

*aus dem*  
Deutschen Krebsforschungszentrum Heidelberg  
(Vorstandsvorsitzender: Prof. Dr. Michael Baumann)  
Abteilung Computer-assistierte Medizinische Interventionen  
(Abteilungsleiterin: Prof. Dr. Lena Maier-Hein)

---

# **Data-driven Quantitative Photoacoustic Imaging**

---

Inauguraldissertation  
zur Erlangung des Doctor scientiarum humanarum (Dr. sc. hum.)  
*an der*  
Medizinischen Fakultät Heidelberg  
*der*  
Ruprecht-Karls-Universität

*vorgelegt von*  
Janek Matthias Gröhl

*aus*  
Hamburg

2020



Dekan: Prof. Dr. Hans-Georg Kräusslich  
Doktormutter: Prof. Dr. Lena Maier-Hein





# Contents

<b>Acronyms</b>	<b>3</b>
<b>List of Figures</b>	<b>7</b>
<b>List of Tables</b>	<b>9</b>
<b>1 Introduction</b>	<b>11</b>
1.1 Motivation . . . . .	11
1.2 Open challenges . . . . .	13
1.3 Objectives . . . . .	15
1.4 Outline . . . . .	16
<b>2 Materials and Methods</b>	<b>17</b>
2.1 Light-Tissue Interactions . . . . .	18
2.1.1 Scattering . . . . .	19
2.1.2 Absorption . . . . .	20
2.1.3 The near-infrared window . . . . .	20
2.2 Photoacoustic imaging . . . . .	22
2.2.1 The photoacoustic effect . . . . .	22
2.2.2 Clinical applications of photoacoustic imaging . . . . .	23
2.2.3 Physical Principles . . . . .	24
2.2.4 The inverse problems of photoacoustic imaging . . . . .	25
2.2.5 Multispectral photoacoustic imaging . . . . .	28
2.3 Photoacoustic data simulation . . . . .	31
2.3.1 Reference hardware setup . . . . .	31
2.3.2 Optical forward modeling . . . . .	32
2.3.3 Acoustic forward modeling . . . . .	36
2.3.4 Sources of error in photoacoustic imaging . . . . .	37
2.3.5 Image reconstruction . . . . .	38
2.4 Machine learning . . . . .	40
2.4.1 Supervised and unsupervised learning . . . . .	40
2.4.2 Deep learning . . . . .	41
2.4.3 Machine learning models used in this thesis . . . . .	44
2.5 Related work . . . . .	49
2.5.1 Model-based quantitative photoacoustic imaging . . . . .	49

2.5.2	Machine learning-based quantitative photoacoustic imaging . . . . .	52
2.5.3	Functional photoacoustic imaging . . . . .	52
2.5.4	Uncertainty estimation . . . . .	53
2.5.5	Conclusions . . . . .	55
<b>3</b>	<b>Results</b>	<b>59</b>
3.1	Quantitative photoacoustic imaging with context encoding . . . . .	61
3.1.1	Introduction . . . . .	61
3.1.2	Methods . . . . .	62
3.1.3	Results . . . . .	68
3.1.4	Discussion . . . . .	75
3.1.5	Conclusion . . . . .	77
3.2	Quantitative photoacoustic imaging with end-to-end deep learning . . . . .	79
3.2.1	Introduction . . . . .	79
3.2.2	Methods . . . . .	80
3.2.3	Results . . . . .	85
3.2.4	Discussion . . . . .	91
3.2.5	Conclusion . . . . .	97
3.3	Functional photoacoustic imaging with learned spectral decoloring . . . . .	99
3.3.1	Introduction . . . . .	99
3.3.2	Methods . . . . .	101
3.3.3	Results . . . . .	106
3.3.4	Discussion . . . . .	112
3.3.5	Conclusion . . . . .	119
3.4	Uncertainty estimation for quantitative photoacoustic imaging . . . . .	121
3.4.1	Introduction . . . . .	121
3.4.2	Methods . . . . .	123
3.4.3	Results . . . . .	127
3.4.4	Discussion . . . . .	139
3.4.5	Conclusion . . . . .	141
<b>4</b>	<b>Discussion</b>	<b>143</b>
4.1	Summary of Contributions . . . . .	143
4.2	Discussion of results . . . . .	144
4.3	Solving inverse problems with data-driven methods . . . . .	147
4.4	Data-driven quantitative photoacoustic imaging . . . . .	149
4.5	Uncertainty-aware parameter estimation . . . . .	150
4.6	Conclusion . . . . .	151
<b>5</b>	<b>Summary</b>	<b>153</b>
<b>6</b>	<b>Zusammenfassung</b>	<b>155</b>
<b>7</b>	<b>Own contributions</b>	<b>175</b>
7.1	Own share in data acquisition and data analysis . . . . .	175
7.2	Own publications . . . . .	176
	<b>Statutory Declaration</b>	<b>185</b>

# Acronyms

<b>2D</b>	2-Dimensional
<b>3D</b>	3-Dimensional
<b>BOLD MRI</b>	Blood Oxygen Level Dependent MRI
<b>CE-qPAI</b>	Context Encoding Quantitative Photoacoustic Imaging
<b>CI</b>	Context Image
<b>CNR</b>	Contrast-to-Noise Ratio
<b>CNN</b>	Convolutional Neural Network
<b>CT</b>	Computed Tomography
<b>DAS</b>	Delay-And-Sum
<b>DE</b>	Diffusion Equation
<b>DMAS</b>	Delay-Multiply-And-Sum
<b>eMSOT</b>	eigenspectra Multispectral Optoacoustic Tomography
<b>fCE-qPAI</b>	Functional CE-qPAI
<b>FCM</b>	Fluence Contribution Map
<b>FEM</b>	Finite Element Model
<b>FFT</b>	Fast Fourier Transform
<b>fNIRS</b>	functional Near-Infrared Spectroscopy
<b>GB</b>	GigaBytes
<b>ICG</b>	Indocyanine Green
<b>IQR</b>	Interquartile Range
<b>INN</b>	Invertible Neural Network

<b>LSD-qPAI</b>	Learned Spectral Decoloring for quantitative Photoacoustic Imaging
<b>LU</b>	Linear Unmixing
<b>MCX</b>	Monte Carlo eXtreme
<b>MRI</b>	Magnetic Resonance Imaging
<b>OCT</b>	Optical Coherence Tomography
<b>PA</b>	Photoacoustic
<b>PAI</b>	Photoacoustic Imaging
<b>PC</b>	Principal Component
<b>PCA</b>	Principal Component Analysis
<b>PML</b>	Perfectly Matched Layer
<b>qPAI</b>	quantitative Photoacoustic Imaging
<b>RAM</b>	Random Access Memory
<b>ROI</b>	Region Of Interest
<b>ReLU</b>	Rectified Linear Unit
<b>RTE</b>	Radiative Transfer Equation
<b>sDMAS</b>	signed Delay-Multiply-And-Sum
<b>SNR</b>	Signal-to-Noise Ratio
<b>US</b>	Ultrasound

# List of Figures

1.1	Number of photoacoustic papers over time . . . . .	12
2.1	Visualization of the fundamental light-tissue interactions . . . . .	18
2.2	Visualization of the photon deflection angle . . . . .	19
2.3	Overview over the different chromophores in biological tissue . . . . .	21
2.4	Visualization of the near-infrared window . . . . .	21
2.5	The photoacoustic effect . . . . .	23
2.6	Photoacoustic signal generation and recording . . . . .	25
2.7	The frequency spectrum of photoacoustic waves . . . . .	27
2.8	Photoacoustic device detection geometries . . . . .	27
2.9	Multispectral photoacoustic imaging . . . . .	29
2.10	Steps in PA image simulation . . . . .	31
2.11	Simulated PA device hardware . . . . .	32
2.12	Visualization of the photon scattering model in Monte Carlo simulations. . . . .	35
2.13	Illustration of the Delay-And-Sum principle . . . . .	39
2.14	Visualization of supervised machine learning . . . . .	40
2.15	The basic building blocks of deep learning networks . . . . .	41
2.16	The deep learning training workflow . . . . .	43
2.17	Visualization of a simple regression tree . . . . .	44
2.18	A feed-forward neural network design . . . . .	45
2.19	The U-Net . . . . .	46
2.20	Visualization of the distribution mapping of an invertible neural network. . . . .	47
2.21	Visualization of the affine coupling layers that make up the Invertible Neural Network architecture . . . . .	47
2.22	Visualization of the distribution mapping of a conditional invertible neural network architecture . . . . .	48
2.23	Iterative photoacoustic reconstruction schemes . . . . .	50
2.24	Overview over different uncertainty methods . . . . .	54
3.1	Overview over the CE-qPAI method . . . . .	63
3.2	Generation of context images . . . . .	64
3.3	CE-qPAI random forest noise level overview . . . . .	68
3.4	CE-qPAI random forest example baseline set estimations . . . . .	69
3.5	CE-qPAI random forest example most complex data set estimations . . . . .	70

3.6	CE-qPAI deep learning noise overview . . . . .	71
3.7	CE-qPAI deep learning example baseline set estimations . . . . .	72
3.8	CE-qPAI deep learning example most complex data set estimations . . . . .	73
3.9	Qualitative results of functional CE-qPAI . . . . .	74
3.10	Overview over the different DL-qPAI approaches . . . . .	81
3.11	DL-qPAI phantom data set setups . . . . .	84
3.12	DL-qPAI realistic data set setups . . . . .	84
3.13	Violin plots and box plots visualizing the distribution of the relative estimation errors of the approach and data set combinations. . . . .	86
3.14	Violin plots of the direct approach versus a sequential application of the acoustic and optical inverse models . . . . .	87
3.15	Representative results for the acoustic inverse model . . . . .	88
3.16	Representative results for the optical inverse model . . . . .	89
3.17	Representative results for the combined inverse model . . . . .	90
3.18	Performance of the two-stage approach regarding the generic tissue data set . . . . .	92
3.19	Error versus imaging depth analysis . . . . .	93
3.20	Illustration of spectral coloring in a simulation experiment . . . . .	100
3.21	Method overview of LSD-qPAI . . . . .	102
3.22	Schematic representation of the flow phantom data set . . . . .	103
3.23	Original schematic of the flow phantom setup . . . . .	104
3.24	In silico LSD-qPAI results . . . . .	106
3.25	In vitro LSD-qPAI results (graph) . . . . .	107
3.26	In vitro LSD-qPAI result images on the human_2 data set . . . . .	108
3.27	In vitro LSD-qPAI result images on the Rat_2 data set . . . . .	109
3.28	In vivo LSD-qPAI results on porcine brain . . . . .	110
3.29	Distribution of LU versus LSD-qPAI unmixed forearm spectra . . . . .	111
3.30	In vivo LSD-qPAI results on human forearm . . . . .	112
3.31	Overview visualization of the effect of spectral coloring in the flow phantom data sets . . . . .	114
3.32	Visualization of spectral coloring on the human_2 data set comparing LU and LSD . . . . .	115
3.33	Visualization of spectral coloring on the Rat_2 data set comparing LU and LSD . . . . .	116
3.34	Principal component analysis of simulated and real spectra . . . . .	117
3.35	Overview figure for the uncertainty estimation workflow. . . . .	123
3.36	Visualization of dropout sampling for uncertainty estimation . . . . .	124
3.37	Visualization of the label space transformations to enable estimation of conditional probability densities . . . . .	125
3.38	Visualization of how conditional probability density estimates are transformed back into label space . . . . .	125
3.39	Visualization of using an observing model for uncertainty estimation . . . . .	126
3.40	Visualization of using a conditional invertible neural network for uncertainty estimation . . . . .	127
3.41	The results of applying the uncertainty method proposed by Feindt to the CE-qPAI method . . . . .	128
3.42	The results of using an external model to estimate the uncertainty of various 2D quantification methods . . . . .	129
3.43	Scatter plots of sO <sub>2</sub> estimates with annotated dropout uncertainty . . . . .	131

3.44	Detailed progression of the $sO_2$ estimation error when excluding an increasing amount of estimates based on their dropout uncertainty. . . . .	131
3.45	Scatter plots of $sO_2$ estimates with annotated dropout uncertainty . . . . .	132
3.46	Detailed progression of the $sO_2$ estimation error when excluding an increasing amount of estimates based on their reconstructed uncertainty. . . . .	132
3.47	Scatter plots of $sO_2$ estimates with annotated dropout uncertainty . . . . .	133
3.48	Detailed progression of the $sO_2$ estimation error when excluding an increasing amount of estimates based on their dropout uncertainty. . . . .	133
3.49	Scatter plots of $sO_2$ estimates with annotated dropout uncertainty . . . . .	134
3.50	Detailed progression of the $sO_2$ estimation error when excluding an increasing amount of estimates based on their dropout uncertainty. . . . .	134
3.51	In vitro LSD-qPAI result images on the "human_2" data set . . . . .	135
3.52	In vitro LSD-qPAI result images on the "Rat_2" data set . . . . .	136
3.53	Uncertainty calibration curves . . . . .	137





# List of Tables

1.1	Overview of the methodological sections contained in the <i>Results</i> chapter of this thesis. . . . .	16
2.1	Visualization of the gap in the qPAI state of the art . . . . .	56
3.1	CE-qPAI <i>in silico</i> data set design . . . . .	65
3.2	Descriptive statistics of the random forest CE-qPAI results . . . . .	68
3.3	All deep learning CE-qPAI results . . . . .	71
3.4	All functional CE-qPAI results . . . . .	74
3.5	Feature and label domains for the end-to-end deep learning approaches . . . . .	81
3.6	Optical forward model parameters . . . . .	82
3.7	Acoustic forward model parameters . . . . .	83
3.8	Descriptive results of deep learning absorption estimation . . . . .	86
3.9	Fractional changes of the estimation error with different uncertainty thresholds . . . . .	130



If the absorption spectrum of a sample is known then it is often possible to work out the **concentrations of the constituent molecules** using a spectroscopic analysis. Surely multiwavelength photoacoustic imaging can, **by analogy**, provide maps of the concentrations of the constituent chromophores?

[Cox et al., 2009b]

# 1 | Introduction

## 1.1 Motivation

Photoacoustic Imaging (PAI) is a comparatively young and emerging imaging modality that promises real-time, noninvasive, and radiation-free imaging of optical tissue properties. With commonly available optical imaging techniques, obtaining information on optical tissue properties is only possible on the surface or in *ex vivo* samples, due to high absorption and scattering of light in the visible wavelength range of light. In contrast to other optical imaging modalities, the most significant advantage of PAI is that it can visualize optical tissue properties up to several centimeters deep in tissue. PAI is based on the *photoacoustic effect*, which enables spatially resolved imaging of chromophores, which are tissue molecules that absorb light, by sending very short pulses of near-infrared laser light into tissue. Where the incident light is absorbed, sound waves emerge in the audible and ultrasound regimes. These are created by a process called *thermoelastic expansion*, which is the effect of changes in the volume of material induced by changes in its temperature and vice versa. These sound waves can be measured with appropriate detection devices such as hydrophones - i.e., piezoelectric transducers - and the measured data can be reconstructed into a spatial image that represents the absorbed energy distribution.

When measuring Photoacoustic (PA) signals of the same structure at different wavelengths of light, knowledge on functional tissue properties can potentially be inferred. This can be achieved with *spectral unmixing* techniques that are capable of computing the fractional composition of chromophores that contribute to the measured multispectral signal. With these ratios of chromophore concentrations, functional parameters such as blood oxygenation  $sO_2$  can be derived locally in tissue on multispectral PA images. *Linear* unmixing algorithms are the primary choice to solve for the concentrations of the individual chromophores. These algorithms are tasked to identify the best fitting linear combination of *a priori* multispectral data of the absorption properties of the chromophores. As such, a fundamental prerequisite for the applicability of these methods to PA images is that the measured signal  $S$  is proportional only to the absorption coefficients.

However, in reality, the measured signal  $S$  reflects the initial pressure distribution  $p_0$  and is therefore proportional to the temperature-dependent Grüneisen coefficient  $\Gamma$ , the optical absorption coefficient  $\mu_a$ , as well as the fluence  $\phi$ .  $\phi$ , in turn, is dependent on both  $\mu_a$  and optical scattering  $\mu_s$ . Due to this interdependency of  $\mu_a$  and  $\phi$ , the quantification of the PA signal is an ill-posed inverse problem. Research in the field of quantitative Photoacoustic Imaging (qPAI) tries to quantify the absolute values of the underlying optical tissue properties from PA images, however, despite ongoing research in the field, no general solution to the problem has been found yet.

While there has been a swift acceleration of research regarding PAI over the last three decades, with exponential growth of the number of publications in the field (cf. Figure 1.1), current approaches to solve the quantification problem have a more-or-less steady number of publications over the last years. So far, the developed methods have only been shown to work in simulation studies or precise and highly controlled experimental settings and are generally not applicable in a robust and general manner to clinically acquired data. Reasons for this might, for example, be ill-chosen model assumptions, dependence on *a priori* knowledge, lack of robustness to noise, and long computation times.

### Number of papers in the field 1991 - 2018

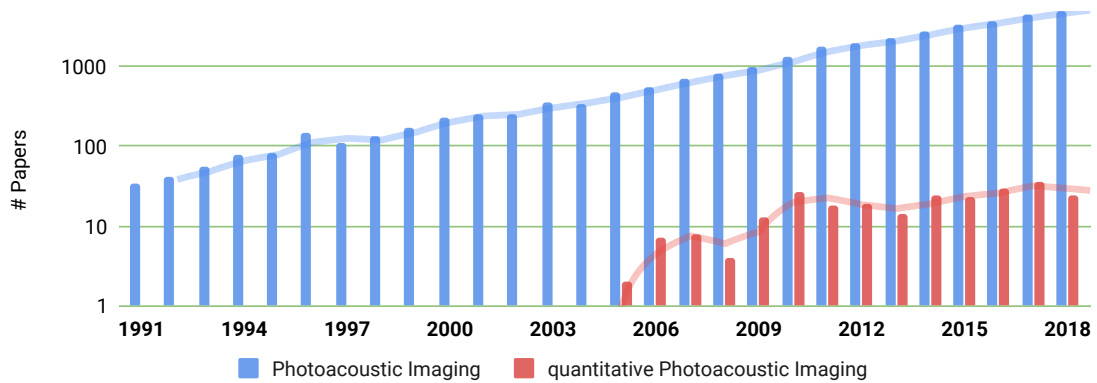


Figure 1.1: Visualization of the number of PA papers over time. It shows an exponential increase of PAI and an overall increase in publications on qPAI. The data shown were obtained using Google Scholar queries isolated to single years. Search terms for blue: ("photoacoustic imaging" OR "optoacoustic imaging" OR "photoacoustic tomography" OR "optoacoustic tomography") and for red: (intitle:"quantitative" AND intitle:"(photoacoustic OR optoacoustic)" AND "(imaging OR tomography)")

## 1.2 Open challenges

Several fundamental problems need to be tackled to achieve a successful quantification of PA signals. These include:

### C1: Solving the acoustic inverse problem:

The first open challenge is to perform quantitative PA image reconstruction from recorded time series pressure data. The objective is to accurately reconstruct an image  $S$  showing the initial pressure distribution  $p_0$  from measured time series pressure data  $p(t)$ . This inverse process of reconstructing  $p_0$  from  $p(t)$  is commonly referred to as the *acoustic inverse problem*. There are several key aspects to this challenge, such as the compensation for limited view detection geometries and accurate calibration of the detection system.

### C2: Solving the optical inverse problem:

This challenge describes the need to account for the light fluence distribution  $\phi$  and the Grüneisen parameter distribution  $\Gamma$  to obtain information on the optical absorption coefficient  $\mu_a$  from the initial pressure distribution  $p_0$ . This inverse process of quantitative  $\mu_a$  measurement from  $p_0$  is usually referred to as the *optical inverse problem* because knowledge on  $\phi$  needs to be obtained. The key problem is that the light propagation through tissue is dependent on the  $\mu_a$  distribution, which means that ideally, both  $\mu_a$  and  $\phi$  have to be estimated in parallel by the inverse process.

### C3: Performing quantitative spectral unmixing:

The process of *spectral unmixing* refers to the estimation of absolute concentrations of chromophores (molecules that have distinct and measurable light absorption behaviors) by using multispectral PA images. Different chromophores usually differ in their magnitude of light absorption, which enables spectral unmixing techniques to find the ratios of the contribution of the different chromophores to the measured signal  $S$ . The key problem is that the fluence is non-linearly dependent on the wavelength of the light and leads to *spectral coloring* of the multispectral signal that has to be accounted for by the unmixing algorithm.

### C4: Extraction of meaningful confidence estimations:

Algorithms developed for the calculation of clinically relevant parameters of the health status of a patient should be accompanied by methods that know when the model fails. In the best case, these also know *why* the model fails, to minimize the risk for the patient. Because of the ill-posed nature of the described inverse problems of qPAI, it is necessary to develop methods that can estimate uncertainty-metrics that correlate to the expected errors of the respective inverse model and that are capable of resolving ambiguities of the inversion process at the same time. The key objective is to identify uncertainty estimation methods that can be used for the qPAI inverse problems and that yield interpretable results.

A lot of different PA device geometries have been proposed so far, and the listed challenges are a lot more pronounced when considering PAI with freehand clinical handheld devices compared

to pre-clinical tomography systems. Tomographic systems having at least a field of view of  $180^\circ$  usually can achieve much higher image quality and, as such, provide more information when having to solve the acoustic inverse problem (details see section 2.3.5). However, they can currently only be used for small animal imaging or particular use cases because of practical constraints. Because of these limitations, methods have to be developed that enable quantitative imaging for handheld devices such that they can be translated into clinical practice.

A major obstacle when validating qPAI methods is the fact that there exists no gold standard method to obtain information on the light fluence or optical tissue properties *in vivo* or *in vitro* in large or complex structures. Currently, obtaining reference measurements is only possible either on the tissue surface by using, for example, diffuse reflectance measurements (cf. [van Veen et al., 2004]) or with *ex vivo* tissue samples when the analysis is performed with complex setups such as a double integrating sphere (cf. [Pickering et al., 1993]). As such, methods that have been developed typically rely on mathematical models of the PA effect or the presence of simulated PA data (e.g. as presented in [Akhlaghi et al., 2019]) and as such, they are often not suitable for application *in vivo*.

This lack of ground truth data has led to an under-representation of machine learning techniques towards qPAI when compared to other imaging modalities. Considering the vast parameter space of geometrical shapes, distributions of chromophores and scattering molecules, as well as the illumination and detection geometries that are possible when simulating PA data (amounting to millions of unknowns), techniques need to be established that can create data-driven models from relatively few data samples while being robust against noise, artifacts, and unknown situations.

The primary objective of the work presented in this Ph.D. thesis is to investigate the feasibility of developing data-driven methods to tackle the listed challenges and to explore their capabilities towards accurate and reliable quantification of PA signals based on *in silico* training data. Due to the computational efficiency of data-driven methods during inference, the successful development of such a method could enable many clinical applications of PAI. This is true, especially in the area of interventional monitoring of functional tissue parameters such as  $\text{sO}_2$  locally in tissue.

## 1.3 Objectives

The work presented in this thesis contributes to several qPAI hypotheses that relate to the open challenges C1-C4:

### **H1: Data-driven methods can be used to solve the optical and acoustic inverse problem.**

The purpose of work towards  $H1$  is to examine the feasibility of using data-driven approaches for the optical and acoustic inverse problem. The hypothesis includes the question of how *in silico* simulated PA images and *a priori* knowledge can be used for training of the developed algorithms.

