

Ultrasound-enhanced Unet model for quantitative photoacoustic tomography of ovarian lesions

Yun Zou^a, Eghbal Amidi^a, Hongbo Luo^b, Quing Zhu^{a,c,*}

^a Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA

^b Department of Electrical and System Engineering, Washington University in St. Louis, St. Louis, MO, USA

^c Department of Radiology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA

ARTICLE INFO

Keywords:

Quantitative photoacoustic tomography
Blood oxygenation saturation
Unet
Ovarian cancer

ABSTRACT

Quantitative photoacoustic tomography (QPAT) is a valuable tool in characterizing ovarian lesions for accurate diagnosis. However, accurately reconstructing a lesion's optical absorption distributions from photoacoustic signals measured with multiple wavelengths is challenging because it involves an ill-posed inverse problem with three unknowns: the Grüneisen parameter (Γ), the absorption distribution, and the optical fluence (ϕ). Here, we propose a novel ultrasound-enhanced Unet model (US-Unet) that reconstructs optical absorption distribution from PAT data. A pre-trained ResNet-18 extracts the US features typically identified as morphologies of suspicious ovarian lesions, and a Unet is implemented to reconstruct optical absorption coefficient maps, using the initial pressure and US features extracted by ResNet-18. To test this US-Unet model, we calculated the blood oxygenation saturation values and total hemoglobin concentrations from 655 regions of interest (ROIs) (421 benign, 200 malignant, and 34 borderline ROIs) obtained from clinical images of 35 patients with ovarian/adnexal lesions. A logistic regression model was used to compute the ROC, the area under the ROC curve (AUC) was 0.94, and the accuracy was 0.89. To the best of our knowledge, this is the first study to reconstruct quantitative PAT with PA signals and US-based structural features.

1. Introduction

Photoacoustic imaging (PAI) is an emerging imaging modality for non-invasive, non-ionizing, real-time measurement of the optical properties of biological tissue [1]. Compared with other optical imaging modalities, PAI can image deeper because acoustic scattering in tissue is an order of magnitude smaller than that of optical scattering [2]. Additionally, PAI inherently possesses ultrasound resolution because received photoacoustic waves are used to form images [3–5]. Photoacoustic tomography (PAT) uses a broad laser beam for illumination and an array of ultrasound transducers to measure the photoacoustic waves generated by the targeted biological tissue [6]. Imaging is typically done either by standard delay-and-sum beamforming if a linear array is used as a receiver or by tomographic reconstruction if a circular receiver array is employed [7]. For clinical applications, linear array transducers in the frequency range of 3–10 MHz are often used because they can easily access an organ through a limited biological window and are widely available at a low cost [8,9]. In this transducer frequency range, the depth of penetration can reach several centimeters [10]. Oncology

applications of PAT include imaging and diagnosis of breast cancer [11, 12], ovarian cancer [13,14], cervical cancer [15], and thyroid cancer [16].

The PA effect is induced by short laser pulses which generate initial pressure rises inside the tissue. The generated initial pressure (p_0) in the tissue is proportional to Grüneisen parameter (Γ), target absorption coefficient (μ_a), and light fluence ($\phi(r)$) ($p_0 = \Gamma \bullet \mu_a \bullet \phi$) [17–20]. It is a technical challenge to quantitatively recover or reconstruct the optical absorption coefficient using PAT data because the received pressure is the product of tissue absorption and light fluence. Quantitative PAT approaches involve an ill-posed, non-unique, and non-linear inverse problem which is difficult to solve reliably for clinical applications [21, 22]. Moreover, PAT data can be noisy and background tissue absorption can produce artifacts, both of which make the reconstruction more challenging [23].

Blood oxygenation saturation (%sO₂) and total hemoglobin concentration (HbT) are functional biomarkers of the malignancy of ovarian tissue [24–26]. A %sO₂ map can be calculated point by point from the ratio of the estimated relative oxygenated hemoglobin to the relative

* Corresponding author at: Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA.
E-mail address: zhu.q@wustl.edu (Q. Zhu).

total hemoglobin concentration, assuming the unknown laser fluence at each imaging point at all wavelengths used is a constant and can be canceled out. However, the accuracy of %sO₂ depends on the PAT data quality received at all wavelengths.

There have been several approaches to reconstructing photoacoustic images, for example, iterative methods [27,28], and machine learning-based approaches [29–32]. Cai used an end-to-end machine learning model to achieve fast reconstruction [33]. Luke et al. proposed a double Unet (named ‘O-net’) to simultaneously estimate the oxygen saturation in blood vessels and segment the vessels from the surrounding background tissue [34], and Bench achieved 3D %sO₂ reconstruction of blood vessels under the skin by a 3D fully convolutional network [35].

However, these studies about machine learning-based PAT reconstruction methods were only using photoacoustic signals as input. Here, based on the knowledge of ovarian lesions, structural morphologies seen by US images were characteristics to classify benignancy or malignancy [41–43] and can be incorporated into the PAT U-net to improve the diagnostic accuracy of ovarian lesions.

Machine learning models, such as the ResNet model and Unet model, have demonstrated the ability to classify or reconstruct biomedical images. ResNet-18 is one of the most popular machine learning models used in image classification [36,37]. The developers of ResNet-18 pre-trained it on the ImageNet dataset [38]. Transfer learning can be adapted to train a ResNet-18 model for ovarian lesion feature extraction, because ImageNet objects such as plates (round shapes), stones (masses), and flags (edges) have features similar to ovarian lesions, such as nodules (round shapes) and solid masses with edges. The Unet model is suitable for PAT reconstruction because it can maintain PA high resolution by using its skip-connection feature to copy the input to output. In contrast to Unet, CNN models are less suitable for PAT reconstruction because they use pooling layers that downgrade the resolution [39,40].

In this manuscript, based on the observed characteristics of ovarian lesions in US images, structural morphologies seen in US images are extracted by a ResNet-18 model. These US features are incorporated into a Unet model structure to perform quantitative PAT reconstruction of absorption coefficient maps and then the calculation of total hemoglobin and %sO₂, referred to as US enhanced Unet (US-Unet). To the best of our knowledge, this is the first report of co-registered US being used to enhance the machine learning performance of PAT reconstruction. Both phantom experiments and clinical studies showed the superiority of the proposed method.

2. Methodology and materials

2.1. Ultrasound-enhanced Unet model

To extract ovarian tissue morphology features, a pre-trained ResNet-18 model was implemented, as shown in Fig. 1B. Ultrasound images of ovarian lesions were used to fine-tune the ResNet-18. Since we only had 1200 ultrasound images after augmentation, we chose to use a pre-trained ResNet-18 to avoid training from scratch with random-initial weights. Three cross-validations were used to evaluate the performance of the model.

To reconstruct PAT images, we designed a US-enhanced Unet model whose structure is shown in Fig. 1. Fig. 1A shows the Unet model used to reconstruct the optical absorption coefficient maps. Fig. 1B shows the ResNet-18 model implemented to extract features from US images. To enhance the Unet model with extracted morphological features, we implemented the ResNet-18 model with 18 convolutional layers. After the final convolutional layer, we changed the fully connected layer to 256 neurons, which was first resized to 16×16 , then concatenated to a bottleneck block in the Unet model (as indicated by the dashed red arrow in Fig. 1.) [45]. Thus, the US-enhanced Unet utilizes the US image features extracted by ResNet and the initial pressure calculated from our co-registered US and PAT system [24], to quantitatively reconstruct optical absorption coefficient maps.

The Unet model was trained using simulation and phantom data to learn the PAT reconstruction process, implementing the mean square error loss function. The ResNet-18 model was trained with clinical US images for classification, utilizing binary cross entropy loss. After training for 100 epochs, the ResNet-18 model could extract morphology features from US images well and classify lesions accurately. The features extracted from the ResNet-18 model were incorporated into the Unet model to perform PAT reconstruction. Thus, the US-enhanced Unet model could reconstruct PAT images with ovarian morphology features incorporated.

