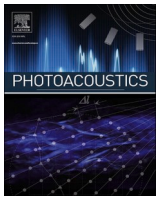
Photoacoustics 40 (2024) 100645

Contents lists available at ScienceDirect

Photoacoustics

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



Joint segmentation and image reconstruction with error prediction in photoacoustic imaging using deep learning

Ruibo Shang [a], Geoffrey P. Luke [b], Matthew O'Donnell [a,*]

[a] uWAMIT Center, Department of Bioengineering, University of Washington, Seattle, WA 98195, USA
[b] Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA

ARTICLE INFO

Keywords:
Photoacoustic Imaging
Deep Learning
Error Prediction
Segmentation
Image Reconstruction
Validation
Quantitation

ABSTRACT

Deep learning has been used to improve photoacoustic (PA) image reconstruction. One major challenge is that errors cannot be quantified to validate predictions when ground truth is unknown. Validation is key to quantitative applications, especially using limited-bandwidth ultrasonic linear detector arrays. Here, we propose a hybrid Bayesian convolutional neural network (Hybrid-BCNN) to jointly predict PA image and segmentation with error (uncertainty) predictions. Each output pixel represents a probability distribution where error can be quantified. The Hybrid-BCNN was trained with simulated PA data and applied to both simulations and experiments. Due to the sparsity of PA images, segmentation focuses Hybrid-BCNN on minimizing the loss function in regions with PA signals for better predictions. The results show that accurate PA segmentations and images are obtained, and error predictions are highly statistically correlated to actual errors. To leverage error predictions, confidence processing created PA images above a specific confidence level.

1. Introduction

A key step in quantitative photoacoustics (PA) is accurately reconstructing the initial pressure distribution (IPD) throughout the imaging volume from data collected on an ultrasound array transducer [1,2]. With this IPD and the estimated fluence at each pixel, the optical absorption coefficient can be imaged [3–5]. Optical absorption images at different wavelengths can then be combined for spectroscopic imaging, bringing molecular profiling to deep structures within the body [6–8]. By combining quantitative spectroscopic PA imaging with real-time ultrasound (US) in a handheld probe, true molecular sensitivity can be added to clinical US (i.e., real-time PAUS imaging) [9]. Therefore, improving IPD reconstruction is key to quantitative PA. Unfortunately, conventional reconstruction methods produce poor IPD estimates, especially when conventional limited-bandwidth, limited-view handheld US probes are used for PA data collection [10,11]. Here we focus on reconstruction methods to improve PA image quality for this geometry as a key step in delivering quantitative PA methods to the clinic.

Deep learning (DL) can significantly enhance the state-of-the-art in image reconstruction compared with conventional algorithms [12–20]. In particular, deep neural networks trained on large image datasets can learn to reconstruct images by optimizing weights in each layer from gradient descent [13]. Many recent studies have demonstrated that DL methods can greatly improve PA image quality [10,21–32]. Generally, the accuracy of DL-reconstructed PA images can be quantified by comparing predicted with ground-truth images (e.g., calculating the peak signal-to-noise ratio (PSNR) [33,34]). However, a major limitation to date has been the absence of validation tools in practical applications where ground-truth images are unknown. In particular, errors in DL reconstructions from conventional convolutional neural networks (CNNs) cannot be quantitatively estimated.

Researchers have developed deep-learning-based error (uncertainty) prediction methods for PA imaging. A CNN with a loss function based on the mean squared error was developed to predict different outputs for the same input using Monte Carlo dropout to estimate the uncertainty of the DL process [35]. However, it only estimated model uncertainty [17,36], and data uncertainty [17,36] was not considered. A different DL approach was developed to estimate errors in computed optical absorption coefficients for quantitative PA imaging [37]. An error-estimation neural network was trained with computed initial pressure images as network inputs and relative errors of the estimated optical absorption coefficients as outputs. In testing, this network predicts errors in the testing dataset or real experiments. However, error estimation is not based on a more accurate statistical approach; that is,

* Corresponding author.
E-mail address: odonnel@uw.edu (M. O'Donnell).

https://doi.org/10.1016/j.pacs.2024.100645
Received 30 June 2024; Received in revised form 16 August 2024; Accepted 10 September 2024
Available online 11 September 2024

the error-estimation neural network only predicts the specific error instead of the statistical variance. In addition, to estimate errors in computed initial pressure images (network inputs), model uncertainty should be calculated and included in the overall estimated error. Finally, the accuracy of the estimated error was not evaluated quantitatively.

A potentially more accurate way to validate PA images reconstructed from DL networks is to use a Bayesian convolutional neural network (BCNN) [38–42] with a probability-distributed likelihood function as the loss function. Compared to conventional CNNs, a BCNN is an effective approach to approximate the uncertainty (pixel-wise confidence level of the image reconstruction) without knowing the ground truth. It assumes that each image pixel represents the parameter defining a probability distribution (e.g., Laplacian or Bernoulli distribution), rather than a deterministic intensity value [39]. Then, model error (model uncertainty) can be quantified by Monte Carlo dropout [43] or Deep Ensembles [44], and data error (data uncertainty) can be quantified by the variance in the predictions for given data. Since the parameters of the probability distribution for each image pixel are completely predicted by the BCNN, both data and model uncertainties are estimated, and the overall uncertainty (error) is the combination of the two. It has been applied to different imaging tasks including image segmentation [42], phase imaging [17], single-pixel imaging [36], optical metrology [45], and image classification [46].

Here, we propose a hybrid BCNN (Hybrid-BCNN) providing joint PA segmentation (Bernoulli distribution) and image (Laplacian distribution) reconstructions with error prediction (i.e., overall uncertainty combining data and model uncertainties). Segmentation here denotes structural segmentation excluding artifacts and background noise. In addition to identifying PA signals, segmentation focuses the Hybrid-BCNN on minimizing the loss function only in regions of PA signals since images from a typical PAUS architecture are usually sparse, and a single probability distribution is not efficient to describe both background and PA (non-background) pixels. The proposed Hybrid-BCNN was trained on simulated PA data and makes predictions on both simulated and experimental data. In addition to predicting PA segmentation and image, both data and model uncertainty can be computed statistically with Monte Carlo dropout to completely estimate the error (i.e., overall uncertainty) in predicted outputs. The proposed Hybrid-BCNN was compared to another BCNN without segmentation to verify its importance and the inefficiency of a single probability distribution. As demonstrated below, the proposed Hybrid-BCNN reconstructs the PA segmentation and image with high accuracy; the segmentation helps BCNN train efficiently; and predicted PA segmentation and image errors (uncertainties) are highly statistically correlated to actual reconstruction errors from the two proposed statistical methods, indicating that it is a promising tool to make and validate DL reconstructions.

To demonstrate one potential way to leverage uncertainty predictions, confidence processing is proposed by computing the relative uncertainty to improve PA image quality (e.g., artifact removal) for both simulations and measurements. Although the proposed Hybrid-BCNN is only applied here to predictions of the IPD and its error, it can also be used in other areas of quantitative PA imaging (e.g. spectroscopic imaging).

2. Materials and methods

2.1. BCNN theory

Rather than simple weights, BCNNs use distributions over network parameters and the training dataset [39]. That is, BCNNs assume stochastic rather than deterministic network processes (e.g., dropout [47], weight initialization [48] etc.). Denote the training dataset as $(\boldsymbol{X}, \boldsymbol{Y}) = \{\boldsymbol{x}_n, \boldsymbol{y}_n\}_{n=1}^{N}$ with $\boldsymbol{X}$ and $\boldsymbol{Y}$ representing network inputs and ground-truth images, respectively. N is the total number of training images. To approximate prediction variability in y given a specific input $\boldsymbol{x}_{test,t}$ within $(\boldsymbol{X}_{test}, \boldsymbol{Y}_{test}) = \{\boldsymbol{x}_{test,\ t},\ \boldsymbol{y}_{test,\ t}\}_{t=1}^{T}$ (T is the total number of images in the testing dataset), the predictive distribution $p(\boldsymbol{y}|\boldsymbol{x}_{test,\ t},\ \boldsymbol{X}, \boldsymbol{Y})$ over all possible learned weights (with marginalization) [43] is used:

$$p(\boldsymbol{y}, |, \boldsymbol{x}_{test,t},\ \boldsymbol{X}, \boldsymbol{Y}) = \int p(\boldsymbol{y}|\boldsymbol{x}_{test,t}, \boldsymbol{W})\ p(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y}) d\boldsymbol{W} \tag{1}$$

where $p(\boldsymbol{y}|\boldsymbol{x}_{test,t}, \boldsymbol{W})$ denotes the predictive distribution including all possible output predictions given the learned weights $\boldsymbol{W}$ and the input $\boldsymbol{x}_{test,t}$. It represents data uncertainty [17]. $p(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y})$ denotes all possible learned weights given the training dataset, representing model uncertainty [17].

To jointly reconstruct the PA segmentation and image with uncertainty quantification, we choose the joint multivariate Bernoulli-distributed (for PA segmentation) and Laplacian-distributed (for PA image) likelihood functions to model data uncertainty, creating a Hybrid-BCNN. The specific choice of a Laplacian distribution for the image is discussed in the Supplementary Document. In this joint distribution, the Laplacian distribution is only assigned to the segmentation region, where the segmentation value is 1 under the Bernoulli distribution (detailed in Eqs. S6-S8 in the Supplementary Document).

