have performed two experiments. In the 1st experiment we have evaluated the needle segmentation approach (stage-1), and assessed effect of the compressed CNN on the performance of the method. In the 2nd experiment we have applied the method on two 3D US sequences and investigated tracking performance of the method.

Details of the data used and the evaluation criteria will be addressed first and then the experiments and results will be described.

3.1. Data

Six sequences of 3D US images at frame rate of 6 Hz were acquired using an US system (iU22, Philips Healthcare, Netherlands) with a 3D transducer (X6-1), when puncturing phantoms with a commercially available ablation needle (HS-14 G, 20 cm, HS Hospital Services S.P.A, Aprilia Italy). The ablation needle had Teflon coating. From those six sequences, four sequences were acquired in a triple modality 3D abdominal phantom, Model 057A (CIRS, Norfolk, Virginia, USA), while two were acquired from a soft-tissue (liver) mimicking PolyVinyl Alcohol (PVA) phantom (Pluymen, 2016; Arif et al., 2017). The phantom was made from a 5% aqueous solution of polyvinyl alcohol (PVA). Silica gel particles (1%) were added to the solution to act as US scatters. The size of the PVA phantom was 30 cm x 15 cm x 16 cm. Each 3D US image had a dimension of 192 × 256 × 128 voxels, with voxel size 1.144 mm x 0.594 mm x 1.193 mm. During acquisition of these sequences, the needle was punctured using the free-hand method at different angles ranging from 25° to 50°. After needle insertion, the transducer was moved translationally and rotationally with respect to needle. We also acquired seventeen 3D US liver biopsy images (378 × 512 × 22 voxels) from ten patients who underwent a diagnostic liver biopsy. A commercially available biopsy needles without any special coating (HS-18 G, 20 cm, HS Hospital Services S.P.A, Aprilia Italy) was used during the procedures. Patient's biopsy images were down-sampled to 192 × 256 × 128 voxels to have the same size as that of phantom data. Informed consent was obtained and the Institutional Review Board (IRB) of the Erasmus MC stated that IRB approval was not required for this study. In liver biopsy procedure, guided needle insertion technique was used.

3.2. Reference standard

During the puncturing procedure of phantoms, needle movement was not very fast and there was no tangible change in needle position between sequential frames. Whereas during the transducer motion, the apparent change in needle position was substantial between sequential frames. Therefore, every 8th image was selected during the puncturing procedure of needle and every 2nd image was selected during transducer motion from the 3D sequences. These selected images were associated with their ground truth, i.e. a binary image (mask) where the needle voxels had value of 1 and the background pixel value 0 for training of network. The ground truth (mask) of the needle was annotated by an experienced US observer, having six years of US imaging and processing experience. The annotation was done by clicking (visible) distal and proximal end of needle, which defined the line of needle in the 3D US image. For training of the network, a binary mask was created from line between these two points, which was subsequently dilated with a $5 \times 5 \times 5$ kernel. Seventeen 3D US liver biopsy images from five patients were also used for experiments. 3D US images were divided into two sets (Set1, Set2) as described in Table 1 and used in experiments.

Table 1
Data sets of 3D US images used in experiments.

Set 1	Set 2
Seq1+Seq2 (Abdominal Phantom)	Seq4+Seq5 (Abdominal Phantom)
Seq3 (PVA Phantom)	Seq6 (PVA Phantom)
Patient (Images = 8)	Patient (Images = 9)
Total 3D US image = 80	Total 3D US image = 69

The performance of the method was evaluated in terms of needle detection time, needle position distance (E_{np}) and needle orientation angle (E_θ). The needle position distance was calculated as the average point-line distances between the points on the manually annotated needle (reference ground truth) axis and the automatically detected needle axis. The needle orientation angle was calculated between the direction vectors of the reference ground truth and the automatically segmented needle.