The model was implemented in Pytorch and trained for a total of 200 epochs on an Nvidia 2080 Ti GPU, using the ADAM optimizer with a learning rate of $1e-4$. A batch size of 256 images was used, and the training process was monitored for early stopping. To speed up training, the initial pressure maps and US images were down sampled from $700 \times 1400 - 128 \times 128$. We used 2000 simulation data and 480 phantom data for training, 540 blood tube datasets for validation, and 621

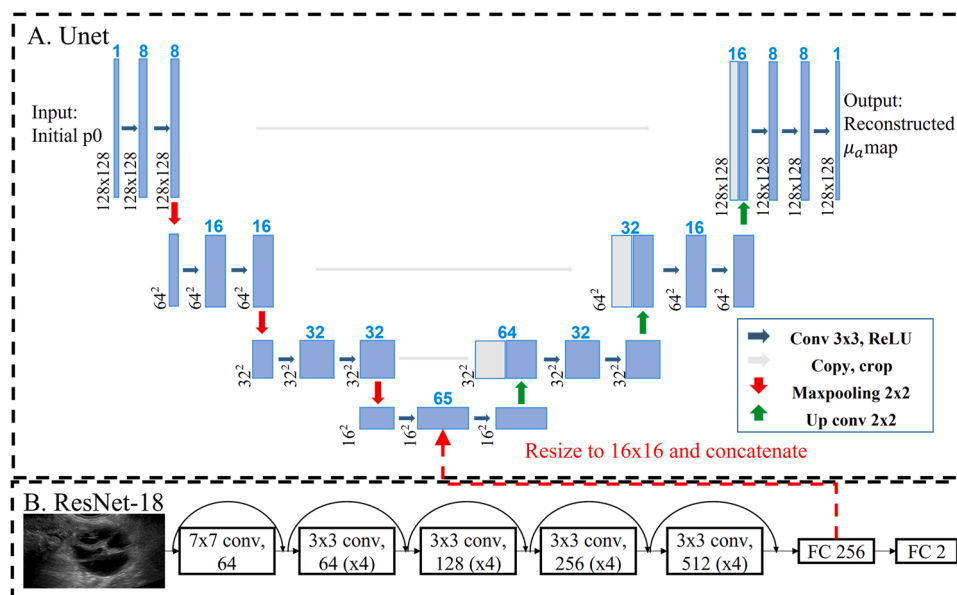


Fig. 1. Ultrasound-enhanced Unet model structure.

clinical datasets for testing.

2.2. %sO2 and total hemoglobin concentration (HbT) calculation

The ultrasound-enhanced Unet model reconstructed the optical absorption coefficient from raw data collected by the transducer. Then, using Eq. (1), we could compute the molar concentrations of oxy- and deoxy- hemoglobin [46]:

$$\mu_a(\lambda) = \ln(10)\epsilon_{oxy}(\lambda)C_{oxy} + \ln(10)\epsilon_{de}(\lambda)C_{de}. \quad (1)$$

Here, λ represents the wavelength. Further, μ_a denotes the optical absorption coefficient, and ϵ_{oxy} and ϵ_{de} are the molar extinction coefficients of oxy- and deoxy-hemoglobin, respectively. C_{oxy} and C_{de} are the molar concentrations of oxy- and deoxy-hemoglobin. For spectrum unmixing, we used the non-negative least square method.

The HbT and the oxygen saturation (%sO2) can be computed from Eq. (2):

$$\begin{aligned} \text{HbT} &= C_{oxy} + C_{de} \\ \%sO2 &= \frac{C_{oxy}}{C_{oxy} + C_{de}} \times 100\%. \end{aligned} \quad (2)$$

2.3. Co-registered US and PAT system

To collect initial pressure in this study, we used the real-time co-registered PAT and US system demonstrated in Ref. [47]. The co-registered PAT system has three parts, 1) a tunable Ti:sapphire laser (700–900 nm) pumped by a Q-switched Nd: YAG laser (Symphotics, Camarillo, CA), which generates 10 ns laser pulses at a 15 Hz repetition rate; 2) an optical fiber-based light delivery system [48]; and 3) a commercial US system (EC-12R, Alpinion Medical Systems, Republic of Korea). In this study, the system acquired PAT data at four wavelengths (730, 780, 800, 830 nm) [49]. In principle, two optical wavelengths are needed to compute C_{oxy} and C_{de} , and ideally, these two wavelengths are on each side of isosbestic wavelength at ~ 800 nm. However, two wavelengths of 730 nm and 780 nm give a more robust estimate of C_{de} , and the use of 800 nm and 830 nm give a more robust estimate of C_{oxy} . Also, these four wavelengths are close to one other, and so the change of wavelength-dependent fluence is minimal for spectral unmixing.

2.4. Simulations and phantoms (Unet training)

Simulation data were used to train the U-net model. To generate the simulation data, we defined digital phantoms with a homogenous background and targets with different shapes and optical properties (absorption and scattering coefficients). The optical fluence inside each digital phantom was calculated using Monte Carlo simulation [50]; the maximum depth was 6 cm, the grid and voxel sizes were $0.1 \times 0.1 \times 0.1$ cm, and the total number of photons used was $1e7$. We then computed the initial pressure value at each pixel of the phantom by multiplying the Grüneisen parameter at that pixel by its absorption coefficient and optical fluence. This initial pressure map was then used as the input of the k-wave toolbox [51], which generated the raw photoacoustic signal at each transducer element. The transducer configuration was the same as our co-registered US and PAT system. The simulation sampling frequency was 41 MHz, and input grid size was 700×1400 , which was $6\text{cm} \times 12\text{cm}$ in real size. Finally, the delay-and-sum algorithm was employed to reconstruct the initial pressure from the raw data, here the Grüneisen parameter was assumed to be 1 [52]. When training the US-Unet model, we used the reconstructed initial pressure map as the input and the optical absorption coefficient map as the output.

Finally, the delay-and-sum algorithm was employed to reconstruct the initial pressure from the raw data. In simulations, the Grüneisen parameter was assumed as 1 [52]. In phantom studies, the Grüneisen parameter was assumed to be constant and the light fluence profiles (ϕ

(r)) were assumed to be wavelength independent. When training the US-Unet model, we used the reconstructed initial pressure map as the input and the optical absorption coefficient map as the output.

We simulated the initial pressure with two different target shapes (a sphere and an ellipsoid), 25 different target absorption coefficients (ranging from 0.1 cm^{-1} to 4.9 cm^{-1} , with a step of 0.2 cm^{-1}), three different target depths (2.5, 3, and 3.5 cm), three different radii (1, 1.5, and 2 cm), and four different wavelengths (730, 780, 800, and 830 nm). In total, 1800 simulation data were generated.

Five ovary-mimicking gelatin-based phantoms were made, all with the same amount of gelatin, diluted intralipid (to provide the desired optical scattering), and glass beads (to enhance ultrasound scattering). Different amounts of ink were added to each phantom to achieve absorption coefficients of 0.2, 0.4, 0.6, 0.8, and 1.0 cm^{-1} . After being refrigerated for 24 h, each of the five solidified phantoms was placed by turns at three different imaging depths (1 cm, 1.5 cm, and 2 cm) in an intralipid solution with calibrated absorption and reduced scattering coefficients of 0.02 and 6 cm^{-1} , similar to the optical properties of soft tissue, and PAT data were acquired at each depth. Two sets of experiments were done for each phantom at each depth, and each measurement contained 4 frames, which generated 480 data sets at the previously mentioned four wavelengths, three depths, and five absorption coefficients.

Besides gelatin phantoms, we made blood tubes to mimic the blood vessels in tissues with different %sO2 contrast. Five calibrated %sO2 blood samples of 24.9 %, 44.2 %, 64.9 %, 83.9 %, and 97.6 % were made by controlling the amount of oxygen and nitrogen in the tube. The blood within the tube was collected from volunteers and mixed with saline water in a temperature/humidity-controlled chamber. The blood tube % sO2 values were calibrated with an ABL90 FLEX Radiometer. More details can be found in Ref. [24]. These blood tubes were placed at 9 different depths (1 cm, 1.5 cm, 2 cm, 2.5 cm, 3 cm, 3.5 cm, 4 cm, 4.5 cm, and 5 cm from the transducer) in a tank filled with intralipid solution. The experiments were repeated three times, generating 540 blood tube datasets in total.

2.5. Patients and ovarian lesion data

The study protocol was approved by the institutional review board of the Washington University Medical School and was compliant with the Health Insurance Portability and Accountability Act. A total of 35 patients with ovarian/ adnexal masses signed the informed consent and participated in this study from February 2017 to November 2018 [24]. The patients' ages ranged from 33 to 87, and the ovarian lesions' diameters ranged from 1.8 cm to 11.5 cm. The final data set consisted of 10 malignant ovarian/adnexal masses, 36 benign masses, and normal ovaries, and 3 borderline tumors of low malignancy potential. A total of 655 regions of interest (ROIs) (421 benign, 200 malignant, and 34 borderline tumor ROIs) were selected and used in this study (Table 1).

Ultrasound images from a commercial US machine were collected from the same group of patients. We selected 600 ovarian lesion images, and after augmentation (rotation and flipping), the resulting 1200 ultrasound images were used in ResNet-18 fine-tune training and validation.

2.6. Ovarian cancer morphologies revealed in US

As shown in Fig. 2, there were five typical morphologies in the

Table 1
Clinical data.

