

Research article

Improving needle visibility in LED-based photoacoustic imaging using deep learning with semi-synthetic datasets

Mengjie Shi^a, Tianrui Zhao^a, Simeon J. West^b, Adrien E. Desjardins^{c,d}, Tom Vercauteren^a, Wenfeng Xia^{a,*}

^a School of Biomedical Engineering and Imaging Sciences, King's College London, London SE1 7EH, United Kingdom

^b Department of Anaesthesia, University College Hospital, London NW1 2BU, United Kingdom

^c Wellcome/EPSCRC Centre for Interventional and Surgical Sciences, University College London, London W1 W 7TY, United Kingdom

^d Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, United Kingdom

ARTICLE INFO

Keywords:

Photoacoustic imaging

Needle visibility

Light emitting diodes

Deep learning

Minimally invasive procedures

ABSTRACT

Photoacoustic imaging has shown great potential for guiding minimally invasive procedures by accurate identification of critical tissue targets and invasive medical devices (such as metallic needles). The use of light emitting diodes (LEDs) as the excitation light sources accelerates its clinical translation owing to its high affordability and portability. However, needle visibility in LED-based photoacoustic imaging is compromised primarily due to its low optical fluence. In this work, we propose a deep learning framework based on U-Net to improve the visibility of clinical metallic needles with a LED-based photoacoustic and ultrasound imaging system. To address the complexity of capturing ground truth for real data and the poor realism of purely simulated data, this framework included the generation of semi-synthetic training datasets combining both simulated data to represent features from the needles and *in vivo* measurements for tissue background. Evaluation of the trained neural network was performed with needle insertions into blood-vessel-mimicking phantoms, pork joint tissue *ex vivo* and measurements on human volunteers. This deep learning-based framework substantially improved the needle visibility in photoacoustic imaging *in vivo* compared to conventional reconstruction by suppressing background noise and image artefacts, achieving 5.8 and 4.5 times improvements in terms of signal-to-noise ratio and the modified Hausdorff distance, respectively. Thus, the proposed framework could be helpful for reducing complications during percutaneous needle insertions by accurate identification of clinical needles in photoacoustic imaging.

1. Introduction

Ultrasound (US) imaging is widely used for guiding minimally invasive percutaneous procedures such as peripheral nerve blocks [1], tumour biopsy [2] and fetal blood sampling [3]. During these procedures, a metallic needle is inserted percutaneously into the body towards the target under real-time US guidance. Accurate and efficient identification of the target and the needle are of paramount importance to ensure the efficacy and safety of the procedure. Despite a number of prominent advantages associated with US imaging such as its real-time imaging capability, high affordability and accessibility, it suffers from intrinsically low soft tissue contrast that sometimes results in insufficient visibility of critical tissue structures such as nerves and small blood vessels. Moreover, visibility of clinical needles with US imaging is strongly dependent on the insertion angle and depth of the needle.

With steep insertion angles, US reflections can be readily outside the transducer aperture, leading to poor needle visibility. Loss of visibility of tissue targets or the needle can cause significant complications [4].

Various methods have been developed for enhancing needle visualisation with US imaging, including echogenic needles [5], Doppler imaging [6], electromagnetic tracking [7], and ultrasonic needle tracking [8–10]. Although promising results have been reported, these methods usually require specialised equipment. Image-based needle tracking algorithms leveraging linear features of needles in US images have also been investigated such as random sample consensus [11,12], Hough Transform [13,14], line filtering [15,16], and graph cut [17]. However, it is challenging to automate these algorithms on US images with a variety of tissue backgrounds and needle contrasts. Deep learning (DL) based models especially convolutional neural networks have demonstrated competitive robustness and accuracy [18–20], however, large

* Corresponding author.

E-mail address: wenfeng.xia@kcl.ac.uk (W. Xia).

<https://doi.org/10.1016/j.pacs.2022.100351>

Received 5 January 2022; Received in revised form 29 March 2022; Accepted 30 March 2022

Available online 7 April 2022

2213-5979/© 2022 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

clinical datasets with fine annotations are usually required for clinical applications but difficult to obtain.

Photoacoustic (PA) imaging has been of growing interest in the past two decades for its various potential preclinical and clinical applications, owing to its unique ability to resolve spectroscopic signatures of tissues at high spatial resolution and depths [21–23]. In recent years, several research groups have proposed the combination of US and PA imaging for guiding minimally invasive procedures by offering complementary information to each other, with US imaging providing tissue structural information and PA imaging identifying critical tissue structures and invasive surgical devices such as metallic needles [24–27]. Recently, laser diodes (LDs) and light emitting diodes (LEDs) have shown promising results as an alternative to solid-state lasers that are commonly used as PA excitation sources due to their favourable portability and affordability, which is of advantage to clinical translation [28–30].

DL has been demonstrated as a powerful tool for signal and image processing, leading to remarkable successes in medical imaging [31–34]. DL-based approaches have been proposed for PA imaging enhancement, especially photoacoustic tomography [35], where they process raw channel data for image reconstruction [36–38] and enhancement [39,40] as well as reconstructed images for image segmentation or classification [41,42].

DL has been recently used by several research groups for improving the imaging quality of LED-based PA/US imaging systems. Anas et al. [43] exploited the use of a combination of a convolutional neural network (CNN) and a recurrent neural network (RNN) to enhance the quality of PA images by leveraging both the spatial features and temporal information in repeated PA image acquisitions. Kuniyil Ajith Singh et al. [29] proposed a U-Net model to improve the SNR by training a neural network using PA images acquired by an improved PA imaging system with a higher laser energy and broadband ultrasound transducers. The pre-trained model was proven effective with LED-based PA images acquired from phantoms. Hariri et al. [44] proposed a multi-level wavelet-CNN to enhance noisy PA images associated with low fluence LED illumination by learning from PA images acquired with high fluence illumination sources. Enhancements were achieved on unseen *in vivo* data with improved image contrast. Most recently, Kalloor Joseph et al. [45] developed a generative adversarial network (GAN)-based framework for PA image reconstruction to mitigate the impact of the limited aperture and bandwidth of the ultrasound transducer. The proposed model was trained on simulated images from artificial blood vessels and tested on *in vivo* measurements of the human forearm. The proposed approach was able to remove artefacts caused by the limited bandwidth and detection view.

Although prominent attention has been given to improving the visualisation of tissue structures, notably, not much effort has been made to improve the visualisation of invasive medical devices in PA imaging. In this work, we proposed a DL-based framework to enhance the visibility of clinical needles with PA imaging for guiding minimally invasive procedures. As clinical needles have relatively simple geometries whilst background biological tissues such as blood vessels are complex, as opposed to using purely synthetic data [46–49], a hybrid method was proposed for generating semi-synthetic datasets [50]. The DL model was trained and validated using such semi-synthetic datasets and blind to the test data obtained from tissue-mimicking phantoms, *ex vivo* tissue and human fingers *in vivo*. The applicability of the proposed model on diverse *in vivo* image data was further assessed on PA video sequences and compared with the standard Hough Transform. To the best of our knowledge, this is the first work that exploits DL for improving needle visualisation with PA imaging as well as utilises semi-synthetic datasets for DL in PA imaging.

2. Material and methods

2.1. System description

A commercial LED-based PA/US imaging system (AcousticX, CYBERDYNE INC, Tsukuba, Japan) was used for acquiring experimental data. Detailed description of it can be found elsewhere [51]. Briefly, PA excitation is provided by two LED arrays that sandwich a linear array US probe at a fixed angle. Each array consists of four rows of LEDs with 36 elements of 1 mm × 1 mm on each row. The LED arrays can be driven at different pulse repetition frequencies from 1 kHz to 4 kHz, and the maximum pulse energy from each array is 200 μJ. The LED pulse duration is controllable between 30 ns to 150 ns. In this study, a pulse width of 70 ns at 850 nm was selected for optimal energy efficiency [52,53]. The illumination area formed by the LED arrays was approximately a rectangle (50 mm × 7 mm), resulting in an optical fluence of 0.11 mJ/cm² at the maximum pulse energy of 400 μJ. The US probe has 128 elements over a linear array distance of 40.3 mm, with each element having a pitch of 0.315 mm, a central frequency of 7 MHz, and a –6 dB fractional bandwidth of 80.9%.