3.3. Implementation

The training of the network was implemented in Keras (version 2.0.2) with Tensor Flow (version 1.0.1) as backend in Python (version 3.5.3). During training, each 3D US image was divided into two volume patches (192 × 256 × 64 voxels). The training was performed on a GPU (GeForce GTX 980Ti). Prediction using trained network was performed on a GeForce GTX TITAN Xp GPU (The Titan Xp used for this research was donated by NVIDIA Corporation). The loss function during training was the Dice overlap metric calculated between the ground truth masks and the output segmented image of the model (Milletariet al., 2016). The loss function of CCN model was optimized using the Adam optimizer (Kingma and Ba, 2014) with a learning rate of 0.001. To prevent overfitting, we added dropout in the last three layers in down-sampling part of the network (Srivastava et al., 2014). As the number of annotated data was limited, we implemented data augmentation; rotation (0–15°, along *x,y,z*-axes), flipping (along *x,y*-axes) with rigid transformation with 50% probability to augment each image during training. This allowed the network to learn invariance to such deformations and also helped to prevent overfitting and generalize better. The output images from the trained model contained voxel values from 0 to 1, where each predicted needle voxel had a value closer to 1.

We used a threshold (T_{seg}) of 5 voxels on the output images from the CNN to remove smaller segments and a small distance (d) of 10 mm, to group the segments as described in needle candidate extraction. In the needle tracking process, the initial needle position N_t was chosen by selecting the largest segmented part of the needle from a 3D US image I_t. For subsequent images I_{t+1} offline registration was performed between each pair of images (I_t, I_{t+1}).

3.4. Experiment 1: optimization and evaluation of the method

In this experiment we investigated the effect of compression and the depth of the CNN on the performance of the method. A

Table 2
Performance of the method at different CNN depth.

Proposed network	Feature channels in 1st layer	Needle detection time per image (msec)	Average needle position distance (E_{np} mm)		Average needle orientation angle (E_θ)	
			Set 1	*Set 2*	*Set 1*	*Set 2*
CNN4_2	2	165	2.01	2.50	2.34°	6.04°
CNN4_4	4	180	1.56	1.14	1.59°	2.24°
CNN5_2	2	165	1.07	1.16	1.61°	2.58°
CNN5_4	4	180	0.99	1.11	1.56°	2.07°
CNN6_2	2	165	0.87	1.06	1.49°	2.32°
CNN6_4	4	180	0.78	0.90	1.29°	1.79°

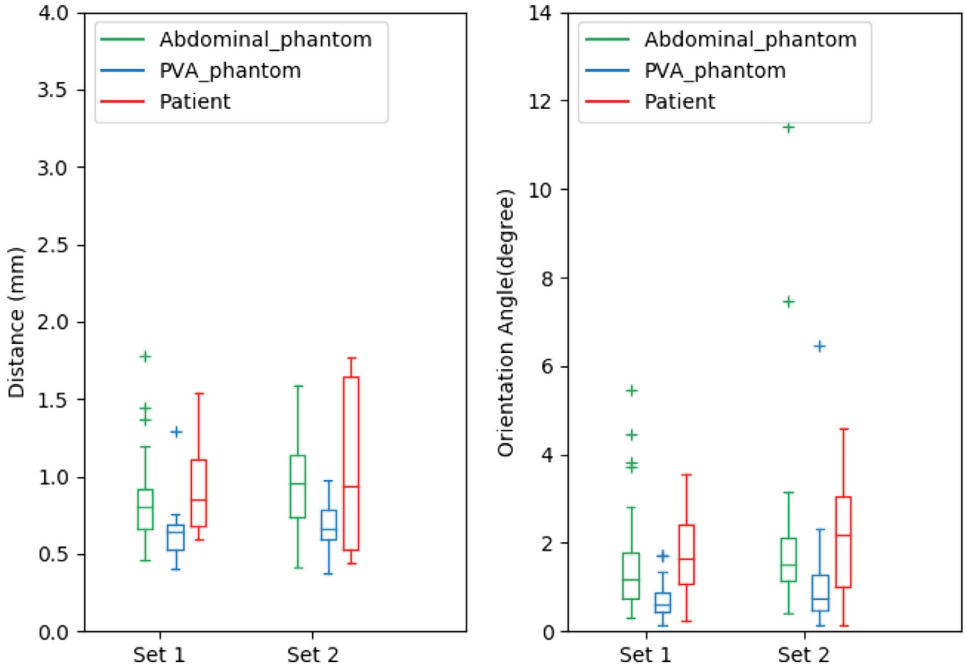

Fig. 3. (a) Needle position distance (mm) from reference ground truth to detected needle axis for two data sets. (b) Needle orientation angle between reference ground truth and detected needle axis.