	Malignant	Benign	Borderline
Number of Patients	6	26	3
Number of Ovarian Lesions	10	36	3
Number of Data Sets	200	421	34

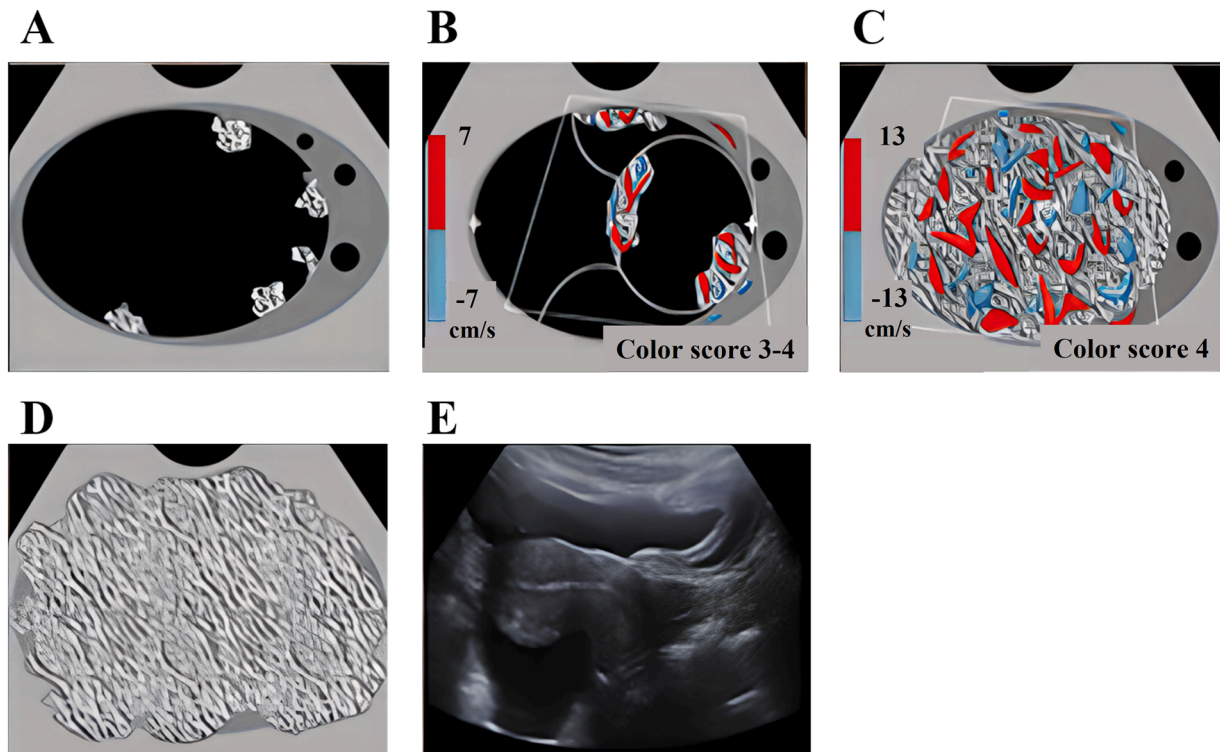


Fig. 2. Five typical morphologies of ovarian cancer [53]. A) Unilocular cyst with 4 or more papillary projections. B) Multinodular cyst with solid component. C) Solid with smooth contour and high blood flow. D) Solid with irregular contour. E) Ascites and peritoneal nodules.

malignant ovarian cancers characterized by US images [53]. Some ovarian cancers appeared as solid papillary projections from the wall of cystic lesions (Fig. 2A and B). Other ovarian cancers had significant blood content, as shown by Doppler (Fig. 2C). Still other ovarian cancers appeared as solid lesions with irregular contours (Fig. 2D and E).

Compared with benign lesions, malignant lesions often have more numerous solid nodules growing from the cyst wall, more irregular solid components both inside and/or at the lesion periphery, and more blood vessels revealed by Doppler US. Although these features are predictive of malignancy, however, they are not sensitive and specific [54]. Here, we sought to improve overall ovarian cancer diagnosis by using additional functional information obtained from co-registered photoacoustic imaging, such as the %sO₂ contrast and the quantitative hemoglobin concentration.

3. Results and analysis

3.1. Phantom results

As shown in Fig. 3, the reconstructed μ_a was calculated for the phantom experiments, each box represents the reconstructed μ_a distribution from the proposed method, and the dashed line is the ground truth. The mean square estimated error for the phantoms, with absorption coefficients of 0.2, 0.4, 0.6, 0.8, and 1.0 cm⁻¹, were 0.018, 0.011, 0.019, 0.011, and 0.013 cm⁻¹, respectively.

3.2. Blood tube validation results

After we trained the Unet model with simulation and phantom data, blood tube data was used to validate the performance of the US-Unet model. Blood tube reconstruction examples are shown in Figs. 4 and 5. Both the calculated %sO₂ and HbT maps were in the similar range of calibrated values, and the shapes were close to the tube shape.

All the blood tube reconstructions were aggregated for quantitative display. As shown in Fig. 6, each box represents the calculated mean %

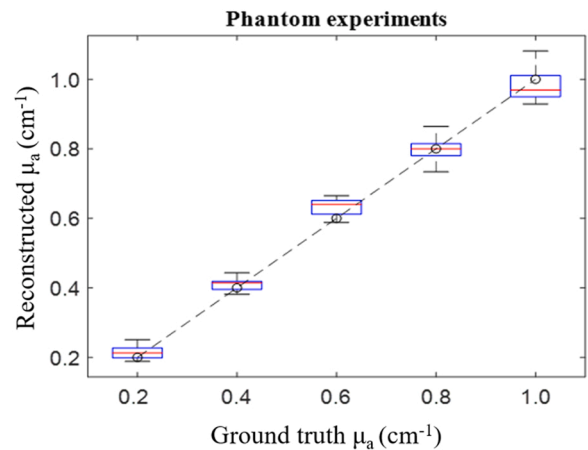


Fig. 3. Reconstructed μ_a vs ground truth μ_a for phantom experiments.

sO₂ for a tube at a different depth, from 1 cm to 5 cm, in steps of 0.5 cm. The dashed line is the ground truth based on calibration.

In our method, the machine learning model was trained with the ground truth, and the blood tubes were homogenous. Therefore, the model could reconstruct the real shape and size of the targets. In other words, the machine learning model can mitigate boundary buildup due to interference [55]. Besides, we downsampled the reconstruction images, which also could yield a more homogenous result.

The normalized mean square estimated error for blood tubes with % sO₂ of 24.9 %, 44.2 %, 64.9 %, 83.9 %, and 97.6 % were 2.9 %, 3.5 %, 2.7 %, 2.6 %, and 3.0 %, respectively. Compared with standard delay-and-sum results of 5.6 %, 3.8 %, 9.8 %, 8.7 % and 17.7 %, respectively [24], as shown in Fig. 6, the errors were much smaller. Fig. 7 shows the Unet reconstructed quantitative HbT data of blood tube. The calibrated values were given in black circles.

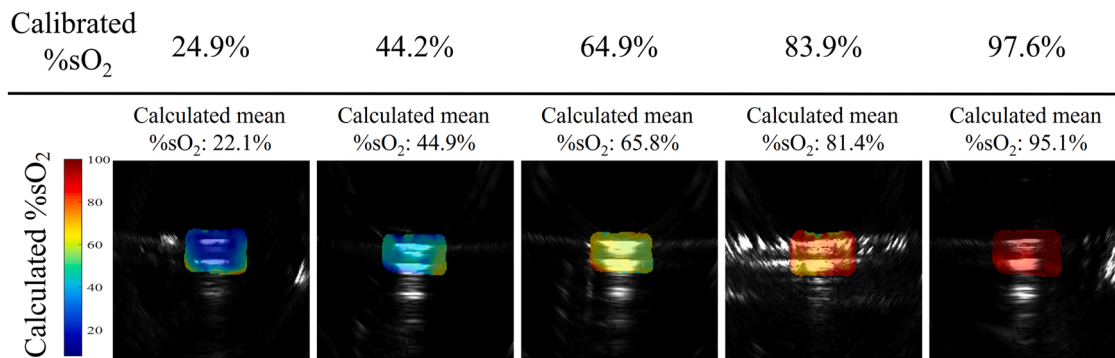


Fig. 4. Reconstructed blood tubes' %sO₂ co-registered with US images for five different %sO₂ concentrations located 2 cm deep.