Radio-frequency (RF) data for PA and US imaging were collected simultaneously from 128 channels on the probe with sampling rates of 40 MHz and 20 MHz, respectively. Interleaved PA and US imaging can be performed in real-time with image reconstruction performed on a graphics processing unit (GPU). Meanwhile, a maximum of 1536 PA frames and 1536 US frames corresponding to a total duration of 20 s could be saved in memory at one time, available for offline reconstruction.

2.2. Semi-synthetic dataset generation

The process of semi-synthetic dataset generation comprised three main steps as shown in Fig. 1: (1) acquisition of *in vivo* data to account for background PA signals originated from biological tissue; (2) generation of synthetic sensor data from needles; (3) image reconstruction with raw channel data combining synthetic and measurement data.

In vivo data for background vasculature were collected by imaging the fingers of 13 healthy human volunteers using AcousticX. Experiments on human volunteers were approved by the King's College London Research Ethics Committee (study reference: HR-18/19-8881). For each measurement, a total of 1536 PA frames and 1536 US frames were saved to the hard drive of the system's workstation for offline reconstruction. The RF data of one PA or US frame was a 2D matrix with dimensions of 1024 × 128. The first 150 (out of 1024) time steps were zeroed to remove strong LED-induced noise that spanned across the upper 5 mm depth in PA and US images. Averaging over 128 frames was implemented for suppressing random noise in the background. Reconstructed PA and US images corresponded to a field-of-view of 40.3 mm (X) × 39.4 mm (Z) according to the geometry of the linear transducer and the total number of time steps (1024) at a 40 MHz sampling rate.

Simulations of the sensor data originated from the needle were performed using the k-Wave toolbox [54]. Initial pressure distribution maps were created by simulating the optical fluence distributions on the needle shaft using Monte Carlo simulations [55]. A 40.0 mm (X) × 40.0 mm (Z) region with a grid size of 0.1 mm was constructed to represent the background tissue. A uniform refractive index, optical scattering coefficient and anisotropy of 1.4, 10 mm⁻¹ and 0.9, respectively, were assigned to this region [56]. Three optical absorption coefficients of 1 mm⁻¹, 1.5 mm⁻¹, 2 mm⁻¹ accounted for the variations in standard tissue. A homogeneous photo beam with a finite size of 38.4 mm was applied to the surface of the simulation area. Each simulation was run for around 10 min with approximate 100,000 photon packets.

A linear array of 128 US transducer elements (with a pitch of 0.315 mm over a total length of 40.3 mm) were assigned to a forward model in k-Wave to receive the generated PA signals from the

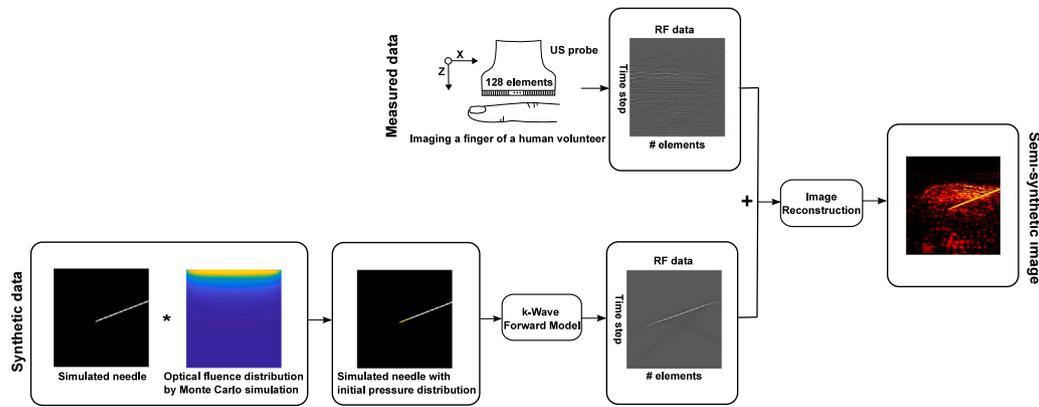


Fig. 1. Flowchart illustration of the process of semi-synthetic training dataset generation. Top row: acquisition of sensor data from human finger vasculature *in vivo* as background. Bottom row: synthetic radio-frequency (RF) sensor data generation from a simulated needle.

initial pressure distributions maps. The US transducer was assigned a central frequency of 7 MHz and a fractional -6 dB bandwidth of 80.9% according to the specifications of AcousticX. RF data collected by the transducer elements were successively down-sampled to 40 MHz to match the sampling rate of the measured data. Considering the variations of needle insertions, simulations were conducted to account for clinically-relevant needle insertion depths and angles, spanning from 5 mm to 25 mm with an increment of 5 mm, and from 20 degrees to 65 degrees with a step of 5 degrees, respectively.

To form a semi-synthetic image, the simulated RF data were normalised to maximum amplitude of *ex vivo* needle signals collected by AcousticX. Subsequently, a pair of 2D data matrices (1024×128) consisted of RF data from a simulation on a needle and a measurement on a human finger were added to form a single 2D data matrix and then fed to a Fourier domain algorithm for image reconstruction [57]. The reconstructed images based on the semi-synthetic data were then interpolated to 578×565 pixels with a uniform pixel size of $70 \mu\text{m} \times 70 \mu\text{m}$. To facilitate network implementations, the images were cropped to 512×512 pixels by removing the corresponding rows from top to bottom and the same number of columns from left to centre and right to centre respectively.

Finally, a total number of 2000 semi-synthetic images with substantial variations on both the needle and background were used for model training with the corresponding initial pressure distributions as the ground truths.

2.3. Acquisition of phantom, *ex vivo* and *in vivo* data for evaluation

Evaluation of the trained neural network was performed on PA images acquired with in-plane needle (20G, BD, USA) insertions into blood-vessel-mimicking phantoms, pork joint tissue *ex vivo* and human fingers *in vivo* (needle outside of tissue; see Supplementary Materials; Video S1). It is noted that the fingers with needle insertions were used to obtain representative real *in vivo* data, but there is no corresponding clinical scenario.

The blood-vessel-mimicking phantoms were created by affixing several carbon fibre bundles in a plastic box filled with 1% Intralipid dilution (Intralipid 20% emulsion, Scientific Laboratory Supplies, UK) that had an estimated optical reduced scattering coefficient of 0.96 mm^{-1} at 850 nm [58]. The fibre bundles were randomly positioned to mimic different orientations of blood vessels. The acquired PA images were prepared following the pipeline used for processing the semi-synthetic data.

2.4. Network implementation

The network architecture implemented in this work was derived from the U-Net architecture proposed in Ref. [35]. In general, this

model followed the original U-Net architecture [59] but had fewer scales and a reduced number of filters at each scale to accommodate a small input size. Experiments about the model capacity (see Supplementary Materials; Section 5) manifested that the model shown in Fig. 2 was able to learn the regularities of the training data and generalise well to unseen data. Besides, this model was built smaller and lighter which could contribute to size and latency reduction that are beneficial for real-time applications.

In Fig. 2, following an encoder path, each scale consisted of two convolutional layers followed by a 2×2 max pooling layer. For the decoder path, similarly, each scale contained two convolutional layers but followed by a transposed convolutional layer with an up-sampling factor of 2. The model was trained using the input pairs with a smaller resolution of 128×128 pixels that adapted well to the receptive field of the model by resizing from the initial size of 512×512 pixels via bicubic interpolation and evaluated on the real images with a size of 256×256 (see 4. Discussion & Conclusions regarding the choice of the image resolution).

Our network was implemented in Python using PyTorch v1.2.0. The semi-synthetic dataset was randomly split into training, test, and validation sets with a ratio of 8:1:1. Training was performed for 5000 iterations with a batch size of 4 that minimised the mean square error (MSE) loss in the validation set using the ADAM optimiser [60] (initial learning rate: 0.001) and NVIDIA Tesla V100 GPUs. The CosineAnnealingLR learning rate scheme [61] was employed to steadily decrease the learning rate during the training.

2.5. Post-processing

A post-processing algorithm based on maximum contour selection [62] was employed for further improving the outcomes of the trained neural network and fitting the needle trajectory. It was assumed that for all the experiments only in-plane placements with one single needle was performed. The isolated outliers in the outputs of the U-Net could be discriminated based on the region size differences from that of the enhanced needle. Thus, the post-processing algorithm detected all contours in the outputs of the proposed model and saved the maximum contour as the one from the needle by counting the number of pixels on each contour boundary.

