Invertible Neural Networks for Uncertainty Quantification in Photoacoustic Imaging

Jan-Hinrich Nölke[1,2], Tim Adler[1,3], Janek Gröhl[1], Thomas Kirchner[4], Lynton Ardizzone[5], Carsten Rother[5], Ullrich Köthe[5], Lena Maier-Hein[1,3,6]

[1]Division of Computer Assisted Medical Interventions, German Cancer Research Center, Heidelberg, Germany
[2]Faculty of Physics and Astronomy, Heidelberg University, Germany
[3]Faculty of Mathematics and Computer Science, Heidelberg University, Germany
[4]Institute of Applied Physics, University of Bern, Switzerland
[5]Visual Learning Lab, HCI, IWR, Heidelberg, Germany
[6]Medical Faculty, Heidelberg University, Germany
j.noelke@dkfz-heidelberg.de

Abstract. Multispectral photoacoustic imaging (PAI) is an emerging imaging modality that enables the recovery of functional tissue parameters such as blood oxygenation. However, the underlying inverse reconstruction problems are potentially ill-posed, meaning that radically different tissue properties may - in theory - yield comparable measurements. In this work, we present a new approach for handling this specific type of uncertainty using conditional invertible neural networks. We propose going beyond commonly used point estimates for tissue oxygenation and convert single-pixel initial pressure spectra to the full posterior probability density. This way, the inherent ambiguity of a problem can be encoded with multiple modes in the output. Based on the presented architecture, we demonstrate two use cases that leverage this information to not only detect and quantify but also to compensate for uncertainties: (1) photoacoustic device design and (2) optimization of photoacoustic image acquisition. Our *in silico* studies demonstrate the potential of the proposed methodology to become an important building block for uncertainty-aware reconstruction of physiological parameters with PAI.

1 Introduction

Photoacoustic Imaging (PAI) is an emerging medical imaging modality that enables the recovery of optical tissue properties with a "light-in-sound-out" approach [1]: Tissue is illuminated using light pulses, which leads to the absorption of photons and subsequent heating of the tissue. The resulting thermoelastic expansion generates pressure waves, which can then be detected by broadband ultrasonic transducers. The initial pressure distribution p_0, determined for multiple wavelengths, can then be used to determine physiological tissue properties like blood oxygenation sO_2. However, the non-linear effect of the so-called light

fluence makes the optical inverse problem ill-posed [2]. This can potentially lead to ambiguous solutions of the tissue properties. Prior work has addressed related problems with different approaches to uncertainty quantification [3,4,5,6], yet explicitly representing ambiguities by full posterior distributions has not been attempted in the context of machine learning-based image analysis. In this work, we address this gap in the literature with conditional invertible neural networks (cINNs) [7]. In contrast to conventional neural networks, the INN architecture enables the computation of the full posterior density function (rather than a simple point estimate), which naturally enables the encoding of various types of uncertainty, including multiple solutions (modes). The contribution of this paper is two-fold: (1) We adapt the concept of cINNs to the specific problem of quantifying tissue parameters from PAI data. (2) We demonstrate the value of our approach with two use-cases, namely PAI device design and optimization of photoacoustic image acquisition.

2 Materials and methods

2.1 Virtual photoacoustic imaging environment

The virtual environment created for testing the proposed approach to uncertainty quantification is based on a digital PAI device. With it, 3D representations of the optical and acoustic properties of tissue can be generated, which are used to simulate synthetic PAI data for a given probe design, pose and ground truth tissue properties. The data is simulated using the Monte Carlo eXtreme framework [8]. For this study, each simulation is performed with 10^7 photons originating from a pencil-like source and a grid spacing of 0.34 mm. Each volume is simulated at 26 equidistant wavelengths between 700 nm and 950 nm.