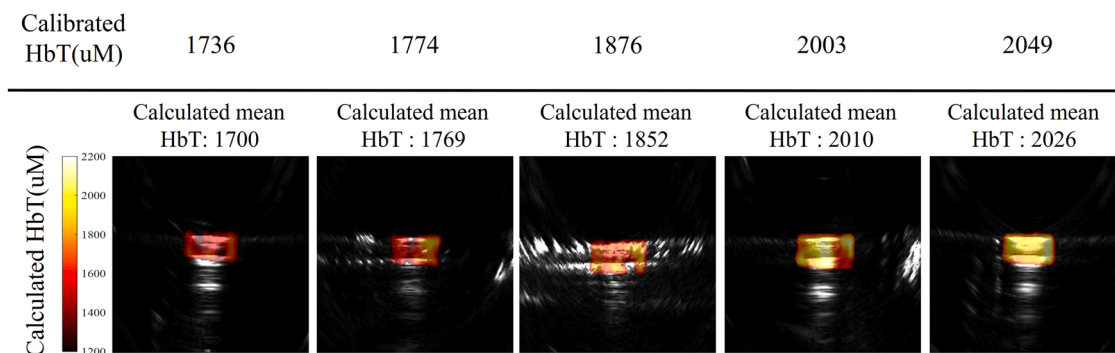


Fig. 5. Reconstructed blood tubes' HbT co-registered with US images for 5 different HbT concentrations located at 2 cm depth.

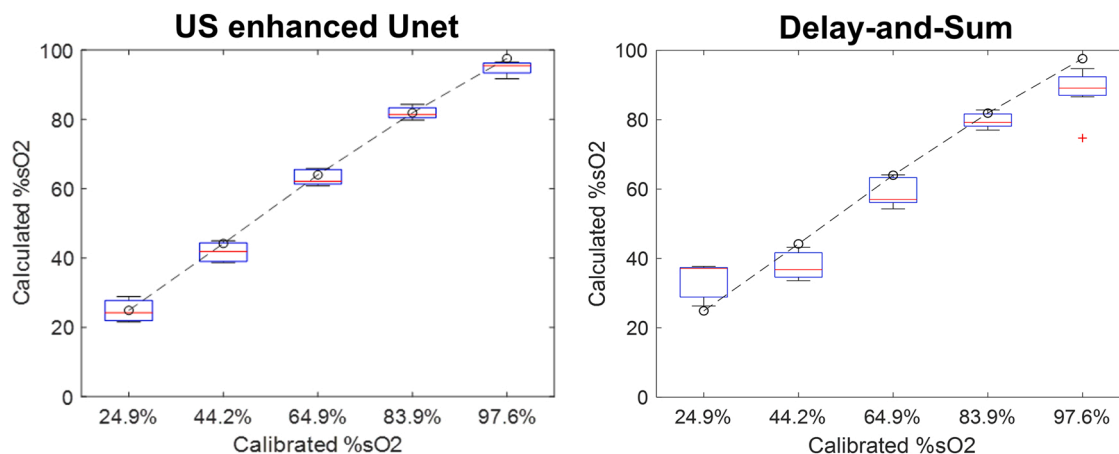


Fig. 6. Calculated %sO₂ vs calibrated %sO₂ from the Unet model and delay-and-sum beamforming. The boxes represent tubes containing blood with different %sO₂ values, located at depths from 1 to 5 cm.

3.3. ResNet-18 results

Three cross-validations were used to test the accuracy of the ResNet-18 model. The accuracies were 0.80 ± 0.02 , 0.76 ± 0.04 , and 0.81 ± 0.04 , which suggests that about 80 % accuracy can be achieved with morphology features only.

We implemented Grad-CAM (gradient-weighted class activation mapping) to learn what features the three ResNet-18 models had learned. Grad-CAM uses the “gradients of any target” concept, back-flowing into the final convolutional layer to produce a coarse localization map highlighting important regions in the image for predicting the diagnostic result [56,57]. In other words, if we have a trained model, we can use specific data and its label to perform back propagation to find the regions the model is “looking at” to make its prediction.

From Fig. 8, we can see the ResNet-18 model learned the “nodules” and boundaries (hot spots) that were essential to benign or malignant lesion classification based on ultrasound images. The results also suggest that, although morphological features could distinguish between benign and malignant lesions, they are not specific or sensitive (moderate accuracy). So, to improve the diagnostic results, these features extracted from ResNet-18 could be incorporated into the US-Unet model.

3.4. Clinical results

For clinical tests, total hemoglobin and %sO₂ are important biomarkers in analyzing ovarian lesions. Total hemoglobin and %sO₂ were calculated from the optical absorption coefficient maps of the US-enhanced Unet model output.

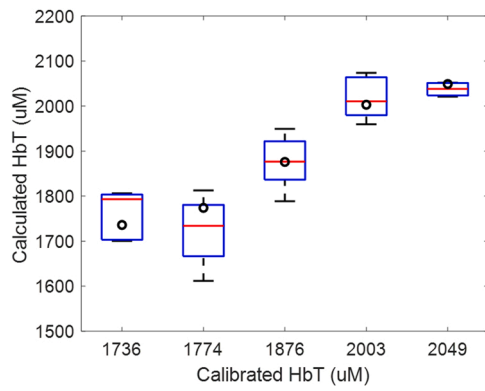


Fig. 7. Calculated HbT vs calibrated HbT from the Unet model. The boxes represent tubes containing blood with different HbT. The normalized mean square estimated error for blood tubes with HbT of 1736, 1774, 1876, 2003, and 2049 were 3.14 %, 4.88 %, 2.45 %, 2.73 %, and 0.89 %, respectively.

Fig. 9 shows the %sO₂ and the HbT results for a malignant and a benign lesion. In **Fig. 9A**, the mean %sO₂ of the ROI is 30.3 %, which is much lower than the mean %sO₂ of 80.6 % in the benign lesion shown in **Fig. 9B**. Besides the mean %sO₂ value, we could also see peripheral %sO₂ distribution around the ultrasonically seen lesion. And according to mean HbT value, the malignant case (377 μ M) is 1.5 times higher than the benign case (255 μ M).

To test the performance of the US-Unet model, three reconstruction methods were implemented: US-Unet, delay-and-sum reconstruction, and Unet reconstruction. As seen in **Fig. 10**, all three methods could distinguish benign from malignant lesions, with contrasts of 1.68, 1.33, 1.44, respectively. However, the US-Unet model was superior to the other two methods in contrast. As for contrast between borderline tumors and benign lesions, the delay-and-sum is better than that US-Unet and Unet alone, however, due to the smaller sample size of borderline tumors, we could not make any conclusion at this point.

Fig. 11 shows the reconstructed quantitative total hemoglobin

concentrations using the US-Unet (A), delay-and-sum (B), and Unet (C). For US-Unet, the median HbT of the malignant cases is 2.15 times higher than that of the benign cases. For delay-and-sum beamforming, the relative HbT is 1.51 times higher. For Unet, it is 1.71 times higher. The HbT contrast between borderline tumors and benign lesions is different for the delay-and-sum than that of US-Unet and Unet alone. Again, due to the smaller sample size of borderline tumors, we could not make any conclusion at this point. More data are needed.

Logistic regression was implemented to evaluate the diagnostic performance of %sO₂ and HbT data and %sO₂ data only. The %sO₂ values, with and without HbT testing data from the output of the US-Unet model, with the input of PAT data and US images, were used to calculate the receiver operating characteristic (ROC) curves and the area under the curve (AUC) (See **Fig. 12**). Accuracies were calculated for the US-enhanced Unet model, the delay-and-sum model, and the Unet model (**Table 2**). Here, we distinguished benign and malignant cases only, because the three borderline cases are not large enough for a category to be classified. The results show the superiority of US-enhanced Unet with %sO₂ and HbT data.

4. Summary and discussion

In this study, a novel ultrasound-enhanced Unet model was proposed to reconstruct the optical coefficient maps of ovarian/adnexal lesions from photoacoustic data. The experimental results showed that the ultrasound-enhanced Unet model reconstructed the target's %sO₂ and HbT maps more accurately than the delay-and-sum method and provided an improved diagnosis of ovarian cancers from benign lesions.