2.6. Comparison method

As a further evaluation step considering processing multiple PA frames from video sequences during dynamic needle insertions, the performance of the trained neural network on needle identification was compared with the standard Hough Transform (SHT), which is a classical baseline method for line detection [63]. The SHT is designed

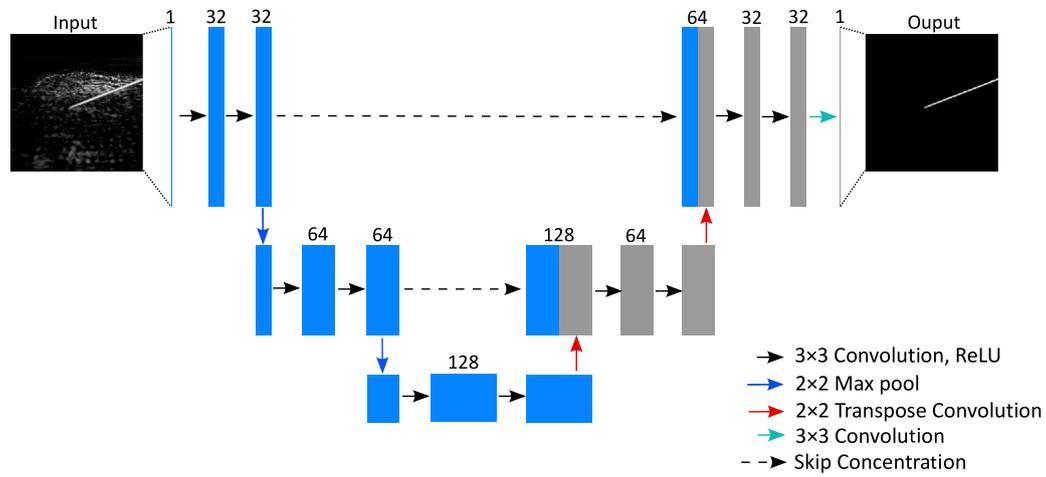


Fig. 2. Architecture of the proposed network for improving needle visibility in photoacoustic imaging.

to identify straight lines in images. It employs the parametric representation of a straight line, which is also called the Hesse normal form and can be expressed as:

$$r = x \sin \theta + y \cos \theta \quad (1)$$

where r is the shortest distance from the origin to the line. θ measures the angle between the x -axis and the perpendicular projection from the origin point to the line. Therefore, a straight line can be associated with a pair of parameters (r, θ) , corresponding to a sinusoidal curve in Hough space. A few points on the same straight line will produce a set of sinusoidal curves that cross the same point (r, θ) which exactly represents that line. In this study, the SHT was implemented by a two-dimensional matrix whose columns and rows were used to save the r and θ values, respectively. For each point in the image, r was calculated for each θ , leading to increments of that bin in the matrix. Finally, the potential straight lines in the image were extracted by selecting the local maxima from the accumulator matrix.

2.7. Evaluation protocol and performance metrics

The needles in the acquired PA images from three different media (phantoms, pork joint tissue *ex vivo*, and human fingers) were manually labelled as line segments by an experienced observer. The line needle segment was generated by connecting two points in the needle: the needle tip and the farthest point to the tip on the needle shaft that was visualised in a PA image. The needle tip had a good contrast on US images when it was surrounded by water in the liquid-based tissue-mimicking phantoms, but had poor visibility for *ex vivo* and *in vivo* measurements. Therefore, solid glass spheres (0–63 μm , Boud Minerals Limited, UK) were injected through the needle after being diluted with water to enhance the contrast of the tip, thus improving the accuracy and precision of manual labelling. For each medium, 20 representative measurements with different backgrounds and needle locations were used for metrics calculation.

To access the accuracy of needle extraction using the proposed DL-based approach, a metric called the modified Hausdorff distance (MHD) was employed [64]. The MHD was adapted from the generalised Hausdorff distances proposed for object matching with improved discriminatory power and greater robustness to outliers. Considering two point sets $\mathcal{A} = \{a_1, \dots, a_{N_a}\}$ and $\mathcal{B} = \{b_1, \dots, b_{N_b}\}$. The distance between a point a from \mathcal{A} and a set of points from \mathcal{B} was defined as $d(a, \mathcal{B}) = \min_{b \in \mathcal{B}} \|a - b\|$. The directed distance measure $d(\mathcal{A}, \mathcal{B})$ was defined as:

$$d(\mathcal{A}, \mathcal{B}) = \frac{1}{N_a} \sum_{a \in \mathcal{A}} d(a, \mathcal{B}) \quad (2)$$

Then, the directed distance measures $d(\mathcal{A}, \mathcal{B})$ and $d(\mathcal{B}, \mathcal{A})$ were combined in the following way, resulting in the definition of the MHD as:

$$\text{MHD} = \max(d(\mathcal{A}, \mathcal{B}), d(\mathcal{B}, \mathcal{A})) \quad (3)$$

The MHD was defined in the unit of pixels which could be converted to real distance considering the pixel size of 70 μm per pixel. Signal-to-noise ratio (SNR) was also used to assess the performance of the needle enhancement and was defined as $\text{SNR} = S / \sigma$, where S is the mean amplitude over the needle region and σ is the standard deviation of the background. The mean amplitude of the needle region was calculated by taking the average of the pixel values over the line segment. The background was defined as one of the largest rectangular regions that excluded the needle pixels.

3. Results

3.1. Blood-vessel-mimicking phantoms

The results of imaging on blood-vessel-mimicking phantoms are shown in Fig. 3. Compared to conventional reconstructions, noticeable improvements in terms of removing background noise and artefacts can be observed in the outputs of the U-Net. The proposed model successfully identified the needle insertion without being perturbed by the background vessels that shared similar line-shape features. The false positives in the U-Net enhanced images were further suppressed by the post-processing. Besides, the proposed model demonstrated robustness to noise and strong artefacts (e.g., Fig. 3(d)). The composite images indicated that the proposed model was able to detect the needle insertion with good correspondence to the conventional reconstructions and the US images.

The performance of the proposed model was quantified with SNR and MHD (Table 1). Compared to the conventional reconstruction, the proposed model achieved a significant improvement in SNR by a factor of 8.3 ($p < .0001$). The MHD had large values as the noise level and artefacts increased. The proposed U-Net led to an initial 2.4 times decrease in MHD (from 63.2 ± 15.9 to 26.4 ± 23.3) and was further optimised by the successive post-processing to achieve the smallest MHD of 1.4 ± 1.3 .

3.2. Pork joint tissue *ex vivo*

The proposed model was also applied to images acquired from needle insertions into *ex vivo* tissue. Fig. 4 shows the representative results comparing the conventional reconstruction and the proposed model. The proposed model robustly enhanced the needle visibility

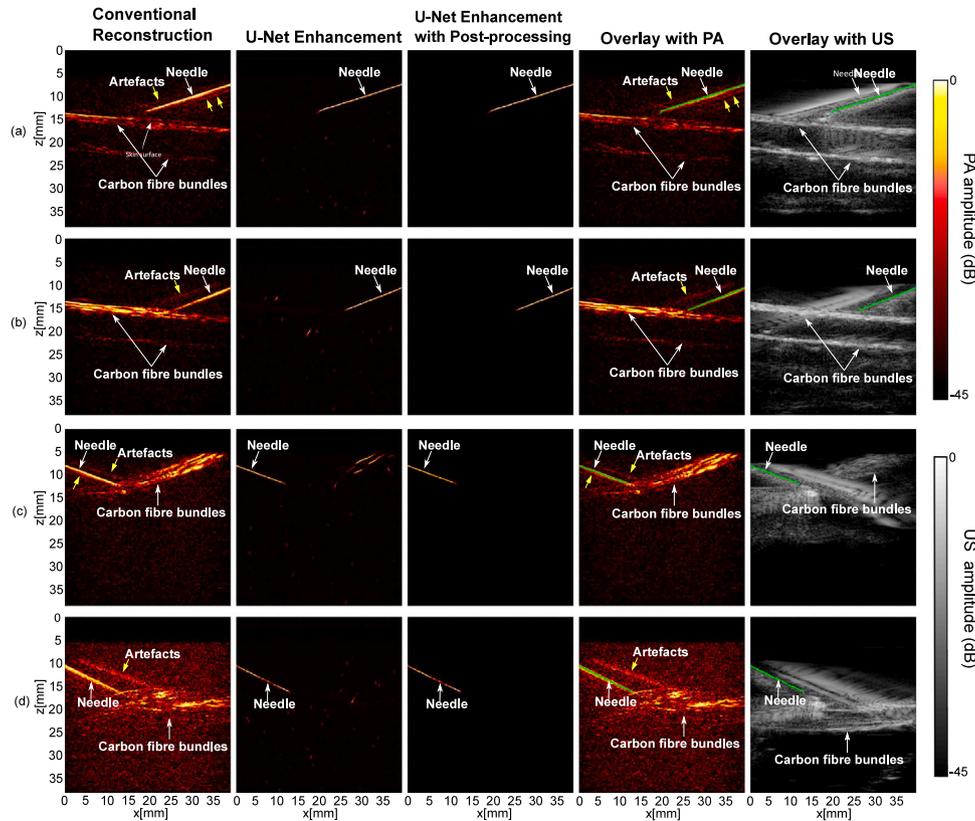


Fig. 3. Photoacoustic imaging with needle insertions into a blood-vessel-mimicking phantom with conventional reconstruction, U-Net enhancement and U-Net enhancement with post-processing.