### **H2: Data-driven methods can be used for spectral unmixing in a realistic context.**

The purpose of work towards  $H2$  is to examine the feasibility of employing data-driven methods trained on simulated multispectral PA data for accurate spectral unmixing in realistic scenarios. The main question is to determine whether the effect of *spectral coloring* is systematic and whether it can be accounted for during the unmixing process.

### **H3: Confidence estimation methods be used to gain insight into qPAI methods.**

The purpose of work towards  $H3$  is to examine whether the use of methods for uncertainty and confidence estimation can yield improvement on the accuracy or robustness of the methods developed towards  $H1$  and  $H2$ . A critical aspect of this hypothesis is to identify state-of-the-art techniques for uncertainty quantification and to investigate their applicability to the qPAI inverse problems.

## 1.4 Outline

The thesis starts with an overview of the principle knowledge necessary to follow the results of this thesis in chapter 2 *Materials and methods*. It gives an overview of the PA effect (section 2.2), the simulation of *in silico* PA images (2.3), machine learning techniques (section 2.4), and the current state of the art (section 2.5) for model-based qPAI, functional PAI, as well as uncertainty estimation methods. The section concludes in an overview of the open questions in qPAI that have not been sufficiently solved yet.

The sections listed in chapter 3 *Results* give a detailed presentation of the work that has been conducted towards the hypotheses  $H_1$ ,  $H_2$ , and  $H_3$  outlined in the previous section. Table 1.1 summarises this relationship for each main chapter and also refers to the relevant papers that have been published.

Section	Challenges	Hypotheses	Contributions
3.1	C2	$H_1$	[Kirchner et al., 2018a]
3.2	C1, C2	$H_1$	[Gröhl et al., 2018a] [Waibel et al., 2018]
3.3	C3	$H_2$	[Gröhl et al., 2019]
3.4	C4	$H_3$	[Gröhl et al., 2018b] [Gröhl et al., 2018a]

Table 1.1: Overview of the methodological sections contained in the *Results* chapter of this thesis.

Work towards  $H_1$  is presented in sections 3.1 and 3.2. Section 3.1 introduces the concept of *context images* to tackle the optical inverse problem with data-driven methods, and section 3.2 presents the application of end-to-end deep learning methods towards both the acoustic and optical inverse problem directly on 2-Dimensional (2D) PA images. Section 3.3 introduces work towards  $H_2$  and presents a method that is capable of estimating functional tissue parameters *in vivo* while exclusively being trained on synthetic training data. Finally, in section 3.4, several methods are compared for estimating the uncertainty of data-driven qPAI methods. The section also presents a technique to exploit these estimates to increase the performance of the respective inversion method.

Chapter 4 *Discussion* discusses the methods and results presented in this thesis and contains a general discussion of the advantages and limitations of using data-driven methods for the qPAI inverse problems. It also contains a summary of the contributions of this thesis and an outlook of the next steps that could be made towards qPAI.

A summary of the findings presented in this thesis is given in chapters 5 (English version) and 6 (German version).



*The purpose of this paper is to establish a **theoretical and conceptual framework** for the papers to follow. To do so we will review the **current understanding of physical processes** involved in photoacoustic [...] techniques.*

[McDonald, 1986]

# 2 | **Materials and Methods**

This chapter gives an overview of the principles needed to be able to follow the content of this thesis. It is divided into five main sections. The first two sections are an introduction to the physics of light tissue interactions and the photoacoustic effect, including a short historical overview from its discovery until today. The third section is about the simulation of PA images by modeling the physical principles involved in PAI. The fourth section gives an overview of the principle ideas and methods of machine learning. Finally, the fifth section covers a brief review of the current state of the art in qPAI and explicitly points out the gaps that are addressed with the work presented in this thesis.

## 2.1 | Light-Tissue Interactions

Further details on light-tissue interactions can be found in the textbook of Wang and Wu [Wang and Wu, 2012], and a great description of optical properties can be found on the omlc website by Steven L. Jacques and Scott A. Prahl (<https://omlc.org/classroom/ece532/class3/>, as visited on 24.11.2019). Most of the information presented in this section is extracted from these two resources and the references contained within.

When light enters biological tissue, it can undergo several fundamental interactions with the medium: (1) Light *refraction* and *reflection* at the medium surface, which happens at the interface between two media when there is a mismatch in the respective refractive indexes. This effect is dependent on the incident angle of the light on this interface (cf. Fresnel Equations [Skaar, 2006]) and can even lead to the total reflection of the incoming light. (2) Photon *absorption* by a molecule leads to the elevation of an electron from the ground state to a higher energy state. (3) *Scattering* of photons, which is the change of direction of photons caused by structures in the medium. Scattering happens when the refractive index of the structure mismatches that of the surrounding medium. Finally, (4) there is *diffuse reflection*, which happens when light escapes the tissue surface after having undergone scattering events in the tissue. For certain imaging techniques - such as transmitted light microscopy - the *transmittance* of light through the sample is of great importance, but this is usually only relevant for very thin media, where the probability of photons not being subject to any interactions within the sample is still very high [Pluta and Maksymilian, 1988]. Figure 2.1 visualizes the explained effects in a comprehensive but simplified overview.

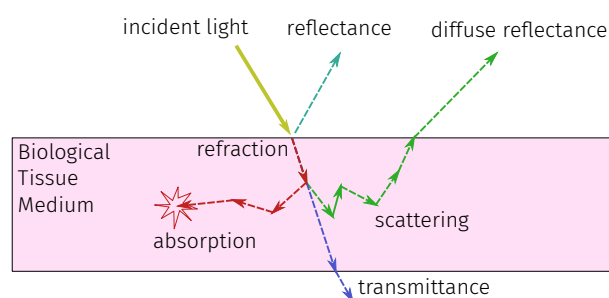


Figure 2.1: Visualization of the fundamental light-tissue interactions: refraction, reflectance, absorption, scattering, transmittance, and diffuse reflectance.

The two most dominant mechanisms relevant for imaging techniques in the visible to near-infrared wavelength ranges (400-1300 nm) are optical absorption and scattering. The realistic imaging depth of optical imaging is limited by the mean free path that photons can travel before

being subject to an interaction event. The mean free path is simply the multiplicative inverse of the absorption and scattering interaction coefficients. In the mentioned wavelength range, the more dominant effect in biological tissue is optical scattering, where the mean free path between photon scattering events is in the order of 0.1 mm (also referred to as the *optical diffusion limit*). On the other hand, the mean path length before a photon absorption event can be about 1-10 cm in normal tissues [Wang and Wu, 2012].

### 2.1.1 Scattering

Scattering of light is caused by changes in the refractive index of *scatterers* compared to their surrounding medium or surrounding structures. Some of the most common scatterers in biological tissue are cells, nuclei, mitochondria, lysosomes, collagen fibers, and cell membranes. Scattering is a wavelength-dependent property because the incident light is scattered the most by structures whose size matches the wavelength of the photons. Scattering is differentiated in *elastic* and *inelastic* scattering that differ in the fact that during elastic scattering, no energy is lost during the scattering process, but during inelastic scattering, there is a change in the particle energy. The change of energy in inelastic scattering leads to a change of the light wavelength and is exploited for example in Raman spectroscopy (cf. e.g. [Choo-Smith et al., 2002]). In the near-infrared regime, elastic scattering is the dominating effect.

The scattering property of a medium is represented by the *scattering coefficient*  $\mu_s$ , which is defined as the probability of a photon to be subject to a scattering event in a medium per unit path length. The scattering coefficient is determined by the volume density of the particles  $\rho_s$ , their geometrical size  $A_s$ , and their specific scattering efficiency  $Q_s$ :  $\mu_s = \rho_s \cdot Q_s \cdot A_s$ . In biological tissue, it is of the order of magnitude of  $100 \text{ cm}^{-1}$  and its multiplicative inverse is referred to as the *scattering mean free path*. The ratio of particles that will have undergone no scattering events (equivalent to the transmittance  $T(x)$  in a medium only subject to scattering) after photon propagation over  $x$  unit path lengths can be calculated using Beer-Lambert law:

$$T(x) = e^{-\mu_s \cdot x}. \quad (2.1)$$

#### Anisotropy

The anisotropy  $g$  is a dimensionless quantity that describes the directionality of the scattering events. Generally speaking, a certain amount of the "forward direction" of the photon is retained during scattering.

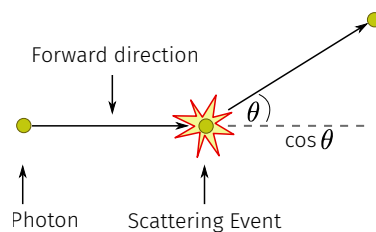


Figure 2.2: Visualization of the photon deflection angle relative to its forward direction.

Figure 2.2 visualizes the change of this forward direction by the deflection angle  $\theta$ . When considering an average deflection angle for a certain scatterer, the expected value of  $\cos(\theta)$  is defined

as  $g$ :

$$g = \int_0^{\pi} p(\theta) \cos(\theta) 2\pi \sin(\theta) d\theta, \quad (2.2)$$

with  $p(\theta)$  being the scattering pattern and where  $\int_0^{\pi} p(\theta) 2\pi \sin(\theta) d\theta = 1$ .

### 2.1.2 Absorption

Photon absorption occurs as a quantum event (all or nothing principle), where the entire energy of a photon is transferred to a molecule, changing the electric state of the molecule. This is because the absorption of a photon raises an electron from the ground state to an excited state. Depending on the wavelength of the absorbed photon and the absorbing molecule, this excitation can be relieved by mechanisms such as non-radiative relaxation, fluorescence, or phosphorescence. For this effect to occur, the photon frequency must match the energy needed for a possible (usually rotational or vibrational) energy transition of the molecule - as such, photon absorption is also a wavelength-dependent phenomenon. The following relationship defines this energy:

$$\text{Energy} = \frac{h \times c}{\lambda}, \quad (2.3)$$

with  $h$  being Planck's constant,  $c$  being the speed of light in vacuum and  $\lambda$  being the wavelength of the light.

Typical absorbers (*chromophores*) that are present in the human body are melanin, oxygenated and deoxygenated hemoglobin, fat, water, and collagen. Due to the wavelength-dependent nature of optical absorption, and the different molecular compositions of the chromophores, each chromophore has its characteristic spectrum (cf. figure 2.3).

The absorption property of a medium is represented by the *absorption coefficient*  $\mu_a$ , which - analogous to the scattering property - is defined as the probability of a photon to undergo absorption in a medium per unit path length. The absorption property is determined by the volume density of the particles  $\rho_a$ , their geometrical size  $A_a$ , and their specific absorption efficiency  $Q_a$ :  $\mu_a = \rho_a \cdot Q_a \cdot A_a$ . In biological tissue, it is of the order of magnitude of  $0.1 \text{ cm}^{-1}$ , and its multiplicative inverse is referred to as the *absorption mean free path*. The ratio of photons that will not have been absorbed (equivalent to the transmittance  $T(x)$  in a medium only subject to absorption) after photon propagation over  $x$  unit path lengths can be calculated using Beer-Lambert law:

$$T(x) = e^{-\mu_a \cdot x}. \quad (2.4)$$

### 2.1.3 The near-infrared window

Both scattering and absorption limit the mean free path of photon propagation and, as such, limit the effective penetration and imaging depth. However, there is a broad window in the wavelength range from  $\approx 600$ - $1300$  nm, in which the combined attenuation (and hence the optical

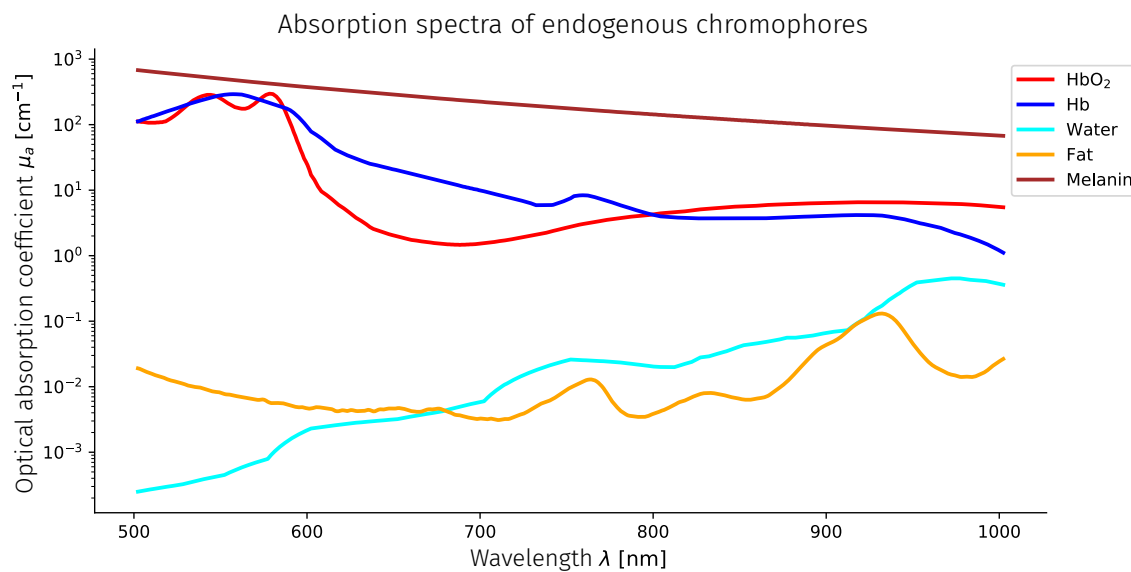


Figure 2.3: Overview over the wavelength-dependent absorption spectra of different endogenous chromophores in biological tissue. The spectra were compiled from [Jacques, 2013] and the references within. Hb and HbO<sub>2</sub> absorption coefficients are calculated under the assumption of a hemoglobin concentration of 150g/l.

and acoustic diffusion limit) is at a local minimum, and the potential imaging depth (mean free path) is at a local maximum; the near-infrared window (cf. figure 2.4).

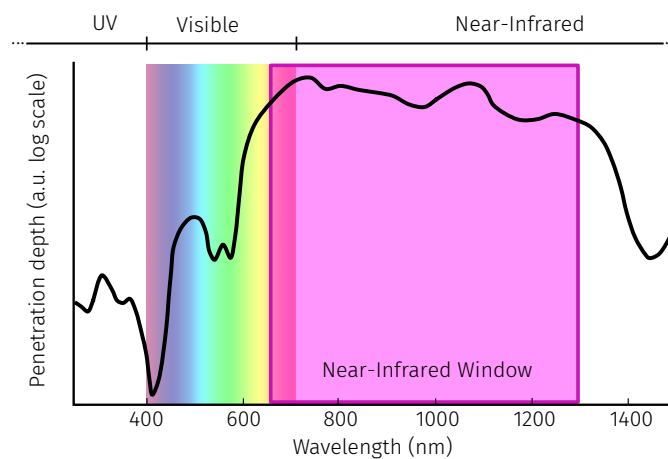


Figure 2.4: Stylized visualization of the near-infrared window. The represented penetration depth values were adapted from [Pezzi et al., 2019] and were scaled and smoothed for better visualization.

## 2.2 | Photoacoustic imaging

PAI is a comparatively young and emerging imaging modality that promises real-time, noninvasive, and radiation-free imaging of optical tissue properties. With commonly used optical imaging techniques, such as multispectral imaging (e.g. [Freeman et al., 1997, Wirkert et al., 2016]) or diffuse optical spectroscopy (e.g. [van Veen et al., 2004]), obtaining information of these optical tissue properties was only possible on the surface because of the aforementioned *optical diffusion limit* due to high absorption and scattering of light in the visible and near-infrared wavelengths of light. The most significant advantage of PAI is, that - in contrast to other optical imaging modalities - it can visualize optical tissue properties up to several centimeters deep in tissue by exploiting the *photoacoustic effect* to ultrasonically break the diffusion limit [Wang, 2012].

### 2.2.1 The photoacoustic effect

A simplified visualization of the PA effect is shown in figure 2.5. The PA effect is based on the principle of thermoelastic expansion, where an extremely short light pulse is able to induce a local pressure wave inside the tissue. It enables structural imaging of molecules that absorb light (so-called chromophores) because the acoustic scattering of the emerging sound waves is orders of magnitude smaller than the optical scattering of the incident light in biological tissue. The molecules that cause PA signals to arise can be categorized into two categories. Firstly, *endogenous* chromophores that are naturally present in living organisms, of which the main contributors are melanin, blood (or rather the red blood cells: oxygenated and deoxygenated hemoglobin), fat, water, and collagen. And secondly, *exogenous* contrast agents such as nanoparticles and dyes (for example gold, Indocyanine Green (ICG), methylene blue, and fluor)[Cox et al., 2005, Yang et al., 2009, Wu et al., 2014, Zackrisson et al., 2014]. The physical principles of the PA effect are detailed in section 2.2.3.

The PA effect was first described in 1880 by Alexander Graham Bell in his pioneering article "Upon the production and reproduction of sound by light" [Bell, 1880]. After already having invented the telephone, Bell was very fond of the idea of the *photophone*, a device to transmit speech using light. He formulated this idea after observing that certain substances, such as selenium, would produce sound when being exposed to a beam of modulated light. He later also proposed the *spectrophone* [Bell, 1881] to examine the absorption characteristics of an illuminated sample. Many years later, the development of modern laser technology paved the way for the exponential increase of PA research, as seen today (see figure 1.1). The PA effect regained the attention of the scientific community (e.g. Kerr and Atwood [Kerr and Atwood, 1968] and Allen Rosencwaig [Rosencwaig, 1975]), also regarding the application of PA for medical imaging [Bowen, 1983]. Breakthroughs in PA image reconstruction (e.g. [Oraevsky et al., 1999, Xu and Wang, 2005]) and the identification of *in vivo* applications of PAI for clinical and pre-clinical imaging (e.g. [Oraevsky et al., 2001, Wang et al., 2003a]) followed at the beginning of the new millennium. A comprehensive historical overview of the development of PAI and the emergence of the PAI nomenclature can be found in the work of Manohar and Razansky [Manohar and Razansky, 2016].

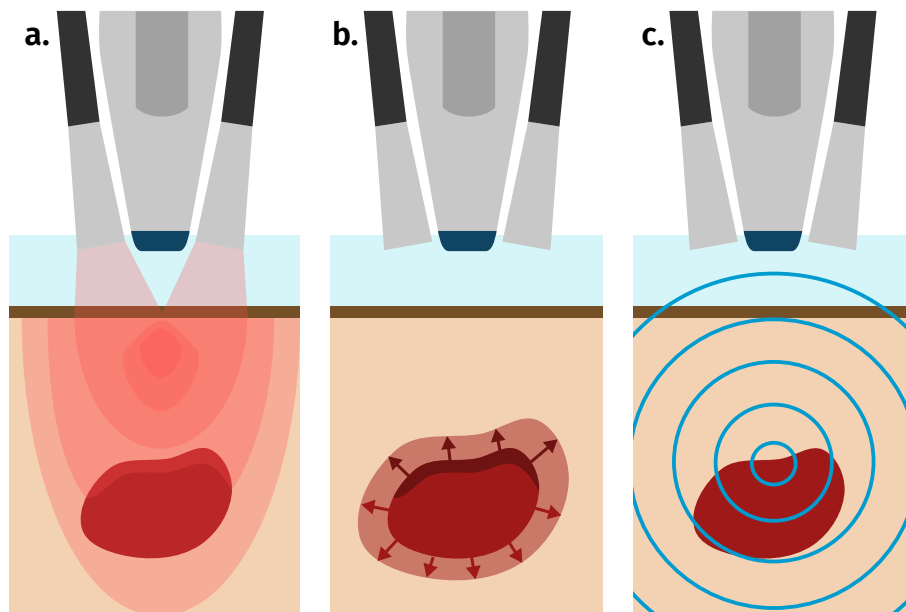


Figure 2.5: Stylized images that display the mechanisms of the PA effect. The individual images depict (from top to bottom) a PA transducer, a coupling agent (light blue), epidermis (dark brown), dermis (creme), and an absorbing structure (red). **a.** illustrates the fluence of laser light in tissue, **b.** demonstrates thermoelastic expansion of the areas where the light was absorbed, and **c.** shows the propagation of the emerging sound waves.

### 2.2.2 Clinical applications of photoacoustic imaging

Numerous groups are working with PAI in some form or another. As such, many fields of application of PAI have been established over the last years. Cancer research is a field where PAI shows great potential [Mallidi et al., 2011]. Here, blood is the enabling endogenous chromophore. One of the hallmarks of cancer is sustained angiogenesis [Hanahan and Weinberg, 2011], which is the formation of new blood vessels to supply the tumor site with blood, inducing an increase in the total hemoglobin concentration at the tumor site, and a second hallmark is an increased metabolism, potentially inducing a decrease in local blood oxygenation [Wang and Wu, 2012]. Much work is conducted towards using PAI for cancer research, including cancerous tissue classification [Li et al., 2019, Zhang et al., 2018a], cancer detection [Jnawali et al., 2019], and delineation of tumor sites with multispectral PAI [Quiros-Gonzalez et al., 2018]. Skin melanoma and metastasis detection and imaging [Oh et al., 2006, Weight et al., 2006, Zhang et al., 2010a, Zhang et al., 2010b], as well as the mapping of sentinel lymph nodes [Song et al., 2008, Erpelding et al., 2010, Garcia-Urbe et al., 2015] are further applications of PAI in the context of cancer research.

Furthermore, because inflammatory processes change the hemodynamic behavior of tissue, PAI is also used for visualization of inflammatory processes such as inflamed joints [Wang et al., 2007, Rajian et al., 2012, Jo et al., 2018] or staging of patients who have Crohn's disease [Knieling et al., 2017, Waldner et al., 2016]; also using endoscopic PAI devices [Lei et al., 2019]. However, different chromophores can be visualized for clinical value as well, such as the content of

collagen in muscle as an imaging biomarker for Duchenne muscular dystrophy [Regensburger et al., 2019].

PAI is also extensively being used in combination with other imaging modalities, especially Ultrasound (US) imaging, in order to enhance image quality and patient care [Niederhauser et al., 2005, Aguirre et al., 2011, Needles et al., 2013, Garcia-Urbe et al., 2015, Elbau et al., 2017, Mandal et al., 2019]. Developed techniques are then for example used for brain imaging [Wang et al., 2003b, Ku et al., 2005, Hu et al., 2009, Yao and Wang, 2014, Mohammadi et al., 2019] or surgical and interventional imaging applications such as needle tracking [Kim et al., 2010, Su et al., 2010, Allman et al., 2019, Johnstonbaugh et al., 2019].

More applications and more details on the listed examples can be found in the numerous comprehensive literature reviews in the field for example in those of Luke et al. (2012) [Luke et al., 2012], Taruttis and Ntziachristos (2015) [Taruttis and Ntziachristos, 2015], Liu and Zhang (2016) [Liu and Zhang, 2016], Omar et al. (2018) [Omar et al., 2019], and Zhu et al. (2018) [Zhu et al., 2018].

### 2.2.3 Physical Principles

For the PA effect to take place, the laser pulse width  $\tau_l$  has to be sufficiently small to be within the confinements of two timescales: The thermal relaxation time and the stress relaxation time. The thermal relaxation time is estimated by  $\tau_{th} = d_c^2/\alpha_{th}$ , with  $d_c$  being the target resolution and  $\alpha_{th}$  being the thermal diffusivity of tissue. The stress relaxation time can be calculated using  $\tau_s = d_c/v_s$ , with  $v_s$  being the speed of sound. From these conditions, it is evident that the stress relaxation time is more strict than the thermal relaxation time and  $\tau_l$  meets the stress and thermal confinements if  $\tau_l < \tau_s \ll \tau_{th}$ . For example, when considering the properties of typical soft tissue ( $\alpha_{th} = 0.13 \text{ mm}^2/\text{s}$  and  $v_s = 1500 \text{ m/s}$ ) and for a target resolution of  $15 \text{ }\mu\text{m}$ , this would result to  $\tau_{th} = 1.7 \text{ ms}$  and  $\tau_s = 10 \text{ ns}$  [Wang and Wu, 2012].