The loss function $L_{Hybrid}(\boldsymbol{W}|\boldsymbol{x}, \boldsymbol{y})$ ($\boldsymbol{W}$ combines network weights $\boldsymbol{W}_1$ for PA segmentation and network weights $\boldsymbol{W}_2$ for PA image) for the joint distributed likelihood function given the training data pair $(\boldsymbol{x}_n,\ \boldsymbol{y}_n)$, where $\boldsymbol{y}_n = (\boldsymbol{y}_{seg,n},\ \boldsymbol{y}_{image,n})$, is,

$$L_{Hybrid}(\boldsymbol{W}|\boldsymbol{x}_n, \boldsymbol{y}_n) = \sum_{m=1}^{M} \left[\left(y_{seg,n}^{m} - 1\right) \log\ \left(1 - \mu_1^m\right) - y_{seg,n}^{m} \log\left(\mu_1^m\right) + y_{seg,n}^{m} \left(\frac{\left| y_{image,n}^{m} - \mu_2^m \right|}{\sigma^m} + \log\ \left(2\sigma^m\right) \right) \right] \tag{2}$$

where $\boldsymbol{y}_{seg,n}$ and $\boldsymbol{y}_{image,n}$ are the nth ground-truth PA segmentation and image in the training dataset, respectively. $y_{seg,n}^{m}$ and $y_{image,n}^{m}$ are the mth pixel of $\boldsymbol{y}_{seg,n}$ and $\boldsymbol{y}_{image,n}$, respectively. μ_1^m is the probability of $\left(y_{seg,n}^{m} = 1 \middle| \boldsymbol{x}_n, \boldsymbol{W}\right)$, which is also the mean of the Bernoulli-function for $y_{seg,n}^{m}$. μ_2^m and σ^m are the mean and standard deviation of the Laplacian-distributed likelihood function for $y_{image,n}^{m}$. M is the total pixel number in $\boldsymbol{y}_{seg,n}$ or $\boldsymbol{y}_{image,n}$. Complete derivations of Eq. (2) are presented in Supplementary Section 1.

The loss function in Eq. (2) is minimized during training. This Hybrid-BCNN has three output channels, where one ($\boldsymbol{\mu}_1$) is for the Bernoulli-distributed likelihood function and two ($\boldsymbol{\mu}_2$ and $\boldsymbol{\sigma}$) are for the Laplacian-distributed likelihood function.

Model uncertainty is measured with a dropout network [43]. A distribution $q(\boldsymbol{W})$ (defined in ref. [43]) is learned to approximate $p(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y})$ (minimizing the Kullback-Leibler divergence between $q(\boldsymbol{W})$ and $p(\boldsymbol{W}|\boldsymbol{X}, \boldsymbol{Y})$) by applying a dropout layer before every layer with learnable weights. During the prediction process, model uncertainty is approximated by Monte Carlo dropout [43]. With Monte Carlo integration, the predictive distribution $p(\boldsymbol{y}|\boldsymbol{x}_{test,\ t},\ \boldsymbol{X}, \boldsymbol{Y})$ in Eq. (1) can be approximated as:

$$p(\boldsymbol{y}, |, \boldsymbol{x}_{test,\ t},\ \boldsymbol{X}, \boldsymbol{Y}) \approx \frac{1}{K} \sum_{k=1}^{K} p(\boldsymbol{y}|\boldsymbol{x}_{test,t},\ \boldsymbol{W}^k) \tag{3}$$

where K is the total number of dropout activations during prediction. In this study K=16, the same as that used in previous publications [17,36].

Finally, the reconstructed PA segmentation and image are represented by the predicted mean $\widehat{\mu}_{test,t}^{m}$(the mth pixel) given test input $\boldsymbol{x}_{test,t}$ under Bernoulli-distributed and Laplacian-distributed likelihood functions, respectively:

$$\widehat{\mu}_{test,t}^{m} = \frac{1}{K} \sum_{k=1}^{K} \widehat{\mu}_{test,t}^{m,k} \tag{4}$$

where μ^m denotes μ_1^m or μ_2^m, and $\widehat{\mu}_{test,t}^{m,k}$denotes the mth pixel of the predicted $\boldsymbol{\mu}_1$ (for PA segmentation) or $\boldsymbol{\mu}_2$ (for PA images) from the kth dropout activation given test input $\boldsymbol{x}_{test,t}$ (i.e., $\widehat{\mu}_{1test,t}^{m,k}$ or $\widehat{\mu}_{2test,t}^{m,k}$).

The corresponding predicted uncertainties (the mth pixel) $\widehat{\sigma}_{test,t(Ber)}^{m}$ (for PA segmentation under Bernoulli function) and $\widehat{\sigma}_{test,t(Lap)}^{m}$ (for PA image under Laplacian function) given the test input data $\boldsymbol{x}_{test,t}$ are,

$$\widehat{\sigma}_{test,t(Ber)}^{m} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}[\widehat{\mu}_{1test,t}^{m,k}(1-\widehat{\mu}_{1test,t}^{m,k})]+\frac{1}{K}\sum_{k=1}^{K}\left(\widehat{\mu}_{1test,t}^{m,k}-\widehat{\mu}_{1test,t}^{m}\right)^2} \tag{5}$$

$$\widehat{\sigma}_{test,t(Lap)}^{m} = \sqrt{\frac{1}{K}\sum_{k=1}^{K}2\left(\widehat{\sigma}_{test,t}^{m,k}\right)^2+\frac{1}{K}\sum_{k=1}^{K}\left(\widehat{\mu}_{2test,t}^{m,k}-\widehat{\mu}_{2test,t}^{m}\right)^2} \tag{6}$$

where $\widehat{\sigma}_{test,t}^{m,k}$ denotes the predicted standard deviation of y_{image}^{m} from the kth dropout activation for test data $\boldsymbol{x}_{test,t}$. Complete derivations of Eqs. (4–6) are presented in Supplementary Section 2.

The reconstructed PA image and its corresponding uncertainty are multiplied with the reconstructed PA segmentation since the Hybrid-BCNN only minimizes the loss function within non-background pixels (given by PA segmentation) and, therefore, the background and its uncertainty are not optimized. In this case, the accuracy (error) of PA segmentation and image predictions can be validated by the predicted uncertainties.

To show the power of the hybrid approach and to prove that a single distribution likelihood function (i.e., Laplacian) is not efficient to describe both background pixels and PA (non-background) pixels, we compared the Hybrid-BCNN to another BCNN with only the Laplacian-distributed likelihood function where there is no PA segmentation (Lap-BCNN). By taking logarithm and negative operations on Eq. (S4) in the Supplementary Document, the loss function $L_{Lap}(\boldsymbol{W}_{Lap}|\boldsymbol{x},\boldsymbol{y})$ for Lap-BCNN given the training data pair $(\boldsymbol{x}_n, \boldsymbol{y}_n)$ where $\boldsymbol{y}_n = \boldsymbol{y}_{image,n}$ is,

$$L_{Lap}(\boldsymbol{W}_{Lap}|\boldsymbol{x}_n,\boldsymbol{y}_n) = \sum_{m=1}^{M}\left[\frac{\left|y_{image,n}^{m}-\mu_2^m\right|}{\sigma^m}+\log(2\sigma^m)\right] \tag{7}$$

where $\boldsymbol{W}_{Lap}$ denotes the network weights of Lap-BCNN and the other variables have the same meanings as those in Eq. (2), and Eqs. (S4) and (S5) in the Supplementary Document.

The predicted PA image and uncertainty from the Lap-BCNN are the same as in Eqs. (4) and (6).

An uncertainty assessment metric [17,49–51] is used to quantify the accuracy of uncertainty predictions in the testing dataset by computing the reliability diagram (*credibility* (***Cred***) vs *empirical accuracy* (***ACC***)). Details of computing this diagram are in Supplementary Section 3. The linear correlation coefficient (CC) between ***Cred*** and ***ACC***, and the slope of the corresponding linear fit, are calculated to quantify the diagonality of the reliability diagram. In this paper, the bound of the credible interval ϵ equals $0.2\widehat{\mu}_{test,t}^{m}$ to provide sufficient sample points from the discrete probability bins to appropriately plot the reliability diagram and evaluate its diagonality. To further test whether predicted uncertainties are related to true reconstruction errors, absolute reconstruction error is plotted versus 2× predicted uncertainty (i.e., 2× standard deviation). Theoretically, the absolute errors of ~95 % of the plotted points should be less than or equal to their corresponding 2× predicted uncertainties.

2.2. BCNN structure, parameters, and PA dataset simulation and preprocessing

The Hybrid-BCNN structure implementing the computational approach presented in the last section is shown in Fig. 1, where a U-Net architecture [52] with an encoder-decoder structure with skip connections between contracting paths and expanding paths is used. The contracting path captures context and the expanding path enables precise localization [52]. Similar to the BCNN in a previous publication [36], dropout layers with a dropout rate of 0.1 appear before each convolution layer to prevent overfitting during training, and L_2 kernel regularizers and bias regularizers with a regularization factor of 1×10^{-6} were included in each convolution layer. Batch normalization (Batchnorm) [53] layers appear after each convolution layer to stabilize the network and make it converge better during training, and LeakyRelu [54] is used as the activation function. For Lap-BCNN, the same architecture is used except that there are two output channels (the predicted PA image and its uncertainty) for only the Laplacian-distributed likelihood function.