Table 1

Quantitative evaluation of the trained neural network using blood-vessel-mimicking phantoms. These performance metrics are expressed as mean \pm standard deviations from 20 measurements acquired from different phantoms and needle positions.

| Metrics | Conventional reconstruction | U-Net enhancement | U-Net enhancement with post-processing |
|---------|-----------------------------|-------------------|--|
| SNR | 8.7 \pm 2.3 | 72.1 \pm 40.1 | – |
| MHD | 63.2 \pm 15.9 | 26.4 \pm 23.3 | 1.4 \pm 1.3 |

Table 2

Quantitative evaluation of the trained neural network using ex vivo needle images. These performance metrics are expressed as mean \pm standard deviations from 20 measurements acquired from different spatial locations of the ex vivo tissue and needle positions.

| Metrics | Conventional reconstruction | U-Net enhancement | U-Net enhancement with post-processing |
|---------|-----------------------------|-------------------|--|
| SNR | 19.0 \pm 9.7 | 91.3 \pm 47.3 | – |
| MHD | 28.7 \pm 16.3 | 6.3 \pm 9.1 | 4.5 \pm 7.0 |

with different insertion depths and angles, and significantly suppressed image artefacts and background noise. It also achieved a 4.8 times improvement in SNR compared to the conventional reconstruction ($p < .00001$) (Table 2). The MHD substantially decreased from 28.7 ± 16.3 to 4.5 ± 7.0 after post-processing the output of the U-Net ($p < .00001$).

It is worth noting that the performance of the proposed model was slightly degraded with visualising the region near the needle tip, which could be attributed to the large depths at around 2.5 cm (Fig. 4(b) and (d)). However, within smaller imaging depths, the proposed model was still effective for increasing the imaging speed by reducing the number of averages required to visualise the needle with a high SNR (see Supplementary Materials; Figure S4).

3.3. In vivo imaging

Fig. 5 shows the results of PA imaging with a 20G needle inserted between two fingers of a human volunteer immersed in a water tank (needle outside of the fingers). A main digital artery which has a two-layered feature is apparent on PA images (marked by hollow triangle wide arrows). A 22 s video consisting of 128 frames was saved during the needle insertions (see Supplementary Materials; Video S2). Fig. 5 shows four frames acquired at different time points with the conventional reconstructions and the overlays of PA and US images after the U-Net enhancement and SHT. The SHT was able to detect the needle, but resulted in excessive false positives. This is because the performance of the SHT is sensitive to the specifications of hyperparameters, such as the detectable length of line segments and the searching resolution of r and θ . Fine-tuning of the hyperparameters based on the observations of the needle on a frame by frame basis is not trivial. In comparison, the proposed model manifested a good ability of robustness and generalisation. The *in vivo* results demonstrated it is insensitive to constantly changing lengths and angles of the needle, and images with excessive noise and artefacts.

Quantitative results are summarised in Table 3. An average of 5.8 times improvement in SNR was observed in the U-Net enhancement versus conventional reconstruction at different time points during the insertion ($p < .00001$). For MHD, the proposed model outperformed the conventional reconstruction and the SHT with a 4.5- and 2.9-fold reduction, respectively ($p < .00001$ for the conventional reconstruction; $p < .0001$ for the SHT). The post-processing algorithm effectively suppressed the outliers in the output of the U-Net, leading to the MHD as small as 0.6.

To further evaluate the model performance, needle detection rate was measured with three *in vivo* PA sequences (Table 4; Supplementary Materials; Video S3). For each sequence containing 128 frames, the number of frames with the identifiable needle was counted by the

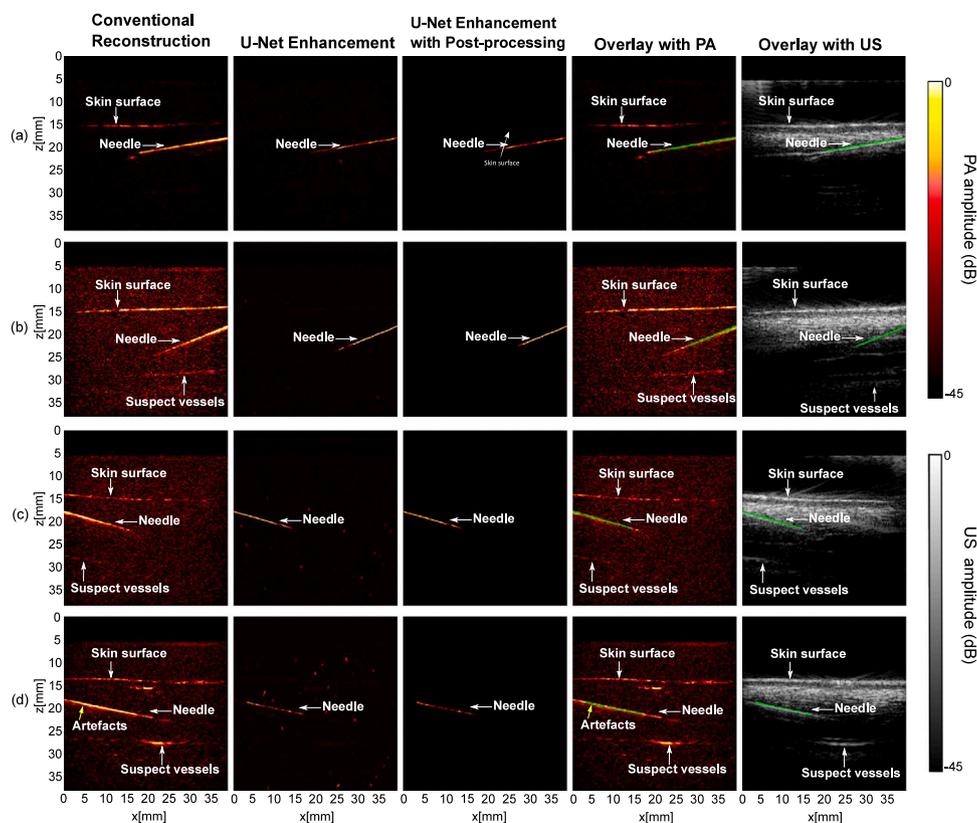


Fig. 4. Photoacoustic imaging with needle insertions into *ex vivo* tissue with conventional reconstruction, U-Net enhancement, and U-Net enhancement with post-processing.

Table 3

Quantitative evaluation of the trained neural network using *in vivo* needle images. These performance metrics are expressed as mean \pm standard deviations from 20 measurements acquired at different time points during the insertion.

| Metrics | CR ^a | U-Net Enhancement | U-Net Enhancement with Post-processing | SHT ^b |
|---------|-----------------|-------------------|--|------------------|
| SNR | 8.8 \pm 1.5 | 51.0 \pm 11.8 | – | – |
| MHD | 87.7 \pm 24.4 | 19.4 \pm 5.0 | 0.6 \pm 0.1 | 55.7 \pm 38.5 |

^aConventional reconstruction.

^bStandard hough transform.

Table 4

Quantitative performance on three PA video sequences *in vivo* with the proposed model.

| | Sequence 1 | Sequence 2 | Sequence 3 |
|------------------------------------|------------|------------|------------|
| Frames with needle/Total frames | 92/128 | 53/128 | 99/128 |
| Frames without needle/Total frames | 36/128 | 75/128 | 29/128 |
| Needle missed | 0 | 5 | 3 |
| True positives | 92 | 48 | 96 |
| False positives | 0 | 1 | 1 |
| True positive rate (%) | 100 | 90.6 | 97.0 |
| False positive rate (%) | 0 | 1.3 | 3.4 |

observer as the reference for calculation. It is worth noting that the proposed model enhanced almost all the frames that contained the needle with a true positive rate up to 100%, 90.6%, and 97.0%, respectively.