shallow and compressed CNN may a segment needle in less time and make needle tracking faster, but potentially at the expense of the needle detection accuracy, as a deeper network with more convolutions can learn more features (Ba and Caruana, 2014). We therefore evaluated networks of three different depths (4, 5 and 6) named CNN4, CNN5 and CNN6 to find out optimal depth and two settings for number of features in first layer (2 and 4) as described in Table 2.

The network was trained and the method was evaluated in a two-fold cross-validation setup using data as described in Table 1. The evaluation results on the two sets in Table 2 show that the needle detection time per image is slightly better in networks with fewer features at the 1st layer. They also show that the addition of layers in the network does not effectively change computation time. The needle position and orientation accuracy decreases with fewer features in 1st layer, and also with the depth of network. As the difference in timings is minor, a six layers deep CNN model with 4 feature channels at the 1st layer (CNN6_4) was selected for our proposed method in our next experiments. The needle position distances and orientation angles for this final network with 6 layers are shown in Fig. 3.

Four examples showing 2D view of the needle are presented in Fig. 4. The needle was segmented correctly in PVA, abdominal and patient data.

3.5. Experiment 2: evaluation of the method on sequence

In experiment 1, we have investigated some of the hyper-parameters with respect to the needle detection accuracy and computational timings in phantoms and in clinical 3D US liver biopsy images acquired from patients. In this experiment, we investigated the method of choice resulting from the previous experiment for real-time needle tracking in 3D US sequence. We trained our network using images from Set1 and tested it on two sequences (4 and 6) acquired in an abdominal and PVA phantom from Set 2. 3D US sequences from patients were not available.

The needle detection (segmentation) and tracking frame rate, including the candidate extraction time in two sequences, is 3 frames per second (See Table 3). The needle visualization process of these two sequences is shown in attached videos. In two videos (Video 1 & Video 2), needle is shown in two planes extracted from original 3D US image based on measured the needle coordinates. This implementation, with the post-processing of results implemented in Python, runs at 3 frames per second (320 ms/frame). Slightly more than half of the time is spend in segmentation (180 ms), and the remaining time in tracking (140 ms). Segmentation time is not much effected by image size, whereas post-processing and tracking depends. We therefore additionally experimented adding a down-sampling step to the network (av-