Ovarian lesion morphologies from US images were predictive of malignancy, therefore, a ResNet-18 was implemented to extract relative features from US images. Grad-CAM demonstrated that the ResNet-18 classification model focused on a pertinent area to make its predictions, which also suggested that the model had extracted related features useful for classification. Compared with both a Unet model without ultrasound extracted features and with the standard delay-and-sum method, the US-enhanced Unet demonstrated superior diagnostic contrast and also achieved diagnostic performance of AUC of 0.94 and

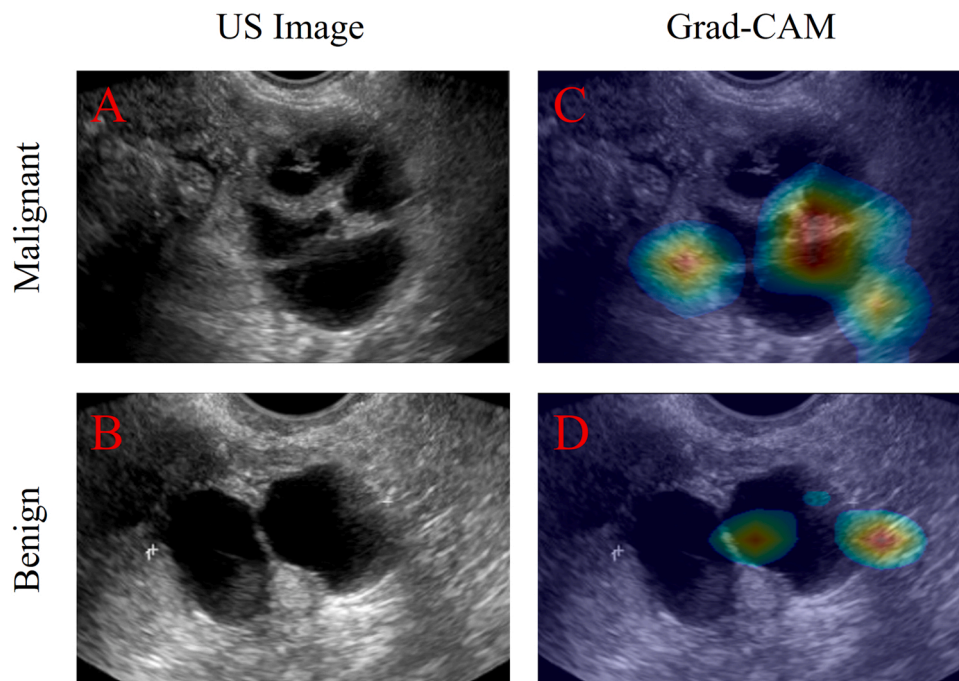


Fig. 8. A is an original US image of a malignant ovarian cancer case. C shows the co-registered US image and grad-CAM image: the hot spots are where the model focused. B and D are respectively a US image and a grad-CAM image of a benign case.

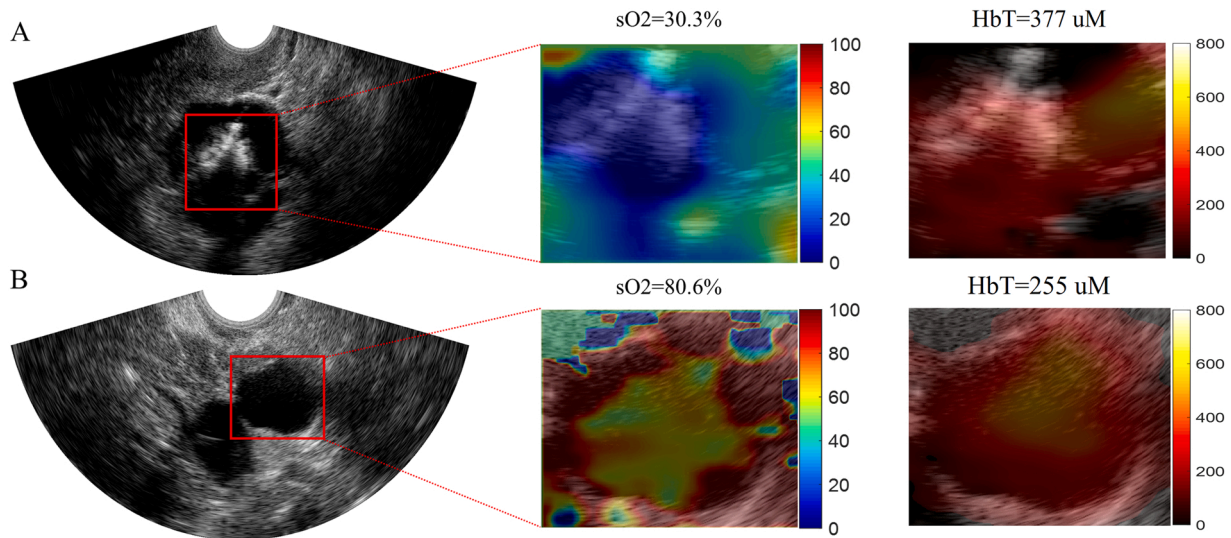


Fig. 9. A. is a malignant lesion US image, and B is a benign lesion US image. For each row, left is the US image, the red square is the ROI. Middle, the %sO2 map of the ROI, with a color bar from 0% to 100 %. Right, the HbT map with a colorbar from 0 to 800 μ M.

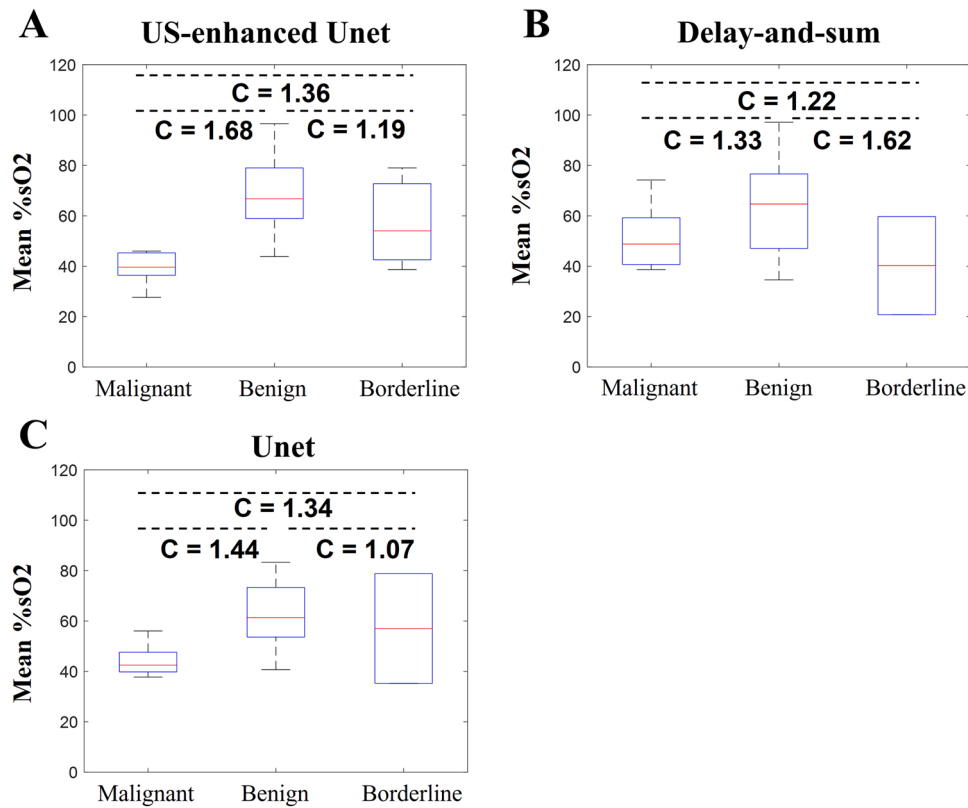


Fig. 10. A. %sO2 results of US-Unet model. B. Delay-and-sum reconstruction results. C. Unet reconstruction results. The label C indicates the ratio of the contrasts of the median.

accuracy of 0.90. Classification results using reconstructions from the Unet model underperformed, even compared with the delay-and-sum method. Without ovarian morphology features, the Unet model could learn only from simulation and phantom data. There is a big difference between clinical and simulation/phantom data, because ovarian tissue is heterogeneous, with complex shapes. This factor could lead to an inferior prediction if using the Unet model only. The delay-and-sum beamforming method directly reconstructed the relative HbT from PAT raw data, with no learning process, and the errors resulted from the

tissue heterogeneity, and the inaccurate assumption that the light fluence was wavelength independent.