When these timescales are met, the absorbed photon energy causes an immediate change in pressure  $p_0$  which is based on the compressibility  $\kappa$  of the tissue, the thermal expansion coefficient  $\beta$ , and the change in temperature  $T$ :

$$p_0 = \frac{\beta}{\kappa} T, \quad (2.5)$$

which - when considering that the temperature rise can be expressed as the absorbed energy  $H$  times the photothermal efficiency  $\eta_{th}$  in relation to the specific heat capacity  $C_V$  and the mass density  $\rho$  - can be rewritten as

$$p_0 = \frac{\beta \eta_{th} H}{\kappa \rho C_V}. \quad (2.6)$$

When combining all temperature-dependent material properties into the Grüneisen coefficient  $\Gamma = \beta\eta_{th}/(\kappa\rho C_V)$  and when considering that  $\kappa = C_P/(\rho v_s^2 C_V)$ , we arrive at  $\Gamma = \beta\eta_{th} v_s^2/C_P$  and can express  $p_0$  as the product of the absorbed energy distribution  $H$  and  $\Gamma$ . Here,  $H$  is the product of the optical absorption coefficient  $\mu_a$  and the fluence  $\phi$ :

$$p_0 = \Gamma \cdot H = \Gamma \cdot \mu_a \cdot \phi. \quad (2.7)$$



The time evolution of the PA waves  $p(t)$  arising from  $p_0$  can, for example, be modeled using linear acoustics. In a lossless medium, the equations of motion, continuity, and state can be written as [Treeby and Cox, 2010, Morse et al., 1969]:

$$\begin{aligned}\frac{\partial u}{\partial t} &= -\frac{1}{\rho_0} \nabla p \\ \frac{\partial \rho}{\partial t} &= -\rho_0 \nabla u \\ p &= v_s^2 \rho,\end{aligned}\tag{2.8}$$

where  $u$  is the acoustic particle velocity,  $\rho$  is the acoustic density, and  $v_s$  is the speed of sound. Using these equations with the initial conditions  $p_0 = \Gamma \mu_a \phi$  and  $\partial p_0 / \partial t = 0$ , one can derive a single combined PA wave equation (for the detailed derivation see [Cox and Beard, 2005]):

$$\nabla^2 p - \frac{1}{v_s^2} \frac{\partial^2 p}{\partial t^2} = \frac{-\beta}{C_V} \frac{\partial H}{\partial t},\tag{2.9}$$

with  $v_s$  being the sound speed,  $\beta$  the thermal expansion coefficient,  $C_V$  the specific heat capacity, and  $H$  the absorbed heat energy (per unit volume and per unit time).

Computing a solution for the PA wave equation for a given  $p_0$  results in the time series pressure data  $p(t)$  and will be denoted as the *acoustic forward operator*  $A$  throughout this thesis:

$$p(t) = A(p_0, \theta).\tag{2.10}$$

Here,  $A$  operates on the initial pressure distribution  $p_0$  taking into consideration the acoustic properties  $\theta$  of the medium that are needed for the acoustic forward operator  $A$  (for example containing the medium speed of sound, density, acoustic attenuation, and the positions of detection elements). For a complete derivation of the PA wave equations confer for example [Cox and Beard, 2005, Treeby and Cox, 2010, Wang and Wu, 2012, Antholzer et al., 2019a]. The steps described in equations 2.5 to 2.10 are visualized in figure 2.6, which illustrates the entire forward process of PA signal generation.

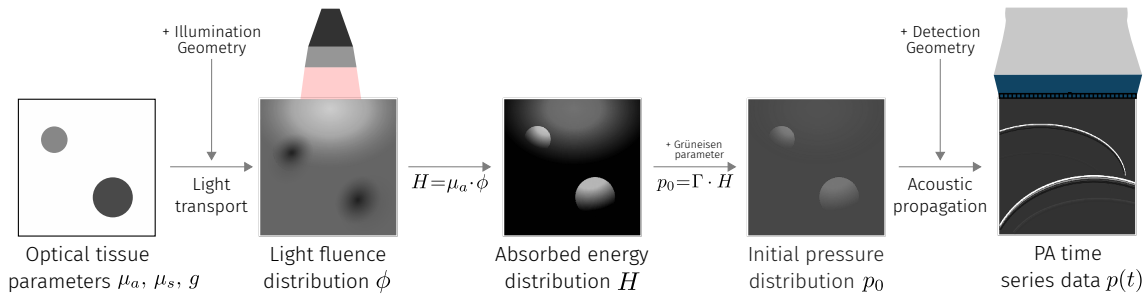


Figure 2.6: Stylized representation of the PA signal generation and recording process.

## 2.2.4 The inverse problems of photoacoustic imaging

When recording PA raw time series data  $p(t)$  with a detection device, two principle inverse problems need to be solved in order to obtain information on the optical properties of the

underlying tissue. The first is the reconstruction of an image  $S$  from recorded raw time series data by estimating the initial pressure distribution  $p_0$  from  $p(t)$  and is referred to as the *acoustic inverse problem*. The second inverse problem is that of recovering the optical tissue properties from  $p_0$ , for example, by calculating the light fluence distribution  $\phi$  and is referred to as the *optical inverse problem*. Descriptions of methods developed in the field to tackle these inverse problems are detailed later in section 2.5.

### Acoustic inverse problem

For the inverse problem of image reconstruction, the task is to infer the initial pressure distribution  $p_0$  from measured time series data of pressure  $p(t)$ . For this, an inverse function  $A^{-1}$  for the acoustic operator  $A$  needs to be calculated in order to reconstruct a signal image  $S$  that is an accurate approximation of  $p_0$ :

$$S = A^{-1}(p(t)) \approx p_0 = \mu_a \cdot \phi \cdot \Gamma. \quad (2.11)$$

Even though - for specific geometries - this inverse problem is well-posed and has a unique solution, several factors lead to considerable difficulties in solving it [Cox et al., 2009b], sometimes making it impossible to find a unique and accurate solution:

**(1) Wrong model assumptions.** The use of realistic tissue models that accurately represent, for example, a heterogeneous speed of sound, tissue density, and acoustic absorption properties complicate the wave equations and lead to an increase in computational complexity. Because of this, homogeneous tissue assumptions are usually made, leading to inaccuracies during the inversion.

**(2) Information gets lost with some detection devices.** The use of detection devices to record  $p(t)$  that do not yield perfect broadband detection can only create imperfect reconstructions of  $p_0$ , as the PA waves are broadband by nature (see figure 2.7). Typical high-frequency PA transducers (such as typical US transducers) are unable to measure the low-frequency components of the PA waves and vice versa for low-frequency transducers. Broadband transducers that are sensitive over the whole frequency range of a PA signal, such as Fabry-Pérot interferometers (cf. e.g. [Zhang et al., 2008, Nuster et al., 2011, Huynh et al., 2017]), are needed to be able to reconstruct a more faithful representation of  $p_0$ .

**(3) Device modeling and calibration.** The frequency and incident angle-dependent responses of detection device elements need to be modeled and accounted for when analyzing the measurements. Furthermore,  $p_0$  needs to be measured in absolute units of pressure, which means that the detection geometry has to be calibrated, and the reconstruction algorithm has to retain the units accurately in the reconstructed image.

**(4) Limited view geometries.** As demonstrated in figure 2.8, many different detection geometries have been developed and proposed for PAI. For 2D image acquisition, these can, for example, be split into these categories: ring arrays [Xia et al., 2012, Xia and Wang, 2013], arc arrays [Kruger et al., 2003], curved linear arrays [Gamelin et al., 2008], and linear arrays [Zeng et al., 2004]. For some of these detection geometries, such as linear US transducers,  $p_0$  has to be reconstructed from incomplete data, leading to non-uniqueness and, therefore, ill-posedness of the inversion, while also giving rise to distinct reconstruction artifacts.

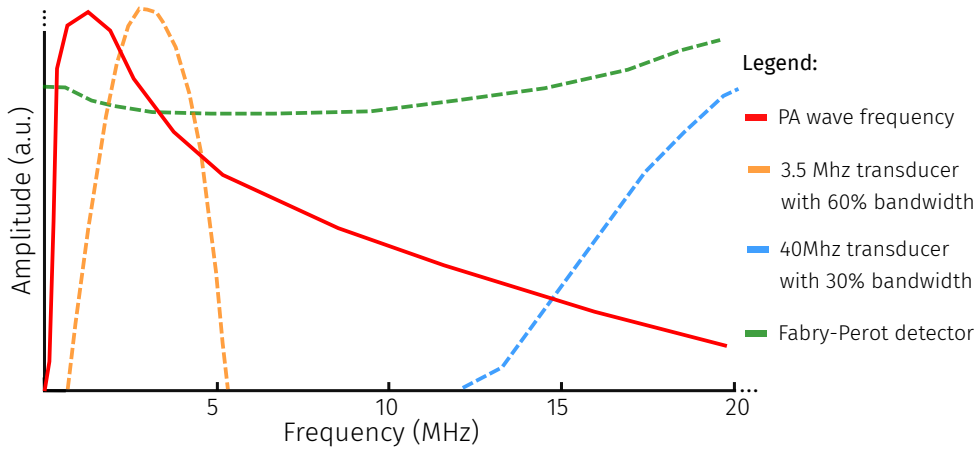


Figure 2.7: Representation of frequency ranges of PA waves and the frequency responses of typical PA detection devices. Data compiled from [Zhang et al., 2009, Treeby et al., 2010, Sethuraman et al., 2007, Ma et al., 2014, Buchmann et al., 2017a, Li et al., 2017]. For accurate and non-stylized values please refer to the cited sources.

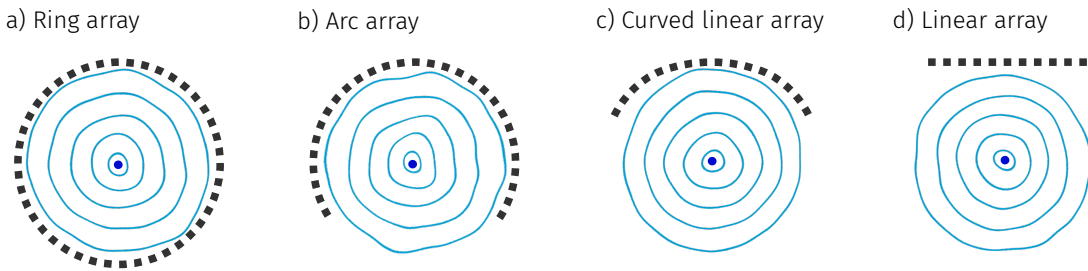


Figure 2.8: PA device detection geometries.

**Optical inverse problem**

Assuming a sufficiently accurate reconstruction of  $p_0$  from  $p(t)$ , the second problem that arises is the task to estimate the underlying optical properties (absorption  $\mu_a$ , scattering  $\mu_s$ , and anisotropy  $g$ ) of the tissue. Of these, the main property of interest is the optical absorption  $\mu_a$ , as it unlocks the potential for functional imaging of tissue (see next section 2.2.5). Principally, to calculate  $\mu_a$  one only needs to account for the Grüneisen parameter  $\Gamma$  and the fluence  $\phi$ :

$$\mu_a = \frac{p_0}{\phi \cdot \Gamma}. \tag{2.12}$$

However, several factors make this problem ill-posed and generally very hard to tackle [Cox et al., 2009b]:

**(1) Wrong model assumptions.** Accounting for a spatially varying Grüneisen parameter  $\Gamma$  and scattering coefficient  $\mu_s$  is a challenging task. As such, most approaches to qPAI have assumed

constant  $\Gamma$  and  $\mu_s$ , which can both spatially vary, and  $\Gamma$  is also dependent on the temperature of the medium, which again, might spatially vary. Thus, an approach for the optical inverse problem that has these homogeneous parameter assumptions may be subject to significant errors.

**(2) The inverse problem is ill-posed and non-linear.** There is assumed to be a non-uniqueness of the optical inverse problem, where different pairs of  $\mu_a$  and  $\mu_s$  will result in the same absorbed energy distribution  $H$ , making it an ill-posed inverse problem, further complicating the inversion [Shao et al., 2011]. At the same time, estimation of the fluence  $\phi$  is difficult, due to the non-linear dependence of fluence on the optical absorption  $\mu_a$  and scattering  $\mu_s$ . Due to this, when trying to estimate  $\mu_a$ , one needs to solve a non-linear inverse problem.

**(3) Spectral coloring.** The fluence is dependent on the wavelength  $\lambda$ , due to wavelength-dependent changes in  $\mu_a$  and  $\mu_s$ , which leads to spectral coloring of  $p_0$ . Furthermore, this effect amplifies for deeper structures, where the light has been subject to more absorption leading to a non-linear change in  $p_0$  spectra.

**(4) Noise and artifacts.** As eluded to in the previous section, the signal detection process is sensitive to reconstruction artifacts and noise from various stochastic or systematic sources. An inversion method needs to be able to cope with or at least identify these noise sources in order to be able to produce reliable quantification results.

With these considerations in mind, equation 2.7 for  $p_0$  should instead be written as:

$$p_0(x, \lambda) = \mu_a(x, \lambda) \cdot \phi(x, \lambda, \mu_a(x, \lambda), \mu_s(x, \lambda)) \cdot \Gamma(x, T(x)) + \epsilon(x, \lambda), \quad (2.13)$$

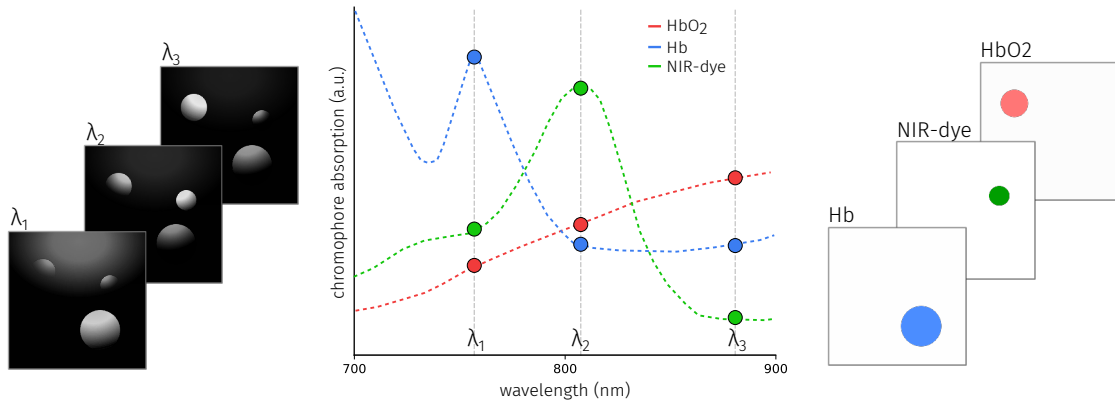
better representing the complexity of the optical inverse problem, with  $x$  being a spatial location in the imaged medium,  $\lambda$  being the wavelength of the laser light,  $T(x)$  being the spatially varying temperature, and  $\epsilon(x, \lambda)$  being a term to represent the noise.

## 2.2.5 Multispectral photoacoustic imaging

One way to remedy the dilemmas of the inverse problems is to use multispectral PAI data, i.e., using the information from PA images acquired at multiple wavelengths. The main goal of multispectral PAI is to gain information on the distribution and concentrations of different chromophores in tissue in order to be able to calculate functional tissue properties [Taruttis and Ntziachristos, 2015, Brunker et al., 2017]. One of the most prominent and promising applications of multispectral PAI for various clinical use cases is the estimation of local blood oxygenation  $sO_2$  from multispectral PA images, as oxygenated hemoglobin  $HbO_2$  and deoxygenated hemoglobin  $Hb$  are two of the main absorbers in the near-infrared.  $sO_2$  can be calculated by computing the relative ratio of the concentrations of  $Hb$  ( $C_{Hb}$ ) and the concentration of  $HbO_2$  ( $C_{HbO_2}$ ). This ratio can be estimated using a process called *spectral unmixing* (see next section):

$$sO_2 = \frac{C_{HbO_2}}{C_{HbO_2} + C_{Hb}}. \quad (2.14)$$

This approach is possible because the absorption coefficient of different chromophores usually distinctively vary with regard to the wavelength (see figure 2.9).



(1) Multispectral PA Images      (2) Chromophore specific absorption spectra      (3) Spectrally unmixed images

Figure 2.9: Idealized visualization of how multispectral PAI works. An area is imaged at multiple different wavelengths (1), and with knowledge of the absorption coefficients of chromophores (2), the images are spectrally unmixed and yield the distribution of the chromophores (3). This figure was primarily inspired by [Taruttis and Ntziachristos, 2015] and [Brunker et al., 2017].

### Spectral unmixing

Spectral unmixing refers to the process of estimating the contribution of different chromophores to a multispectral PA signal. When assuming that every chromophore contributes linearly to the absorption coefficient present in a specific spatial location  $x$ , one can use the following linear mixing model:

$$\mu_a(x, \lambda) = \sum_{k=1}^K C_k(x) \cdot a_k(\lambda), \quad (2.15)$$

where  $C_k(x)$  is the concentration of chromophore  $k$  at the spatial location  $x$ , and  $a_k(\lambda)$  is the absorption coefficient of  $k$  at wavelength  $\lambda$ . The process of unmixing the specific concentrations of chromophores from the total absorption can be tackled by solving a simple system of linear equations when assuming a linear relationship of the mixing components:

$$\begin{pmatrix} \mu_a(x, \lambda_1) \\ \vdots \\ \mu_a(x, \lambda_N) \end{pmatrix} = \begin{pmatrix} a_1(x, \lambda_1) & \dots & a_K(x, \lambda_1) \\ \vdots & \ddots & \vdots \\ a_1(x, \lambda_N) & \dots & a_K(x, \lambda_N) \end{pmatrix} \begin{pmatrix} C_1(x) \\ \vdots \\ C_K(x) \end{pmatrix}. \quad (2.16)$$

Assuming this linear relationship, at least  $N = K$  wavelengths need to be recorded to be able to solve for the concentration coefficients  $C_k$  for  $K$  chromophores.

### Why linear spectral unmixing models do not work

Despite the promising applications of multispectral PAI, PA images of  $H$  or  $p_0$  cannot simply be used for quantitative and accurate spectral unmixing to estimate chromophore concentrations. The reason for this is that one cannot easily solve for  $\mu_a$ , as the fluence  $\phi$  is unknown and has traditionally been immeasurable in complex turbid media. When substituting equation 2.15 into

equation 2.13, which is derived by considering all problems of the optical inverse problem, one can arrive at the following equation for the initial pressure distribution  $p_0$ :

$$p_0(x, \lambda) = \left( \sum_{k=1}^K C_k(x) a_k(x, \lambda) \right) \cdot \phi(x, \lambda, \left( \sum_{k=1}^K C_k(x) a_k(x, \lambda) \right), \mu_s(x, \lambda)) \cdot \Gamma(x, T(x)) + \epsilon(x, \lambda). \quad (2.17)$$

From this equation, it is evident, that knowledge of the chromophore absorption spectra is not enough, in order to use reconstructed  $p_0$  data for spectral unmixing. Instead, a successful application of linear unmixing techniques requires careful modeling or experimental examination of the imaged tissue sample, taking into account, for example, the tissue morphology, likely  $\text{sO}_2$  range, and image formation process [Hochuli et al., 2019]. Alternatively, methods have to be identified that can estimate  $\phi$ ,  $\mu_a$  and the relative concentrations of chromophores at the same time in order to be able to derive quantitative estimations for the concentrations  $C_k$ .

## 2.3 | Photoacoustic data simulation

This section gives an overview of computational methods that can be used to simulate the Photoacoustic (PA) imaging process in order to obtain *in silico* PA images from optical tissue properties. The model representation of the PA device that was used in the simulation experiments is shown in section 2.3.1. The entire simulation process is divided into four major steps (cf. figure 2.10): light transport in tissue (section 2.3.2), sound propagation (section 2.3.3), noise modeling (section 2.3.4), and PA image reconstruction (section 2.3.5).

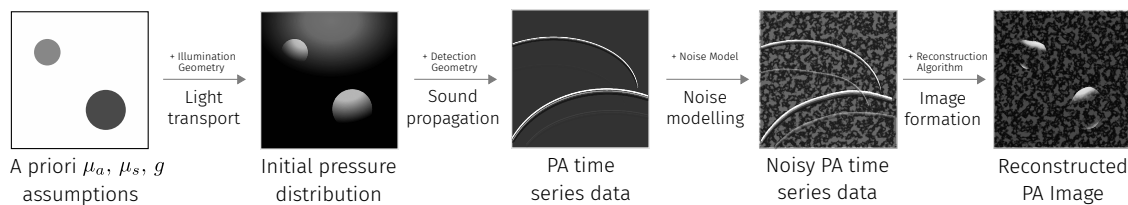


Figure 2.10: Stylized representation of the individual forward simulation steps needed for the generation of *in silico* PA images.

### 2.3.1 Reference hardware setup

This section introduces the specific assumptions for the model of the PAI device geometry that was used for image acquisition and as a template for the forward simulation. To this end, the device was modeled in each of the respective simulation frameworks using the schematics and assumptions detailed in this section.

#### DiPhAs PA probe

The PAI device that is modeled in this thesis is a custom build device based on the combination of a Fraunhofer DiPhAs US research system (Fraunhofer IBMT, St. Ingbert, Germany). It is combined with a fast tuning laser system (Phocus Mobile, Opotek, Carlsbad, USA) which can generate laser light pulses in the wavelength range from 690 nm to 950 nm, with a pulse repetition rate of 20 Hz, and a laser pulse energy of up to 50 mJ [Kirchner et al., 2016, Kirchner et al., 2019]. Thomas Kirchner implemented the communication software for these two major hardware components during his master's thesis and early Ph.D. work in the lab. The device was used for all initial experiments conducted in the lab. The principle setup is shown in figure 2.11 and is built around a 128-element linear US transducer with a 7.5 MHz center frequency, a bandwidth of 80%, and an element pitch of 0.3 mm (L7-Xtech, Vermon, Tours, France). A laser output fiber slit with a width of 2.45 cm is attached to either of the sides of the US transducer at a distance of 0.83 cm, with a tilt of  $22^\circ$ , and an opening angle with a standard deviation of  $12^\circ$ .

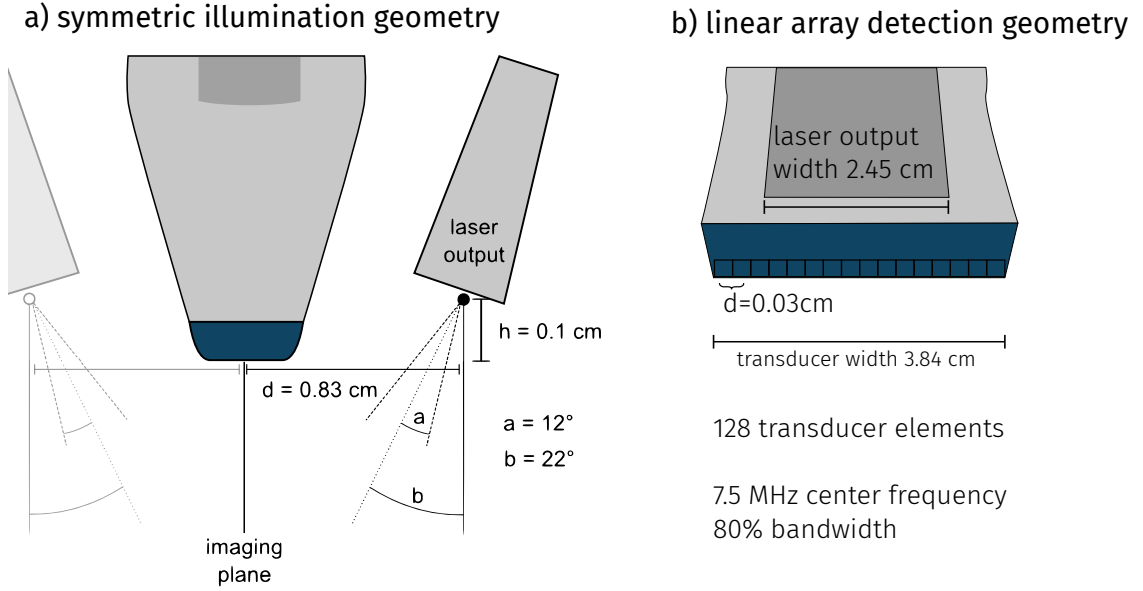


Figure 2.11: Schematic representation of the DiPhAs PA device showing a) the illumination geometry and b) the detection geometry.