In a previous publication on DL-based real-time integrated PAUS imaging [32], two preprocessing approaches were applied to data directly acquired from the transducer. One is delay-and-sum (DAS) beamforming and the other is the multiple-channel transformed array (MC), in which propagation delays were applied between all observation points in the image and each transducer element without summation across array elements [32]. The MC approach has multiple channels, where the channel number equals the transducer element number, preserving more information embedded in the original data [32]. Previous results have shown that DL predictions with MC as network inputs outperform those with DAS inputs [32]. Therefore, only

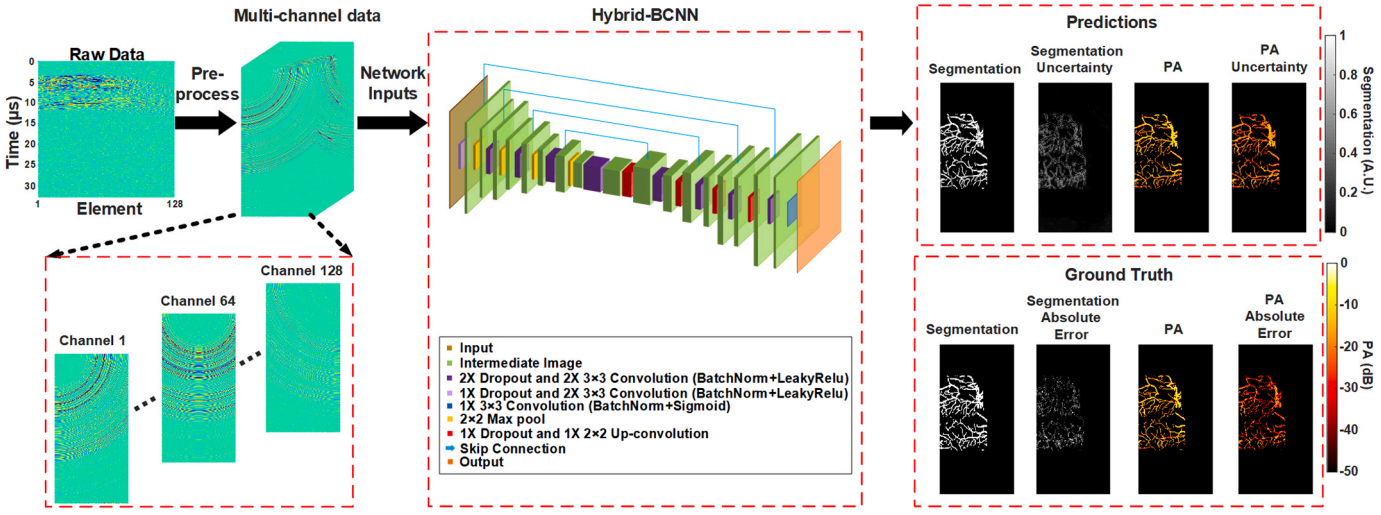


Fig. 1. The Hybrid-BCNN structure with multi-channel data (preprocessed from transducer array data) as network inputs and network predictions.

MC-preprocessed data were used as network inputs in this work.

We trained the network by generating simulated data mimicking typical microvessel images with the method described in a previous publication [32]. The microvessel images were obtained from the fundus oculi drive [55], which are representative of typical microvascular networks expected in many clinical applications. Only the binary images (manually extracted images) [32] were used, where white pixels denote the blood vessel segmentation and black pixels denote background. These images were randomly re-sized, re-located, rotated, partitioned and combined for data augmentation, which makes their signal components widely distributed in both space and frequency domains. These binary images were also used as ground-truth segmentation images.

PA images were generated by modifying binary images to gray-scale images with a dynamic range of 20 dB and amplifying the gray-scale images with different values to obtain measurements for a signal-to-noise ratio (SNR) of 10–35 dB. The simulation parameters are shown in Table S1 in Supplementary Section 4. Note that these parameters closely mimic those for our experimental PAUS imaging system described below. A total of 6.85 % of the overall pixels in the dataset are non-background pixels (with PA signals). After shuffling the overall 16,000 images, 12,800 (80 %) were used for training, 1600 (10 %) were used for validation and the remaining 1600 (10 %) were used for testing (no replicated images among the three datasets).

A PA imaging forward model was created with these parameters to convert PA images to transducer array data. In this model, the forward solution from the Green's function is [56],

$$p(t,\boldsymbol{r}') = \frac{\beta}{4\pi C_p}\frac{\partial}{\partial t}\int \frac{s(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}'|}\delta\left(t-\frac{|\boldsymbol{r}-\boldsymbol{r}'|}{v_s}\right)d\boldsymbol{r} \tag{8}$$

where β is the thermal coefficient of volume expansion, C_p is the specific heat capacity at constant pressure, v_s is the speed of sound, $\boldsymbol{r}'$ is the detection position, and $s(\boldsymbol{r})$ is the spatial absorption function. Transducer array data recorded by the ith element (at the position of $\boldsymbol{r}_i'$) is,

$$y\left(t,\boldsymbol{r}_i'\right) = h(t) * \left[p_{direct}\left(t,\boldsymbol{r}_i'\right) + n(t,\boldsymbol{r}_i')\right] \tag{9}$$

where $h(t)$ is the system impulse response, $n(t,\boldsymbol{r}_i')$ is system noise, '$*$' is the temporal convolution operator, and $p_{direct}\left(t,\boldsymbol{r}_i'\right)$ is the pressure function $p(t,\boldsymbol{r}')$ with the directivity pattern.

$$p_{direct}\left(t,\boldsymbol{r}_i'\right) = p(t,\boldsymbol{r}') \times \frac{\sin\left(\frac{\pi l}{\lambda}\sin\theta_i\right)}{\frac{\pi l}{\lambda}\sin\theta_i}, \quad \theta_i = \tan^{-1}\left(\frac{x-x_i'}{z}\right) \tag{10}$$

where l is the transducer element pitch, λ is the ultrasound wavelength, and θ_i is the incident angle of a wave propagating from position $\boldsymbol{r} = (z,x)$ to the ith element of the transducer (at the position of $\boldsymbol{r}_i' = x_i'$). Then, a PA imaging backward model was created with the same simulation parameters to convert transducer array data to MC data as network inputs.

Table S1 in Supplementary Section 4 presents some training hyperparameters of both BCNNs. They were trained on a NVIDIA GeForce GTX 1080 Ti GPU with 11 GB of memory. In training, the Adam optimizer was used with a constant learning rate of 0.0005 with reference to BCNN training in a previous publication [36], batch size was chosen to be 8 given limitations on the GPU memory, and the maximum epoch number was chosen to be 1000 to guarantee complete training. An early stopping criterion was applied, where training would finish when the validation loss value did not decrease in 50 consecutive epochs (50 epochs were chosen to prevent training from stopping at incorrect places such as a plateau at the beginning of training, a sudden jump of the loss value during training, etc.), and optimal weights with the lowest validation loss would be retrieved retrospectively. The maximum training time was approximately 35 h.

2.3. Confidence processing

The strong statistical correlation between absolute errors and predicted uncertainties suggests that the predicted uncertainty can help determine the confidence in the reconstructed PA image. For example, consider a pixel with a high signal (segmentation or PA value) and low uncertainty such that the relative uncertainty (standard deviation normalized to the mean) is small. There is high confidence in the reconstructed value for this pixel. In contrast, for a pixel with a low mean and a high uncertainty (high relative uncertainty), there is low confidence in the reconstructed value. This information can help identify artifacts and improve image quality. In general, confidence may provide a tool to judge the accuracy of quantitative PA images.

Here we explore a straightforward processing method to enhance PA images using confidence in the reconstruction, where the confidence is simply related to the ratio of the standard deviation to the mean (SD/M) at each pixel output by the Hybrid-BCNN. High confidence corresponds to a low value of this ratio. As a simple example of how to use this information, we set a threshold on the SD/M to eliminate pixels with low confidence. This operation acts like the segmentation threshold but uses relative uncertainties of both the segmentation and PA image, not just the mean value of the segmentation. The threshold can be varied to present images at different confidence levels.

2.4. Experimental PAUS imaging system and phantom preparations

To initially test the performance of the proposed BCNN for practical applications, two simple phantom objects constructed from optically absorbing metal wires with known geometries were imaged using an integrated, experimental PAUS system [9,32] (as shown in Fig. 7 in [32]). The wires were twisted into the shape of the letters 'S' and 'W' and suspended in the x-z imaging plane in a cubical container. These simple shapes were used because ground-truth absorption profiles are reasonably well known. Note that data from the 'W' shaped wire was used in our previous publication of DL-based real-time integrated PAUS imaging [32]. It was reused here to test whether uncertainty predictions from the proposed Hybrid-BCNN can help with additional image interpretation.

The 'S' shaped wire is new; it is used here because conventional DAS images of this object exhibit some clear artifacts and the overall SNR was reduced from that of the original W-phantom (lower laser energy was used) and is more typical of the SNR for *in vivo* PA images. Consequently, it can test whether the proposed Hybrid-BCNN can find, mark and, ultimately, correct these artifacts at typical SNR levels. The container was filled with a 2 % intralipid solution (Fresenius Kabi, Deerfield, USA) acting as a scattering medium with effective attenuation coefficient of ~0.1 mm^{-1}. Finally, a human finger was scanned and imaged to evaluate the performance of Hybrid-BCNN for *in vivo* vascular imaging. This study was approved by the Institutional Review Board of the University of Washington (Study# 00009196) with safe optical and acoustic energies following ANSI (optical) and FDA (US) guidelines.