The US-enhanced Unet has several limitations. The US-enhanced Unet model was limited by low resolution and small reconstructed image size, due to down sampling. However, using a large size of initial pressure will increase the training time exponentially. Additionally, although we reconstructed the quantitative oxy- and deoxy-hemoglobin concentrations, there were outliers in the reconstructions due to ill-posed problems and the diagnostic performance of HbT only was

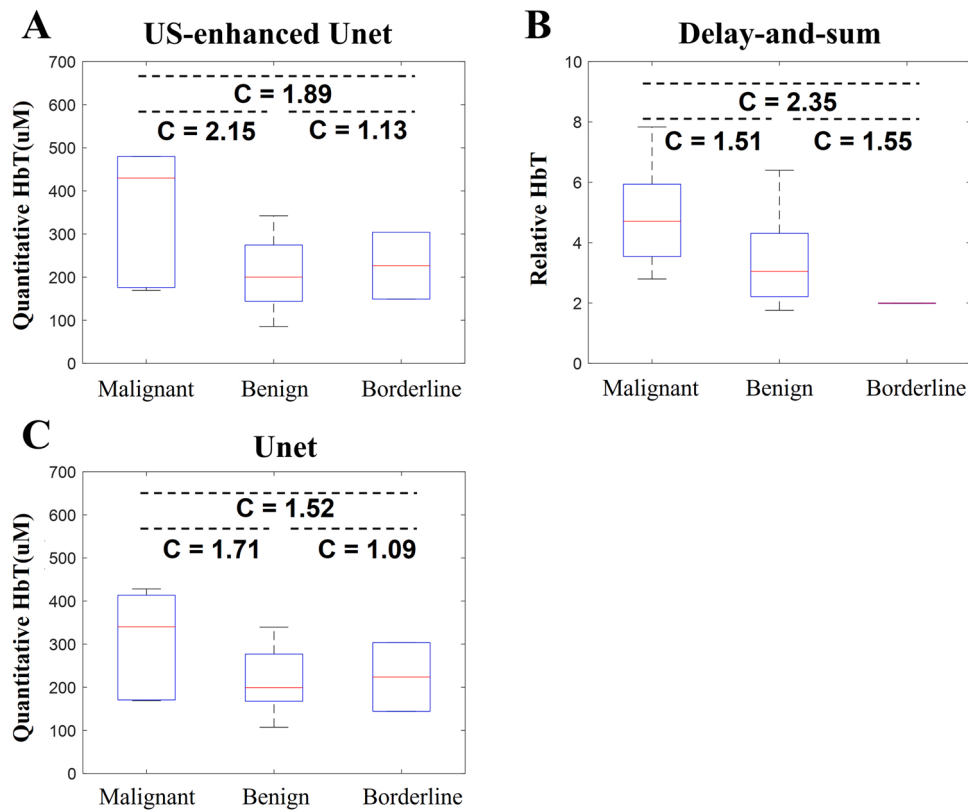


Fig. 11. A. Quantitative total hemoglobin results for all patients' data by US-Unet reconstructions. B. Quantitative total hemoglobin results for delay-and-sum beam-forming reconstruction. C. Quantitative total hemoglobin results for the Unet model. The label C indicates the ratio of the contrasts of the median value of each box.

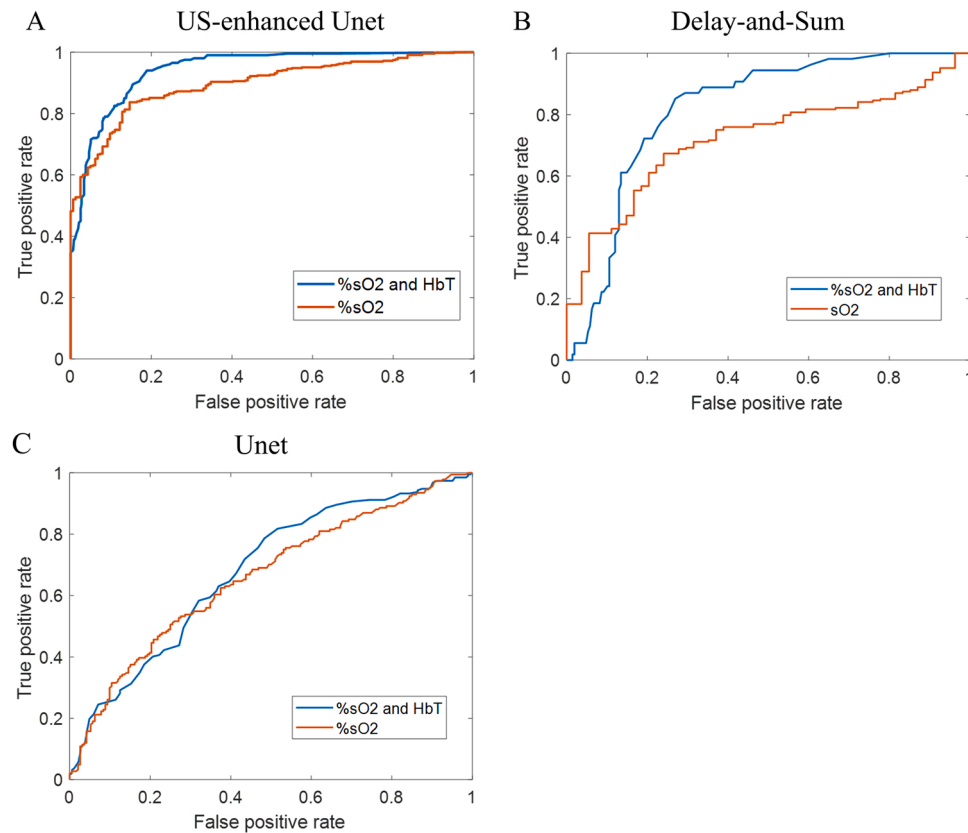


Fig. 12. A. ROC curves with %sO2 and HbT data and with %sO2 data only for the US-enhanced Unet model. B. ROC curves with %sO2 and HbT data and with %sO2 data only for the delay-and-sum beam-forming model. C. ROC curves with %sO2 and HbT data and with %sO2 data only for the Unet model.

Table 2
Logistic regression accuracy and AUC.

	US-enhanced Unet		Unet		Dealy-and-Sum	
	accuracy	AUC	accuracy	AUC	accuracy	AUC
%sO2 and HbT	0.89	0.94	0.62	0.68	0.71	0.82
%sO2	0.80	0.89	0.61	0.66	0.66	0.72

lower than %sO2. Finally, the absorption coefficients used in simulations and gelatin phantoms in training are our best estimates because we found very limited literature reporting ovarian tissue absorption coefficients. One study, Ref. 44, used diffuse optical tomography to image excised ovarian lesions. Among 24 samples measured at 780 nm, 6 malignant samples had a mean absorption coefficient of 0.153 cm^{-1} , and benign samples had a mean coefficient of 0.093 cm^{-1} . However, 4 out of 6 malignant samples were pieces of large malignant lesions, and there was significant blood loss from the samples before imaging. Additionally, the diffuse optical tomography underestimated the target absorption because of the linear assumption of the target perturbation and target absorption using Born approximation. Thus, we believed the ovarian tissue optical absorption coefficients reported in Ref 44 were underestimated. In another *ex vivo* ovarian specimen study using spatial frequency domain imaging, the measured absorption coefficient of ovarian tissue ranged from 0.1 cm^{-1} to 0.7 cm^{-1} [58]. We also measured the absorption coefficient for the blood tubes used by spectrometry and found the absorption coefficients for the four wavelengths ranged from 1.11 cm^{-1} to 2.06 cm^{-1} , more details in Appendix B. From the literature, the optical absorption coefficient of blood is in the range of $1\text{--}5 \text{ cm}^{-1}$ [59,60]. Because malignant ovarian lesions consist of mixed blood vessels and stromal tissues, we used estimated absorption coefficients in the range of $0.2\text{--}1 \text{ cm}^{-1}$ for gelatin phantoms and $0.1\text{--}4.9 \text{ cm}^{-1}$ in simulations.

The %sO2 and HbT data shown in Figs. 10 and 11 reveal that delay-and-sum beamforming performed differently from US-Unet and Unet for three borderline tumor cases. Delay-and-sum showed borderline tumors had lower %sO2 and lower relative HbT than benign lesions, while the US-enhanced Unet and Unet both showed that the %sO2 and HbT values of borderline cases were similar to those of benign lesions. Borderline ovarian lesions are slow-growing tumors with low malignant potential. The tumor angiogenesis development could be slow, and the lesion morphology could overlap between large benign and malignant lesions. There is no reported %sO2 data to support the suggestion that borderline ovarian/adnexal lesions should have the same or lower values than benign lesions. Nevertheless, it is likely that Unet and US-enhanced Unet did not see enough of these cases to robustly predict %sO2 and HbT. It is also likely that delay-and-sum was not sufficiently robust to estimate %sO2 when PAT signals were low. More borderline ovarian/adnexal lesions will be acquired from the ongoing trial to further evaluate the performance of US-Unet and beamforming algorithms in predicting borderline ovarian/adnexal lesions.

For clinical data, the quality of reconstruction was highly affected by the system's SNR. For low SNR data, the reconstruction will be less robust. Future work will be focused on building a larger database from phantoms to enhance training and reduce the effect of a low SNR. Currently, a clinical trial is on-going with the goal of acquiring US and PAT data of diverse ovarian/adnexal lesions of malignancies of various types, benign solid and cystic lesions, and mixed solid and cystic lesions, as well as normal ovaries. This large *in vivo* database will be used to further validate the initial finding that US-enhanced Unet has the superior diagnostic performance to standard delay-and-sum beamforming.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence

the work reported in this paper.

Data Availability

Data will be made available on request.

Acknowledgments

This work was supported by National Institute of Health (R01 CA237664 and R01 CA228047). We thank the entire GYN oncology group for helping identify patients and the Radiology team for helping with US scans. Finally, we thank Professor James Ballard for manuscript editing.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.pacs.2022.100420](https://doi.org/10.1016/j.pacs.2022.100420).