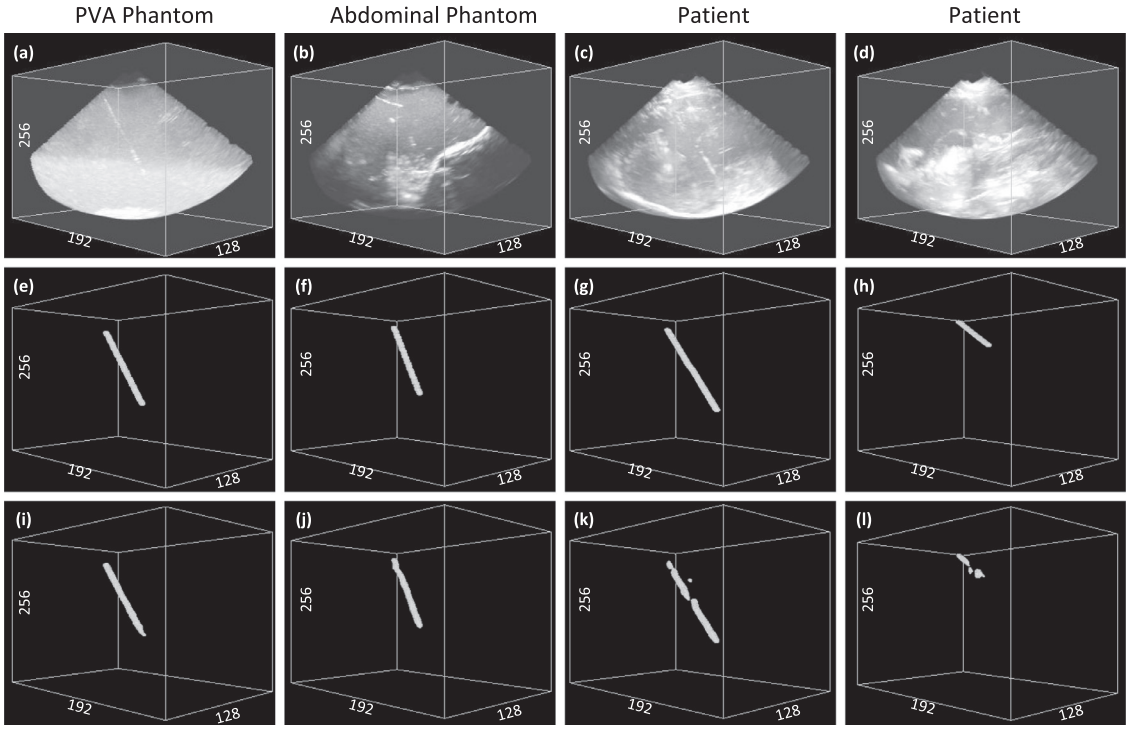

Fig. 4. (a–d) 2D view of 3D US images acquired from PVA, abdominal phantoms and patients. (e–h) Ground truth needle manually annotated by an observer. (i–l) Binary image segmentation of the needle.

Table 3
Needle segmentation and post-processing (candidate extraction and tracking) time for two sequences.

Proposed network	Needle segmentation time per image	Needle tracking time per image	Total time per image	E_{np}	E_{θ}
CNN6_4 (Output=192 × 256 × 128)	180 ms	140 ms	320 ms	0.86 mm	1.72°
CNN6_4 (Output=96 × 128 × 64)	172 ms	18 ms	190 ms	1.47 mm	1.85°

erage pooling layer), which allowed to run the same algorithm at 5 Hz (Table 3). From both sequences 42 images out of 224 images, with different needle position and orientation, were selected by using the reference standard criteria (as described in Section 3.2). Performance of the method was evaluated with original image size (192 × 256 × 128 voxels) and down-sampled image size (96 × 128 × 64 voxels) by calculating the average needle position (E_{np}) and orientation (E_{θ}) accuracy. The results showed that, by down-sampling image size, needle position accuracy decreases more as compared to orientation accuracy.

4. Discussion

Ultrasound image guidance is commonly used for minimally invasive procedures. However, the procedure of needle insertion using 2D US keeping the transducer position to view the needle and reach the target is challenging. Automatic needle detection in 3D US could overcome these limitations and ease image guidance procedures. The objective of this study was to develop a robust needle detection and tracking method in 3D US imaging with high precision and accuracy. We used a CNN as a preprocessing step to obtain an initial needle segmentation in 3D US images; the network is a compressed version of the V-net (Milletari et al., 2016). In a subsequent post-processing, the needle was extracted and tracked in each 3D US image. The average method run time per image was 320 ms (CNN6_4 = 180 ms & tracking = 140 ms) i.e. the method runs at 3 Hz, making it possible to use the algorithm for needle visualization during interventional procedures. Our US system produces 3D US images at 6 Hz, thus the needle position could be updated with this system every other frame. Further improvement in computation time was investigated by down-sampling the output from the CNN. Table 3 shows that by decreasing the image size to half in each dimension, the needle tracking time decreased from 140 ms to 18 ms, which leads to an average needle detection time of 190 ms with E_{np} = 1.59mm & $E_{\theta} = 1.74^0$, which is approximately 5 Hz. Another approach to reduce computational time could be obtained by running the needle detection in a localized region (close to the needle position in the previous frame), combined with an image-based correction for transducer motion. Alternatively, a setup where the needle position is updated every *n*th frame and the position is copied (also possibly adapted by an image-based correction of transducer motion) to in between frames may provide a working solution for 3D imaging rates of 6–10 Hz with same needle detection accuracy.