### 2.3.2 Optical forward modeling

This section gives an overview of different methods to model photon transport in biological tissue. It contains a brief overview of the analytical model Radiative Transfer Equation (RTE) and its approximation with diffusion theory and a more detailed insight into the numerical Monte Carlo method, which is considered to be the gold standard method to model the RTE [Wang and Wu, 2012].

#### Radiative transfer equation

The RTE is derived from the principle of conservation of energy, where four contributions to energy changes in a medium are considered. (1) beam divergence  $dP_{div}$ , (2) photon extinction  $dP_{ext}$ , (3) photon scattering  $dP_{sca}$  and (4) the photon source  $dP_{src}$ . In the RTE it is assumed that there exist no other significant contributions to the total change in energy  $dP$ . The energy change in a volume element is negatively impacted by beam divergence, and by photon extinction, and positively impacted by photon scattering and the existence of photon sources. Hence, it is required that

$$dP = -dP_{div} - dP_{ext} + dP_{sca} + dP_{src}. \quad (2.18)$$

When substituting the analytical solutions to the individual contributions into this equation, one can arrive at the RTE [Wang and Wu, 2012]:

$$\frac{\partial L(\vec{r}, \hat{s}, t)/c}{\partial t} = -\hat{s} \cdot \nabla L(\vec{r}, \hat{s}, t) - \mu_t L(\vec{r}, \hat{s}, t) + \mu_s \int_{4\pi} L(\vec{r}, \hat{s}', t) P(\hat{s}' \cdot \hat{s}) d\Omega' + S(\vec{r}, \hat{s}, t), \quad (2.19)$$



where  $L/c$  is the propagating energy per unit volume per unit solid angle,  $\hat{s}$  is the photon propagation direction,  $d\Omega$  is a differential solid angle element around  $\hat{s}$ ,  $\mu_t$  is the probability of extinction by absorption or scattering per unit path length,  $\mu_s$  is the probability for photon scattering per unit path length, and  $S$  is the energy produced by the source. A detailed derivation of all parts of the equation can be found in [Wang and Wu, 2012].

### Diffusion theory

For most applications, the RTE is approximated using diffusion theory because of the six independent variables of the RTE, which make it difficult and computationally expensive to solve. The Diffusion Equation (DE) assumes that the influence of scattering is orders of magnitudes higher than that of absorption ( $\mu_a \ll \mu_s \cdot (1 - g)$ ) and that scattering can be considered to be approximately isotropic. A general formulation of the DE is derived in [Wang and Wu, 2012]:

$$\frac{\partial \phi(\vec{r}, t)}{c \partial t} + \mu_a \phi(\vec{r}, t) - \nabla \cdot [D \nabla \phi(\vec{r}, t)] = S(\vec{r}, t), \quad (2.20)$$

where  $\phi(\vec{r}, t)$  is the fluence rate,  $S(\vec{r}, t)$  is an isotropic source, and  $D$  is referred to as the diffusion coefficient as is defined as

$$D = \frac{1}{3(\mu_a + \mu_s \cdot (1 - g))}. \quad (2.21)$$

For a point source and when assuming time independence, the solution for the fluence rate simplifies drastically to a Beer-Lambert law [Graaff and Hoenders, 2005]:

$$\phi(r) = \frac{S_0 e^{-\mu_{\text{eff}} r}}{4\pi D r}, \quad (2.22)$$

where  $\phi(r)$  is the fluence rate,  $r$  is the distance from the source,  $\mu_{\text{eff}} = \sqrt{\mu_a/D}$ , and  $S_0$  is the incident light source energy.

### Monte Carlo simulation of light transport in tissue

In general, Monte Carlo methods use a statistical approach, for example, to investigate differential equations prevalent in the natural sciences [Metropolis and Ulam, 1949]. Sometimes, finding a solvable or computable solution to the mathematical formulation of a physical process can be impossible, unfeasible, or computationally too expensive. The idea of Monte Carlo methods is to find a set of rules to model the underlying process by performing numerous repetitions of the same random experiment that will eventually produce the expected value.

For a Monte Carlo method to work for light propagation in tissue, a fitting model has to be constructed that can be repeated a large number of times. Here, the propagation of individual photons - being modeled as particles - through tissue, has been a promising approach. Photon propagation is modeled using a 'Hop-Drop-Spin-Roulette' algorithm [Jacques, 2014]. The purpose of this algorithm is to account for two of the fundamental interactions that photons can undergo in tissue during propagation: absorption and scattering while considering spatially modeled optical properties  $\mu_a$ ,  $\mu_s$ , and  $g$ . Pseudocode for the algorithm is listed in the following:

```

def random():
    return Uniform(0, 1)

def monte_carlo_light_propagation(tissue_model, illumination_geometry):

    ENERGY_THRESHOLD = 1e-4 # small value > 0 and << 1
    SURVIVAL_PERCENTAGE = 0.1 # small value > 0 and << 1

    for number_of_photons:

        photon = launch_photon(illumination_geometry, random())

        while photon.is_alive:

            # (1) Propagation (Hop)
            photon.step()

            # (2) Absorption (Drop)
            photon.deposit_energy(tissue_model)

            # (3) Scattering (Spin)
            photon.change_direction(tissue_model, random())

            # (4) Roulette
            if photon.energy < ENERGY_THRESHOLD:
                if random() < SURVIVAL_PERCENTAGE:
                    photon.energy /= SURVIVAL_PERCENTAGE
                else:
                    photon.is_alive = False

```

A photon is modeled as a point element that has an energy and is launched from a source. Here, the source is modeled based on the illumination geometry. The position vector and direction vector of the photon are initialized with plausible values that are derived from the model representation of the illumination geometry that is created within the Monte Carlo program. After launch, the photon undergoes interactions with the optical tissue model in each iteration of the 'Hop-Drop-Spin-Roulette' loop. The parameters that are needed for these tissue interactions are the optical absorption coefficient  $\mu_a$ , the scattering coefficient  $\mu_s$ , and the scattering anisotropy  $g$ .

**(1) Hop - photon propagation.** In the Monte Carlo model, photons propagate with step size  $s$ . This step size is dependent on its mean free path between scattering events and can be calculated using  $s = -\ln(\text{rnd})/\mu_s$  with  $\text{rnd}$  being a random value subject to  $\text{rnd} \in ]0, 1]$ . Every time the hop method is called, the step size is calculated based on the optical parameters at the photons position vector, and the photon takes one step into its current direction.

**(2) Drop - photon absorption.** In contrary to the real physical interaction of photon absorption, in the Monte Carlo program, the photon does not get absorbed in its entirety and deposits all its energy into the tissue. Instead, the effect of optical absorption is modeled by having the

photon deposit fractions of its energy into the tissue. With this, the number of photons that are needed for a high-resolution simulation can be reduced. The energy deposited at each step  $\Delta E$  is calculated with the local optical absorption coefficient and the remaining energy  $E$  of the photon (for example  $\Delta E = E \cdot (1 - e^{-\mu_a \cdot s})$  in [Jacques, 2014]).

**(3) Spin - photon scattering.** After each step, the effect of scattering in the medium is simulated by changing the photon direction vector by a deflection angle  $\theta$  and an azimuthal angle  $\psi$ . The scattering model is visualized in figure 2.12.

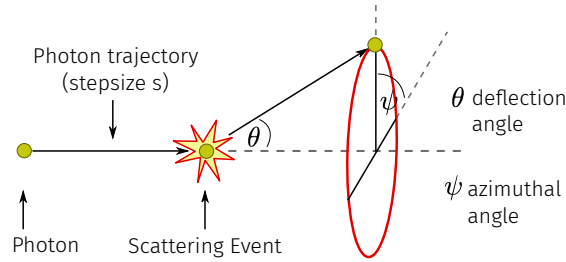


Figure 2.12: Visualization of the photon scattering model in Monte Carlo simulations.

Usually,  $\psi$  is altered by a random amount  $\psi = 2\pi \text{rnd}$ , and the deflection angle is calculated based on the scattering anisotropy using the Henyey-Greenstein-Function and a randomly drawn number  $\text{rnd} \in ]0, 1]$ :

$$\cos \theta = \frac{1 + g^2 - \left( \frac{1 - g^2}{1 - g + 2g \text{rnd}} \right)}{2g}. \quad (2.23)$$

**(4) Roulette - photon termination.** Because the optical absorption is modeled statistically using the Monte Carlo method, photons are not naturally terminated in the medium. To counteract this problem, a roulette method is implemented that terminates the photons once their energy falls below a certain threshold, in order to prevent low energy photons from endlessly propagating without having a significant contribution to the deposited energy. When this happens, a random number between 0 and 1 is drawn. If this random number is below a survival chance of for example 10%, their energy is increased by the inverse of the survival probability to account for the loss in energy by early photon termination (For example, if the chance of survival is 10%, then the photon energy is increased by a factor of ten). If the random number is above this threshold, the photon is terminated.

In order to save valuable computational resources, calculations can be performed using the reduced scattering coefficient  $\mu'_s = \mu_s \cdot (1 - g)$  instead of the scattering coefficient  $\mu_s$ . The theory behind this is that  $N = 1/(1 - g)$  anisotropic steps of length  $1/\mu_s$  through the tissue may result in the same spatial location as a single random step of length  $1/\mu'_s$ . When using the reduced scattering coefficient for simulating the scattering events, all scattering events are assumed to happen isotropically. In this scenario,  $g$  would be set to zero in equation 2.23, and hence, and the deflection angle would be calculated using  $\theta = 2\pi \text{rnd}$ . A detailed derivation of the formulations used in this method, as well as further literature, can be found in the textbook of Wang and Wu [Wang and Wu, 2012], as well as in the Ph.D. thesis of Roman Hochuli [Hochuli, 2016]. Well known

and broadly used implementations of the Monte Carlo method for photon propagation include MCXYZ [Jacques, 2014] and MCX [Fang and Boas, 2009].

### 2.3.3 Acoustic forward modeling

For a pressure wave to travel in a medium, the medium needs to be compressible to allow dynamic changes of, for example, the pressure, density, particle velocity. These changes can be described by a set of first-order partial differential equations as done in section 2.2.1 and are usually combined into a single second-order wave equation:

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = \frac{-\beta}{C_V} \frac{\partial H}{\partial t}, \quad (2.24)$$

with  $c$  being the sound speed,  $\beta$  the thermal expansion coefficient,  $C_V$  the specific heat capacity, and  $H$  the absorbed heat energy (per unit volume and per unit time). In order to simulate the progression of waves through such a medium, numerical methods need to be implemented that solve either the system of first-order equations or the second-order combined equation.

#### k-Wave Toolbox

One of the most utilized frameworks to numerically solve this equation is the MATLAB k-Wave toolbox [Treeby and Cox, 2010]. It uses the “k-space pseudo-spectral” method to solve the acoustic equations [Bojarski, 1982, Bojarski, 1985]. The main advantage of this method is the computational efficiency, especially in terms of memory requirements on expensive 3-Dimensional (3D) calculations, which is mainly achieved by the combination of a spatial Fourier transform and a discretization of the needed time derivative. In the method, the solution of the wave equation is calculated in several steps using a staggered spatio-temporal grid.

The authors numerically solve the set of first-order wave equations instead of the second-order combined equation, as it allows to include mass distributions and force sources into the equations easily. Because the computation of the derivatives using Fast Fourier Transform (FFT) would cause waves reappearing on the other side of the computational grid when they left on one side, k-Wave implements an absorbing boundary, which is referred to as a Perfectly Matched Layer (PML). In the PML, all waves traveling within it and normal to the boundary are absorbed.

A more detailed description of the k-Wave toolbox and the underlying methods can be found in the k-Wave user manual ([http://www.k-wave.org/manual/k-wave\\_user\\_manual\\_1.1.pdf](http://www.k-wave.org/manual/k-wave_user_manual_1.1.pdf), as visited on 24.11.2019) and the respective publications [Treeby and Cox, 2010, Cox and Treeby, 2010].

### 2.3.4 Sources of error in photoacoustic imaging

In general, sources of error can be distinguished to be of *epistemic* or *aleatoric* origin [Kendall and Gal, 2017]. Here, aleatoric uncertainty refers to uncertainty in the observation, such as a noisy data acquisition or motion artifacts. This source of uncertainty cannot be reduced by data-driven approaches, no matter the available amount of data. On the other hand, epistemic uncertainty refers to uncertainty caused by systematic errors in the model of the process that can be reduced by increasing the amount of data available during training.

PAI is subject to noisy and erroneous data acquisition and image reconstruction. These errors have three distinct manifestations: stochastic noise and image artifacts, which can be categorized as aleatoric errors, and systematic errors that can be categorized as epistemic errors. The stochastic signal noise in PAI originates from the physics of the detection geometry, i.e., in case of piezoelectric elements caused by random changes in the measurement voltage. Signal artifacts can be caused by several reasons in PAI. For one, typical US artifacts, such as reverberation artifacts or mirroring artifacts [Kremkau and Taylor, 1986], can sometimes be found in PA images, because the two modalities are so closely related. On the other hand, PAI introduces new potential sources for artifacts, where the most prominent example is the *transducer artifact*, which arises from photons that back-scatter into the detection geometry and induce a pressure wave there [Singh et al., 2017]. Finally, systematic errors emerge from the models of the optical or acoustic forward processes and encompass, for example, spatial light variations, the imperfectness of the theoretical model, or the influence of other components of blood [Sivaramakrishnan et al., 2007].

When simulating PA images, the imaging noise can be modeled as having an additive component (with stochastic origin), which is independent of the measured signal, and a multiplicative component (with systematic cause) that is proportional to the PA signal amplitude. The stochastic noise component on the time series data is usually represented using Gaussian noise models [Tarvainen et al., 2013, Allman et al., 2018c] modeled as zero-mean noise. The standard deviation of these models is usually set to a value that relates to a relevant dataset (for example 2% of the maximum signal value [Antholzer et al., 2019a]).

A more sophisticated approach can also be to create a model of the system-specific noise by measuring an empty measurement setup [Tarvainen et al., 2013] to also capture spatial correlations of the noise and systematic noise sources. In contrast to the other noise models, this might be very useful, as the device-specific noise does not have to be centered around zero and, as such, might introduce a bias into the measurement. For example, one could take a measurement  $I$  of  $N$  images of the empty setup at the same wavelength  $I = \{p(t, \lambda)_0, p(t, \lambda)_1, \dots, p(t, \lambda)_N\}$ , and then create a Gaussian noise model  $\epsilon_{\text{measured}}$  for  $I$  by using the mean signal  $\text{avg}(I)$  and standard deviation of the signal  $\text{std}(I)$ . Such measurements would also be capable of capturing and including the *transducer artifact* into the simulation.

### 2.3.5 Image reconstruction

Many techniques for image reconstruction are currently being investigated. This is true for PAI as well as for other imaging modalities such as US [Solberg et al., 2007], Computed Tomography (CT) [Padole et al., 2015], or Magnetic Resonance Imaging (MRI) [Geethanath et al., 2013]. For the reconstruction of an image from PA time series data, the recorded pressure data needs to be reconstructed into a spatial and human-readable image. Performance metrics for these algorithms include the accuracy of the recovery of initial pressure (for quantitative PA applications) and Signal-to-Noise Ratio (SNR)-based metrics (for qualitative PA applications). There are several comprehensive overview and review papers regarding PA image reconstruction in acoustically homogeneous and heterogeneous media [Hristova et al., 2008, Rosenthal et al., 2013, Haltmeier and Nguyen, 2019, Poudel et al., 2019]. In the context of PAI, one especially needs to distinguish between the well-posed case, where full view time series data is available, and the ill-posed case, where only limited view data is available.

**(1) Full view reconstruction.** For ring arrays with an angle covering  $360^\circ$ , the reconstruction problem is referred to as the full view problem, which is well-posed, as the signal can be sufficiently sampled. In these cases, very similar to CT imaging, an inverse of the radon transformation [Kuchment, 2013], i.e. filtered back-projection, is applied [Gamelin et al., 2008, Xu and Wang, 2005]. There have been many efforts in the field to adapt the reconstruction method, for example, to also take into account variable speed of sound [Haltmeier et al., 2019] or to suppress arising streak artifacts using adaptive back projection [Cai et al., 2019]. In addition, different model-based iterative reconstruction methods have been developed [Ding et al., 2017b, Prakash et al., 2019], sometimes also using deep learning in parts of the reconstruction methods [Antholzer et al., 2018, Boink et al., 2018, Antholzer et al., 2019b], for example, to eliminate artifacts from under-sampling when using less detection elements [Davoudi et al., 2019].

**(2) Limited view reconstruction.** Theoretically, the full view reconstruction methods can also be used for limited view geometries but will introduce reconstruction artifacts due to the ill-posed nature of the problem introduced by the limited angle coverage of the detector elements. Due to this, much effort is put into the development of limited view reconstruction methods. One popular method is the Delay-And-Sum (DAS) beamforming [Thomenius, 1996] as well as its Delay-Multiply-And-Sum (DMAS) adaptation [Matrone et al., 2014], which have been adapted and improved on in many different ways [Mozaffarzadeh et al., 2017, Kirchner et al., 2018b, Rostami et al., 2019, Su et al., 2019, Fournelle and Bost, 2019, Paridar et al., 2019]. In the DAS algorithm, the signal  $S_{\text{DAS}}(x, y)$  for signal originating at position  $(x, y)$  is reconstructed by summing up all pixels corresponding to a specific time delay in the raw time series data as measured by transducer  $j$  ( $\tau(x, y, j)$ ):

$$S_{\text{DAS}}(x, y) = \sum_{j=1}^N S(j, \tau(x, y, j)) \cdot A_x(j), \quad (2.25)$$

where  $A_x(j)$  is an apodization function. The time delay can be calculated using:

$$\tau(x, y, j) = \frac{\sqrt{y^2 + |x - j \cdot \Delta x|^2}}{c}. \quad (2.26)$$

More detailed information on the DAS, DMAS, and one of its derivatives, the signed Delay-Multiply-And-Sum (sDMAS) algorithm, can be found in [Kirchner et al., 2018b].

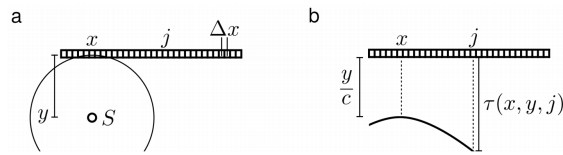


Figure 2.13: **a** Illustration of a signal  $S$  originating at depth  $y$  and lateral position  $x$  relative to a transducer array with example element  $j$  and propagating as a spherical acoustic wave with speed  $c$ ; **b** Illustration of the same scenario in the time series domain, where  $\tau(x, y, j)$  denotes the propagation time of  $S$ . Figure reprinted and caption adapted from the CC-BY licensed publication [Kirchner et al., 2018b].

Model-based iterative image reconstruction schemes are also being developed and optimized [Qin et al., 2019, Mozaffarzadeh et al., 2018, Tick et al., 2016] and quite recently also machine learning techniques have been applied to image reconstruction [Schwab et al., 2019a, Schwab et al., 2019b, Shan et al., 2019, Liu et al., 2019a] and have been used for artifact removal [Allman et al., 2018a].

## 2.4 | Machine learning

This section gives an overview of the basic underlying machine learning principles, as the methods presented in this thesis mostly utilize data-driven machine learning approaches. The term *machine learning* refers to a collection of data-driven clustering, pattern recognition, and inversion techniques and is often categorized as a subfield of *artificial intelligence* [Kononenko, 2001]. This use of nomenclature has gained popularity due to the dominant emergence of deep learning techniques over the last decade. There is a clear distinction, however, because the key objective of artificial intelligence research is to “seek a coherent theory for explaining the nature and mechanisms of both natural and artificial intelligence [Wang, 2009]”, whereas the key objective of machine learning in the context of data science is to “build models of data [VanderPlas, 2016]”, which - at most - can be classified as an application of *weak* or *narrow* artificial intelligence [Goertzel and Pennachin, 2007]. Finally, the term *deep learning* refers to a set of algorithms “that exploit many layers of non-linear information processing for supervised or unsupervised feature extraction and transformation [Deng and Yu, 2014]”.

### 2.4.1 Supervised and unsupervised learning

In the field of machine learning, it is attempted to adjust an algorithm in a way such that it *learns* to make sense of a collection of data. The data usually is separated into the data the algorithm should learn from (referred to as *features*), and the data describing *what* the algorithm should learn from the data (referred to as *labels*). However, there is a principle distinction between *supervised* and *unsupervised* learning. The main driver for this distinction is the nature of the underlying data, where supervised learning uses *labeled* data (cf. figure 2.14), and unsupervised learning uses *unlabeled* data.

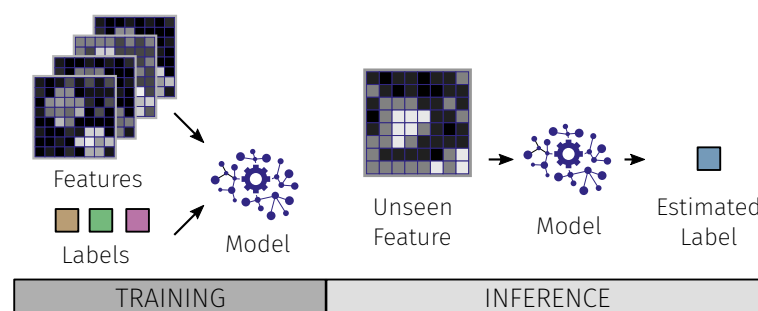


Figure 2.14: Visualization of supervised machine learning. During training, a model is tasked to learn from data features and corresponding labels. During inference, the model gives label estimates for previously unseen features.

The idea behind *supervised learning* is to train a learning algorithm to find a mapping from feature space into label space by being able to see examples of corresponding features and



labels. By learning from a training data set, the algorithm is tasked to learn a function to map previously unseen data from the feature space into the label space. The two most common subcategories of supervised learning are *classification* and *regression* tasks. In classification, the label space usually consists of a finite number of discrete values, whereas in regression, the label space is continuous [VanderPlas, 2016]. An overview of many different classical supervised machine learning techniques can be found in [Caruana and Niculescu-Mizil, 2006].

In *unsupervised learning*, on the other hand, the machine learning algorithm is tasked to find *hidden* patterns in data sets to which no or only limited labeled training data exists. Principal component analysis [Wold et al., 1987], clustering analysis [Wagstaff et al., 2001, Reynolds, 2015], and self-supervised learning [Masood et al., 2015] are some of the main categories of methods that are used in unsupervised learning. Common applications of this technique are for example density estimation [Escobar and West, 1995], dimensionality reduction of data [Dash et al., 1997], and feature extraction [Fleming and Cottrell, 1990].