References

- [1] L.V. Wang, L. Gao, Photoacoustic microscopy and computed tomography: from bench to bedside, *Annu. Rev. Biomed. Eng.* 16 (2014) 155–185, <https://doi.org/10.1146/annurev-bioeng-071813-104553>.
- [2] P. Beard, Biomedical photoacoustic imaging, *Interface Focus* 1.4 (2011) 602–631, <https://doi.org/10.1098/rsfs.2011.0028>.
- [3] A.B.E. Attia, G. Balasundaram, M. Moothanchery, U.S. Dinis, R. Bi, V. Ntziachristos, M. Olivo, A review of clinical photoacoustic imaging: current and future trends, *Photoacoustics* 16 (2019), 100144, <https://doi.org/10.1016/j.pacs.2019.100144>.
- [4] J. Xia, J. Yao, L.V. Wang, Photoacoustic tomography: principles and advances, *Electromagn. Waves* 147 (2014) 1, <https://doi.org/10.2528/PIER14032303>.
- [5] S. Jeon, J. Kim, D. Lee, J.W. Baik, C. Kim, Review on practical photoacoustic microscopy, *Photoacoustics* 15 (2019), 100141, <https://doi.org/10.1016/j.pacs.2019.100141>.
- [6] Y. Zhou, J. Yao, L.V. Wang, Tutorial on photoacoustic tomography, *J. Biomed. Opt.* 21.6 (2016), 061007, <https://doi.org/10.1117/1.JBO.21.6.061007>.
- [7] R. Manwar, K. Kratiewicz, K. Avnaki, Overview of ultrasound detection technologies for photoacoustic imaging, *Micromachines* 11.7 (2020) 692, <https://doi.org/10.3390/mi11070692>.
- [8] T.L. Szabo, P.A. Lewin, Ultrasound transducer selection in clinical imaging practice, *J. Ultrasound Med.* 32.4 (2013) 573–582, <https://doi.org/10.7863/jum.2013.32.4.573>.
- [9] L.G. Montilla, R. Olafsson, D.R. Bauer, R.S. Witte, Real-time photoacoustic and ultrasound imaging: a simple solution for clinical ultrasound systems with linear arrays, *Phys. Med. Biol.* 58.1 (2012) N1.
- [10] L.V. Wang, Tutorial on photoacoustic microscopy and computed tomography, *IEEE J. Sel. Top. Quantum Electron.* 14.1 (2008) 171–179, <https://doi.org/10.1109/JSTQE.2007.913398>.
- [11] X. Li, C.D. Heldermon, L. Yao, L. Xi, H. Jiang, High resolution functional photoacoustic tomography of breast cancer, *Med. Phys.* 42.9 (2015) 5321–5328, <https://doi.org/10.1118/1.4928598>.
- [12] L. Xi, X. Li, L. Yap, S. Grobmyer, H. Jiang, Design and evaluation of a hybrid photoacoustic tomography and diffuse optical tomography system for breast cancer detection, *Med. Phys.* 39.5 (2012) 2584–2594, <https://doi.org/10.1118/1.3703598>.
- [13] S. Nandy, A. Mostafa, I.S. Hagemann, M.A. Powell, E. Amidi, K. Robinson, D. G. Mutch, C. Siegel, Q. Zhu, Evaluation of ovarian cancer: initial application of coregistered photoacoustic tomography and US, *Radiology* 289.3 (2018) 740–747, <https://doi.org/10.1148/radiol.2018180666>.
- [14] H.S. Salehi, H. Li, A. Merkulov, P.D. Kumavor, H. Vavadi, M. Sanders, A. Kueck, M. A. Brewer, Q. Zhu, Coregistered photoacoustic and ultrasound imaging and classification of ovarian cancer: *ex vivo* and *in vivo* studies, *J. Biomed. Opt.* 21.4 (2016), 046006, <https://doi.org/10.1117/1.JBO.21.4.046006>.
- [15] K.S. Valluru, J.K. Willmann, Clinical photoacoustic imaging of cancer, *Ultrasonography* 35.4 (2016) 267, <https://doi.org/10.14366.16035>.
- [16] V.S. Dogra, B.K. Chinni, K.S. Valluru, J. Moalem, E.J. Giampoli, K. Evans, N.A. Rao, Preliminary results of *ex vivo* multispectral photoacoustic imaging in the management of thyroid cancer, *Am. J. Roentgenol.* 202.6 (2014) W552–W558, <https://doi.org/10.2214/AJR.13.11433>.
- [17] M. Xu, L.V. Wang, Photoacoustic imaging in biomedicine, *Rev. Sci. Instrum.* 77 (4) (2006), 041101, <https://doi.org/10.1063/1.2195024>.
- [18] B.T. Cox, J.G. Lauffer, P.C. Beard, S.R. Arridge, Quantitative spectroscopic photoacoustic imaging: a review, *J. Biomed. Opt.* 17.6 (2012), 061202, <https://doi.org/10.1117/1.JBO.17.6.061202>.
- [19] M. Luriakose, M.A. Borden, Microbubbles and nanodrops for photoacoustic tomography, *Curr. Opin. Colloid Interface Sci.* 55 (2021), 101464, <https://doi.org/10.1016/j.cocis.2021.101464>.