3.4. Impact of the needle diameter

To further assess the generalisability of the proposed model, in-plane insertions of needles with different diameters (16G, 18G, 20G, BD, USA; 25G, 30G, Meso-relle, Italy) were imaged with AcousticX during in plane insertions into pork joint tissue *ex vivo*. The results (see

Supplementary Materials; Figure S5) were consistent with the previous results of the 20G needle. For the needles with small diameters, the acquired PA images readily suffered from lower SNRs. However, as expected, the proposed model yielded robust enhancement on the needles at different contrast levels. For SNR and MHD, the proposed model outperformed the conventional reconstruction with an average of 11.2 and 6.5 times improvement, respectively (see Supplementary Materials; Table S1).

4. Discussion

Previous works on DL in PA imaging mainly focused on improving the visualisation of the vasculature by denoising and artefacts removal. However, the performances of DL networks are highly dependent on the training dataset. Networks that are specifically trained to enhance the visualisation of vasculature usually have poor performance on visualising needles due to their different image features. In this work, we are the first to apply DL to specifically improve the needle visualisation with PA imaging for minimally invasive guidance. Considering the relatively simple geometries of clinical needles compared to vasculature, a prominent contribution of this work is that we developed a semi-synthetic approach to address the challenges associated with obtaining ground truth for *in vivo* data as well as the poor realism of purely simulated data.

According to our experimental results (not shown for the sake of brevity), the simulated optical fluence distribution had a minimal effect on the performance of the proposed model when evaluated on unseen real needle images even with deep insertion angles. This is because the DL-based method was able to enhance the visualisation of the needle by learning its relatively simple spatial features that remained largely consistent. In addition, we found that the inference performance of the trained model on the real images did not benefit from a higher resolution input data than that is currently used (128 \times 128 pixels; See Supplementary Materials; Figure S7). The low resolution images

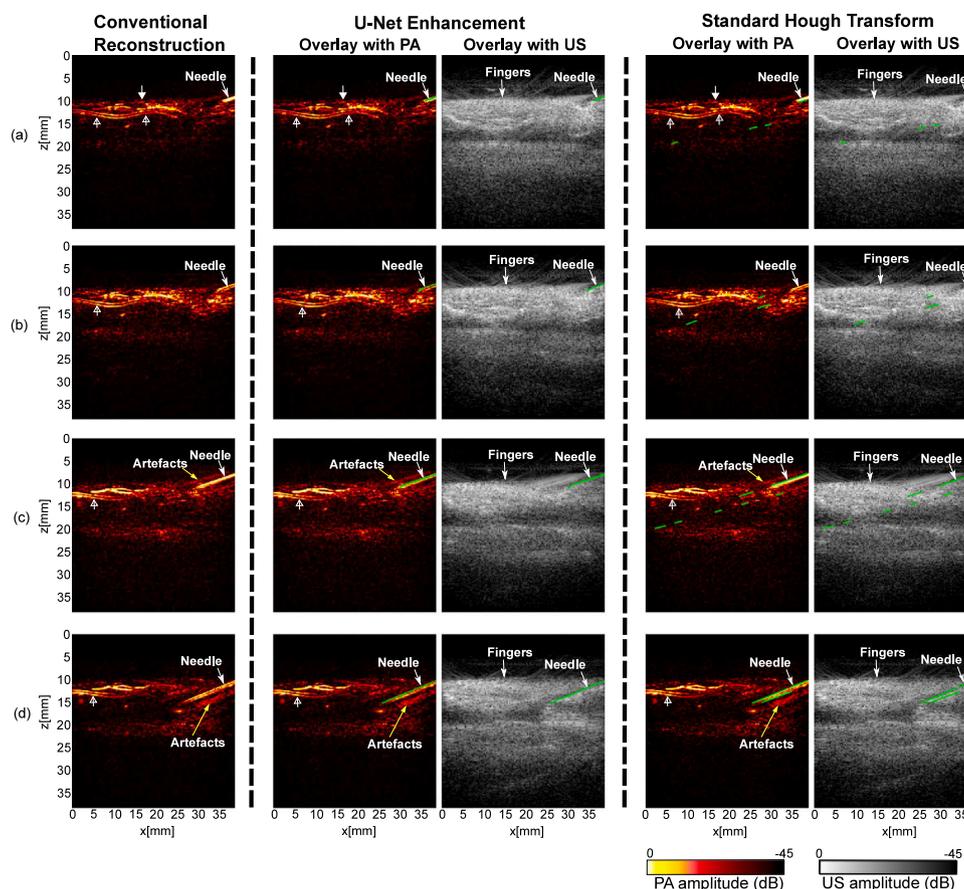


Fig. 5. Photoacoustic (PA) imaging with needle insertions into human fingers *in vivo* with conventional reconstruction, U-Net enhancement, and standard Hough Transform. Signals from the skin surface are indicated by triangle wide arrows, and signals that may be from digital arteries are indicated by hollow triangle wide arrows. The outcomes of U-Net enhancement and standard Hough Transform are denoted by green lines in PA and ultrasound (US) overlays. (a) - (d) are from a reconstructed PA image sequence recorded during needle insertions in real time.

performed sufficiently well considering the lightweight model and simple features of the input data. Compared to a high resolution input, a low resolution input is also advantageous in terms of the computational costs; the inferring time for one image was around 90 ms using one GPU (NVIDIA GeForce RTX 2070 Max-Q). Further reduction on the inferring time could be realised by using more powerful GPUs for real-time applications.

During minimally invasive procedures, accurate and clear visualisation of the needle is essential for successful outcomes. Needle visibility has been greatly improved by PA imaging as compared to US imaging, but the image quality in terms of SNR with the LED light source is still sub-optimum due to the low pulse energy. Frame averaging is effective for reducing background noise, but at the cost of the imaging speed and introduces movement artefacts. Further, blood vessels in the background with similar line-shape structures to the needle are readily regarded as visual disturbances for clinicians to identify the needle trajectory. Finally, line artefacts above or beneath the needle shaft are often non-negligible that can lead to misinterpretation of the needle position. Therefore, in this work, PA images of needle insertions into different types of blood-vessel-mimicking phantoms, *ex vivo* tissue, and *in vivo* human fingers were acquired to evaluate the proposed model.

Qualitative results demonstrated that our proposed model was able to achieve substantial enhancement on the needle visualisation regarding noise suppression, artefacts removal, and needle detection. The enhancement was further quantified by the SNR and MHD. Performance of the proposed model was compared to the conventional reconstruction and the standard Hough transform on images acquired from blood-vessel-mimicking phantoms, *ex vivo* pork joint tissue, and

human fingers as shown in Supplementary Materials (Figure S1; Figure S2; Figure S3). For SNR, our proposed model achieved 8.3, 4.8, and 5.8 times enhancement for phantom, *ex vivo*, and *in vivo* data respectively. The MHD as a measure of similarity of two objects was employed for its great robustness and discriminatory power. It was observed that the MHD had the smallest values with our proposed model compared to the conventional reconstruction and the SHT (1.4, 4.5, and 0.6 for phantom, *ex vivo*, and *in vivo* data respectively). Additionally, it is evidenced that the post-processing method based on maximum contour selection was effective to remove the false positives of the U-Net enhancement while preserving the needle pixels.

We also compared the proposed model with a conventional line detection algorithm, SHT, with *in vivo* video sequences. The SHT performed quite well on some cases with carefully chosen critical hyperparameters, but its performance was readily affected by imperfection errors from the former edge detection step and sensitive to some decision criteria such as empirical values of r and θ that are directly related to the detection efficiency. Fine-tuning of these hyperparameters is impractical for real-time applications where the effective length and angle of the needle placements could constantly vary in each frame. In contrast, our proposed model can efficiently improve the needle visualisation on a variety of PA images from *in vivo* measurements in near real-time.

Nonetheless, the DL-based enhancement was sensitive to the SNRs of the images. More importantly, the visibility of the needle, especially its tip was still limited to a depth of around 1 cm with *in vivo* measurements. In the future, deep neural networks could be applied for real-time denoising [44] as an alternative to frame-to-frame averaging

to improve the imaging depth. For needle tip visualisation, a fibre-optic US transmitter could be integrated within the needle cannula so that the needle tip can be unambiguously visualised in PA imaging with high SNRs [10]. Light-absorbing coatings based on elastomeric nanocomposites could also be applied to the needle shaft for enhancing its visualisation for guiding minimally invasive procedures [65].