### 2.4.2 Deep learning

The term *deep learning* refers to a collection of machine learning methods and is based on artificial neural networks, which are graph models that try to mimic the behavior of biological neurons. The networks are built from neurons that have non-linear activation functions and weighted connections to other neurons (cf. figure 2.15). The output of a single neuron  $\hat{y}$  can be computed as

$$\hat{y} = g\left(\sum_{i=1}^n x_i \cdot w_i + b \cdot w_0\right), \quad (2.27)$$

where  $g$  is the non-linear activation function, and  $b$  is the bias (which is typically either one or zero). Each of the weights  $w_0, \dots, w_n$  represents a tunable parameter that is adjustable during the network training process.

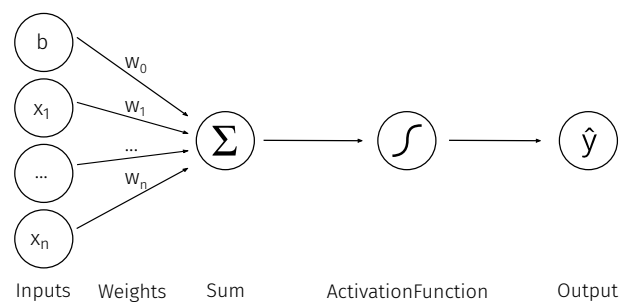


Figure 2.15: The basic building blocks of deep learning networks.

In these building blocks, the activation function of the node has the task of transforming the weighted sum of the inputs to the output value. In nature, a neuron also has a similar principle, the all-or-none law, where the charge along the neuron membrane has to exceed a specific value in order for an action potential to be unleashed and to travel along the axon. This could be modeled with a binary activation function:

$$g(x) = \begin{cases} 0, & \text{if } x \leq \text{threshold} \\ 1, & \text{if } x > \text{threshold} \end{cases} \quad (2.28)$$

When choosing between activation functions for artificial neural networks, non-linear activation functions have to be prioritized. This is because a concatenation of several linear activation functions can also be expressed by a single linear activation function. For example, stacking three activation functions of form  $a_i \cdot x + b_i$ , which would have the form  $(a_3 \cdot (a_2 \cdot (a_1 \cdot x + b_1) + b_2) + b_3)$  would yield the result  $(a_1 \cdot a_2 \cdot a_3) \cdot x + (a_2 \cdot a_3 \cdot b + a_3 \cdot b_2 + b_3)$ , which is still linear. With non-linear activation functions, complex representations can be created even with the use of only a few artificial neurons. Commonly used activation functions include sigmoidal functions such as the tangens hyperbolicus

$$g(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (2.29)$$

or piecewise linear functions, such as the Rectified Linear Unit (ReLU)

$$g(x) = \text{ReLU}(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1, & \text{if } x > 0 \end{cases} . \quad (2.30)$$

State-of-the-art neural networks are typically more complex and are built from multiple layers that can be composed of different functions, such as convolutions, pooling operations, normalizations, or linear transformations. A network that does exclusively contain convolutional operations (that is only learning filter operations) is called a fully Convolutional Neural Network (CNN). These network architectures find huge success in the field of computer vision, where deep learning has become the state of the art to tackle problems such as object segmentation [Qi et al., 2017], multi-instance segmentation (like pedestrian detection in images [Zhang et al., 2018b]), human pose estimation [Toshev and Szegedy, 2014], 3D face modelling from 2D images [Richardson et al., 2016], clothing identification [Ge et al., 2019], or even detection on potentially malicious face swapping on videos [Güera and Delp, 2018].

In the field of medical imaging, the same techniques are being applied with great success [Esteva et al., 2019]. There has been a great variety of applications that include radiology tasks, such as pneumonia detection [Rajpurkar et al., 2017], brain tumor segmentation [Havaei et al., 2017], whole body bone segmentation [Klein et al., 2018, Klein et al., 2019], or free-text report classification [Chen et al., 2017]. Also, a large number of different tasks have been examined, for instance, the building of augmented reality microscopes [Chen et al., 2019], classification of age-related macular degeneration based on Optical Coherence Tomography (OCT) images [Lee et al., 2017], and applications in bioinformatics [Min et al., 2017]. However, the field of medical image computing traditionally has suffered from the lack of high-quality data to learn from due to ethical and data privacy concerns. Generally speaking, only well-trained experts have sufficient expertise to analyze and annotate medical images.

Training a deep learning algorithm and testing its performance on a training data set usually follows the workflow depicted in figure 2.16. All available data is split into a training and a test data set. In an ideal case, to avoid any unnecessary biases, the test data set is left untouched and only evaluated once to calculate the final results. The training data set is then used to adjust the parameters within the neural network, with the goal to minimize the loss function. This is done in an iterative update scheme, where for  $n_{\text{epochs}}$  *epochs*, the network is shown  $n_{\text{batches}}$

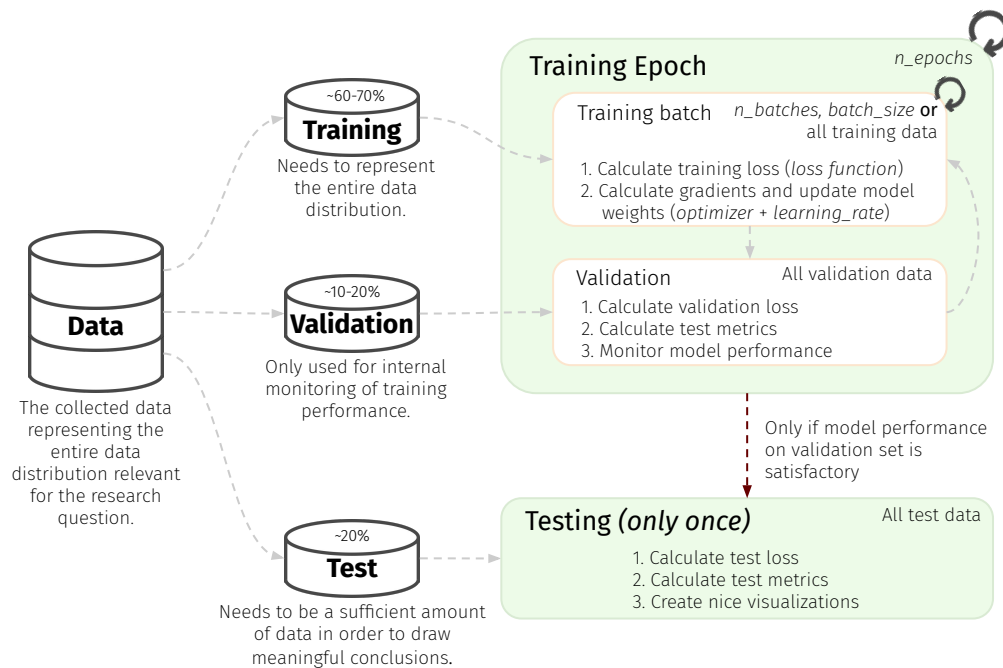


Figure 2.16: The deep learning training workflow.

batches of size  $batch\_size$  that are extracted from the training data. An epoch is usually defined as one iteration through all samples the data set, but can also be defined as a certain number of batches (which can be useful, for example, when drawing random samples from the data set for each batch). With the use of a loss function, a value that describes the performance of the neural network is calculated, and the weights are updated with the gradient of the loss in a process called *backpropagation*. A small percentage of the training data (the validation set) is held out from this iteration process and used, for example, once per epoch to supervise the learning process. This is critical in order to detect *overfitting* of the neural network to the training data. Overfitting refers to a process in which the neural network is too closely fit to the training data points, significantly impeding its ability to generalize to previously unseen data samples. This ability is an attribute of machine learning algorithms referred to as *generalization*. The final results on the validation set can be regarded as a good approximation for the model performance on the test set and can then be used to tune the hyperparameters of the training process to obtain the most optimal results.

One of the more astounding recent developments in the field of deep learning has been the implementation of *all-optical* neural networks, potentially enabling inference at the speed of light (cf. e.g. [Miscuglio et al., 2018]).

### 2.4.3 Machine learning models used in this thesis

This section introduces the specific machine learning algorithms that were used in this thesis.

#### Random forests

One of the strategies to make machine learning methods less prone to overfitting is the use of ensemble methods, which rely on aggregating over a collection of estimators. The random forest is one of these ensemble methods, where these estimators are binary regression trees. To train a regression tree, the entire training data set is split among the leaf nodes in a top-down approach. For each step down the tree, the data is split into two subsets that are distributed onto the two new leaf nodes. When a node has to consider  $n$  sorted data items for the split, it tests all options to split the data and then tries to find the optimal split based on a metric such as the variance, the sum of squared differences, or the mean squared error. This process of data splitting is usually continued until a maximum depth, or a minimum amount of data in a leaf node is reached [Wilkinson, 2004]. A simple visualization of such a regression tree is shown in figure 2.17.

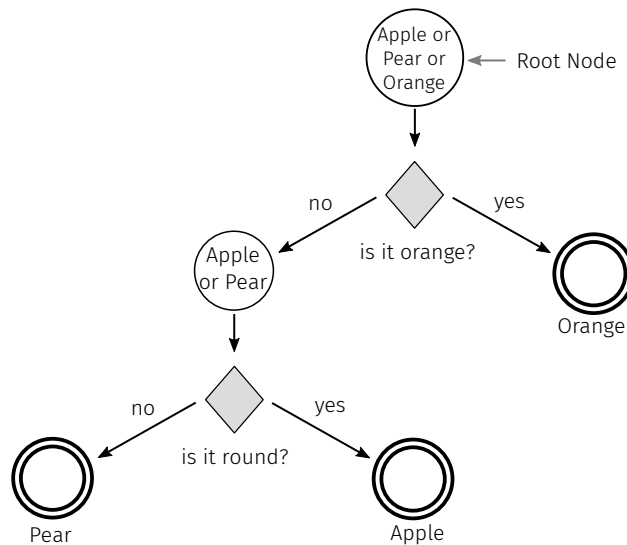


Figure 2.17: Visualization of a simple regression tree.

During inference, the trained tree is traversed to obtain a prediction. For this, the feature vector in question follows the decisions of the tree until it reaches a leaf node. At this point the estimated label  $\hat{y}$  for the feature vector is computed as the average over all  $N$  training labels in the leaf node  $l_n$ :

$$\hat{y} = \frac{1}{N} \sum_{y_i \in l_n} y_i. \quad (2.31)$$

A random forest consists of many of these decision trees. Each of the decision trees is different from the others by being randomized. This is done in a process called bootstrap aggregation, which is mainly performed by using (1) a different subset of the training data for each tree or (2)

by only allowing the tree to consider a certain fraction of the features for each split. Due to this creation of diversity, many regression trees that have significant individual biases or that have overfitted on the data might still be able to generalize quite well on average. In regression tasks, the final result of the forest is computed by the average over all estimates of the individual trees. In case of classification tasks majority voting can be performed.

### Feed-forward neural network

A feed-forward neural network is a network in which the nodes do not form cycles. The most straightforward implementation of this is a multi-layer perceptron, which consists of at least three layers of neurons: an input layer, 1 to  $n$  hidden layers, and an output layer. The energy of each neuron in this network is calculated by the weighted sum of all incoming connections, modified by a non-linear activation function, as described earlier in section 2.4.2. In this thesis, leaky ReLUs are used as the activation functions, that transform the output neuron energy  $\hat{y}$  in the following way:

$$\text{leaky ReLU}(\hat{y}) = \begin{cases} \hat{y}, & \text{if } \hat{y} \geq 0 \\ c \times \hat{y}, & \text{otherwise} \end{cases} \quad (2.32)$$

The feed-forward neural networks implemented throughout the upcoming chapters use fully connected layers, which means that the output of each neuron from a layer is connected to each input of each neuron in the next layer. Furthermore, dropout layers are used after each layer that randomly deactivate the weights of a certain percentage of neurons in a layer during the forward calculation. This is done to counteract overfitting by taking away the opportunity to rely on specific neurons during inference.

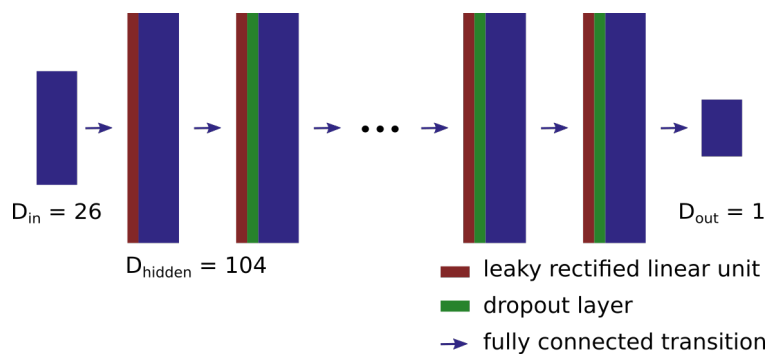


Figure 2.18: A feed-forward neural network design. The figure is reprinted from the CC-BY licensed publication [Gröhl et al., 2019].

### U-Net

The U-Net was originally developed by [Ronneberger et al., 2015] for the segmentation of biomedical images. It has since then exploded in popularity and is the best performing network architecture for numerous different use cases [Isensee et al., 2018]. The principle idea of the architecture is to use the power of convolutional layers to create data representations on increasingly high abstraction levels, while still being able to estimate label space outputs for

each pixel.

The U-Net is a fully convolutional neural network that consists out of a contracting half and an expanding half that goes through a bottleneck. On the contracting side, each layer consists out of  $3 \times 3$  convolutions, followed by a non-linear activation function and  $2 \times 2$  pooling operations for down-sampling. During each pooling operation, the number of feature channels is increased by a factor of two. In each pooling step, the expanding side of the network is also built using  $3 \times 3$  convolutions followed by a non-linear activation function, but before this, up-convolutions are used for up-sampling of the data that also decrease the number of feature channels by a factor of two. With this strategy, with each pooling step, the semantic expressivity of the network increases at the cost of spatial resolution. In order to remedy the loss of high-frequency information, skip connections are used at each pooling step, in which data from the contracting part of the network is concatenated to the data of the expanding half of the network. In the modification of the U-Net used throughout this thesis, the skip connections also contain a convolution layer that enables the input and output to be of different sizes. Furthermore, dropout layers are used throughout the network for “implicit data augmentation [Ronneberger et al., 2015]” and to increase its robustness against overfitting (cf. Figure 2.19).

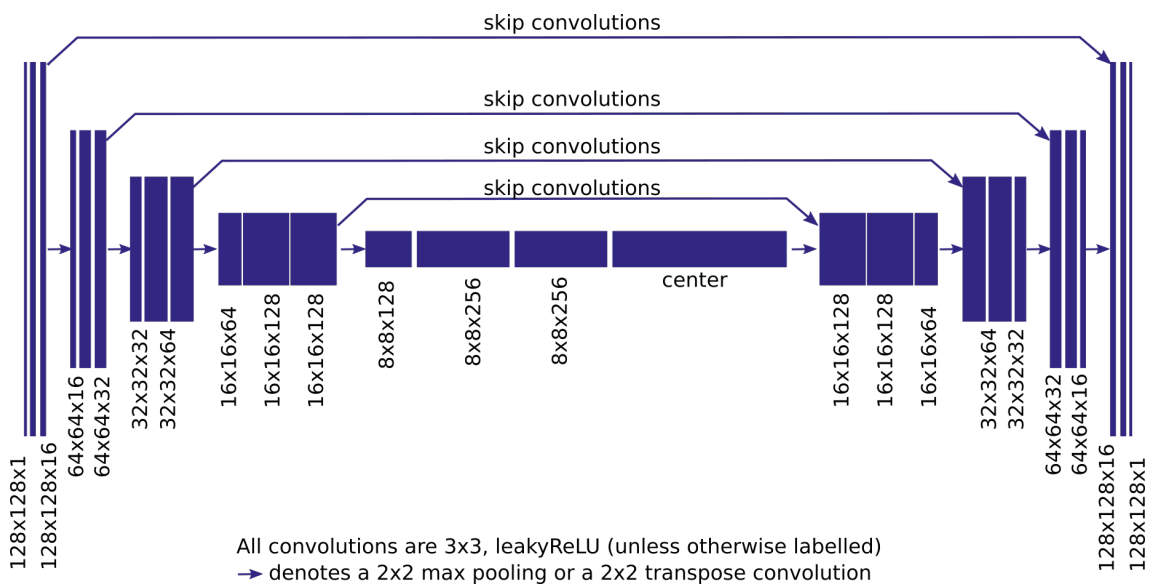


Figure 2.19: The U-Net. Reprinted from the CC-BY licensed publication [Gröhl et al., 2018a].

### Conditional Invertible Neural Networks

Invertible Neural Networks (INNs) are a particular subset of neural networks that are designed to be bijective. Such methods were developed by several groups in the field, for example, [Dinh et al., 2016, Kingma and Dhariwal, 2018, Ardizzone et al., 2018]. As such, they are by nature invertible, and some even have a tractable Jacobian determinant. In principle, INNs represent a function  $f_\theta$  that is subject to the model weights  $\theta$  and is tasked to map an input distribution  $q(x)$  into a Gaussian latent distribution  $p(z)$ . During training, a maximum likelihood loss (see equation 2.33) is used to transform  $q(x)$  into the Gaussian latent space. During inference, Gaussian samples can be taken from  $p(z)$  and can be mapped into the input domain.

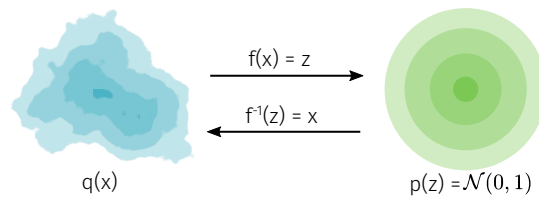


Figure 2.20: Visualization of the distribution mapping of an invertible neural network.

This is achieved with the usage of affine coupling blocks (cf. figure 2.21) in the network architecture, in which the output can be used to reconstruct the input. On a forward pass, the input data is split into two parts  $u_1$  and  $u_2$ , which are transformed by the learned functions  $s_i$  and  $t_i$ . The output is calculated by concatenating the partial results  $v_1$  and  $v_2$ . With some arrangements, the input can be recovered from the output in a backward pass, by computing the inverse operations within the affine coupling layer.

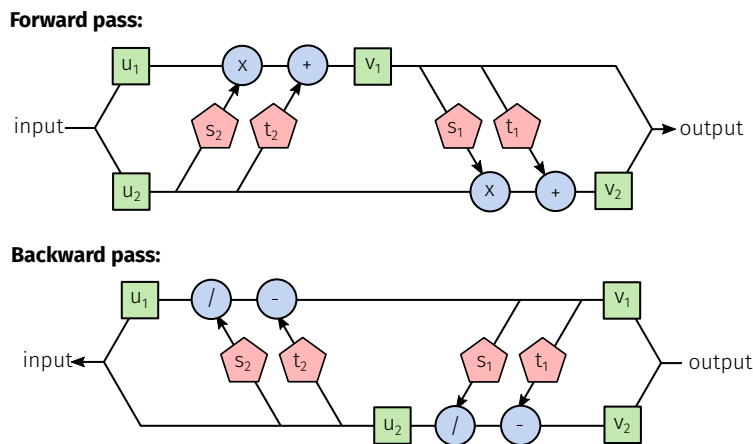


Figure 2.21: Visualization of the affine coupling layers that make up the INN architecture. Green blocks denote parts of the input and output data, blue circles denote mathematical operations, and red pentagons denote learned functions for data transformation. Here, "x" denotes element-wise multiplication and "/" denotes element-wise division.

Conditional INNs are a special class of the INN, where a conditional input distribution  $q(x|c)$  is transformed into a conditional latent distribution  $p(z|c)$ .

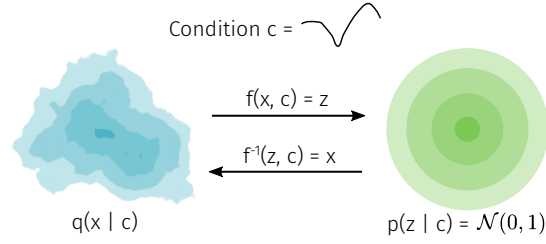


Figure 2.22: Visualization of the distribution mapping of a conditional invertible neural network. The conditional distributions are now subject to a prior condition  $c$ , which is represented by a graph in the figure.

They were introduced by [Ardizzone et al., 2019] and have the distinct advantage that they are very robust and stable during training. The authors concatenate the conditioning data to the internal functions  $s_1$ ,  $s_2$ ,  $t_1$ , and  $t_2$  of the network, which enables the forward and backward usage of the INN with the condition. During inference, when keeping a fixed condition  $c$ , the conditional INN architecture allows random sampling from the latent space  $p(z|c)$ , which enables the reconstruction of the full posterior distribution  $q(x|c)$ .

When assuming a Gaussian distribution for the latent distribution, one can arrive at the following maximum likelihood loss to train the conditional INN using standard coordinate transform theory and Bayes theorem:

$$\begin{aligned}
 q_\theta(x|c) &= p(f_\theta(x, c)) \cdot |\det J_x f_\theta(x, c)| \\
 \log(q_\theta(x|c)) &= \log(p(f_\theta(x, c)) \cdot |\det J_x f_\theta(x, c)|) \\
 &= -\frac{1}{2} \|f_\theta(x, c)\|^2 + \log |\det J_x f_\theta(x, c)|.
 \end{aligned} \tag{2.33}$$

This loss is optimized during training, where the two components of the loss can be interpreted in the way that  $-1/2 \|f_\theta(x|c)\|^2$  pulls the mass of the latent distribution towards zero, and  $\log |J_x f_\theta(x|c)|$  ensures a standard deviation of one. In order to formulate the loss for a minimization problem, one has to change the sign of the derived equation:

$$-\log(q_\theta(x|c)) = \frac{1}{2} \|f_\theta(x, c)\|^2 - \log |\det J_x f_\theta(x, c)|. \tag{2.34}$$

In the thesis, a deep conditional INN was used, in which five affine coupling blocks were concatenated. For the learned functions  $s_1$ ,  $s_2$ ,  $t_1$ , and  $t_2$  within these, shallow fully connected networks with a single 128-dimensional hidden layer were used.



## 2.5 | Related work

This section gives an overview of related state-of-the-art research papers. The goal is to give an insight into the main methodological ideas in the field but not to explain every cited paper in great detail. For further reading, one can follow the cited papers and the references within these. The section's structure is aligned with the primary fields that are studied in this thesis: model-based quantitative Photoacoustic Imaging (qPAI), machine learning-based qPAI, functional PAI, and uncertainty estimation. In the end, it is concluded how the research hypotheses correlate with the gap in the state-of-the-art research progress.

### 2.5.1 Model-based quantitative photoacoustic imaging

The field of quantitative Photoacoustic Imaging (qPAI) strives to obtain the ability to quantitatively estimate the optical properties of tissue from recorded PA images. This is highly relevant for any multispectral PA applications with the goal of estimating chromophore concentrations to yield quantitative, accurate, and reliable results. There have been several review papers describing the objectives and challenges towards qPAI, for example, [Cox et al., 2009b] and [Cox et al., 2012].

Research towards qPAI can broadly be divided into three categories: (1) methods focusing on the optical inverse problem only, (2) methods solving the acoustic and optical inverse problem sequentially, and (3) methods that try to solve both inverse problems simultaneously.

**Optical Inverse problem.** To solve the optical inverse problem, one needs to obtain estimates for  $\mu_a$  from the initial pressure distribution  $p_0$  or the deposited energy distribution  $H$ , when  $\Gamma$  is known. The first proposed approach towards qPAI was an iterative reconstruction scheme for the optical parameters (cf. figure 2.23) [Cox et al., 2005, Cox et al., 2006, Yuan and Jiang, 2006].