The experimental PAUS system employs a unique fiber optic delivery system where each laser pulse is sequentially delivered to one of 20 different fibers. At the probe front side, the fibers are equally distributed along the azimuthal axis of the US array transducer (LA 15/128–1633, Vermon S.A. France) with 10 fibers on each side. A mechanical rotation system sequentially directs the beam from a kHz-rate, wavelength-tunable (700–900 nm) diode-pumped laser (Laser Export, Russia) to a different one of the 20 fibers on each pulse, delivering laser energy to the sample at a PA frame rate of about 50 Hz. Interleaved with PA pulses were US pulses enabling simultaneous acquisition of PA and US image data at the 50 Hz frame rate (detailed in [9]). All data acquisition was controlled by a commercial US scanner (Vantage, Verasonics, WA, USA) using trigger signals created by the motor controller and encoder to ensure accurate synchronization between US and PA scan sequences.

For this study, PA data were acquired using a single wavelength at 795 nm. One PA raw data frame contains 2048 temporal samples × 128

elements × 20 fibers. Each data frame was averaged over 20 fibers to improve SNR. The size of each reconstructed PA image is 512 × 128 with an axial pixel size of 0.05 mm and lateral pixel size of 0.1 mm. Finally, PA segmentation, images and uncertainties were reconstructed from acquired experimental data using BCNNs trained with simulated data.

3. Results

3.1. Simulation results and quantitative evaluation

Hybrid-BCNN and Lap-BCNN results on two representative samples of simulated PA data are shown in Fig. 2. Ground-truth PA segmentations, images and conventional DAS results are shown in Fig. 2(a) and (e). The reconstructed PA segmentations and images together with their corresponding predicted uncertainties from Hybrid-BCNN are shown in Fig. 2(b) and (f), respectively. The PA images and their corresponding predicted uncertainties from Lap-BCNN are shown in Fig. 2(c) and (g), respectively. Separate predicted data and model uncertainties of the PA segmentation and image from both Hybrid-BCNN and Lap-BCNN are shown in Fig. S1 in Supplementary Section 5. Data uncertainty clearly dominates over model uncertainty in both Hybrid-BCNN and Lap-BCNN, indicating that the major error comes from input data imperfection.

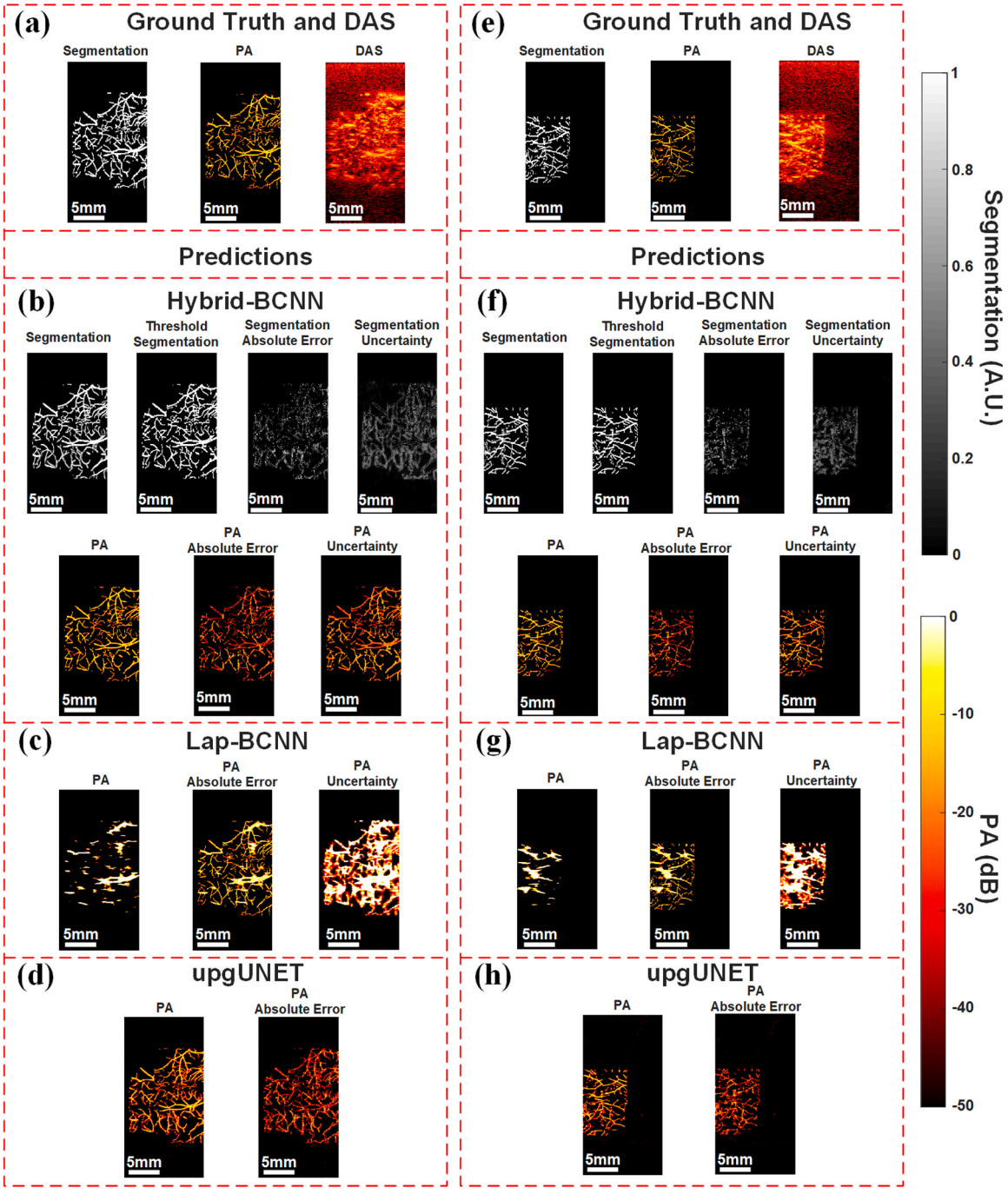


Fig. 2. Hybrid-BCNN and Lap-BCNN results for simulated PA data. (a) A representative ground-truth PA segmentation and image, and the corresponding DAS-reconstructed PA image. (b) Reconstructions and prediction errors from Hybrid-BCNN. (c) Reconstructions and prediction errors from Lap-BCNN. (d) Reconstructions from conventional CNN-based upgUNET (with the U-Net architecture) without error predictions. (e-h) are the same as (a-d) except for a different image sample. The scale bars denote 5 mm.

To compare the PA image prediction of Hybrid-BCNN with a conventional CNN tuned for PA imaging that does not provide uncertainty predictions, the results of upgUNET [32] based on the U-Net architecture are also shown in Fig. 2(d) and (h). The segmentation/PA absolute error in Fig. 2(b-d), (f-h) is the absolute difference between the reconstructed segmentation/PA and the ground-truth. PA reconstructions and uncertainties are presented on a log scale over a 50 dB scale relative to the peak signal of the images in the testing dataset. PA segmentations and predicted segmentation uncertainties are presented on a linear scale.

Fig. 2(b), (d), (f) and (h) show that Hybrid-BCNN has comparable PA image reconstruction as conventional CNN-based upgUNET with the U-Net architecture. The averaged PSNR was calculated within the testing dataset (28.9855 dB for Hybrid-BCNN, 28.4959 for upgUNET and 21.6148 dB for DAS) to quantitatively demonstrate that Hybrid-BCNN has comparable PA image reconstruction as upgUNET, and both outperform DAS. However, Hybrid-BCNN can also provide uncertainty predictions while upgUNET cannot. Fig. 2(c) and (g) clearly show that the Lap-BCNN makes less accurate PA image reconstructions and uncertainty predictions.

The averaged PSNR was calculated within the testing dataset for Lap-BCNN (18.5948 dB, compared to 28.9855 dB for Hybrid-BCNN) to quantitatively demonstrate that Lap-BCNN provides a less accurate PA image reconstruction. The reason is that PA images acquired with linear-array, limited-bandwidth ultrasound transducers are usually sparse (e.g., in this study, only 6.85 % of the overall pixels in the dataset have PA signals). Therefore, the Lap-BCNN learns from pixels where most of their values are 0 when minimizing the loss function for the Laplacian-distributed likelihood function in Eq. (7). In this case, the Lap-BCNN focuses on the majority 0-value pixels and tends to converge to a wrong solution no matter what the network input is. Clearly, the Laplacian-distributed likelihood function alone is not sufficient to describe both background pixels and pixels with PA signals.

Therefore, we propose that a combination of Bernoulli-distributed and Laplacian-distributed likelihood functions to jointly reconstruct the PA segmentation and image (i.e., Hybrid-BCNN) is more appropriate. Segmentation using a Bernoulli-distributed likelihood function can identify background pixels and pixels with PA signals, focusing the network on pixels with PA signals to minimize the value of the loss function for a Laplacian-distributed likelihood function driving PA image reconstructions. Within pixel regions with PA signals there is a range of signal amplitudes, making the network learn the correct Laplacian distribution. Although pixel values (background pixels vs PA pixels) are still unevenly distributed for segmentation, the Laplacian term $y^m_{seg,n}\left(\left|y^m_{image,n}-\mu^m_2\right|/\sigma^m+\log(2\sigma^m)\right)$ will balance the Bernoulli term $\left(y^m_{seg,n}-1\right)\log\left(1-\mu^m_1\right)-\quad y^m_{seg,n}\log\left(\mu^m_1\right)$ in Eq. (2) to learn the correct segmentation.