5. Conclusions

In this work, we provided a DL-based framework for enhancing needle visualisation with PA imaging. The DL-model was built using only semi-synthetic data generated by combining simulated data and *in vivo* measurements. Evaluation was performed on unseen real data acquired by inserting needles into blood-vessels-mimicking phantoms, *ex vivo* tissue and human fingers (needle outside tissue). Compared to the conventional reconstruction, the proposed framework substantially improved the needle visualisation with PA imaging. It also outperformed the standard Hough Transform on PA *in vivo* videos with improved robustness and generalisability. Therefore, our framework could be useful for guiding minimally invasive procedures that involve percutaneous needle insertions by accurate identification of clinical needles.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to <https://doi.org/10.1016/j.pacs.2022.100351>. Adrien E. Desjardins reports a relationship with Echopoint Medical that includes: equity or stocks. Tom Vercauteren reports relationships with Hypervision Surgical Ltd and Mauna Kea Technologies that include: equity or stocks.

Acknowledgements

This work was funded in whole, or in part, by the Wellcome Trust, United Kingdom [203148/Z/16/Z, WT101957, 203145Z/16/Z], the Engineering and Physical Sciences Research Council (EPSRC), United Kingdom (NS/A000027/1, NS/A000050/1, NS/A000049/1), and King's–China Scholarship Council PhD Scholarship Program (K-CSC) (202008060071). For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.pacs.2022.100351>.

References

- [1] K.J. Chin, A. Perlas, V.W.S. Chan, R. Brull, Needle visualization in ultrasound-guided regional anesthesia: Challenges and solutions, *Reg. Anesth. Pain Med.* 33 (6) (2008) 532–544, <http://dx.doi.org/10.1136/rapm-00115550-200811000-00005>.
- [2] T.H. Helbich, W. Matzek, M.H. Fuchsjäger, Stereotactic and ultrasound-guided breast biopsy, *Eur. Radiol.* 14 (3) (2004) 383–393, <http://dx.doi.org/10.1007/s00330-003-2141-z>.
- [3] F. Daffos, M. Capella-Pavlovsky, F. Forestier, Fetal blood sampling during pregnancy with use of a needle guided by ultrasound: A study of 606 consecutive cases, *Am. J. Obstet. Gynecol.* 153 (6) (1985) 655–660, [http://dx.doi.org/10.1016/s0002-9378\(85\)80254-4](http://dx.doi.org/10.1016/s0002-9378(85)80254-4).
- [4] J.P. Rathmell, H.T. Benzon, P. Dreyfuss, M. Huntoon, M. Wallace, R. Baker, K.D. Riew, R.W. Rosenquist, C. Aprill, N.S. Rost, A. Buvanendran, D.S. Kreiner, N. Bogduk, D.R. Fourney, E. Fraifeld, S. Horn, J. Stone, K. Vorenkamp, G. Lawler, J. Summers, D. Kloth, D. O'Brien, S. Tutton, Safeguards to prevent neurologic complications after epidural steroid injections: Consensus opinions from a multidisciplinary working group and national organizations, *Anesthesiology* 122 (5) (2015) 974–984, <http://dx.doi.org/10.1097/ALN.0000000000000614>.
- [5] C.H. Hovgesen, J.E. Wilhjelm, P. Vilmann, E. Kalaitzakis, Echogenic surface enhancements for improving needle visualization in ultrasound: A PRISMA systematic review, *J. Ultrasound Med.* 41 (2) (2022) 311–325, <http://dx.doi.org/10.1002/jum.15713>.
- [6] M.P. Fronheiser, S.F. Idriss, P.D. Wolf, S.W. Smith, Vibrating interventional device detection using real-time 3-D color Doppler, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 55 (6) (2008) 1355–1362, <http://dx.doi.org/10.1109/TUFFC.2008.798>.
- [7] S.M. Klein, M.P. Fronheiser, J. Reach, K.C. Nielsen, S.W. Smith, Piezoelectric vibrating needle and catheter for enhancing ultrasound-guided peripheral nerve blocks, *Anesth. Analg.* 105 (6) (2007) 1858–1860, <http://dx.doi.org/10.1213/01.ane.0000286814.79988.0a>.
- [8] W. Xia, J.M. Mari, S.J. West, Y. Ginsberg, A.L. David, S. Ourselin, A.E. Desjardins, In-plane ultrasonic needle tracking using a fiber-optic hydrophone, *Med. Phys.* 42 (10) (2015) 5983–5991, <http://dx.doi.org/10.1118/1.4931418>.
- [9] W. Xia, S.J. West, M.C. Finlay, J.-M. Mari, S. Ourselin, A.L. David, A.E. Desjardins, Looking beyond the imaging plane: 3D needle tracking with a linear array ultrasound probe, *Sci. Rep.* 7 (1) (2017) 3674, <http://dx.doi.org/10.1038/s41598-017-03886-4>.
- [10] W. Xia, S. Noimark, S. Ourselin, S.J. West, M.C. Finlay, A.L. David, A.E. Desjardins, Ultrasonic needle tracking with a fibre-optic ultrasound transmitter for guidance of minimally invasive fetal surgery, in: M. Descoteaux, L. Maier-Hein, P. Jannin, D.L. Collins, S. Duchesne (Eds.), *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2017*, Vol. 10434, Springer International Publishing, Cham, 2017, pp. 637–645, http://dx.doi.org/10.1007/978-3-319-66185-8_72.
- [11] M. Uherčík, J. Kybic, H. Liebgott, C. Cachard, Model fitting using RANSAC for surgical tool localization in 3-D ultrasound images, *IEEE Trans. Biomed. Eng.* 57 (8) (2010) 1907–1916, <http://dx.doi.org/10.1109/TBME.2010.2046416>.
- [12] M. Waïne, C. Rossa, R. Sloboda, N. Usmani, M. Tavakoli, 3D shape visualization of curved needles in tissue from 2D ultrasound images using RANSAC, in: 2015 IEEE International Conference on Robotics and Automation (ICRA), 2015, pp. 4723–4728, <http://dx.doi.org/10.1109/ICRA.2015.7139855>.
- [13] M. Ding, A. Fenster, A real-time biopsy needle segmentation technique using hough transform, *Med. Phys.* 30 (8) (2003) 2222–2233, <http://dx.doi.org/10.1118/1.1591192>.
- [14] S.H. Okazawa, R. Ebrahimi, J. Chuang, R.N. Rohling, S.E. Salcudean, Methods for segmenting curved needles in ultrasound images, *Med. Image Anal.* 10 (3) (2006) 330–342, <http://dx.doi.org/10.1016/j.media.2006.01.002>.
- [15] M. Kaya, O. Bebek, Needle localization using gabor filtering in 2D ultrasound images, in: 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 4881–4886, <http://dx.doi.org/10.1109/ICRA.2014.6907574>.
- [16] M. Uherčík, J. Kybic, Y. Zhao, C. Cachard, H. Liebgott, Line filtering for surgical tool localization in 3D ultrasound images, *Comput. Biol. Med.* 43 (12) (2013) 2036–2045, <http://dx.doi.org/10.1016/j.compbiomed.2013.09.020>.
- [17] A. Ayvaci, P. Yan, S. Xu, S. Soatto, J. Kruecker, Biopsy needle detection in transectal ultrasound, *Comput. Med. Imaging Graph.* 35 (7–8) (2011) 653–659, <http://dx.doi.org/10.1016/j.compmedimag.2011.03.005>.
- [18] A. Pourtaherian, F. Ghazvinian Zanjani, S. Zinger, N. Mihajlovic, G.C. Ng, H.H.M. Korsten, P.H.N. de With, Robust and semantic needle detection in 3D ultrasound using orthogonal-plane convolutional neural networks, *Int. J. Comput. Assist. Radiol. Surg.* 13 (9) (2018) 1321–1333, <http://dx.doi.org/10.1007/s11548-018-1798-3>.
- [19] M. Arif, A. Moelker, T. van Walsum, Automatic needle detection and real-time bi-planar needle visualization during 3D ultrasound scanning of the liver, *Med. Image Anal.* 53 (2019) 104–110, <http://dx.doi.org/10.1016/j.media.2019.02.002>.
- [20] D.J. Gillies, J.R. Rodgers, I. Gyacskov, P. Roy, N. Kakani, D.W. Cool, A. Fenster, Deep learning segmentation of general interventional tools in two-dimensional ultrasound images, *Med. Phys.* 47 (10) (2020) 4956–4970, <http://dx.doi.org/10.1002/mp.14427>.
- [21] L.V. Wang, J. Yao, A practical guide to photoacoustic tomography in the life sciences, *Nature Methods* 13 (8) (2016) 627–638, <http://dx.doi.org/10.1038/nmeth.3925>.
- [22] V. Ntziachristos, Going deeper than microscopy: The optical imaging frontier in biology, *Nature Methods* 7 (8) (2010) 603–614, <http://dx.doi.org/10.1038/nmeth.1483>.
- [23] P. Beard, Biomedical photoacoustic imaging, *Interface Focus* 1 (4) (2011) 602–631, <http://dx.doi.org/10.1098/rsfs.2011.0028>.
- [24] M. Kuniyil Ajith Singh, W. Steenbergen, S. Manohar, Handheld probe-based dual mode ultrasound/photoacoustics for biomedical imaging, in: M. Olivo, U.S. Dinis (Eds.), *Frontiers in Biophotonics for Translational Medicine: In the Celebration of Year of Light (2015)*, in: *Progress in Optical Science and Photonics*, Springer, Singapore, 2016, pp. 209–247, http://dx.doi.org/10.1007/978-981-287-627-0_7.
- [25] S. Park, J. Jang, J. Kim, Y.S. Kim, C. Kim, Real-time triple-modal photoacoustic, ultrasound, and magnetic resonance fusion imaging of humans, *IEEE Trans. Med. Imaging* 36 (9) (2017) 1912–1921, <http://dx.doi.org/10.1109/TMI.2017.2696038>.