2.2 Approach to uncertainty quantification

Our architecture builds upon the cINN architecture proposed in [7]. Based on a known forward process for converting tissue properties (here: pixel-wise tissue oxygenation x) to resulting measurements (here: pixel-wise initial pressure

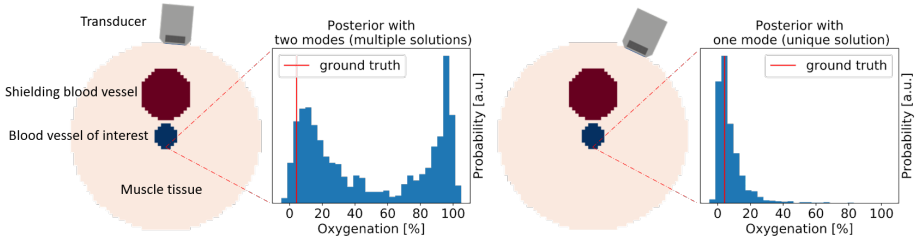

Fig. 1. *In silico* setting illustrating how slight changes in the PAI probe pose can resolve model ambiguity (training on S2b, see sec. 2.3). Left: the posterior corresponding to a pixel of interest features two modes. Right: Owing to an improved acquisition pose, the same pixel features a uni-modal posterior.

spectrum y), the task is to train a neural network to recover x from y while accounting for potential ambiguities. To this end, cINNs are leveraged as follows: Given training data consisting of (simulated) pairs (x, y), a cINN is trained to convert x to a Gaussian distributed latent space z, using y as conditioning input. This is achieved with maximum likelihood training. During inference time, because of the invertible architecture, we can sample the latent distribution and, given a new measurement y used as conditioning input y, generate a conditional probability distribution $p(x|y)$.

The architecture implemented in this work consists of 20 blocks, each with a random permutation and a conditional generative flow coupling block [9] (two fully connected layers of size 512 and rectified linear unit activations). During training, we apply normally distributed random noise with $\sigma = 0.001$ to the normalized input and $\sigma = 0.1$ to the conditioning input. The models are trained for 60 epochs with the AdamW optimizer and weight decay of 0.01. We start with a learning rate of 10^{-3} and reduce it by a factor of 10 after epoch 40 and 50.

To automate the detection of multimodal posteriors, we introduce a multimode score. We perform kernel density estimation on the posterior samples with 21 different bandwidths between 0.01 p.p. and 0.1 p.p.. The score is then the fraction of estimates with more than one maximum relative to all estimates.

2.3 Experiments

The purpose of our experiments was to (1) validate the proposed approach to uncertainty quantification in PAI and to (2) showcase use cases that leverage the posteriors to not only detect and quantify uncertainties but to compensate for them. To this end, we generated four different settings.

S1: Single vessel, single illumination unit (IU): Images (probabilistically) generated for this setting comprise a tube of muscle tissue with 2 cm diameter as background with blood oxygenation uniformly drawn between 0 and 1. In the center, a blood vessel with a radius uniformly drawn between 1 mm and 3 mm and oxygenation between 0 and 1 is placed. A single illumination source is used.

S2: Multiple vessels, single IU: Setting S1 is enhanced by introducing an additional blood vessel randomly placed between the light source and the central vessel of interest. This vessel also has a radius between 1 mm and 3 mm and an oxygenation between 0 and 1. Fig. 1 illustrates the basic setup of the phantoms.

S2b: Multiple vessels, shifted single IU: This setting is identical to S2, but the scene is illuminated from two additional angles ($\pm 45°$). We use the three different illumination setups as independent samples leading to a three times bigger data set. This setting (S2b) was exclusively used to generate Fig. 1, i. e. to demonstrate the effect of probe position on the resulting posterior.

S3: Multiple vessels, multiple IUs: The setting uses the same data as S2b, but we concatenate the three spectra from the different illumination setups (thus

simulating a complex device with three illumination units/detectors) which leads to a conditioning input dimension of $3 \cdot 26$. Fig. 2 gives an overview of the settings S1-S3.