As shown in Fig. 2(b) and (f), both the PA segmentation and image, along with corresponding uncertainties, are reconstructed with Hybrid-BCNN. The final segmentation was computed by thresholding the predicted segmentation at the mid-range value of 0.5 (i.e., pixel values larger than 0.5 are rounded up to 1 and remaining pixel values are rounded down to 0). The low segmentation/PA absolute errors show that these predictions are very accurate. Uncertainty predictions by the Hybrid-BCNN of PA segmentation and image are highly statistically correlated to the segmentation/PA absolute errors, as quantitatively shown below.

The credibility map (Cred Map) and reliability diagram (ACC vs Cred) were computed for the results in Fig. 2 and presented in Fig. 3. They are highly statistically correlated to the PA absolute errors in Fig. 2 (b) and (f), where pixels with low PA absolute errors are marked with high credibility. By computing the CC and slopes of the linear fits, the reliability diagrams from the Hybrid-BCNN in Fig. 3 show high levels of diagonality. These results indicate that the predicted uncertainties from the Hybrid-BCNN match well with the absolute errors and can be a reliable tool to quantify reconstruction errors.

To further demonstrate the strong relationship between predicted uncertainties and reconstruction errors, absolute errors vs 2× predicted uncertainties (i.e., 2× standard deviation) results were computed for Fig. 2 and presented in Fig. 4. Fig. 4(a) and (b) correspond to the results in Fig. 2(b), and Fig. 4(c) and (d) correspond to the results in Fig. 2(f). Fig. 4(a) shows that 98.14 % of the overall pixels (within the predicted segmentation regions) have absolute errors less than or equal to 2× predicted uncertainties (i.e., 2× standard deviation) for this case. As shown in Fig. 4(b), within a specific range of 2× predicted uncertainties [95 % × ½ max(2×uncertainty), 105 % × ½ max(2×uncertainty)], 96.58 % of the pixels have absolute errors less than or equal to 2× predicted uncertainties. This indicates that the predicted PA image uncertainty in Fig. 2(b) can accurately quantify the errors in the reconstructed PA image. The same conclusion can be made from Fig. 4(c) and (d) for the results in Fig. 2(f).

Quantitative metrics (segmentation accuracy, segmentation structural similarity index (SSIM), PA image PSNR, segmentation CC, CC and slopes of the ACC vs Cred reliability diagrams) were computed for all samples in the testing dataset to evaluate the Hybrid-BCNN with Laplacian distribution, as shown in Table S2 in Supplementary Section 6 (average value and standard deviation (in the parentheses) for each metric). The results show that, quantitatively, the Hybrid-BCNN with Laplacian distribution can accurately reconstruct PA segmentations, images and corresponding uncertainties.

The selection of probability distribution functions was also studied, and the results are presented in Supplementary Section 6. For segmentation, the Bernoulli-distributed likelihood function is the only practical option. However, for PA image reconstruction within vessel regions, there are multiple options. Here, we compare Hybrid-BCNNs with Laplacian-distributed and Gaussian-distributed likelihood functions. The loss function for Hybrid-BCNN with the Gaussian-distributed likelihood function is derived in detail in Supplementary Section 6.

Fig. S2 in Supplementary Section 6 shows results for a representative sample in the testing dataset from Hybrid-BCNNs with Laplacian-distributed and Gaussian-distributed likelihood functions. They show

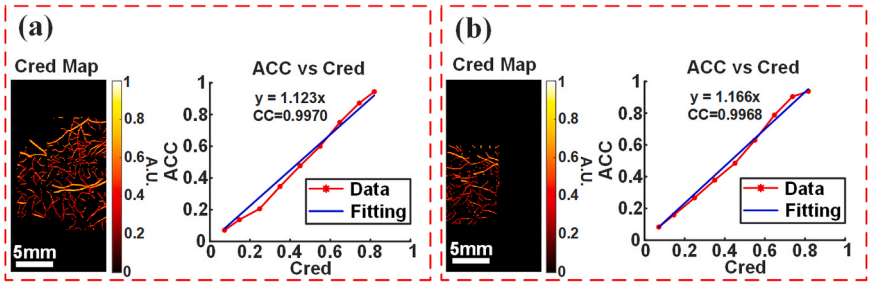


Fig. 3. Evaluation of Hybrid-BCNN PA image uncertainty predictions with the Cred Map (credibility) and ACC vs Cred (reliability diagram) calculation. (a) The Cred Map and ACC vs Cred (with CC and the slope of the linear fitting line) calculation in Hybrid-BCNN for the image sample in Fig. 2. (a-d). (b) is the same as (a) except that the results are for the image sample in Fig. 2(e-h). The scale bars denote 5 mm.

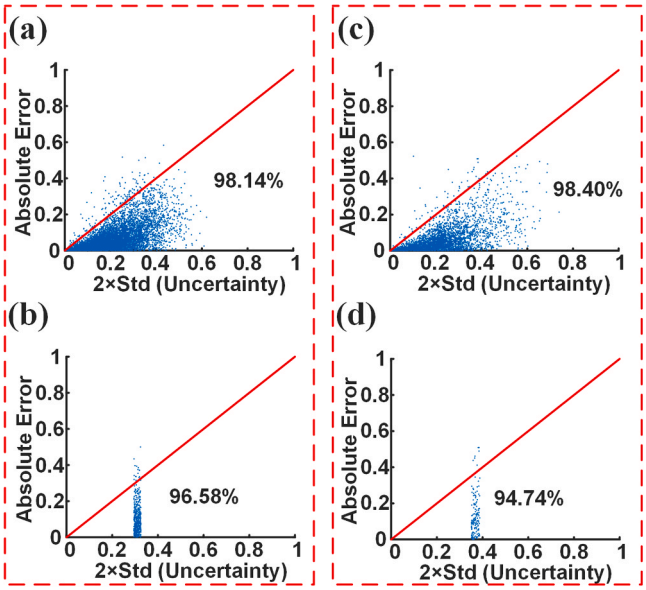


Fig. 4. Absolute Error vs 2× predicted uncertainty (i.e., 2× standard deviation (std)). (a) Absolute Error vs 2× predicted uncertainty for all pixels (within reconstructed segmentation regions) corresponding to the results in Fig. 2(b). (b) Absolute Error vs 2× predicted uncertainty for pixels within a specific range of 2× predicted uncertainties [95 % × ½ max(2×uncertainty), 105 % × ½ max (2×uncertainty)] from (a). (c) Absolute Error vs 2× predicted uncertainty for all pixels (within reconstructed segmentation regions) corresponding to the results in Fig. 2(f). (d) Absolute Error vs 2× predicted uncertainty for pixels within a specific range of 2× predicted uncertainties [95 % × ½ max (2×uncertainty), 105 % × ½ max(2×uncertainty)] from (c). Red lines indicate absolute error = 2×predicted uncertainty (std).

that the two Hybrid-BCNNs both accurately produce the PA segmentation and image, and the corresponding uncertainties, with similar performance. The uncertainty predictions of PA segmentation and image for the two Hybrid-BCNNs are both highly statistically correlated to the segmentation/PA absolute errors. The corresponding credibility maps (Cred Maps) and reliability diagrams (ACC vs Cred) are shown in Fig. S2 (c) and (f), and the corresponding absolute error vs 2× predicted uncertainty is shown in Fig. S2 (d) and (g), where the two Hybrid-BCNNs still exhibit similar performance.

Quantitative metrics (segmentation accuracy, segmentation SSIM, PA image PSNR, segmentation CC, CC and slopes of the ACC vs Cred reliability diagrams) were computed for all samples in the testing dataset to evaluate the two Hybrid-BCNNs, as shown in Table S2 in Supplementary Section 6 (average value and standard deviation (in the parentheses) for each metric). For segmentation uncertainties, since the ground-truth only has two values (0 and 1), CC is sufficient to evaluate the uncertainty accuracy. For PA segmentation, the high segmentation accuracy and CC in Table S2 show that the two Hybrid-BCNNs both make accurate and similar reconstructions of the PA segmentation and its uncertainty. For PA image reconstruction and its uncertainty, the two Hybrid-BCNNs still show accurate and similar performance.

These comparisons show that the Hybrid-BCNN is robust to the specific probability distribution function for PA image reconstruction and its uncertainty if the selected function is reasonable. The choice of the Laplacian-distributed likelihood function in this paper is because of its slightly better performance in simulations.

3.2. *Confidence processing results*

Fig. 5 shows confidence processing results (described in *Section 2.3*) using the image and results from Fig. 2(a) and (b). The segmentation and segmentation uncertainty results are copied from Fig. 2(b) and shown in Fig. 5(a) and (b). The relative segmentation uncertainty in Fig. 5(c) is generated by calculating the pixel-wise SD/M ratio using the results in Fig. 5(b) and (a). Note that the predicted segmentation is not exactly zero at all pixels with no PA signal due to the stochastic nature of the network, and in those pixels the SD/M ratio is quite variable. To eliminate all pixels from further consideration that clearly do not have a PA signal, a soft threshold (0.05 used here – results are not significantly different for any value less than 0.1) is first applied. Pixels with a predicted segmentation larger than the threshold are retained in the final relative segmentation uncertainty image (Fig. 5(c)).