- [26] W. Xia, D.I. Nikitichev, J.M. Mari, S.J. West, R. Pratt, A.L. David, S. Ourselin, P.C. Beard, A.E. Desjardins, Performance characteristics of an interventional multispectral photoacoustic imaging system for guiding minimally invasive procedures, *J. Biomed. Opt.* 20 (8) (2015) 086005, <http://dx.doi.org/10.1117/1.JBO.20.8.086005>.
- [27] T. Zhao, A.E. Desjardins, S. Ourselin, T. Vercauteren, W. Xia, Minimally invasive photoacoustic imaging: Current status and future perspectives, *Photoacoustics* 16 (2019) 100146, <http://dx.doi.org/10.1016/j.pacs.2019.100146>.
- [28] W. Xia, M. Kuniyil Ajith Singh, E. Maneas, N. Sato, Y. Shigeta, T. Agano, S. Ourselin, S.J. West, A.E. Desjardins, Handheld real-time LED-based photoacoustic and ultrasound imaging system for accurate visualization of clinical metal needles and superficial vasculature to guide minimally invasive procedures, *Sensors* 18 (5) (2018) 1394, <http://dx.doi.org/10.3390/s18051394>.
- [29] M. Kuniyil Ajith Singh, W. Xia, Portable and affordable light source-based photoacoustic tomography, *Sensors* 20 (21) (2020) 6173, <http://dx.doi.org/10.3390/s20216173>.
- [30] J. Joseph, M.K. Ajith Singh, N. Sato, S.E. Bohndiek, Technical validation studies of a dual-wavelength LED-based photoacoustic and ultrasound imaging system, *Photoacoustics* 22 (2021) 100267, <http://dx.doi.org/10.1016/j.pacs.2021.100267>.
- [31] A.S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on MRI, *Z. Med. Phys.* 29 (2) (2019) 102–127, <http://dx.doi.org/10.1016/j.zemedi.2018.11.002>.
- [32] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S.X. Li, D. Ni, T. Wang, Deep learning in medical ultrasound analysis: A review, *Engineering* 5 (2) (2019) 261–275, <http://dx.doi.org/10.1016/j.eng.2018.11.020>.
- [33] I. Domingues, G. Pereira, P. Martins, H. Duarte, J. Santos, P.H. Abreu, Using deep learning techniques in medical imaging: A systematic review of applications on CT and PET, *Artif. Intell. Rev.* 53 (6) (2020) 4093–4160, <http://dx.doi.org/10.1007/s10462-019-09788-3>.
- [34] J. Gröhl, M. Schellenberg, K. Dreher, L. Maier-Hein, Deep learning for biomedical photoacoustic imaging: A review, *Photoacoustics* 22 (2021) 100241, <http://dx.doi.org/10.1016/j.pacs.2021.100241>.
- [35] A. Hauptmann, B. Cox, Deep learning in photoacoustic tomography: Current approaches and future directions, *J. Biomed. Opt.* 25 (11) (2020) <http://dx.doi.org/10.1117/1.JBO.25.11.112903>.
- [36] S. Antholzer, M. Haltmeier, R. Nuster, J. Schwab, Photoacoustic image reconstruction via deep learning, in: *Photons Plus Ultrasound: Imaging and Sensing 2018*, 10494, SPIE, 2018, pp. 433–442, <http://dx.doi.org/10.1117/12.2290676>.
- [37] N. Davoudi, X.L. Deán-Ben, D. Razansky, Deep learning optoacoustic tomography with sparse data, *Nat. Mach. Intell.* 1 (10) (2019) 453–460, <http://dx.doi.org/10.1038/s42256-019-0095-3>.
- [38] H. Lan, D. Jiang, F. Gao, F. Gao, Deep learning enabled real-time photoacoustic tomography system via single data acquisition channel, *Photoacoustics* 22 (2021) 100270, <http://dx.doi.org/10.1016/j.pacs.2021.100270>.
- [39] D. Allman, A. Reiter, M.A.L. Bell, Photoacoustic source detection and reflection artifact removal enabled by deep learning, *IEEE Trans. Med. Imaging* 37 (6) (2018) 1464–1477, <http://dx.doi.org/10.1109/TMI.2018.2829662>.
- [40] N. Awasthi, G. Jain, S.K. Kalva, M. Pramanik, P.K. Yalavarthy, Deep neural network-based sinogram super-resolution and bandwidth enhancement for limited-data photoacoustic tomography, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 67 (12) (2020) 2660–2673, <http://dx.doi.org/10.1109/TUFFC.2020.2977210>.
- [41] Y.E. Boink, S. Manohar, C. Brune, A partially-learned algorithm for joint photoacoustic reconstruction and segmentation, *IEEE Trans. Med. Imaging* 39 (1) (2020) 129–139, <http://dx.doi.org/10.1109/TMI.2019.2922026>.
- [42] D. Allman, F. Assis, J. Chrispin, M.A.L. Bell, A deep learning-based approach to identify in vivo catheter tips during photoacoustic-guided cardiac interventions, in: *Photons Plus Ultrasound: Imaging and Sensing 2019*, 10878, SPIE, 2019, pp. 454–460, <http://dx.doi.org/10.1117/12.2510993>.
- [43] E.M.A. Anas, H.K. Zhang, J. Kang, E. Boctor, Enabling fast and high quality LED photoacoustic imaging: A recurrent neural networks based approach, *Biomed. Opt. Express* 9 (8) (2018) 3852, <http://dx.doi.org/10.1364/BOE.9.003852>.
- [44] A. Hariri, A. Hariri, K. Alipour, K. Alipour, Y. Mantri, J.P. Schulze, J.P. Schulze, J.V. Jokerst, J.V. Jokerst, J.V. Jokerst, Deep learning improves contrast in low-fluence photoacoustic imaging, *Biomed. Opt. Express* BOE 11 (6) (2020) 3360–3373, <http://dx.doi.org/10.1364/BOE.11.033603>.
- [45] F. Kallor Joseph, A. Arora, P. Kancharla, M. Kuniyil Ajith Singh, W. Steenbergen, S. S. Channappayya, Generative adversarial network-based photoacoustic image reconstruction from bandlimited and limited-view data, in: A.A. Oraevsky, L.V. Wang (Eds.), *Photons Plus Ultrasound: Imaging and Sensing 2021*, SPIE, Online Only, United States, 2021, p. 54, <http://dx.doi.org/10.1117/12.2577750>.
- [46] T. Heimann, P. Mountney, M. John, R. Ionasec, Real-time ultrasound transducer localization in fluoroscopy images by transfer learning from synthetic training data, Special Issue on the 2013 Conference on Medical Image Computing and Computer Assisted Intervention, *Med. Image Anal.* 18 (8) (2014) 1320–1328, <http://dx.doi.org/10.1016/j.media.2014.04.007>.
- [47] M. Unberath, J.-N. Zaech, C. Gao, B. Bier, F. Goldmann, S.C. Lee, J. Fotouhi, R. Taylor, M. Armand, N. Navab, Enabling machine learning in X-ray-based procedures via realistic simulation of image formation, *Int. J. CARS* 14 (9) (2019) 1517–1528, <http://dx.doi.org/10.1007/s11548-019-02011-2>.
- [48] E. Maneas, A. Hauptmann, E.J. Alles, W. Xia, T. Vercauteren, S. Ourselin, A.L. David, S. Arridge, A.E. Desjardins, Deep learning for instrumented ultrasonic tracking: From synthetic training data to in vivo application, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* (2021) 1, <http://dx.doi.org/10.1109/TUFFC.2021.3126530>.
- [49] Y. Movshovitz-Attias, T. Kanade, Y. Sheikh, How useful is photo-realistic rendering for visual learning? in: G. Hua, H. Jégou (Eds.), *Computer Vision – ECCV 2016 Workshops*, in: *Lecture Notes in Computer Science*, Springer International Publishing, Cham, 2016, pp. 202–217, http://dx.doi.org/10.1007/978-3-319-49409-8_18.
- [50] L.C. Garcia-Peraza-Herrera, L. Fidon, C. D’Ettorre, D. Stoyanov, T. Vercauteren, S. Ourselin, Image compositing for segmentation of surgical tools without manual annotations, *IEEE Trans. Med. Imaging* 40 (5) (2021) 1450–1460, <http://dx.doi.org/10.1109/TMI.2021.3057884>.
- [51] M.K.A. Singh, LED-based photoacoustic imaging: From bench to bedside, 2020.
- [52] T. Agano, M.K.A. Singh, R. Nagaoka, K. Awazu, Effect of light pulse width on frequency characteristics of photoacoustic signal – an experimental study using a pulse-width tunable LED-based photoacoustic imaging system, *Int. J. Eng. Technol.* 7 (4) (2018) 4300–4303, <http://dx.doi.org/10.14419/ijet.v7i4.19907>.
- [53] A. Hariri, J. Lemaster, J. Wang, A.S. Jeevarathinam, D.L. Chao, J.V. Jokerst, The characterization of an economic and portable LED-based photoacoustic imaging system to facilitate molecular imaging, *Photoacoustics* 9 (2018) 10–20, <http://dx.doi.org/10.1016/j.pacs.2017.11.001>.
- [54] B.E. Treeby, B.T. Cox, K-wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, *J. Biomed. Opt.* 15 (2) (2010) 021314, <http://dx.doi.org/10.1117/1.3360308>.
- [55] L. Wang, S.L. Jacques, L. Zheng, MCML—Monte Carlo modeling of light transport in multi-layered tissues, *Comput. Methods Prog. Biomed.* 47 (2) (1995) 131–146, [http://dx.doi.org/10.1016/0169-2607\(95\)01640-F](http://dx.doi.org/10.1016/0169-2607(95)01640-F).
- [56] S.L. Jacques, Optical properties of biological tissues: A review, *Phys. Med. Biol.* 58 (11) (2013) R37–R61, <http://dx.doi.org/10.1088/0031-9155/58/11/R37>.
- [57] M. Jaeger, S. Schüpbach, A. Gertsch, M. Kitz, M. Frenz, Fourier reconstruction in optoacoustic imaging using truncated regularized inverse k-space interpolation, *Inverse Problems* 23 (6) (2007) S51–S63, <http://dx.doi.org/10.1088/0266-5611/23/6/S05>.
- [58] H.J. van Staveren, C.J.M. Moes, J. van Marie, S.A. Prahl, M.J.C. van Gemert, Light scattering in Intralipid-10% in the wavelength range of 400–1100 nm, *Appl. Opt.* AO 30 (31) (1991) 4507–4514, <http://dx.doi.org/10.1364/AO.30.004507>.
- [59] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, 2015, [arXiv:1505.04597](https://arxiv.org/abs/1505.04597) [cs].
- [60] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, 2017, [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) [cs].
- [61] I. Loshchilov, F. Hutter, SGDR: Stochastic gradient descent with warm restarts, 2017, [arXiv:1608.03983](https://arxiv.org/abs/1608.03983) [cs, math].
- [62] P. Arbeláez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (5) (2011) 898–916, <http://dx.doi.org/10.1109/TPAMI.2010.161>.
- [63] R.O. Duda, P.E. Hart, Use of the Hough transformation to detect lines and curves in pictures, *Commun. ACM* 15 (1) (1972) 11–15, <http://dx.doi.org/10.1145/361237.361242>.
- [64] M.-P. Dubuisson, A.K. Jain, A modified hausdorff distance for object matching, in: *Proceedings of 12th International Conference on Pattern Recognition*, Vol. 1, IEEE, 1994, pp. 566–568.
- [65] W. Xia, S. Noimark, E. Maneas, N.M. Brown, M.K.A. Singh, S. Ourselin, S.J. West, A.E. Desjardins, Enhancing photoacoustic visualization of medical devices with elastomeric nanocomposite coatings, in: *Photons Plus Ultrasound: Imaging and Sensing 2019*, Vol. 10878, International Society for Optics and Photonics, 2019, p. 108783G, <http://dx.doi.org/10.1117/12.2508658>.