We simulated 2,000 volumes for each of the settings and trained cINN models as described in sec. 2.2 on each of them with 85% of the data. The remaining 15% of the data was used for testing. To validate the accuracy of the posteriors, we computed the calibration curves for scenarios S1-S3 as proposed in [10]. We processed the results to analyze the capability of our method to reveal ambiguous problems (multiple modes) and to determine the effect of device pose and design.

3 Results

As can be seen in Fig. 3 all calibration curves are close to the identity (median calibration error < 1.5 p.p.). This implies that the width of the posteriors is reliable. For the setting with a single vessel and single illumination (S1), the model is slightly underconfident.

In Fig. 4 we compare the distribution of IQRs, absolute errors, and the multi-mode score for the scenarios S1-S3 described in sec. 2.3. Our results demonstrate that not only the accuracy but also the likelihood for ambiguity of the problem depends crucially on the characteristics of the probe (e.g., number of illumination/detection units). For all three metrics, the performance for the setting with multiple vessels, but only one illumination (S2) is clearly the worst. In particular, this setting includes a non-negligible fraction of multimodal posteriors.

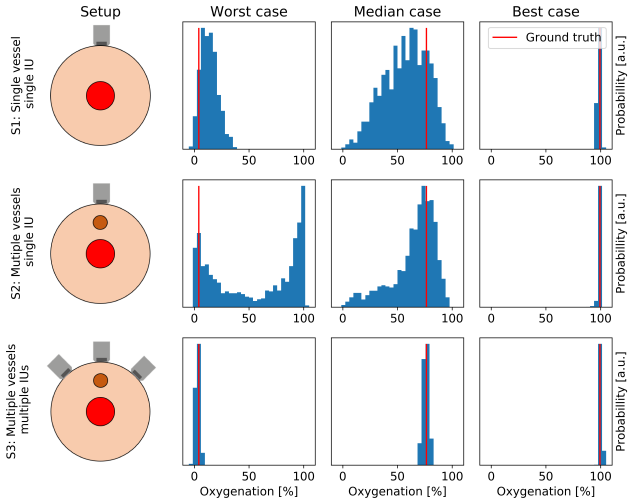

Fig. 2. Worst, median, and best case with respect to the IQR of S2 for three investigated settings. In contrast to a single vessel scenario (top), ambiguities are likely to occur in a multi-vessel scenario (middle) when using a single light source. These can be compensated for with multiple light sources (bottom).

Fig. 1 and Fig. 2 further show that the accuracy at a given pixel depends crucially on the pose and the illumination geometry of the PAI device. Moreover, the ambiguity of the inverse problem can be potentially resolved by performing the acquisition from a different position/angle or by using a multiple illumination setting (S3). Our approach could thus serve as a basis for optimizing the measurement process and photoacoustic device design.

4 Discussion

To our knowledge, this is the first work exploring the concept of INNs in the context of PAI. Specifically, we have demonstrated the capabilities of cINNs to represent and quantify uncertainties in the context of physiological parameter estimation. Based on our initial experiments, we believe that our approach could serve as a basis for optimizing PAI probe design and image acquisition.

This work is similar to that proposed by Adler et al. [11] in the context of multispectral optical imaging. However, it differs in that we used cINNs instead of the original INN architecture, which comes along with several major advantages, including (1) no zero-padding needed, leading to smaller network size, (2) maximum likelihood training, and (3) no hyperparameters in the loss function.

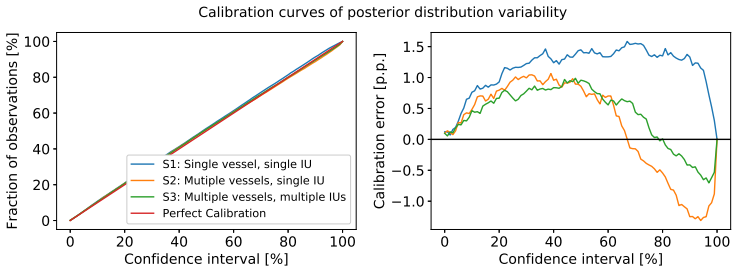

Fig. 3. Calibration curves of the posterior distributions of the settings S1-S3 as described in sec. 2.3. Fraction of observations (left) and calibration error (right) as a function of the confidence interval on the test set.