The final confident segmentation presented in Fig. 5(d) was then generated by thresholding the segmentation based on the SD/M ratio (all pixels with an SD/M < 1 are retained for the example shown here) and setting the final value to one if the segmentation value is > 0.5 and to zero if the value is ≤ 0.5. This approach follows the procedure described in the last section but adds a level of confidence to the final

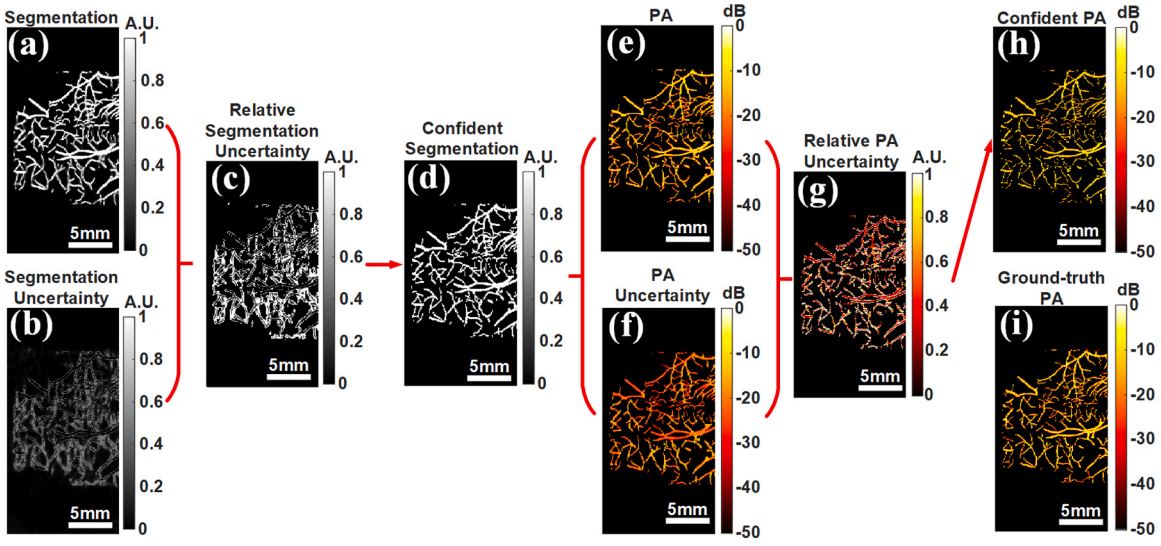


Fig. 5. Illustration of confidence processing. (a) Predicted segmentation. (b) Predicted segmentation uncertainty. (c) Relative segmentation uncertainty (SD/M) generated from (a, b). (d) Confident segmentation generated from (a, c) with a specific confidence level (i.e., relative segmentation uncertainty < 1). (e) Predicted PA image (multiplied with (d)). (f) Predicted PA uncertainty (multiplied with (d)). (g) Relative PA uncertainty (SD/M) generated from (e, f). (h) Confident PA image generated from (e, g) with a specific confidence level (i.e., relative PA uncertainty ≤ 0.9). (i) Ground-Truth PA image for reference. The scale bars denote 5 mm.

segmentation.

Similarly, the relative PA uncertainty in Fig. 5(g) is the pixel-wise SD/M ratio using Fig. 5(f) and (e) within the non-zero pixel regions in Fig. 5(d). Finally, the confident PA image in Fig. 5(h) is generated by eliminating (i.e., set to 0) pixels in Fig. 5(e) with corresponding SD/M above the threshold (a value of 0.9 is used here). The ground-truth PA image in Fig. 5(i) (copied from Fig. 2(a)) is compared with the confident PA image. This processing can eliminate obvious image artifacts but the final reconstructed image is a function of the selected threshold. Clearly, a user can vary the threshold in real-time to see images at different confidence levels to aid image interpretation. This approach is explored further using the experiments presented below.

3.3. Experimental results

Experimental results for the two wire phantom objects and the *in vivo*

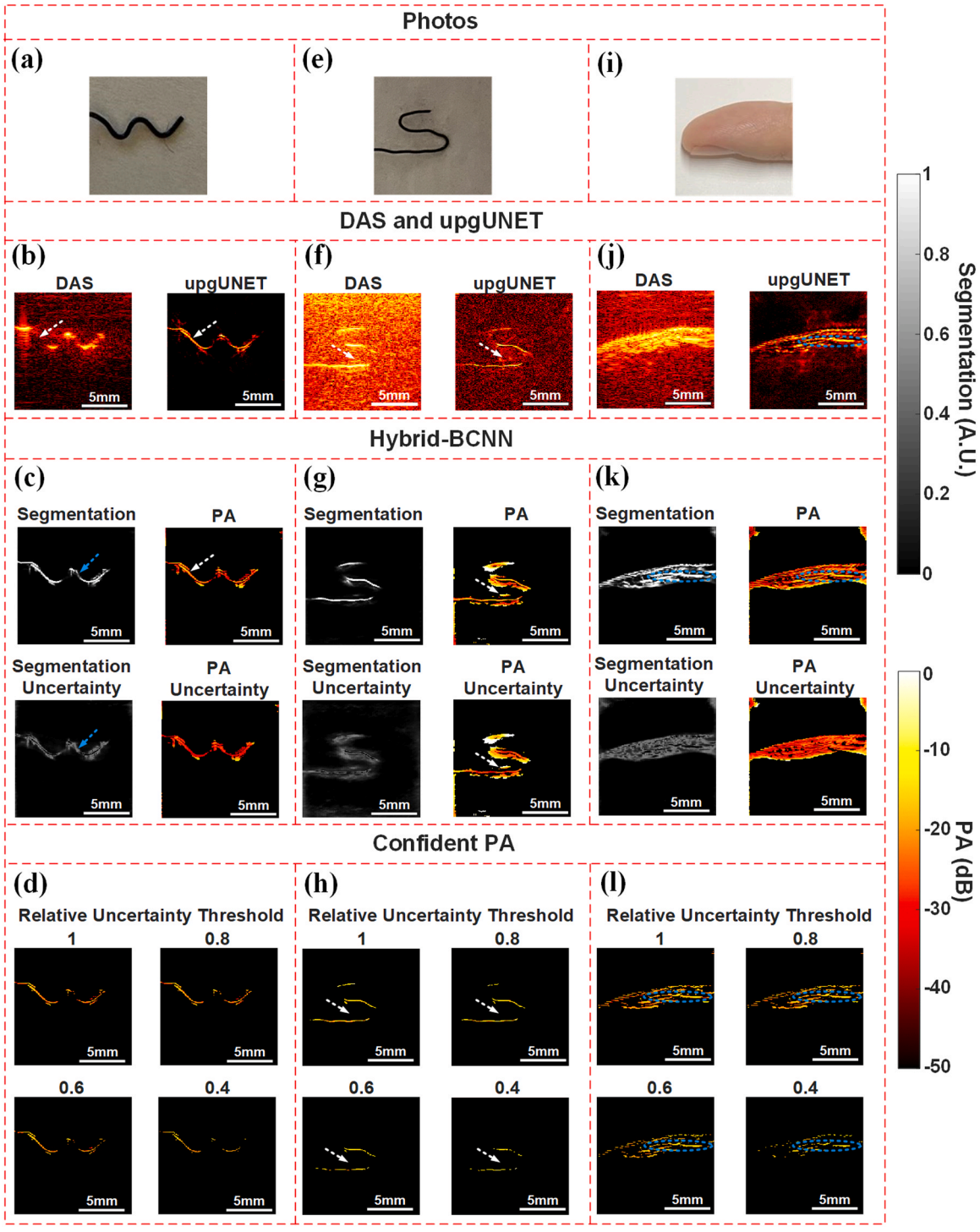


Fig. 6. Hybrid-BCNN results for experimental PA data ('S' and 'W' shaped objects, and the *in vivo* human finger). (a) Photo of the 'W' shaped object. (b) DAS and upgUNET reconstructions of the 'W' shaped object. (c) Results from Hybrid-BCNN for the 'W' shaped object. (d) Confident PA images with different confidence levels (i.e., relative uncertainty thresholds) for the 'W' shaped objects. (e-h) are the same as (a-d) except for the 'S' shaped object. (i-l) are the same as (a-d) except for the *in vivo* human finger. The scale bars denote 5 mm.

human finger described in *Section 2.4* are presented here. Fig. 6 shows photos, the conventional DAS reconstruction results, upgUNET results and Hybrid-BCNN results for these objects using exactly the same procedure as that used to produce the simulation results in Fig. 2 except that a soft segmentation threshold of 0.05 is first used in Hybrid-BCNN to generate the corresponding PA and PA uncertainty images. The Lap-BCNN results are not shown since, as demonstrated above, they provide less accurate PA images. PA reconstructions and uncertainties are presented on a 50 dB log scale relative to the peak signal in the image. PA segmentations and predicted segmentation uncertainties are presented on a linear scale.

Similar to the findings in the simulation, Hybrid-BCNN has comparable performance as upgUNET in PA image predictions, and both outperform the DAS approach, as demonstrated in Fig. 6(b) and (c) where the local structure of the W-shaped wire indicated by the white dashed arrows can be reconstructed by Hybrid-BCNN and upgUNET but is absent in the DAS image. However, the added uncertainty information from Hybrid-BCNN highlights most missing features while upgUNET cannot. For the human finger, both Hybrid-BCNN and upgUNET can reconstruct the microvessel indicated by the dashed ellipse in Fig. 6(j) and (k). Note that the upgUNET and DAS results for the 'S' shaped object and human finger have higher noise in the background due to lower SNR compared to the W-phantom. However, the segmentation in Hybrid-BCNN can distinguish vessel regions from the background, thus eliminating background noise.