The results showed that for computational time, number of features in first layer was more relevant than the depth of network. The computational time for needle segmentation did not increase by adding extra convolutional layers to the network. The reason could be that at deeper layers, convolution with smaller image sizes is not computationally expensive as compared to the initial network layers. However, number of feature channels extracted at initial layers can affect computation time. Therefore, the needle detection time was improved slightly by learning fewer features at initial layers, but accuracy decreased simultaneously. By selecting

comparatively less deep network, the results became unsatisfactory. The performance of the method was improved by addition of extra layers; however, the image size at the deepest (sixth) layer ($6 \times 8 \times 2$ voxels) was very small. Further addition of layers would not improve the results.

The results in Table 2 also showed that the proposed CNN could be used to achieve an accuracy similar to other methods employing conventional techniques (Pourtaherian et al., 2014, 2015). The proposed method was evaluated on several 3D US images acquired in two types of phantoms (PVA and Abdominal) and also on patient data. Average needle position and orientation accuracy achieved were 0.86 mm and 1.54° respectively. In (Pourtaherian et al., 2017), a method based on orthogonal-plane convolutional network was used to detect needles in 3D images. The authors achieved a computation time of 2.2 s on average for needle detection in 3D US image using GeForce GTX TITAN X GPU. The method presented runs approximately six-times faster on the same hardware. The needle position and orientation results of both approaches are in the same range (less than 1 mm and 2°) (Pourtaherian et al., 2015). Thus, the method presented, using a data driven approach, seems to be as good as a more conventional image processing approach, but runs at a higher speed, an observation which is common for data driven segmentation approaches. Note that data used is obtained from different sources, i.e. 3D US images acquired in two types of phantoms (PVA and Abdominal) and patient data, whereas in (Pourtaherian et al., 2017) chicken breast phantom data was used. One should, hence, be careful in the drawing strong conclusions from this comparison.

In our experiment with needle tracking, the transformation between two subsequent image frames was computed offline using a rigid image registration (Banerjee et al., 2015). However, the authors (Banerjee et al., 2015) have reported that registration could be implemented at 8 Hz on a GPU. Therefore, by performing the registration process parallel to needle segmentation on another GPU, the needle detection and tracking process should not take more time than reported on a dual-GPU system. The purpose of image registration was to automatically compensate transducer motion, which would be required for free hand needle insertion. The needle detection and tracking after registration can be viewed in two attached videos. The videos show that the transducer motion was compensated correctly and that the needle was tracked accurately. For visualization, exact needle tip location is not relevant, as it can be deduced by observer automatically. Correct location of needle plane is of more importance using 3D US for image guidance procedures.

Our study has some limitations. First, we only had a limited data set for training of the network, among which 17 patient's images were used. By using data augmentation, we were able to train the network such that it was able to detect the needle with sufficient accuracy. Retraining with more images may further improve segmentation results and may possibly supersede subsequent needle detection. Secondly, we did not evaluate complete tracking on patient live image sequences, as these were not available. However, we perceived that similar results could be obtained with live data provided that transducer and needle motion was similar to the motion in our experiment. Main reason for this was that the needle segmentation, which was the first crucial step, was evaluated on patient data.

5. Conclusion

We have presented a robust approach for automatic needle detection, tracking and visualization using 3D US for image guidance during interventional procedures in the liver. The method uses a combination of a convolutional neural network and a post-processing step to track the needle in 3D US images. The evaluation of the method shows that technique detects the needle position and orientation with mean errors of 1 mm and 2° respectively. Needle tracking with original 3D US volumes ($192 \times 256 \times 128$ voxels) and down-sampled 3D US volumes ($96 \times 128 \times 64$ voxels) is achieved at a rate of 3 Hz and 5 Hz respectively.

Conflict of interest

None.