- [20] V.P. Nguyen, Y.M. Paulus, Photoacoustic ophthalmoscopy: principle, application, and future directions, *J. Imaging* 4.12 (2018) 149, <https://doi.org/10.3390/jimaging4120149>.
- [21] P. Shao, B. Cox, R.J. Zemp, Estimating optical absorption, scattering, and grüneisen distributions with multiple-illumination photoacoustic tomography, *Appl. Opt.* 50 (19) (2011) 3145–3154.
- [22] M. Haltmeier, L.V. Nguyen, Analysis of iterative methods in photoacoustic tomography with variable sound speed, *SIAM J. Imaging Sci.* 10.2 (2017) 751–781.
- [23] A. Kazakeviciute, C.J.H. Ho, M. Olivo, Multispectral photoacoustic imaging artifact removal and denoising using time series model-based spectral noise estimation, *IEEE Trans. Med. Imaging* 35 (9) (2016) 2151–2163, <https://doi.org/10.1109/TMI.2016.2550624>.
- [24] E. Amidi, G. Yang, K.M.S. Uddin, H. Luo, W. Middleton, M. Powell, C. Siegel, Q. Zhu, Role of blood oxygenation saturation in ovarian cancer diagnosis using multi-spectral photoacoustic tomography, *J. Biophotonics* 14 (4) (2021), e202000368, <https://doi.org/10.1002/jbio.202000368>.
- [25] Srivalleesha Mallidi, Geoffrey P. Luke, Stanislav Emelianov, Photoacoustic imaging in cancer detection, diagnosis, and treatment guidance, *Trends Biotechnol.* 29.5 (2011) 213–221, <https://doi.org/10.1137/16M1104822>.
- [26] J.L. Su, B. Wang, K.E. Wilson, C. Bayer, Y. Chen, S. Kim, K.A. Homan, S. Y. Emelianov, Advances in clinical and biomedical applications of photoacoustic imaging, *Expert Opin. Med. Diagn.* 4.6 (2010) 497–510, <https://doi.org/10.1517/17530059.2010.529127>.
- [27] B.T. Cox, S.R. Arridge, K.P. Kostli, P.C. Beard, Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method, *Appl. Opt.* 45 (8) (2006) 1866–1875.
- [28] C. Huang, K. Wang, L. Nie, L.V. Wang, M.A. Ananstasio, Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media, *IEEE Trans. Med. Imaging* 32.6 (2013) 1097–1110, <https://doi.org/10.1109/TMI.2013.2254496>.
- [29] S. Guan, K. Hsu, M. Eyassu, P.V. Whitnis, Dense dilated UNet: deep learning for 3D photoacoustic tomography image reconstruction, *arXiv Prepr. arXiv* (2021), <https://doi.org/10.48550/arXiv.2104.03130>, 2104.03130.
- [30] C. Yang, H. Lan, H. Zhong, and F. Gao, Quantitative photoacoustic blood oxygenation imaging using deep residual and recurrent neural network, 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019. (<https://doi.org/10.1109/ISBI.2019.8759438>).
- [31] N. Awasthi, G. Jain, S.E. 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, <https://doi.org/10.1109/TUFFC.2020.2977210>.
- [32] 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) 3852–3866, <https://doi.org/10.1364/BOE.9.003852>.
- [33] C. Cai, K. Deng, C. Ma, J. Luo, End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging, *Opt. Lett.* 43.12 (2018) 2752–2755, <https://doi.org/10.1364/OL.43.002752>.
- [34] G.P. Luke, K.H. Hawlik, A.C.V. Namen, R. Shang, O-Net: a convolutional neural network for quantitative photoacoustic image segmentation and oximetry, 1911.01935, *arXiv Prepr. arXiv* (2019), <https://doi.org/10.48550/arXiv.1911.01935>, 1911.01935.
- [35] C. Bench, A. Hauptmann, B.T. Cox, Toward accurate quantitative photoacoustic imaging: learning vascular blood oxygen saturation in three dimensions, *J. Biomed. Opt.* 25.8 (2020), 085003, <https://doi.org/10.1117/1.JBO.25.8.085003>.
- [36] G. Liang, L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, *Comput. Methods Prog. Biomed.* 187 (2020), 104964, <https://doi.org/10.1016/j.cmpb.2019.06.023>.
- [37] A. Mahajan, S. Chaudhary, Categorical image classification based on representational deep network (RESNET), 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2019. (<https://doi.org/10.1109/ICECA.2019.8822133>).
- [38] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, F. Li, Imagenet: A large-scale hierarchical image database, 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009. (<https://doi.org/10.1109/CVPR.2009.5206848>).
- [39] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, 2015.
- [40] M. Drozdal, E. Vorontsov, G. Chartrand, S. Kadoury, C. Pal, The importance of skip connections in biomedical image segmentation, *Deep learning and data labeling for medical applications*, Springer, Cham, 2016, pp. 179–187, https://doi.org/10.1007/978-3-319-46976-8_19.
- [41] U. Menonm, A.G. Maharaj, R. Hallett, A. Ryan, M. Burnell, A. Shama, S. Lewis, S. Davies, S. Philpott, A. Lopes, K. Godfrey, D. Oram, J. Herod, K. Williamson, M. W. Seif, I. Scott, T. Would, R. Woolas, J. Murdoch, S. Dobbs, N.N. Amso, S. Leeson, D. Cruickshank, L. Fallowfield, N. Singh, A. Dawney, S.J. Skates, M. Parmar, I. Jacobs, Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), *Lancet Oncol.* 10.4 (2009) 327–340, [https://doi.org/10.1016/S1470-2045\(09\)70026-9](https://doi.org/10.1016/S1470-2045(09)70026-9).
- [42] D. Timmerman, A.C. Testa, T. Bourne, L. Aमेyе, D. Jurkovic, C.V. Holsbeke, D. Paladini, B.V. Calster, I. Vergote, S.V. Huffel, L. Valentin, Simple ultrasound-based rules for the diagnosis of ovarian cancer, *Ultrasound Obstet. Gynecol.* 31.6 (2008) 681–690, <https://doi.org/10.1002/uog.5365>.
- [43] R. Kamal, S. Hamed, S. Mansour, Y. Mounir, S.A. Sallam, Ovarian cancer screening—ultrasound; impact on ovarian cancer mortality, *Br. J. Radiol.* 91.1090 (2018) 20170571, <https://doi.org/10.1259/bjr.20170571>.
- [44] A. Aguirre, Y. Ardeshirpour, M.M. Sanders, M. Brewer, Q. Zhu, Potential role of coregistered photoacoustic and ultrasound imaging in ovarian cancer detection and characterization, *Transl. Oncol.* 4 (1) (2011) 29–37, <https://doi.org/10.1593/tlo.10187>.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [46] H. Liu, D.A. Boas, Y. Zhang, A.G. Yodh, B. Chance, Determination of optical properties and blood oxygenation in tissue using continuous NIR light, *Phys. Med. Biol.* 40.11 (1995) 1983, <https://doi.org/10.1088/0031-9155/40/11/015>.
- [47] S. Nandy, A. Mostafa, I.S. Hagemann, M.A. Powell, E. Amidi, K. Robinson, D. G. Mutch, C. Siegel, Q. Zhu, Evaluation of ovarian cancer: initial application of coregistered photoacoustic tomography and US, *Radiology* 289 (3) (2018) 740–747, <https://doi.org/10.1148/radiol.2018180666>.
- [48] H.S. Salehi, T. Wang, P.D. Kumavor, H. Li, Q. Zhu, Design of miniaturized illumination for transvaginal coregistered photoacoustic and ultrasound imaging, *Biomed. Opt. Express* 5 (9) (2014) 3074–3079, <https://doi.org/10.1364/BOE.5.003074>.
- [49] G. Yang, E. Amidi, W. Champman, S. Nandy, A. Mostafa, H. Abdelal, Z. Alipour, D. Chatterjee, M. Mytch, Q. Zhu, Co-registered photoacoustic and ultrasound real-time imaging of colorectal cancer: ex-vivo studies, *International Society for Optics and Photonics*, 2019, *Photons Ultrasound: Imaging Sens.* Vol. 10878 (2019), <https://doi.org/10.1117/12.2507638>.
- [50] S. Yan, and Q. Fang, Hybrid mesh and voxel based Monte Carlo algorithm for accurate and efficient photon transport modeling in complex bio-tissues, *Biomed. Opt. Express*, 11(11) pp. 6262–6270. (<https://doi.org/10.1364/BOE.409468>).
- [51] 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, <https://doi.org/10.1117/1.3360308>.
- [52] M. Klemm, L.J. Craddock, J.A. Leendertz, A. Preece, R. Benjamin, Improved delay-and-sum beamforming algorithm for breast cancer detection, *Int. J. Antennas Propag.* 2008 (2008), <https://doi.org/10.1155/2008/761402>.
- [53] R.F. Andreotti, D. Timmerman, L.M. Strachowski, W. Froyman, B.R. Benacerragf, G.L. Bennett, T. Bourne, D.L. Brown, B.G. Coleman, M.C. Frates, S.R. Goldstein, U. M. Hamper, M.M. Horrow, M.H. Schulman, C. Reinhold, S.L. Rose, B.P. Whitcomb, W.L. Wolfman, P. Glanc, O-RADS US risk stratification and management system: a consensus guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee, *Radiology* 294.1 (2020) 168–185, <https://doi.org/10.1148/radiol.2019191150>.
- [54] M. Shetty, Imaging and differential diagnosis of ovarian cancer, *Seminars in Ultrasound, CT and MRI*, Vol. 40, WB Saunders, 2019, <https://doi.org/10.1053/j.sult.2019.04.002>.
- [55] Z. Guo, L. Li, L.V. Wang, On the speckle-free nature of photoacoustic tomography, *Med Phys.* 36 (9) (2009) 4084–4088, <https://doi.org/10.1118/1.3187231>.
- [56] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, Grad-cam: Visual explanations from deep networks via gradient-based localization, *Proceedings of the IEEE international conference on computer vision*. 2017.
- [57] L. Chen, J. Chen, H. Hajimirsadeghi, and G. Mori, Adapting grad-cam for embedding networks, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020.
- [58] S. Nandy, A. Mostafa, P.D. Kumavor, M. Sanders, M. Brewer, Q. Zhu, Characterizing optical properties and spatial heterogeneity of human ovarian tissue using spatial frequency domain imaging, *J. Biomed. Opt.* 21 (10) (2016), 101402.
- [59] A. Roggan, M. Friebel, K. Dörschel, A. Hahn, A. G.J. Mueller, Optical properties of circulating human blood in the wavelength range 400–2500 nm, *J. Biomed. Opt.* 4 (1) (1999) 36–46.
- [60] N. Bosschaert, G.J. Edelman, M.C. Aalders, T.G. van Leeuwen, D.J. Faber, A literature review and novel theoretical approach on the optical properties of whole blood, *Lasers Med. Sci.* 29 (2) (2014) 453–479.



Yun Zou received B.S. and M.S. in Engineering Physics from Tsinghua University. He is a PhD candidate in Biomedical Engineering Department of Washington University in St. Louis. His research interests are diffuse optical tomography and photoacoustic tomography with novel machine learning methods. He is also interested in classification and detection using Ultrasound images for cancer diagnosis.



Eghbal Amidi received his BS and MS in Electrical engineering. In 2021, He graduated with his Ph.D. in Biomedical Engineering from Washington University in St. Louis. He is currently working as a Data Scientist at Caris Life Sciences, Irving, TX. His research interests include computer vision, image processing, and machine learning in radiology and histology images.



Quing Zhu is the Edwin H. Murty Professor of Biomedical Engineering, Washington University in St. Louis. She is also an associate faculty in Radiology, Washington University in St. Louis. Professor Zhu has been named a Fellow of OSA, a Fellow of SPIE, a Fellow of AIMBE. Her research interests include multi-modality ultrasound, diffuse light, photoacoustic imaging, and Optical Coherence Tomography for breast, ovarian, colorectal cancer applications.



Hongbo Luo received his B.S. in Opto-Information Science and Technology from Huazhong University of Science and Technology, China and M.S. in Optical Engineering from the same university. He is currently pursuing his PhD in Electrical and System Engineering at Washington University in St. Louis, USA. His research interest includes Optical Coherence Tomography and Photoacoustic imaging.