Although the ground-truth is only approximately known from the photos (Fig. 6(a), (e) and (i)) for these objects, most obvious artifacts and missing features can be easily identified and compared to uncertainty predictions to test whether they can be used to identify these errors even when the ground truth is unknown (e.g., in clinical imaging). As presented in Fig. 6(c), the missing feature marked by the blue dashed arrow in the predicted PA segmentation is also marked with high uncertainties by the corresponding blue dashed arrow in the predicted PA segmentation uncertainty. The artifacts noted with the white dashed arrow in the predicted PA image shown in Fig. 6(g) are marked with high uncertainties by the white dashed arrow in the predicted PA image uncertainty. Separate data and model uncertainty results are shown in Fig. S3 in Supplementary Section 7. They clearly show that data uncertainty dominates model uncertainty, indicating that the major error comes from input data imperfection.

Fig. 6 also shows confidence processing results when ground truth is not precisely defined using the same approach as in Fig. 5. The relative segmentation uncertainty threshold is 1, as in the simulations. The confident PA images of the 'W' and 'S' shaped wire objects and human finger at different PA confidence levels (i.e., SD/M thresholds) are presented in Fig. 6(d), (h) and (l). As the confidence level increases (i.e., decreasing SD/M), more artifacts are removed, as expected. For the 'S' shaped object, in particular, artifacts in the original predicted PA image as shown by the white arrows in Fig. 6(g) are removed in Fig. 6(h) at all 4 confidence levels. For the human finger, the microvessel indicated by the dashed ellipse in Fig. 6(k) is mostly preserved in Fig. 6(l) at all 4 confidence levels. However, for all three objects, at very low levels of the SD/M threshold, useful image information is also removed. Therefore, there is a tradeoff between artifact removal and loss of real image features using this simple confidence processing approach. In general, a user can tune the threshold to see images at different confidence levels for different imaging tasks.

Overall, these experimental results suggest that the proposed Hybrid-BCNN has the potential to be a reliable tool to make comparable DL predictions as conventional CNNs (e.g., upgUNET) and, at the same time, at least partially validate DL reconstructions from predicted uncertainties when the ground-truth is unknown (e.g., DL-reconstructed clinical images). The proposed approach can help evaluate reconstruction confidence, identify artifacts and, potentially, remove them with simple processing. This can be a very useful tool to validate quantitative PA methods and provide a measure of confidence in quantitative results. Clearly, more sophisticated approaches can be used, but the relative uncertainties output by this network can help separate artifacts from image features.

4. Discussion

We presented predicted uncertainties to quantify errors in PA segmentations and images and used these predictions to generate confident PA images. However, we have not fully used these uncertainties to improve network predictions themselves. Therefore, our future work will focus on using predicted uncertainties as feedback to tune network structures or parameters and better select training datasets. We will also focus on applying the proposed BCNN to more *in vivo* PA studies to explore how uncertainty predictions can help clinical applications of DL-based quantitative PA imaging.

Although the proposed BCNN is used to predict the IPD and its error in this paper, it can also improve quantitative PA imaging since the error prediction of the IPD can help produce a more reliable reconstruction of optical absorption. As a further step, we can extend the BCNN approach to directly predict the optical absorption coefficient and its error, and, therefore, the confidence in the predicted absorption coefficient.

The simple confidence processing approach presented here leverages the strong correlation between predicted uncertainties and actual errors to improve PA image quality (e.g., artifact removal). Although it is very effective at removing artifacts, it is far from perfect because it can also eliminate real image features. A human observer can vary the confidence level of the PA image to tune the tradeoff between artifact and true feature removal for a given imaging task, but more sophisticated schemes should be explored to optimize this tradeoff. We also will investigate other ways to use confidence measures for different applications. For example, confidence processing can help select high confidence pixels in quantitative PA images of blood vessels to determine blood oxygenation levels. Given the typical artifacts in limited view and bandwidth imaging, high confidence pixels are much more likely to produce more accurate blood oxygenation values for real-time PAUS imaging with handheld US arrays.

Finally, the MC approach was used to preprocess measurement data to facilitate network training. We would also like to explore unprocessed measurement data as BCNN inputs. We expect that a much deeper neural network and a much larger training dataset are needed since this BCNN also must learn to map from measurement to image domains. However, the added information in unprocessed signals may lead to more robust images and uncertainty predictions.

5. Conclusions

In summary, we proposed a Hybrid-BCNN for limited-view-and-bandwidth PA imaging to jointly reconstruct PA segmentations (structural segmentation excluding artifacts and background noise) and images with uncertainty quantification by minimizing the loss function in Eq. (2), which jointly combines Bernoulli-distributed and Laplacian-distributed likelihood functions. Compared to conventional CNNs (e.g., upgUNET), BCNNs not only reconstruct images with comparable performance but also predict uncertainties in these images without knowing the ground truth. Due to the sparsity of PA images acquired in limited-view-and-bandwidth PA imaging (e.g., with conventional handheld linear-array transducers), a single probability distribution is not sufficient to describe both background and PA (non-background) pixels. Therefore, two probability distributions (Bernoulli and Laplacian) are used, and the segmentation focuses the network on minimizing the loss function in regions with PA signals. In addition, it can identify background regions and PA signal regions.

Simulation and experimental results show that the Hybrid-BCNN accurately reconstructs PA segmentations, images and corresponding uncertainties. Compared to the poor results from Lap-BCNN, which only reconstructs PA images without segmentation, the importance of

segmentation in Hybrid-BCNN is verified in simulations. Predicted uncertainties were also shown to strongly statistically correlate with actual reconstruction errors with two statistical methods. Thus, uncertainty predictions provide a potentially reliable tool to validate PA segmentations and images by marking incorrect reconstruction areas (e.g., missing features and artifacts) with high uncertainty. This can be very important in practical applications (e.g., clinical imaging) where the ground truth is unknown. Confidence processing is effective at removing artifacts to further improve PA image quality. With some adjustments, the proposed Hybrid-BCNN and confidence processing can be applied directly to quantitative PA imaging. In particular, confidence processing may be a powerful tool to accurately estimate blood oxygenation from spectroscopic PA data acquired with conventional handheld ultrasound array transducers.

CRediT authorship contribution statement

Geoffrey P. Luke: Writing – review & editing, Methodology, Conceptualization. **Matthew O'Donnell:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Ruibo Shang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

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.

Acknowledgements

This work was supported by the National Institutes of Health (R01EB030484). The authors acknowledge Professor MinWoo Kim from Pusan National University for providing simulated and experimental PA datasets, and Professor Ivan M. Pelivanov from the University of Washington for valuable discussions related to the results.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.pacs.2024.100645.