When attempting to solve the optical inverse problem with this reconstruction scheme, one creates an initial guess for the optical tissue properties. For each step  $k$  in the process, a forward simulation of the initial pressure or the deposited energy distribution  $H^{(k)}$  is computed. The simulation result is compared to the measured data  $\hat{H}$ . As long as the difference between these two images above a certain threshold, the optical parameter maps for simulation step  $k + 1$  are updated using  $H^{(k)}$  and  $\hat{H}$ .

Usually, solutions for the Radiative Transfer Equation (RTE) are used as the optical forward model. Alternatively, the forward simulation is conducted using the Diffusion Equation (DE) [Tarvainen et al., 2017, Pulkkinen et al., 2015, Wang and Zhou, 2017]. Because the RTE is very difficult to solve, Finite Element Models (FEMs), or other numerical solutions to the RTE are commonly seen [Yuan and Jiang, 2009, Naser et al., 2018]. In these, the image medium is discretized and divided into piecewise polynomial functions [Richling et al., 2001]. Alternatively, Monte Carlo methods have been used as the optical forward model [Hochuli et al., 2016, Liu et al., 2016, Fonseca et al., 2017, Buchmann et al., 2017b, Kaplan et al., 2017, Buchmann et al., 2019b]. Monte Carlo methods

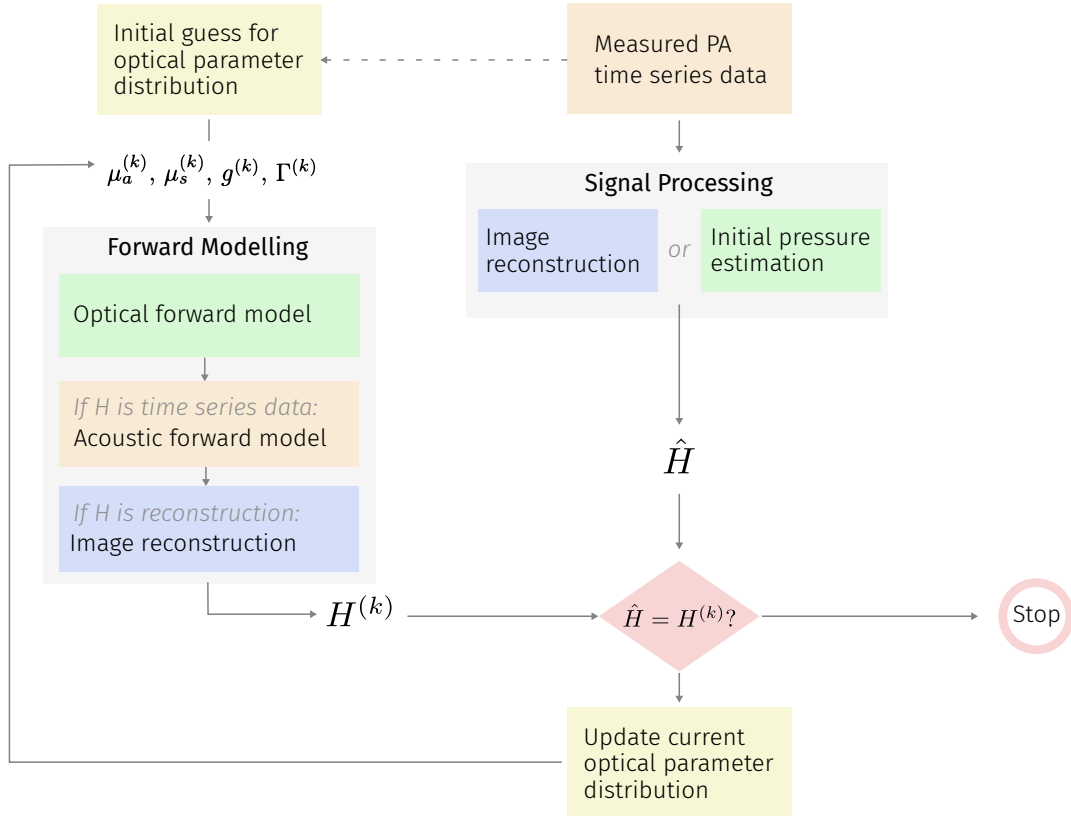


Figure 2.23: The basic principle of iterative PA reconstruction schemes. From a PA measurement, an initial guess for the optical properties is created. A synthetic image  $H^{(k)}$  is simulated using a forward model, and the initial guess is then updated with a gradient that is computed using  $H^{(k)}$  and  $\hat{H}$ .

are considered the *gold standard* for forward modeling [Cox et al., 2012], however, they can be computationally expensive for a high number of simulated photons and a large number of voxels in the computational grid. Only very recently has an implementation of this method been applied to *in vitro* phantom measurements [Buchmann et al., 2019a]. This was possible with the use of a Fabry-Pérot interferometer as a broadband detection device and careful choice of the *a priori* assumptions.

**Sequential approaches.** In these, first, the energy deposition  $H$  is calculated from the time series data. Then the absorption coefficients are estimated by estimating the fluence  $\phi$  and calculating  $\mu_a = H/\phi$ . [Banerjee et al., 2008] propose an approach to solve for  $\phi$  analytically using the Diffusion Equation (DE) by employing several simplifying assumptions (such as a diffusion coefficient that is only dependent on  $\mu_s$ , and a known and constant  $\mu_s$  in tissue). This approach was tested *in silico*, and as long as their assumptions hold, they even showed that it is theoretically feasible to get rid of the costly iteration process. Other common assumptions include modeling of the absorption distribution as a piecewise constant functional [Naetar

and Scherzer, 2014, Beretta et al., 2015]. However, the derived methods from these assumptions have also only been tested *in silico*. One also needs to be mindful of the assumptions that are made, as in some cases, the expected result has been used as the *a priori* initial guess for the reconstruction scheme [Liu et al., 2019b].

**Single-stage approaches.** Single-stage methods directly estimate optical properties from the raw time series data. The main advantage is that a direct calculation does not rely on intermediate results and, as such, might be less affected by uncertainty in the input data (for example, caused by noise). These approaches have been thoroughly investigated [Haltmeier et al., 2015, Pulkkinen et al., 2016, Pulkkinen et al., 2017, Javaherian and Holman, 2019]. The results of this body of work show that *in silico* it is feasible to directly reconstruct optical absorption and scattering distributions directly from raw time series data. However, the lack of successful translations of the proposed techniques into practice is evident.

Apart from iterative reconstruction schemes, different ideas have also been employed to try to solve the inverse problems of qPAI. A simple but effective method that is applicable *in vitro* is to account for the changes of the fluence depending solely on the depth inside the tissue sample. One approach to this is to perform fluence correction with a pre-simulated fluence map that is obtained assuming homogeneous parameters [Bu et al., 2012, Zhao et al., 2017, Perekatova et al., 2017, Vogt et al., 2019]. When the homogeneous assumptions for the simulated medium match the average optical properties of the actual imaged samples, these approaches have shown *in vitro* to be of benefit. Furthermore, this strategy has been improved on by manually creating a segmentation map of the imaged tissue first and then calculating a customized fluence distribution based on *a priori* assumptions on the segmented tissue properties [Brochu et al., 2016]. Using this strategy, the authors showed that they were able to compensate for the drop of the fluence inside of a tubular phantom *in vitro*.

Other approaches try to remedy the ill-posedness of the problem by acquiring more information for the inversion algorithm. For example, diffuse reflectance measurements can be included into the reconstruction algorithms [Nykänen et al., 2017, Bauer et al., 2011], or the influence of multiple light sources is investigated [Zemp, 2010, Song et al., 2014, Mastanduno and Gambhir, 2016, Alessandrini et al., 2017, Wang et al., 2018b]. One of these approaches - *multiple irradiation sensing* - uses several measurements with different source-detector distances to estimate the optical properties of the background tissue using the DE [Held et al., 2016].

Another direction of research is the application of acousto-optics, which is ultrasonic tagging of light and has been postulated to be able to measure the light fluence for subsequent fluence correction of the recorded signal [Hussain et al., 2018]. Here, photons that traverse the modulated region have a small probability of undergoing inelastic scattering, which causes a shift in their wavelength that is related to the sound wave frequency [Mahan et al., 1998]. This causes a shift in the speckle pattern of the light that escapes the medium, making it possible to isolate the amount of light coming from the modulated region.

### 2.5.2 Machine learning-based quantitative photoacoustic imaging

Machine learning methods have only recently become part of the research in the field of qPAI, essentially starting in mid-2017 with the pre-print publication of the paper “Context encoding enables machine learning-based quantitative photoacoustics [Kirchner et al., 2018a]” (cf. section 3.1). In addition to the efforts of data-driven image reconstruction, machine learning-based approaches have since then been used towards solving the optical inverse problem and towards more accurate spectral unmixing.

In late 2018, Cai et al. [Cai et al., 2018] have proposed to use a residual U-Net implementation to estimate the absolute concentrations of ICG and to estimate  $sO_2$  maps *in silico*. In a residual architecture, the input of a particular convolutional substructure is again concatenated to its output. To validate the method, the authors create a multispectral *in silico* data set with a circular structure of 1 cm radius at 5 nm spaced wavelengths from 700 nm to 800 nm. The method is also validated on another *in silico* data set extracted from a digital mouse phantom (Digimouse [Dogdas et al., 2007]).

In work from Hoffer et al. [Hoffer-Hawlik and Luke, 2019] in mid-2019, a U-Net has been used for the estimation of tissue oxygenation *in silico* as well. Here, an *in silico* data set with 125 phantoms is generated at two wavelengths (700 nm and 900 nm), and the method is validated on another *in silico* data set. Later, in November 2019, an adaptation of this method was published [Luke et al., 2019] for simultaneous  $sO_2$  estimation and vessel classification by using two adjacent U-Net pathways. This architecture is referred to as the *O-Net* by the authors.

Yang et al. [Yang et al., 2019] have also presented a neural network approach to *in silico*  $sO_2$  estimation with an Encoder-Decoder architecture (EDA-Net) in October 2019. Here, the authors use a numerical breast phantom to simulate initial pressure distributions using Monte Carlo methods and raw time series data with the k-Wave toolbox. The method is validated on a held-out test set of the simulated data set.

### 2.5.3 Functional photoacoustic imaging

As already eluded to in section 2.2.5 and as evident in the previous section, the recovery of functional tissue parameters is one of the primary objectives of multispectral PAI. The linear spectral unmixing method for  $sO_2$  recovery from PA signals is analogous to spectral unmixing techniques from the field of hyperspectral imaging and remote sensing [Keshava and Mustard, 2002, Keshava, 2003]. Here, one tries to find the ratio of known extinction spectra of chromophores [Laufer et al., 2006] by calculating the best fitting linear mixture of the *a priori* assumptions to the measured signal. In the past decade, the linear unmixing method has been rigorously investigated, and many improvements to the linear unmixing method have been proposed [Tzoumas et al., 2013, Li et al., 2018].

One of these improvements is to impose the constraint to the unmixing algorithm that the result can only feature non-negative contributions [Ding et al., 2017a] because negative chromophore concentrations are physiologically impossible. A multispectral application of the classical iterative qPAI reconstruction methods for direct or indirect  $sO_2$  recovery has also been investigated [Cox et al., 2009a], as well as the application of non-linear unmixing models for example based

on minimizing the mutual information between the independent chromophores [An et al., 2017]. Mutual information is an approach that measures the degree of dependence  $I(A, B)$  of two distributions  $A$  and  $B$ , which can, for example, be calculated by  $I(A, B) = H(B) - H(B|A)$ , where  $H(B)$  is the entropy of  $B$  and  $H(B|A)$  is the conditional entropy of  $B$  given  $A$  [Maes et al., 1997]. Another approach presented by [Radrić and Ntziachristos, 2016] uses measurements of isosbestic points of chromophores to account for the effects of scattering in tissue.

Other methods try to include different sources of information into the unmixing minimization algorithm. For instance, Gao et al. [Gao et al., 2015] propose to estimate tissue  $sO_2$  while using only one wavelength by also using information obtained from simultaneous diffuse reflectance measurements. Ulrich et al. [Ulrich et al., 2019] published a method for spectral correction that is based on using near-infrared optical imaging in reflection mode as additional information for the inversion process.

A promising approach to spectral unmixing was the eigenspectra Multispectral Optoacoustic Tomography (eMSOT) approach proposed by [Tzoumas et al., 2016] and [Olefir et al., 2018]. Here, the principle idea is to calculate the eigenvectors of pixel-wise fluence distributions in tissue and to add these eigenvectors to the inversion process to represent the fluence term in  $p_0 = \mu_a \cdot \phi$ . However, because the solution space of this formulation of the problem has a lot of local minima, the minimization needs to be heavily constrained in order to predict correct results.

#### 2.5.4 Uncertainty estimation

Uncertainty quantification and its incorporation in medical use cases is an essential research objective in computer sciences and has been studied extensively in various fields, including image-guided navigation [Maier-Hein et al., 2011, Alterovitz et al., 2008], multi-modal image registration [Sykes et al., 2009, Risholm et al., 2013], and lesion detection [Nair et al., 2019].

Some of the principal goals of uncertainty estimation in science are (1) to have the ability to assign a quantity to a measured value that represents the standard (expected) error of this value and (2) to resolve ambiguity in the measured quantity and gain the ability to identify multiple modes in the value distribution [Adler et al., 2019a]. In a standard, highly calibrated measurement setup, the goal is to measure all possible sources of uncertainty independently and then use these to calculate the standard error of the final result (for example, using error propagation). With data-driven approaches, knowledge on the uncertainty of the input data acquisition process and of ambiguity in the label space is not always present. As such, methods need to be investigated, which can yield such metrics of uncertainty for the estimated values.

Current popular methods for obtaining confidence intervals or uncertainty estimated for neural network predictions are all principally related and include ensemble methods, Bayesian models, dropout sampling, or the sampling from a latent space (see figure 2.24).

**Ensemble methods.** Ensemble methods combine many machine learning models to derive single estimates (e.g. [Lakshminarayanan et al., 2017, Smith and Gal, 2018]). The random forest (cf. section 2.4.3) is a prominent example of an ensemble technique, where the predictions of many differently trained decision trees are combined into a single more robust estimate. Having access to various different estimates also allows for the calculation of uncertainty statistics, such as the

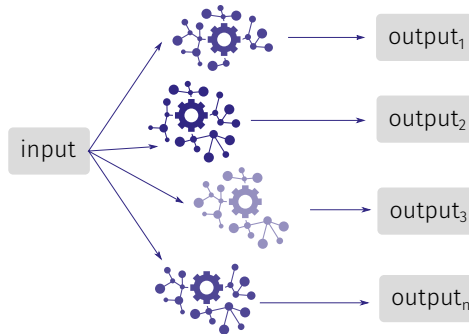
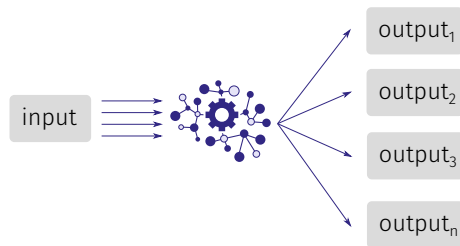
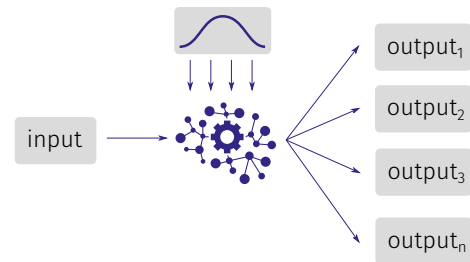
**a. Ensemble methods****b. Bayesian models****c. Dropout sampling****d. Latent space sampling**

Figure 2.24: Overview of how diverse output samples are generated with different uncertainty estimation methods. **a.** shows ensemble methods, where each model in the ensemble generates an output, **b.** is a visualization of how Bayesian models estimate a probability function for the output, **c.** shows how many samples are generated by deactivating different neurons on each pass through the network, and **d.** shows how the drawing of random samples from a latent space can lead to diverse output estimates.

standard deviation of the estimates.

**Bayesian models.** For Bayesian inference, the goal is to model the network weights as probability distributions and not as a deterministic variable. Creating such machine learning models that can be trained and evaluated in polynomial time has proven to be difficult and, as such, approximations (such as dropout sampling) are needed [Miller et al., 2018]. Related literature includes e.g. [Feindt, 2004, Zhu and Zabararas, 2018, Kohl et al., 2018].

**Dropout sampling.** In dropout sampling, a certain number of neurons get deactivated during each pass through the network (see e.g. [Srivastava et al., 2014, Li and Gal, 2017, Lebig et al., 2017]). Gal et al. [Gal and Ghahramani, 2016] suggested that performing dropout during training is “mathematically equivalent to an approximation to the probabilistic deep Gaussian process [i.e. Bayesian inference]”, which has been met with criticism towards the technical correctness of this view (cf. e.g. [Hron et al., 2017]). In practice, one of the major benefits of using dropout

is that it also forces the network to learn a robust and more generalized representation of the input data and to create redundancies within the network, as it cannot rely on the individual weights of a single neuron.

**Latent space sampling.** The latent space refers to a feature representation in the neural network. In Autoencoders [Ng, 2011], the term latent space typically refers to the bottleneck layer, but it might refer to the feature representation at any level. In some work, for example, in variational autoencoders [Kingma and Welling, 2013], it is attempted to subdivide this bottleneck layer into a representation for the mean estimate and the standard deviation of the estimate. This allows taking different samples from the latent representation of the standard deviation for the decoding part of the network and, as such, to obtain different inference results for the same input feature vector. This principle idea of introducing variation in a specific feature representation of the network for reconstruction is also used in other approaches, such as [Kingma and Welling, 2013, Mescheder et al., 2017, Ardizzone et al., 2018, Ardizzone et al., 2019, Adler et al., 2019b].

### Uncertainty estimation in photoacoustic imaging

The sensitivity, stability, and uncertainty of qPAI inversion schemes have been investigated [Fonseca et al., 2016, Bonnetier et al., 2019, Tick et al., 2016], mostly in the framework of Bayesian modeling of the optical or acoustic inversion [Tarvainen et al., 2013, Tick et al., 2018, Hänninen et al., 2018]. Bayesian techniques yield a posterior distribution for the properties in question and thus enable the recovery of uncertainty estimates. These investigations show that uncertainties in *a priori* assumptions can negatively impact model-based inversion schemes and that adding more degrees of freedom to the inversion method can significantly improve inversion accuracy. It is further shown that inversion stability deteriorates exponentially with increasing depth within the medium.

### 2.5.5 Conclusions

Qualitative PAI has had great success featuring numerous clinical applications. However, most of this work either correlate metrics with arbitrary unit PA signal intensities or use linear spectral unmixing techniques to obtain estimates of the oxygen saturation of tissue. For the clinical translation of PAI, this poses a problem, as these practices fall short of the theoretical potential of the imaging technique and might cause sub-optimal, if not wrong, diagnosis and treatment of patients.

As evident from the literature overview in the previous sections, much effort has been put into the development of techniques for qPAI. The most promising of these is the iterative approach proposed by Cox et al. in 2005 [Cox et al., 2005] which has been extensively validated *in silico* and also *in vitro* [Buchmann et al., 2019b]. Other approaches try to circumvent the quantification problem by directly working with multispectral data. With the use of sophisticated unmixing algorithms that, for example, take into account natural constraints on the physiological value ranges or the underlying fluence distributions, these methods show great promise in achieving more accurate  $sO_2$  estimation results compared to naïve linear unmixing techniques. Nevertheless, despite the efforts, the translational gap of application of qPAI in clinical practice has yet to be overcome. Table 2.1 summarizes this translational gap of the developed techniques for both  $\mu_a$  quantification as well as direct  $sO_2$  unmixing techniques.

	<i>in silico</i>	<i>in vitro</i>	<i>in vivo</i>	
			small animal	clinical
accurate $sO_2$ estimation	✓ cf. e.g. [Laufer, 2006] [Yang 2019]	✓ cf. e.g. [Held, 2016] [Gehring, 2019] [Li, 2018] [Vogt, 2019] [Ulrich, 2019]	(✓)	✗
accurate $\mu_a$ quantification	✓ cf. e.g. [Cox, 2005] [Cox, 2009b] [Pulkkinen 2015] [Yuan, 2006] [Bannerjee, 2008] [Bu, 2012] [Yang, 2019]	(✓) cf. e.g. [Buchmann, 2017] [Buchmann, 2019]	✗	✗

Table 2.1: Tabulated, impertinent view of a selection of papers presenting methods to tackle the  $\mu_a$  quantification and  $sO_2$  estimation problems at the different translational stages from simulated data (*in silico*) over phantom measurements (*in vitro*) to measurement in living organisms (*in vivo*).

While very first advances have been made to apply model-based quantification schemes to *in vitro* settings, the accuracy with which the unknown parameters have to be determined in order to obtain accurate results has hindered fast advancement of these methods. Other methods for functional PA imaging (for example, the eMSOT method) have been proven to work exceptionally well in certain scenarios, limited by the need for specific constraints for the inversion and have yet to be shown to work in a general fashion. The potential use of uncertainty estimation methods towards the inverse problems of qPAI is also vastly unexplored, and it should be investigated how estimations of the inversion uncertainty or the respective standard error could be used to either improve the inversion results or to improve patient care.

Data-driven approaches, on the other hand, show great promise to remedy at least some of the shortcomings of the model-based approaches but suffer from a lack of ground truth data, and the absence of well-curated (even unlabeled) data sets to learn from reliably. As such, all work in this direction has exclusively been conducted *in silico*.

Being faced with this gap, and keeping in mind the key challenges of qPAI (C1-C4 presented in



section 1.2), the key objectives of this thesis are (1) to explore the potential of data-driven methods towards bridging the translational gap of qPAI and (2) to also explore the use of state-of-the-art uncertainty estimation techniques in this context.



Absolute **quantitative PA measurement** of oxygenation and temperature in deep tissue **remains a challenge**, mainly due to the **unknown local optical fluence** [...]. New imaging **methods and mathematical models** are needed to better map the optical properties of the tissue.

[Wang and Yao, 2016]

# 3 | Results

This results chapter is the central part of the thesis. It comprises four sections where each introduces experiments that were conducted in pursuit of the three hypotheses that were constructed in the introduction (chapter 1):

- H1:** Data-driven methods can be used to solve the optical and acoustic inverse problem.
- H2:** Data-driven methods can be used for spectral unmixing in a realistic context.
- H3:** Confidence estimation methods be used to gain insight into qPAI methods.

Section 3.1 is aligned to research objective **H1** and introduces a data-driven method for the optical inverse problem of qPAI by estimating the local light fluence in tissue. The core idea of the method is to encode the acquired signal and *a priori* information of the illumination geometry in a hand-crafted feature vector individually for each voxel in the imaging plane. The assumption is that this feature vector can be used to generate many unique training samples from a single 3D volume and that it can easily be calculated for a newly measured signal during inference.

Section 3.2 is also aligned research objective **H1** and investigates the feasibility of employing state-of-the-art end-to-end deep learning algorithms towards the inverse problems of qPAI. The core idea is to use the U-Net and apply it to investigate the optical and acoustic inverse problems in isolation *and* as a combined problem. In contrast to the previous section, the methods developed here do not consider single voxel feature vectors, but instead, process entire 2D images at once.

Section 3.3 introduces a method towards research objective **H2**: the accurate spectral unmixing of multispectral PA images. The goal is to achieve quantitative blood oxygenation measurements

directly from multispectral initial pressure distributions that suffer from *spectral coloring*. The core idea of the method is to create a data set on which a machine learning model can learn to account for the wavelength-dependent non-linear variations of the fluence on the multispectral initial pressure distribution.