Acknowledgment

This research project Navigate is supported by the ZonMW project number 104003015.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2019.02.002.

References

Arif, M., Moelker, A., et al., 2017. Needle tip visibility in 3D ultrasound images. Cardiovasc. Interv. Radiol.

Ba, L.J., Caruana, R., 2014. Do deep nets really need to be deep? In: Proceedings of the 27th International Conference on Neural Information Processing Systems – Vol. 2. Montreal, Canada. MIT Press, pp. 2654–2662.

Baka, N., Leenstra, S., et al., 2017. Ultrasound aided vertebral level localization for lumbar surgery. IEEE Trans. Med. Imaging 36 (10), 2138–3147.

Banerjee, J., Klink, C., et al., 2015. Fast and robust 3D ultrasound registration – block and game theoretic matching. Med. Image Anal. 20 (1), 173–183.

Barratt, D.C., Davies, A.H., et al., 2001. Optimisation and evaluation of an electromagnetic tracking device for high-accuracy three-dimensional ultrasound imaging of the carotid arteries. Ultrasound Med. Biol. 27 (7), 957–968.

Chapman, G.A., Johnson, D., et al., 2006. Visualisation of needle position using ultrasonography. Anaesthesia 61 (2), 148–158.

De Lorenzo, D., Vaccarella, A., et al., 2011. Accurate calibration method for 3D free-hand ultrasound probe using virtual plane. Med. Phys. 38 (12), 6710–6720.

Fronheiser, M.P., Idriss, S.F., et al., 2007. Vibrating interventional device detection using real-time 3D color Doppler. In: 2007 IEEE Ultrasonics Symposium Proceedings, Vols 1–6, pp. 868–871.

Guo, J., Zhou, B., 2018. Model compression for faster structural separation of macromolecules captured by Cellular Electron Cryo-Tomography. ArXiv:1801.10597.

Ioffe, S., Szegedy, C., 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd International Conference on International Conference on Machine Learning, 37. JMLR.org, Lille, France, pp. 448–456.

Kingma, D.P., Ba, J., 2014. Adam: a Method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR).

Milletari, F., Navab, N., et al., 2016. V-Net: fully Convolutional neural networks for volumetric medical image segmentation. In: Proceedings of 2016 Fourth International Conference on 3d Vision (3dv), pp. 565–571.

Pluymen, L.H., 2016. Liver Tissue Mimicking Materials for Image-Guided Needle Interventions Master Thesis. TU Delft.

Pourtaherian, A., Ghazvinian Zanjani, F., et al., 2017. Improving Needle Detection in 3D Ultrasound Using Orthogonal-Plane Convolutional Networks. Springer International Publishing, Cham.

Pourtaherian, A., Zinger, S., et al., 2014. Gabor-based needle detection and tracking in three-dimensional ultrasound data volumes. In: 2014 IEEE International Conference on Image Processing (Icip), pp. 3602–3606.

Pourtaherian, A., Zinger, S., et al., 2015. Benchmarking of state-of-the-art needle detection algorithms in 3D ultrasound data volumes. Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling 9415.

Ronneberger, O., Fischer, P., et al., 2015. U-Net: convolutional networks for biomedical image segmentation. Med. Image Comput. Comput.-Assist. Interv. 9351, 234–241 Pt Iii.

Slakey, D.P., Simms, E., et al., 2013. Complications of liver resection: laparoscopic versus open procedures. JSLS 17 (1), 46–55.

Srivastava, N., Hinton, G., et al., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Uhercik, M., Kybic, J., et al., 2013. Line filtering for surgical tool localization in 3D ultrasound images. Comput. Biol. Med. 43 (12), 2036–2045.

Wei, Z.P., Gardi, L., et al., 2004. Oblique needle segmentation for 3D trus-guided robot-aided transperineal prostate brachytherapy. In: 2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano, Vols 1 and 2, pp. 960–963.

Zhao, Y., Cachard, C., et al., 2013. Automatic needle detection and tracking in 3D ultrasound using an ROI-based RANSAC and Kalman method. Ultrason. Imaging 35 (4), 283–306.