References

[1] A. Rosenthal, V. Ntziachristos, D. Razansky, Acoustic inversion in optoacoustic tomography: a review, Curr. Med. Imaging 9 (4) (2013) 318–336.
[2] A. Rosenthal, D. Razansky, V. Ntziachristos, Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography, IEEE Trans. Med. Imaging 29 (6) (2010) 1275–1285.
[3] L.V. Wang, S. Hu, Photoacoustic tomography: in vivo imaging from organelles to organs, Science 335 (6075) (2012) 1458–1462.
[4] X. Wang, Y. Pang, G. Ku, X. Xie, G. Stoica, L.V. Wang, Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain, Nat. Biotechnol. 21 (7) (2003) 803–806.
[5] M. Xu, L.V. Wang, Photoacoustic imaging in biomedicine, Rev. Sci. Instrum. 77 (4) (2006) 041101.
[6] B. Cox, J.G. Laufer, S.R. Arridge, P.C. Beard, Quantitative spectroscopic photoacoustic imaging: a review, J. Biomed. Opt. 17 (6) (2012), 061202-061202.
[7] P. Beard, Biomedical photoacoustic imaging, Interface Focus 1 (4) (2011) 602–631.
[8] J. Weber, P.C. Beard, S.E. Bohndiek, Contrast agents for molecular photoacoustic imaging, Nat. Methods 13 (8) (2016) 639–650.
[9] G.-S. Jeng, M.-L. Li, M. Kim, S.J. Yoon, J.J. Pitre Jr, D.S. Li, I. Pelivanov, M. O'Donnell, Real-time interleaved spectroscopic photoacoustic and ultrasound (PAUS) scanning with simultaneous fluence compensation and motion correction, Nat. Commun. 12 (1) (2021) 716.
[10] M. Kim, I. Pelivanov, M. O'Donnell, Review of deep learning approaches for interleaved photoacoustic and ultrasound (PAUS) imaging, IEEE Trans. Ultrason. Ferroelectr. Freq. Control (2023).
[11] M.K.A. Singh, W. Steenbergen, Photoacoustic-guided focused ultrasound (PAFUSion) for identifying reflection artifacts in photoacoustic imaging, Photoacoustics 3 (4) (2015) 123–131.
[12] G. Barbastathis, A. Ozcan, G. Situ, On the use of deep learning for computational imaging, Optica 6 (8) (2019) 921–943.
[13] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.
[14] Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, A. Ozcan, Deep learning microscopy, Optica 4 (11) (2017) 1437–1443.
[15] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, A. Ozcan, Phase recovery and holographic image reconstruction using deep learning in neural networks, Light. Sci. Appl. 7 (2) (2018) 17141.
[16] Y. Li, Y. Xue, L. Tian, Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media, Optica 5 (10) (2018) 1181–1190.
[17] Y. Xue, S. Cheng, Y. Li, L. Tian, Reliable deep-learning-based phase imaging with uncertainty quantification, Optica 6 (5) (2019) 618–629.
[18] A. Sinha, J. Lee, S. Li, G. Barbastathis, Lensless computational imaging through deep learning, Optica 4 (9) (2017) 1117–1125.
[19] A. Goy, K. Arthur, S. Li, G. Barbastathis, Low photon count phase retrieval using deep learning, Phys. Rev. Lett. 121 (24) (2018) 243902.
[20] A. Goy, G. Rughoobur, S. Li, K. Arthur, A.I. Akinwande, G. Barbastathis, High-resolution limited-angle phase tomography of dense layered objects using deep neural networks, Proc. Natl. Acad. Sci. 116 (40) (2019) 19848–19856.
[21] H. Deng, H. Qiao, Q. Dai, C. Ma, Deep learning in photoacoustic imaging: a review, J. Biomed. Opt. 26 (4) (2021), 040901-040901.
[22] J. Gröhl, M. Schellenberg, K. Dreher, L. Maier-Hein, Deep learning for biomedical photoacoustic imaging: A review, Photoacoustics 22 (2021) 100241.
[23] C. Yang, H. Lan, F. Gao, F. Gao, Review of deep learning for photoacoustic imaging, Photoacoustics 21 (2021) 100215.
[24] P. Rajendran, A. Sharma, M. Pramanik, Photoacoustic imaging aided with deep learning: a review, Biomed. Eng. Lett. (2022) 1–19.
[25] 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.
[26] H. Lan, D. Jiang, C. Yang, F. Gao, F. Gao, Y-Net: Hybrid deep learning image reconstruction for photoacoustic tomography in vivo, Photoacoustics 20 (2020) 100197.
[27] T. Vu, A. DiSpirito III, D. Li, Z. Wang, X. Zhu, M. Chen, L. Jiang, D. Zhang, J. Luo, Y. S. Zhang, Deep image prior for undersampling high-speed photoacoustic microscopy, Photoacoustics 22 (2021) 100266.
[28] M. Shi, T. Zhao, S.J. West, A.E. Desjardins, T. Vercauteren, W. Xia, Improving needle visibility in LED-based photoacoustic imaging using deep learning with semi-synthetic datasets, Photoacoustics 26 (2022) 100351.
[29] A. Hauptmann, B. Cox, Deep learning in photoacoustic tomography: current approaches and future directions, J. Biomed. Opt. 25 (11) (2020), 112903-112903.
[30] J. Gröhl, T.R. Else, L. Hacker, E.V. Bunce, P.W. Sweeney, S.E. Bohndiek, Moving beyond simulation: data-driven quantitative photoacoustic imaging using tissue-mimicking phantoms, IEEE Trans. Med. Imaging (2023).
[31] A. DiSpirito, D. Li, T. Vu, M. Chen, D. Zhang, J. Luo, R. Horstmeyer, J. Yao, Reconstructing undersampled photoacoustic microscopy images using deep learning, IEEE Trans. Med. Imaging 40 (2) (2020) 562–570.
[32] M. Kim, G.-S. Jeng, I. Pelivanov, M. O'Donnell, M. O'Donnell, Deep-learning image reconstruction for real-time photoacoustic system, IEEE Trans. Med. Imaging 39 (11) (2020) 3379–3390.
[33] A. Hore, D. Ziou, Image Quality Metrics: PSNR Vs. SSIM. 2010 20th international conference on pattern recognition, IEEE, 2010, pp. 2366–2369.
[34] D.J. Hemanth, D. Gupta, V.E. Balas, Intelligent Data Analysis for Biomedical Applications: Challenges and Solutions, Academic Press, 2019.
[35] G. Godefroy, B. Arnal, E. Bossy, Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties, Photoacoustics 21 (2021) 100218.
[36] R. Shang, M.A. O'Brien, F. Wang, G. Situ, G.P. Luke, Approximating the uncertainty of deep learning reconstruction predictions in single-pixel imaging, Commun. Eng. 2 (1) (2023) 53.
[37] J. Gröhl, T. Kirchner, T. Adler, L. Maier-Hein, Confidence estimation for machine learning-based quantitative photoacoustics, J. Imaging 4 (12) (2018) 147.
[38] E. Goan, C. Fookes, Bayesian neural networks: an introduction and survey, case studies in Applied Bayesian data science: CIRM Jean-Morlet Chair, Fall 2018 (2020) 45–87.
[39] L.V. Jospin, H. Laga, F. Boussaid, W. Buntine, M. Bennamoun, Hands-on Bayesian neural networks—a tutorial for deep learning users, IEEE Comput. Intell. Mag. 17 (2) (2022) 29–48.
[40] J. Lampinen, A. Vehtari, Bayesian approach for neural networks—review and case studies, Neural Netw. 14 (3) (2001) 257–274.
[41] D.M. Titterington, Bayesian methods for neural networks and related models, Stat. Sci. (2004) 128–139.
[42] A. Kendall, Y. Gal, What uncertainties do we need in bayesian deep learning for computer vision? Adv. Neural Inf. Process. Syst. 30 (2017).
[43] Y. Gal, Z. Ghahramani, Dropout as A Bayesian Approximation: Representing Model Uncertainty in Deep Learning, International Conference on Machine Learning, PMLR, 2016, pp. 1050–1059.
[44] B. Lakshminarayanan, A. Pritzel, C. Blundell, Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles, arXiv preprint arXiv:1612.01474 (2016).

[45] S. Feng, C. Zuo, Y. Hu, Y. Li, Q. Chen, Deep-learning-based fringe-pattern analysis with uncertainty estimation, Optica 8 (12) (2021) 1507–1510.
[46] Y. Kwon, J.-H. Won, B.J. Kim, M.C. Paik, Uncertainty quantification using Bayesian neural networks in classification: application to biomedical image segmentation, Comput. Stat. Data anal. 142 (2020) 106816.
[47] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.
[48] B. Hanin, D. Rolnick, How to start training: the effect of initialization and architecture, Proc. 32nd Int. Conf. Neural Inf. Process. Syst. (2018) 569–579.
[49] V. Kuleshov, N. Fenner, S. Ermon, Accurate Uncertainties for Deep Learning Using Calibrated Regression. International conference on machine learning, PMLR, 2018, pp. 2796–2804.
[50] A. Niculescu-Mizil, R. Caruana, Predicting good probabilities with supervised learning, Proc. 22nd Int. Conf. Mach. Learn. (2005) 625–632.
[51] M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, Content-aware image restoration: pushing the limits of fluorescence microscopy, Nat. Methods 15 (12) (2018) 1090–1097.
[52] 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, 2015, pp. 234–241.
[53] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, Int. Conf. Mach. Learn. pmlr (2015) 448–456.
[54] B. Xu, N. Wang, T. Chen, M. Li, Empirical Evaluation of Rectified Activations in Convolutional Network, arXiv preprint arXiv:1505.00853 (2015).
[55] J. Staal, M.D. Abràmoff, M. Niemeijer, M.A. Viergever, B. Van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Trans. Med. Imaging 23 (4) (2004) 501–509.
[56] J. Xia, J. Yao, L.V. Wang, Photoacoustic tomography: Principles and advances, Electromagn. Waves 147 (2014) 1.

Ruibo Shang, Ph.D. is a postdoctoral scholar supervised by Prof. Matthew O'Donnell and Prof. Ivan M. Pelivanov in the Department of Bioengineering at the University of Washington. He completed his Ph.D. supervised by Prof. Geoffrey P. Luke in the Thayer School of Engineering at Dartmouth College. He is the recipient of Alma Hass Milham Fellowship in Biomedical Engineering for the academic year 2020–2021 and has won Neukom Prize for Outstanding Graduate Research in 2021 at Dartmouth College. His research focuses on photoacoustic imaging, computational imaging with sparsity-based optimization and deep learning, and ultrafast optical imaging.

Geoffrey P. Luke, PhD is an Associate Professor of Engineering Sciences in the Thayer School of Engineering at Dartmouth College. He earned a BS in Computer Engineering and Mathematics and a MS in Electrical Engineering from the University of Wyoming. He went on to complete his PhD at the University of Texas at Austin. He currently leads the Functional and Molecular Imaging Laboratory, with research focusing on the intersection of ultrasound, optics, and nanotechnology. Current projects seek to leverage computational and deep-learning approaches for improved biomedical imaging.

Matthew O'Donnell Following undergraduate, graduate, and post-doctoral training at Notre Dame and Washington University in St. Louis, Dr. O'Donnell joined General Electric Corporate Research and Development Center in Schenectady, NY in 1980, where he worked on medical electronics, including MRI and ultrasound imaging systems. In 1990, he moved to University of Michigan in Ann Arbor, MI where he held appointments in Electrical Engineering & Computer Science and in Biomedical Engineering. In 1998, he was named the Jerry W. and Carol L. Levin Professor of Engineering. From 1999–2006 he also served as Chair of the Biomedical Engineering Department. In 2006 he moved to the University of Washington in Seattle, WA where he was the Frank and Julie Jungers Dean of Engineering from 2006 to 2012. He is now Frank and Julie Jungers Dean Emeritus and a Professor of Bioengineering at the University of Washington. His most recent research has explored new imaging modalities, including elasticity imaging, *in vivo* microscopy, optoacoustic devices, photoacoustic contrast agents for molecular imaging and therapy, laser ultrasound systems, and catheter-based devices. He has won numerous awards, including the Distinguished Alumni Award from the University of Notre Dame, the Achievement and Rayleigh Awards from the IEEE-UFFC Society, the William J. Morlock Award for Excellence in Biomedical Technology from the IEEE-EMBS Society, and the IEEE Biomedical Engineering Award. He is a fellow of the IEEE and AIMBE and is an elected member of the Washington State Academy of Sciences and the National Academy of Engineering.