Finally, in line with research objective **H3**, section 3.4 investigates the applicability of state-of-the-art uncertainty estimation methods to the estimation of parameters from PA images. Several different methods are introduced and systematically compared. The core idea is to investigate the capability of the methods to improve on the accuracy of the aggregated result when strategically removing an increasing amount of the most uncertain estimates from a predefined region of interest.

[Photoacoustic] quantification efforts focus on retrieving the **light fluence distribution**, which is unknown and depends on the distribution of **both absorption and scattering** coefficients within the sample.

— [Brochu et al., 2016]

## 3.1 | Quantitative photoacoustic imaging with context encoding

### Disclosures to this work:

The core idea of this work has been developed by Lena Maier-Hein, Thomas Kirchner, and myself during my master's thesis [Gröhl, 2016]. During my Ph.D., a more rigorous validation scheme was implemented, and more diverse data sets were simulated. Furthermore, the feasibility experiments were also conducted using deep learning for the inversion. Parts of this work have been published in the *Journal of Biomedical Optics* [Kirchner et al., 2018a].

### 3.1.1 Introduction

The optical absorption coefficient of tissue is a highly interesting property, as knowledge of it allows for spectroscopic analysis of functional tissue parameters, such as the blood oxygenation  $sO_2$ , which is a clinically relevant parameter of the health status of a patient. While the Photoacoustic (PA) signal  $S$  is dependent on the optical absorption  $\mu_a$  and as such - in theory - enables measurements of it,  $S$  is also dependent on the Grüneisen parameter  $\Gamma$  and the fluence  $\phi$  leading to  $S \propto \mu_a \cdot \Gamma \cdot \phi$ . The fluence, however, is also non-linearly influenced by  $\mu_a$ , which makes the estimation of  $\phi$  or  $\mu_a$  from  $S$  an ill-posed inverse problem.

When the Grüneisen parameter is assumed to be constant throughout tissue, the robust and accurate estimation of  $\phi$  is the most crucial aspect for achieving quantification of PA signals. Challenges related to the associated inverse problem include *spectral coloring* and the *absorption-scattering non-uniqueness* [Brochu et al., 2016]. Here, *spectral coloring* is caused by the non-linear wavelength dependent and spatially varying behavior of  $\phi$ . The *absorption-scattering non-uniqueness* refers to the inherent ambiguity in PAI, where different combinations of optical absorption and scattering can lead to the same fluence distribution or the same measured PA signal.

In order to be able to estimate  $\phi$  based on a measured PA signal, the entire 3D image context has to be considered in order to be able to account for the distribution of light within the medium correctly. In the case of 2D handheld clinical PA transducers, this is a major challenge, as only 2D tomographic slices can be acquired. Out-of-plane chromophores can have a profound influence

on the local  $\phi$  in the imaging plane, which can be amplified depending on the nature of the illumination geometries.

To tackle this problem, a data-driven method Context Encoding Quantitative Photoacoustic Imaging (CE-qPAI) is presented that can estimate the light fluence by using a 3D signal image created from a series of 2D scans of the medium. The core principle of CE-qPAI is the development of a hand-crafted feature vector (the Context Image (CI)) that encodes the signal and with respect to *a priori* knowledge of the illumination geometry for each specific voxel in the imaged medium.

Given the lack of ground truth knowledge of the optical parameters in a complex medium *in vivo*, the difficulty of fabricating acoustically and optically realistic tissue phantoms, the complexity of obtaining reliable reference measurements for the optical properties *in vitro*, an *in silico* study to determine the feasibility of the envisioned approach is conducted. In a series of *in silico* experiments, it is examined whether it is feasible to use CE-qPAI to estimate local fluence on a voxel-wise level with only a few hundred simulated tissue structures.

#### Hypothesis investigated in this chapter

*H1: Data-driven methods can be used to solve the optical and acoustic inverse problem.*

The main focus of the work presented in this section is to examine the general feasibility of the application of data-driven approaches to the optical inverse problem of qPAI. To this end, two main research questions are addressed:

- 1) Is it possible to encode the voxel-specific 3D context into a single low-dimensional feature vector to derive a voxel-wise estimate of the local fluence?
- 2) Is the calculation of voxel-wise feature vectors from only a limited number of simulated volumes sufficient for generalization given the large possible parameter space of the distribution of optical properties?

### 3.1.2 Methods

The proposed method is the first to apply machine-learning methods to the optical inverse problem of qPAI. It uses hand-crafted feature vectors with corresponding labels to learn the light fluence  $\phi$  on a voxel level. The inferred knowledge on the local  $\phi$  can then be used to correct the initial pressure distribution  $p_0$  to deduce the corresponding optical absorption  $\mu_a$ .

The core principle of CE-qPAI is the introduction of a hand-crafted voxel-based low-dimensional feature vector, the Context Image (CI), that encodes the complete context information of the target signal voxel together with *a priori* characteristics of the imaging system in a single feature vector.

This is done by relating the recorded initial pressure distribution with *a priori* created simulations on the distribution of light in a homogeneous medium, represented by Fluence Contribution

Maps (FCMs) (see section 3.1.2). Because a CI is calculated for each voxel of the central imaging plane of a 3D sweep, this enables tackling the challenge of fluence estimation as single voxel regression within the framework of machine learning techniques and allows for the generation of large amounts of training samples from the relatively few simulations that can be computed in a reasonable amount of time.

During the training phase, a regressor is presented the voxel-specific context images  $CI(v)$  as feature vectors that are labeled with the local fluence  $\phi(v)$ . These data pairs are calculated for every voxel  $v$  for each volume in the training data. During inference, the regressor is then presented a context image  $CI(v)$  and is tasked to estimate the corresponding local fluence  $\phi(v)$ .

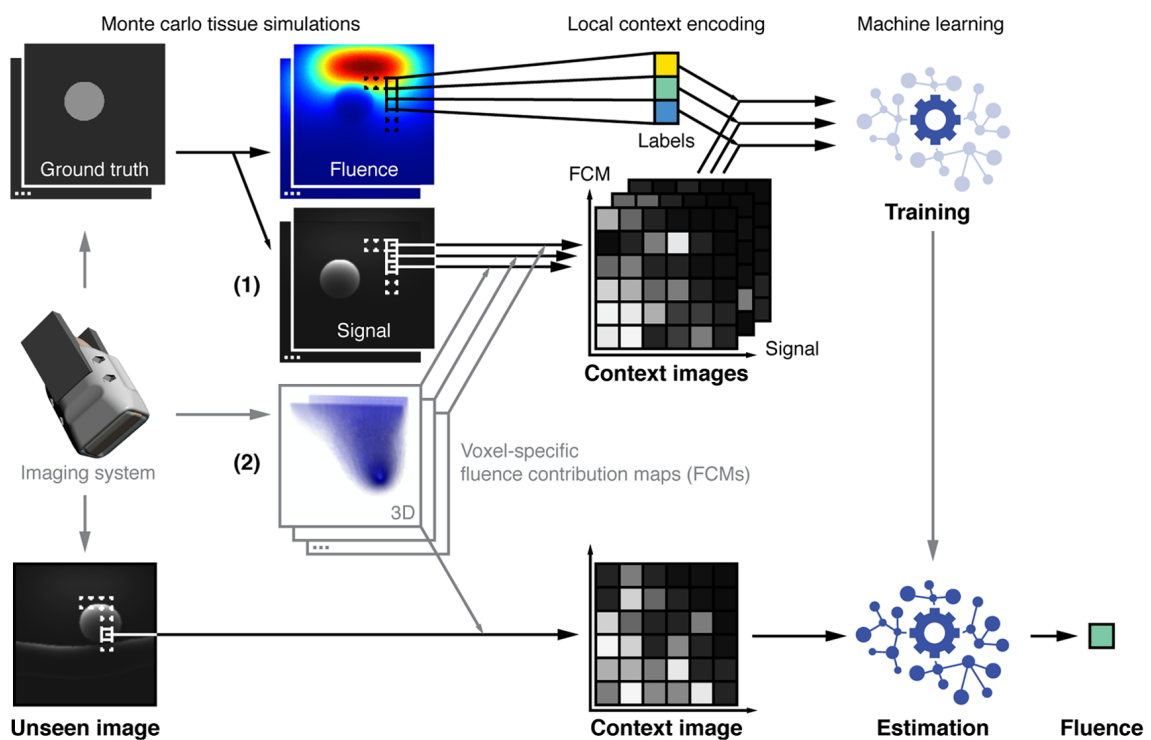


Figure 3.1: Context encoding approach to qPAI. CIs are calculated for each voxel in the imaging plane using both (1) the 3D signal context and (2) *a priori* knowledge on the imaging system. A regression algorithm is trained with tuples of CIs and corresponding fluence values. Reprinted with permission from the CC-BY licensed publication [Kirchner et al., 2018a].

Figure 3.1 explains the core principle of the CE-qPAI method.

For training of the algorithm and using a realistic PA illumination geometry as a template, 3D signals of initial pressure and fluence are generated using a Monte Carlo simulation of light propagation for many different absorption coefficient distributions (1). Furthermore, using the same illumination geometry, voxel-specific Fluence Contribution Maps (FCMs) are simulated for each voxel in the imaging plane (2). For each of the simulated 3D volumes, the signal context is combined with all FCMs to create Context Images (CIs) for each voxel of the imaging plane.

The CIs are then labeled with the corresponding fluence values obtained from the Monte Carlo simulation and presented to the machine learning algorithm for training. With this process, numerous training examples (as many as there are voxels in the imaging plane) can be extracted from each of the 3D signals.

During inference, a previously unseen 3D image is measured with the PA device that was used as the template in the simulation. Using the same FCMs from the training process, CIs are generated for each voxel of the central imaging plane of the measured volume. These CI are then presented to the trained machine learning algorithm in order to obtain estimates for the light fluence in the medium of the unseen image.

### Fluence Contribution Map

An Fluence Contribution Map (FCM) is an artificial construct that is designed to contain *a priori* knowledge on how light propagates in a homogeneous medium. An FCM always corresponds to a specific voxel  $v$  in the volume and stores information on the fluence of all photons remaining when discarding those that do not traverse voxel  $v$  in their path through the medium. It can also be interpreted as the probability map for  $v$ , which contains the probability of a photon to traverse any given voxel  $v'$  in the volume before it reaches the detector voxel  $v$ . As such,  $FCM_v(v')$  is the probability that a photon arriving in  $v$  has traversed  $v'$ .

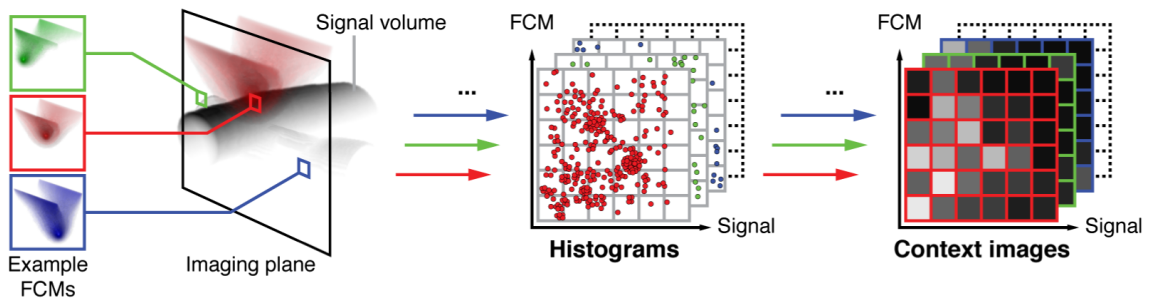


Figure 3.2: Visualization of the calculation of CIs for three voxels ( $v_1, v_2, v_3$ ). The FCMs contains information on how light propagates in a homogeneous medium. For each voxel  $v_i$ , all tuples of 3D signal and FCM are determined and combined into the CI histogram. This figure was reprinted with permission from the CC-BY licensed publication [Kirchner et al., 2018a].

### Context Image

The Context Image (CI) corresponding to a voxel  $v$  is a histogram that contains value pairs of corresponding FCMs and PA signals  $S$ . In other terms, the CI is a set of the data pairs of  $S$  and FCM for all voxel positions  $v'$ :  $\{(S(v'), FCM(v')) | v' \in N(v)\}$ . Here, it becomes evident that  $S$  is the same for each CI for the volume. The FCMs, however, are individual for each voxel location, leading to variations in the calculated CIs.  $N(v)$  is the neighborhood around  $v$  where the FCM probability that photons traversed it before reaching  $v$  is above a certain threshold  $\epsilon$ . As such,  $N(v)$  is defined as  $\{v' | FCM_v(v') > \epsilon\}$ , and its purpose is to save valuable computational



resources by excluding voxels  $v'$  that would only have a negligible contribution to the fluence in  $v$ . Figure 3.2 illustrates how the FCMs and signal volume are used to create the CIs.

### Experimental design

The presented CE-qPAI method was tested in an *in silico* experiment, which had the goal to estimate the fluence distribution of 3D volumes composed from 2D initial pressure  $p_0$  slices simulated with a handheld transducer illumination geometry as described in section 2.3.1. To this end, *in silico* volumes were simulated that consisted out of vascular structures in a homogeneous background medium. To approximate partial volume effects, the distribution of absorption coefficients was post-processed using a Gaussian blur with a sigma of 0.6 mm. Every simulation was performed using the Monte Carlo method for photon transport (in this case using mcxyz by Steven L. Jacques [Jacques, 2014]) with  $2 \times 10^6$  photons for all training sets and  $10^8$  photons for the respective test and validation sets.

**Training data simulation** Variations on four different parameters were examined: the vessel radius, number of vessels, absorption coefficient of the vessel structures, and the absorption coefficient of the background medium. From these parameters, six different data sets were created: one with no variation, one with variations of all of the parameters, and four where each of the parameters was varied in isolation (see Table 3.1). For each volume, the values were randomly sampled from a uniform distribution in the specified range. The optical absorption ( $\mu_a$ ) parameter ranges used in the respective data sets were chosen to reflect the entire range of absorption coefficients found in physiological tissue [Jacques, 2013]. A constant scattering ( $\mu_s$ ) parameter of  $15 \text{ cm}^{-1}$  and an anisotropy ( $g$ ) of 0.9 were used throughout all data sets. Furthermore, a constant Grüneisen parameter of 1 was assumed in the data set, effectively disregarding it for the conducted experiment.

**Tissue and Geometry Parameters**

data set	radius [mm]	vessel $\mu_a$ [ $\text{cm}^{-1}$ ]	num vessels	backg. $\mu_a$ [ $\text{cm}^{-1}$ ]
DS <sub>base</sub>	3	4.7	1	0.1
DS <sub>radius</sub>	<b>0.5 - 6</b>	4.7	1	0.1
DS <sub>absorb</sub>	3	<b>1 - 12</b>	1	0.1
DS <sub>vessel</sub>	3	4.7	<b>1 - 7</b>	0.1
DS <sub>backg.</sub>	3	4.7	1	<b><math>10^{-4}</math> - 0.2</b>
DS <sub>multi</sub>	0.5 - 6	1 - 12	1 - 7	$10^{-4}$ - 0.2

Table 3.1: The design parameters of the *in silico* data sets (DS) for the quantitative evaluation of the CE-qPAI method. DS<sub>base</sub> is the baseline data set with minimal variations, whereas DS<sub>radius</sub>, DS<sub>absorb</sub>, DS<sub>vessel</sub>, and DS<sub>backg.</sub> each introduce variations in one of the design choice parameters. Finally, DS<sub>multi</sub> contains variations of all of the parameters.

The simulated data were divided into a training set with  $N_{\text{train}} = 150$  volumes for training of the respective regression algorithm, a validation set with  $N_{\text{validation}} = 25$  volumes for tuning of

the hyperparameters and supervision of the convergence of the training process, and a test set with  $N_{\text{test}} = 25$  volumes on which the final results were calculated. Each of these 3D volumes was composed of 31 2D scans resulting in a size of  $64 \times 47 \times 62(x \times z \times y)$  with an isotropic spacing of 0.6 mm per voxel. Here  $x$  is the axis along the US transducer elements,  $z$  is the depth axis, and  $y$  is the axis orthogonal to the imaging plane. To reduce the number of needed forward simulations, each simulated 2D x-z plane was added twice into the composed 3D volume. Furthermore, the number of training volumes was increased to 400 for this data set to account for the high complexity of variation in  $DS_{\text{multi}}$ .

The background signal level of the simulated data was  $(4.2 \pm 2.8)$  [a.u.]. Several different noise levels were applied to the data sets to investigate the robustness to noise. In this context, three different Gaussian noise models were used:

1. 2% multiplicative and  $(0.125 \pm 0.125)$  a.u. additive noise
2. 10% multiplicative and  $(0.625 \pm 0.625)$  a.u. additive noise
3. 20% multiplicative and  $(1.25 \pm 1.25)$  a.u. additive noise

In cases where negative signals were produced during the application of the Gaussian noise, these values were set to 0. The test set results were evaluated twice: first on all voxels of the image and then only in an Region Of Interests (ROIs) containing voxels corresponding to vessel structures and having a signal higher than a noise equivalent threshold (calculated using a Contrast-to-Noise Ratio (CNR)  $> 2$  [Welvaert and Rosseel, 2013]).

**CI and FCM parameters** For all experiments, FCMs were simulated using the background optical parameters of  $\mu_a = 0.1$ ,  $\mu_s = 15$ , and  $g = 0.9$ . To achieve a similar photon count over all detector voxels  $v$ , the number of simulated photons was varied depending on the depth of  $v$ . The CIs were created using 12 logarithmically scaled bins corresponding to the signal and FCM axes. The value ranges of the bins were set to be  $0 < \log(S) < \log(255)$  and  $\log(\epsilon) < \log(\text{FCM}) < -1$ , with  $\epsilon$  being set to  $10^{-5}$ .  $S$  and FCM values larger than the upper boundary were included in the last bin, whereas  $S$  and FCM below the lower threshold were excluded.

**Regression methods** In the experiments, two machine learning methods are examined. The first method was a *random forest* [Breiman, 2001], which, as detailed in section 2.4.1, is a supervised learning method, in which a collection of decision trees is formed based on random subsets of the training data. For this implementation, the python scikit-learn random forest implementation [Pedregosa et al., 2011] with default hyperparameters and 100 regression trees was used.

The second method was a fully connected feed forward neural network, as detailed in section 2.4.2, with a total of four layers containing leaky rectified linear units and dropout connections to prevent overfitting. For training, the network was given the same computational budget of 100 epochs, with 1000 batches per epoch and 1000  $(CI_s, \phi)$  pairs per batch on all data sets. The pytorch [Paszke et al., 2017] Adam optimizer [Kingma and Ba, 2014] with a learning rate of  $10^{-4}$  and an  $L_1$ -loss function were used as the training hyperparameters.

### CE-qPAI for functional parameter estimation

A second experiment was designed to determine whether it would be feasible to use the proposed CE-qPAI method to directly estimate blood oxygenation ( $sO_2$ ) from CIs. In order to extend the

CE-qPAI method to a functional use case, multispectral PA images have to be acquired. For this experiment, the wavelengths 750, 800, and 850 nm were chosen.

A new data set  $DS_{oxy}$  was simulated that consisted of 240 multispectral training volumes with homogeneous oxygenation values randomly drawn from a uniform  $sO_2$  distribution  $U(0\%, 100\%)$ . The test data consisted of 11 volumes, each simulated with 11 homogeneous oxygenation levels (0%, 10%, ..., 90%, 100%). The optical absorption coefficients of the background and the vessels were calculated based on blood volume fraction, the wavelength, and the tissue oxygenation, as described by [Jacques, 2013]. The blood volume fraction was chosen to be 0.5% in the background medium and 100% in blood vessels. Hemoglobin concentration was assumed to be 150 g/L. A constant scattering ( $\mu_s$ ) parameter of  $15 \text{ cm}^{-1}$  and an anisotropy ( $g$ ) of 0.9 were used throughout all volumes.

Blood oxygenation was estimated using three different methods: (1) linear spectral unmixing (cf. section 2.2.5) of the initial pressure data as a baseline. (2) Linear spectral unmixing of the signal after correction by the CE-qPAI estimated fluence. (3) Direct estimation of  $sO_2$  using an adaptation of CE-qPAI, where CIs at the same spatial location of all three wavelengths were stacked into one feature vector and assumed contain the multispectral information needed for the inversion. During training, the algorithm was then presented tuples  $((CI_{750}(v), CI_{800}(v), CI_{850}(v)), \phi(v))$ .

### 3.1.3 Results

#### Monospectral random forest regression

Table 3.2 summarizes the descriptive statistics of the relative fluence estimation errors for the experiments on absorption estimation using random forest regression. As the relative fluence estimation error does not follow a normal distribution, the median and Interquartile Ranges (IQRs) are reported for all data sets. The results show that even for the most complex data set  $DS_{\text{multi}}$  with variations of multiple parameters, CE-qPAI yields a median overall relative fluence estimation error below 3%.

data set	All voxels		ROI voxels	
	Median	IQR	Median	IQR
$DS_{\text{base}}$	1.1	(0.5, 2.0)	4.4	(2.0, 8.0)
$DS_{\text{radius}}$	1.5	(0.6, 3.4)	5.8	(2.5, 11.9)
$DS_{\text{absorb}}$	1.3	(0.6, 2.9)	13.8	(5.2, 31.8)
$DS_{\text{vessel}}$	1.9	(0.7, 6.0)	6.9	(3.1, 13.5)
$DS_{\text{backg.}}$	0.8	(0.4, 1.6)	4.3	(2.0, 7.6)
$DS_{\text{multi}}$	2.3	(0.8, 19.8)	14.0	(6.1, 30.8)

Table 3.2: Descriptive statistics of all CE-qPAI results using a random forest as the regressor. For each of the data sets, the median error, as well as the IQR are shown for all voxels and exclusively on voxels in the ROI.

Previously proposed qPAI approaches reveal high drops in estimation performance when dealing with noisy data (cf. e.g. [Beretta et al., 2015]). To remedy this, methods have been proposed to incorporate more accurate noise representations into model-based reconstruction algorithms [Tarvainen et al., 2013, Tarvainen et al., 2017]. The validation of CE-qPAI under different levels of noise shows that it yields stable accuracy even under unrealistically high noise levels of up to 20% multiplicative noise (cf. figure 3.3).

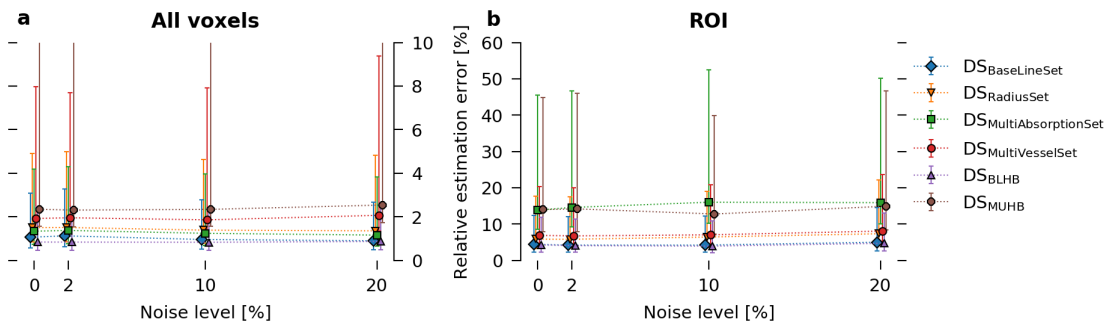


Figure 3.3: Overview of the influence of different noise levels on the relative fluence estimation errors when using a random forest as the inference model. Each data set is represented by a different color/symbol combination.