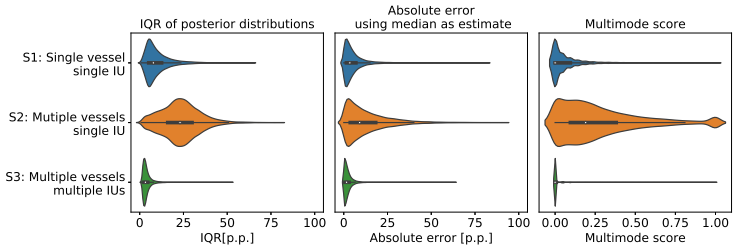

Fig. 4. The violin plots show the interquartile range (IQR) of the posterior distribution on the test set, the absolute error when using the median as an estimate, and the multimode score, introduced in sec. 2.2. We differentiate between the settings S1-S3 as described in sec. 2.3.

Our findings for device design are in line with Shao et al. [12] where a multi-illumination setup is suggested to improve image reconstruction. Our initial experiments indicate that our method may help in the optimization of the acquisition process. As a next step, our approach has to be extended such that it not only shows the current ambiguities but also proposes possible poses to resolve them. This might be achieved through the application of reinforcement learning.

In conclusion, we have demonstrated the potential of cINNs to reconstruct tissue parameters from PAI data while systematically representing and quantifying uncertainties.

Acknowledgement. This project has received funding from the European Union's Horizon 2020 research and innovation programme (grant agreement No. ERC2015-StG-37960 and No. 647769) and from the Federal Ministry of Education and Research of Germany project High Performance Deep Learning Framework (No. 01IH17002).

References

1. Zackrisson S, van de Ven SMWY, Gambhir SS. Light In and Sound Out: Emerging Translational Strategies for Photoacoustic Imaging. Cancer Res. 2014;74(4):979–1004.
2. Yang C, Lan H, Gao F, et al. Deep learning for photoacoustic imaging: a survey. arXiv:200804221 [cs, eess]. 2020;.
3. Tarvainen T, Pulkkinen A, Cox BT, et al. Bayesian Image Reconstruction in Quantitative Photoacoustic Tomography. IEEE Trans Med Imaging. 2013 Dec;32(12):2287–2298.
4. Tick J, Pulkkinen A, Tarvainen T. Image reconstruction with uncertainty quantification in photoacoustic tomography. J Acoust Soc Am. 2016;139(4):1951–1961.
5. Gröhl J, Kirchner T, Adler T, et al. Confidence Estimation for Machine Learning-Based Quantitative Photoacoustics. J Imaging. 2018;4(12):147.
6. Godefroy G, Arnal B, Bossy E. Solving the visibility problem in photoacoustic imaging with a deep learning approach providing prediction uncertainties. arXiv:200613096 [physics]. 2020;.
7. Ardizzone L, Lüth C, Kruse J, et al. Guided Image Generation with Conditional Invertible Neural Networks. arXiv:19070233092 [cs]. 2019;.
8. Fang Q, Boas DA. Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units. Opt Express. 2009;17(22):20178–20190.
9. Kingma DP, Dhariwal P. Glow: Generative Flow with Invertible 1x1 Convolutions. arXiv:180703039 [cs, stat]. 2018;.
10. Ardizzone L, Kruse J, Wirkert S, et al. Analyzing Inverse Problems with Invertible Neural Networks. arXiv:180804730 [cs, stat]. 2019;.
11. Adler TJ, Ardizzone L, Vemuri A, et al. Uncertainty-aware performance assessment of optical imaging modalities with invertible neural networks. Int J Comput Assist Radiol Surg. 2019;14(6):997–1007.
12. Shao P, Cox B, Zemp R. Estimating optical absorption, scattering, and Grueneisen distributions with multiple-illumination photoacoustic tomography. Appl Opt. 2011;50:3145–54.