Mengjie Shi is a Ph.D. student in the School of Biomedical Engineering & Imaging Sciences at King's College London, UK. She completed her bachelor's degree in Optoelectricity Information of Science and Technology at Nanjing University of Science and Technology, China in 2019 and master's degree in Communications and Signal Processing at Imperial College London in 2020. Her research interests focus on improving photoacoustic imaging with affordable light sources for guiding minimally invasive procedures.



Tianrui Zhao is a Ph.D. student in the School of Biomedical Engineering & Imaging Sciences at King's College London, UK. He received his B.Sc. in Materials Science and Engineering from Northwestern Polytechnical University, China, and M.Sc. degree in Materials for Energy and Environment from University College London, UK, in 2015 and 2016, respectively. His research interests include developing minimally invasive imaging devices based on photoacoustic imaging.



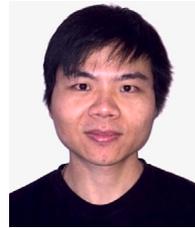
Dr. Sim West is a consultant anaesthetist at UCLH. He graduated from Sheffield in 2000, and completed his training in anaesthesia in North London, spending 2012 as the Smiths Medical Innovation Fellow. He was appointed to UCLH in 2013 and is lead for regional anaesthesia and the orthopaedic hub. His research interests include improving visualisation of needles, catheters and nerves.



Dr. Adrien Desjardins is a Professor in the Department of Medical Physics and Biomedical Engineering at the University College London, where he leads the Interventional Devices Group. His research interests are centred on the development of new imaging and sensing modalities to guide minimally invasive medical procedures. He has a particular interest in the application of photoacoustic imaging and optical ultrasound to guide interventional devices for diagnosis and therapy.



Tom Vercauteren is a Professor of Interventional Image Computing at King's College London since 2018 where he holds the Medtronic/Royal Academy of Engineering Research Chair in Machine Learning for Computer-assisted Neurosurgery. From 2014 to 2018, he was Associate Professor at UCL where he acted as Deputy Director for the Wellcome/EPSRC Centre for Interventional and Surgical Sciences (2017–18). From 2004 to 2014, he worked for Mauna Kea Technologies, Paris where he led the research and development team designing image computing solutions for the company's CE- marked and FDA-cleared optical biopsy device. His work is now used in hundreds of hospitals worldwide. He is a Columbia University and Ecole Polytechnique graduate and obtained his Ph.D. from Inria in 2008. Tom is also an established open-source software supporter.



Dr. Wenfeng Xia is a Lecturer in the School of Biomedical Engineering & Imaging Sciences at King's College London, UK. He received a B.Sc. in Electrical Engineering from Shanghai Jiao Tong University, China, and a M.Sc. in Medical Physics from University of Heidelberg, Germany, in 2005 and 2007, respectively. In 2013, he obtained his Ph.D from University of Twente, Netherlands. From 2014 to 2018, he was a Research Associate / Senior Research Associate in the Department of Medical Physics and Biomedical Engineering at University College London, UK. He currently leads the Photons+ Ultrasound Research Laboratory (<https://www.purkcl.org/>). His research interests include non-invasive and minimally invasive photoacoustic imaging, and ultrasound-based medical devices tracking for guiding interventional procedures.