Qualitative random forest results on the baseline data set  $DS_{\text{base}}$  with 20% noise:

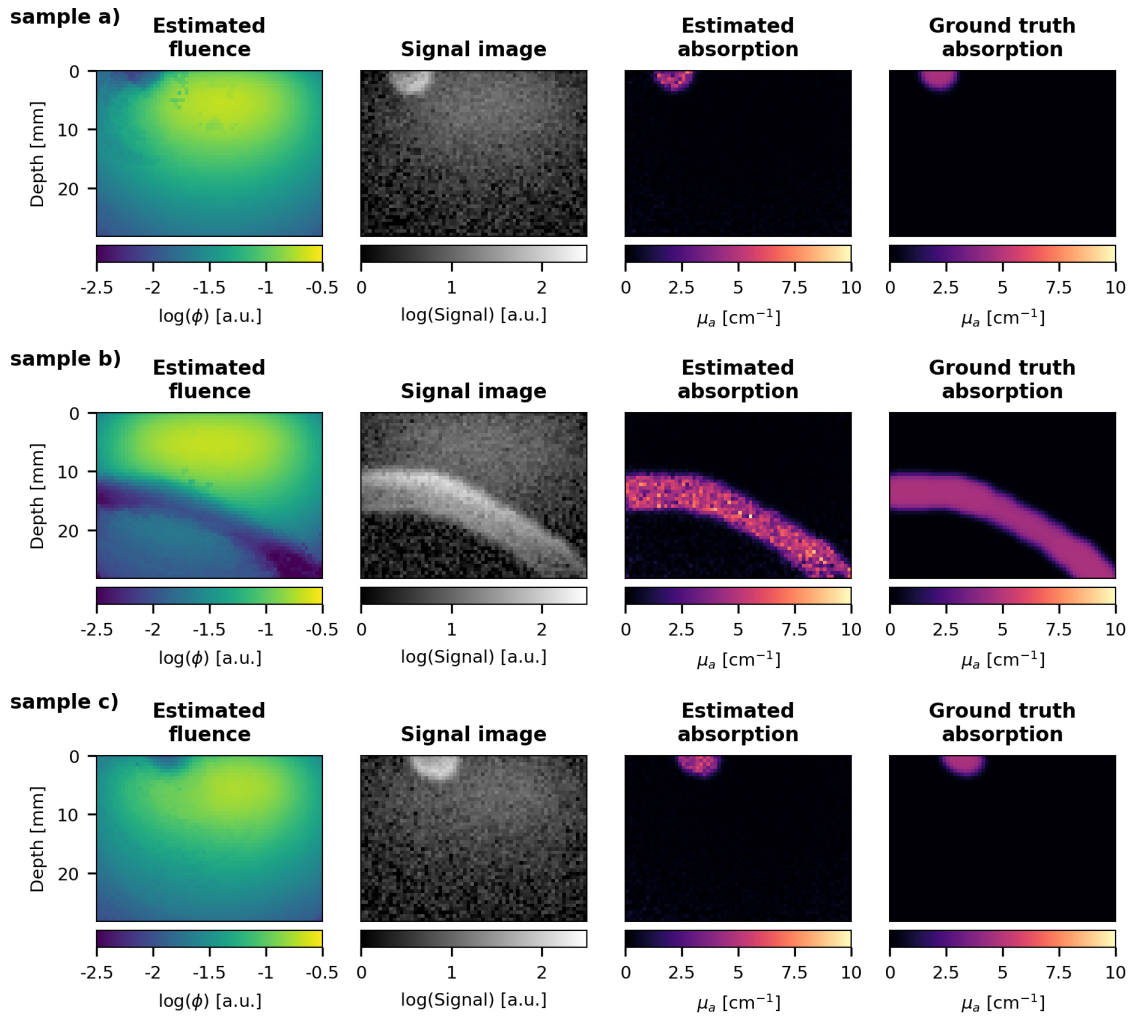


Figure 3.4: Visualization of the performance of CE-qPAI when using a random forest as the inference model. Three representative examples are randomly chosen from the  $DS_{\text{base}}$  data set and show the estimated fluence, the simulated signal, the estimated absorption calculated by fluence correction of the signal, and the ground truth absorption coefficients. For images corresponding to the worst, best, and average case on the 0% noise data set, refer to the journal publication [Kirchner et al., 2018a].

Qualitative random forest results on the completely randomized data set  $DS_{\text{multi}}$  with 20% noise:

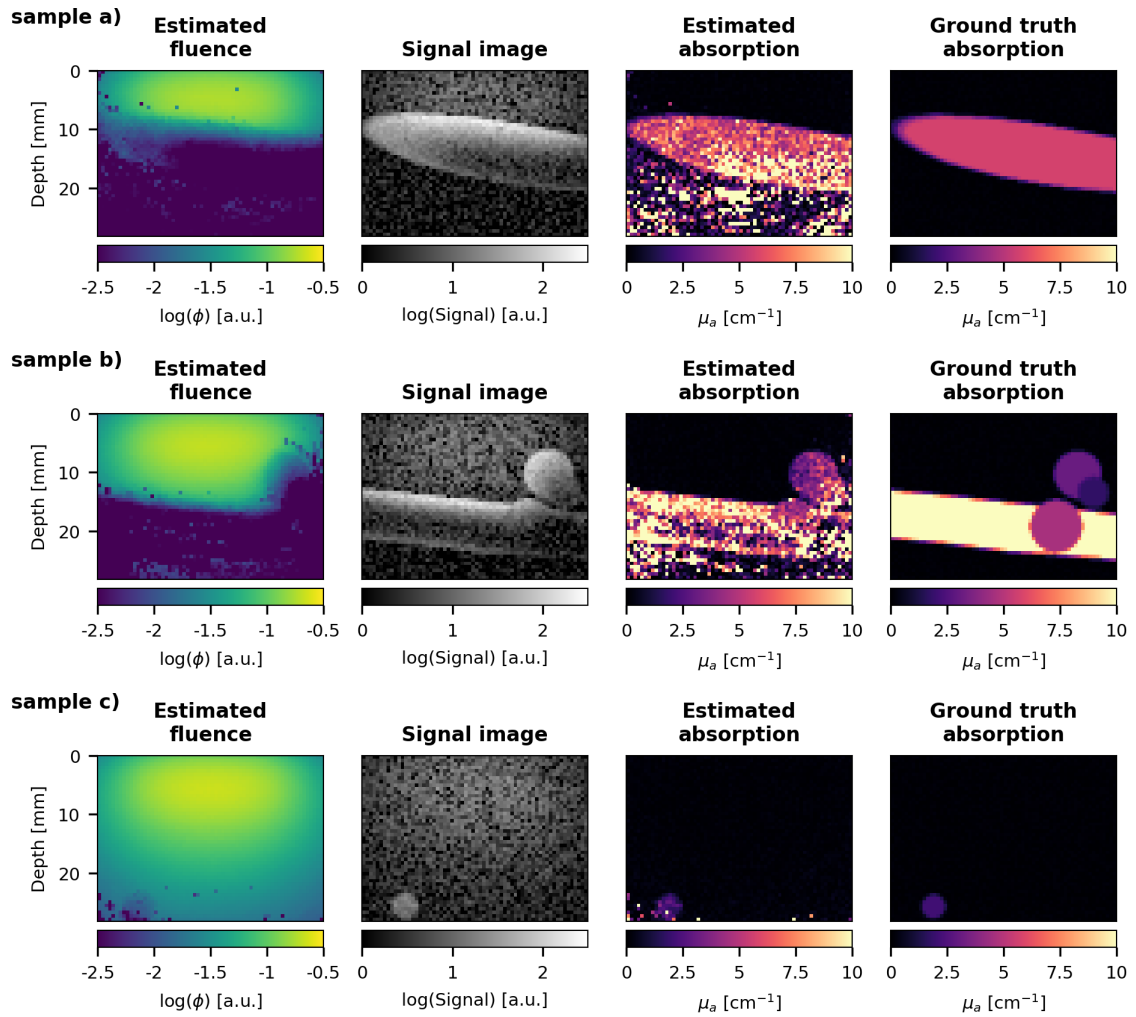


Figure 3.5: Visualization of the performance of CE-qPAI when using a random forest as the inference model. Three representative examples are randomly chosen from the  $DS_{\text{multi}}$  data set and show the estimated fluence, the simulated signal, the estimated absorption calculated by fluence correction of the signal, and the ground truth absorption coefficients.

### Monospectral deep learning regression

Table 3.3 summarizes the descriptive statistics of the relative fluence estimation errors for the experiments on absorption estimation using deep learning regression. In line with section 3.1.3, the median error and IQRs are reported for all data sets. The results show that even for the most complex data set  $DS_{\text{multi}}$  with variations of multiple parameters, CE-qPAI yields a median overall relative fluence estimation error below 5% when regarding all voxels, except for the most complicated data set, where the median error regarding all voxels is 7.0%.

data set	All voxels		ROI voxels	
	Median	IQR	Median	IQR
$DS_{\text{base}}$	2.5	(1.2, 4.4)	7.7	(3.6, 14.1)
$DS_{\text{radius}}$	3.1	(1.4, 5.9)	12.7	(5.7, 23.8)
$DS_{\text{absorb}}$	3.5	(1.6, 6.5)	21.8	(8.9, 43.9)
$DS_{\text{vessel}}$	4.8	(2.1, 10.5)	15.3	(6.7, 28.7)
$DS_{\text{backg.}}$	2.5	(1.2, 4.4)	7.8	(3.6, 14.0)
$DS_{\text{multi}}$	7.0	(3.0, 19.6)	21.9	(10.5, 36.7)

Table 3.3: Descriptive statistics of all CE-qPAI results using a fully connected feed forward neural network as the regressor. For each of the data sets, the median error, as well as the IQR, are computed for all voxels and exclusively on voxels in the ROI.

The validation of deep learning-based CE-qPAI with various noise levels on the data shows that even though there are variations in the results, there is no systematic increase of the estimation error even under unrealistically high noise levels of up to 20% (cf. figure 3.6).

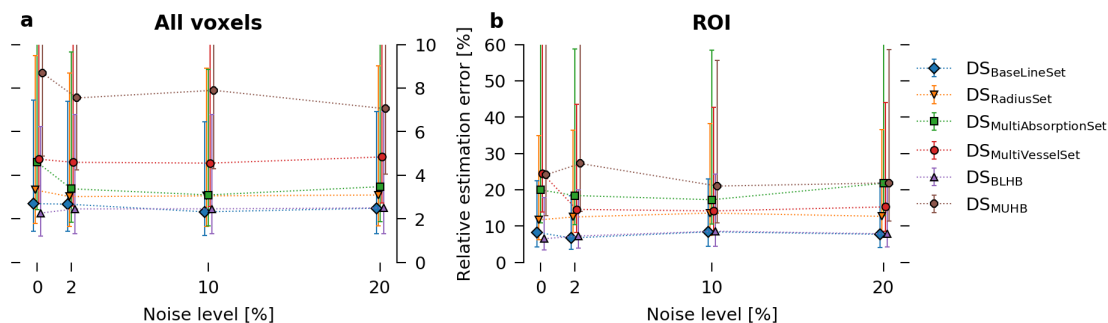


Figure 3.6: Overview on the influence of different noise levels on the relative fluence estimation errors when using a fully connected feed forward neural network as the inference model. Each data set is represented by a different color/symbol combination.

Qualitative deep learning results on the baseline data set  $DS_{\text{base}}$  with 20% noise:

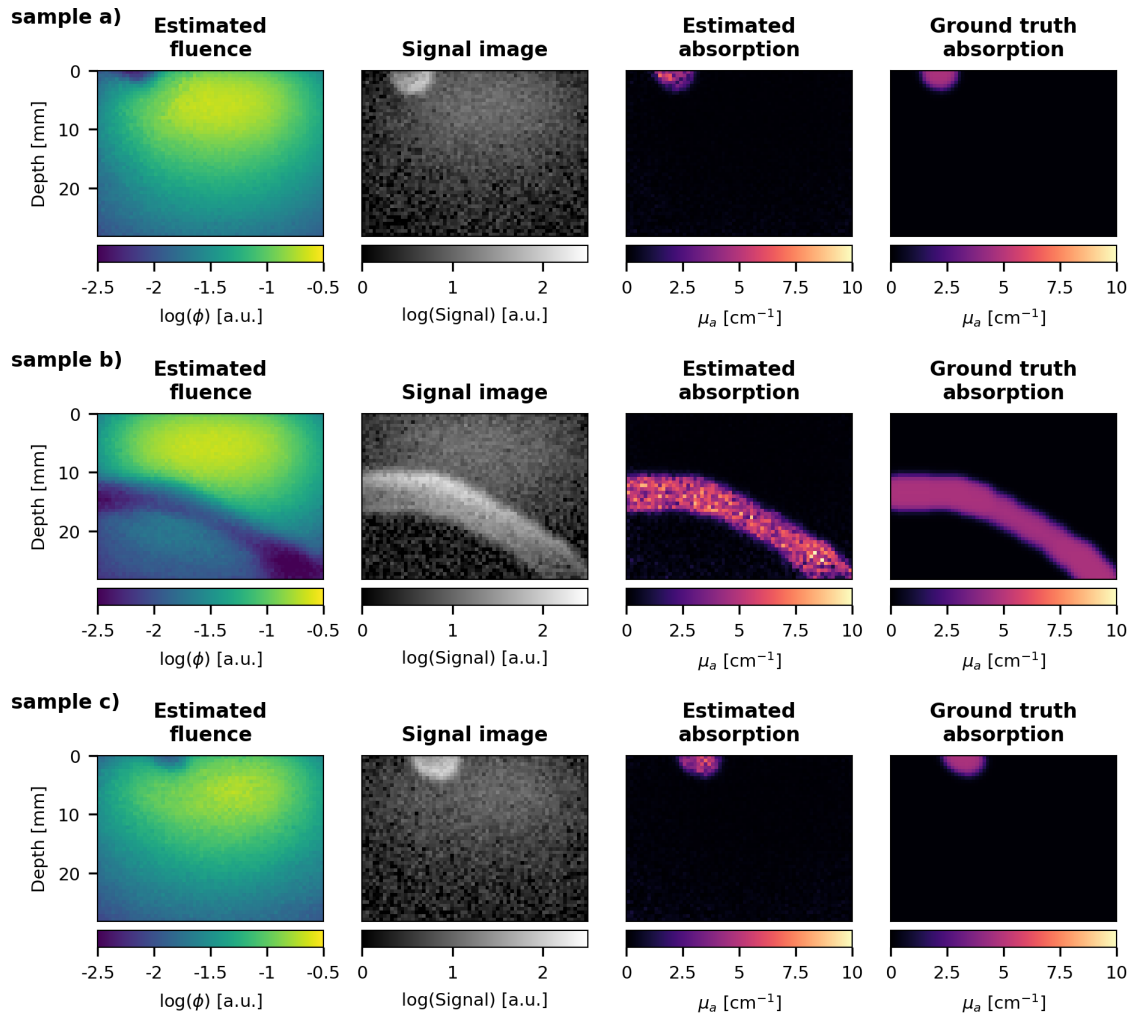


Figure 3.7: Visualization of the performance of CE-qPAI when using a feed forward neural network as the inference model. The same examples as in the random forest regression are taken from the  $DS_{\text{base}}$  data set and show the estimated fluence, the simulated signal, the estimated absorption calculated by fluence correction of the signal, and the ground truth absorption coefficients.



Qualitative deep learning results on the completely randomized data set  $DS_{\text{multi}}$  with 20% noise:

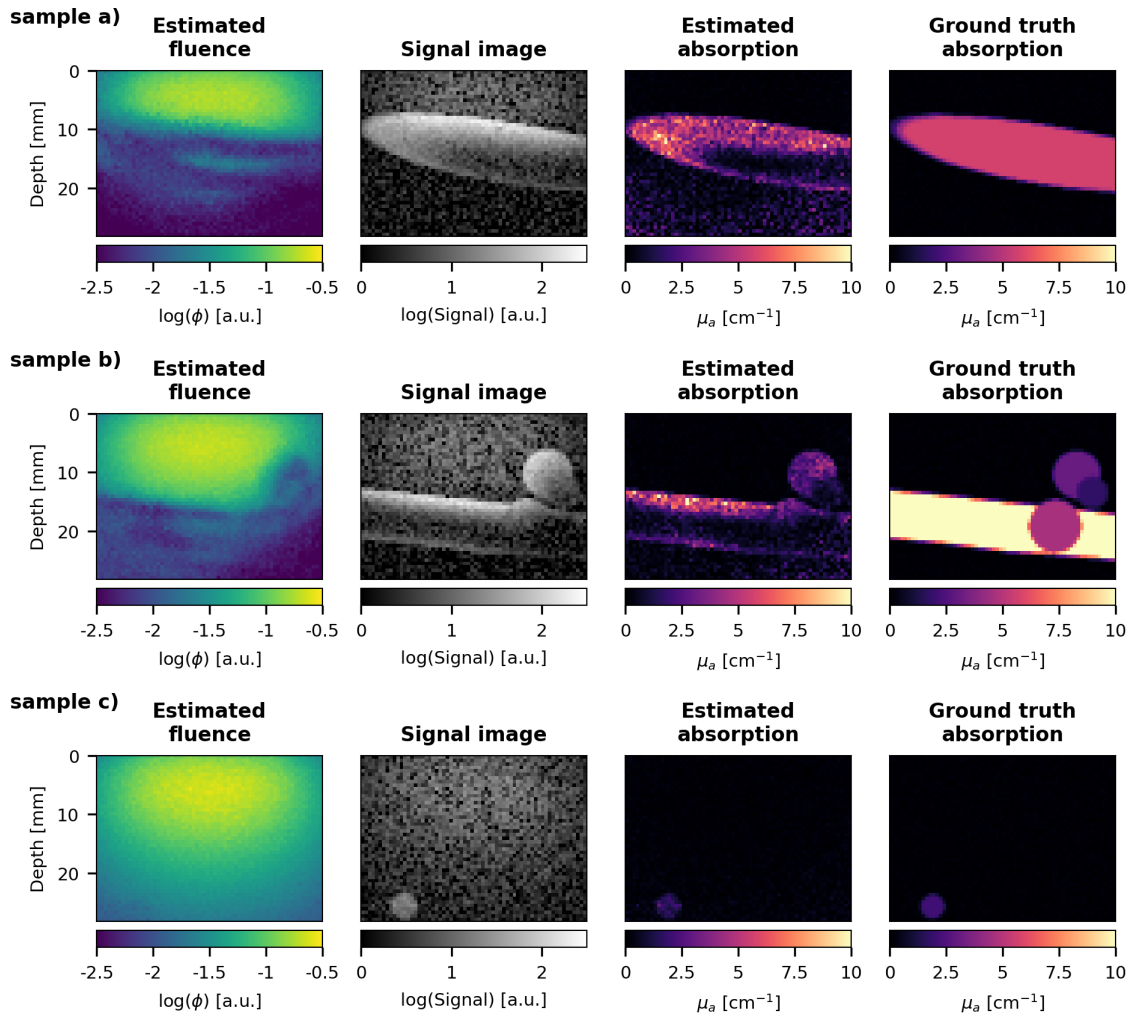


Figure 3.8: Visualization of the performance of CE-qPAI when using a feed forward neural network as the inference model. The same examples as in the random forest regression are taken from the  $DS_{\text{multi}}$  data set and show the estimated fluence, the simulated signal, the estimated absorption calculated by fluence correction of the signal, and the ground truth absorption coefficients.

### Functional CE-qPAI

For the multispectral PAI data set, the fluence estimation error performance yielded a median error in the ROI of 5.7% for random forest regression and 5.8% for deep learning. The absolute oxygenation saturation estimation error for the direct  $sO_2$  estimation method was 3.8% when using random forest regression and 2.7% when using deep learning. The complete descriptive statistics of the results for both methods are shown in table 3.4.

	Random Forest		Deep learning	
	Median	IQR	Median	IQR
rel. $\phi$ error (all) [%]	1.8	(0.8, 3.4)	3.5	(1.7, 6.2)
rel. $\phi$ error (ROI) [%]	5.7	(2.6, 11.0)	5.8	(2.6, 11.0)
abs. $sO_2$ error (all) [%]	2.1	(0.7, 6.1)	2.3	(1.0, 4.9)
abs. $sO_2$ error (ROI) [%]	3.8	(1.3, 10.2)	2.7	(1.2, 5.8)

Table 3.4: Functional CE-qPAI results. The  $\phi$  estimation errors are shown as relative errors in percent, and the  $sO_2$  estimation results show the absolute errors in percentage points.

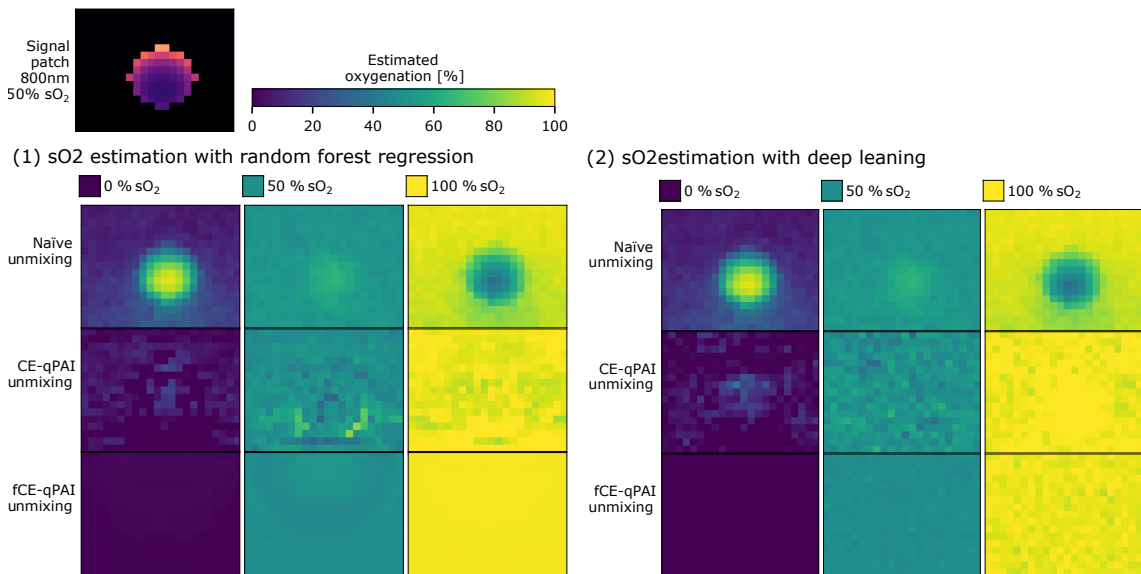


Figure 3.9: Qualitative results of functional CE-qPAI. The results are shown for (1) random forest regression and (2) deep learning for an example signal patch containing a vascular structure with oxygenation levels of 0%, 50%, and 100%. Naïve unmixing is the application of least square regression directly on the multispectral signal image, CE-qPAI unmixing refers to the application of least squares regression after CE-qPAI fluence correction, and fCE-qPAI unmixing refers to the direct estimation of  $sO_2$  based on multispectral CIs.

Figure 3.9 shows qualitative  $sO_2$  estimation results for three different spectral unmixing methods: naïve linear unmixing based on the multispectral input images, linear unmixing after fluence

correction of the input images using CE-qPAI, and direct  $\text{sO}_2$  estimation using Functional CE-qPAI (fCE-qPAI). While the data set contained eleven distinct oxygenation samples, only three of them are shown in the figure (0%, 50%, and 100%). The CE-qPAI methods were applied using (1) random forest regression and (2) a fully connected neural network.

### 3.1.4 Discussion

This work addresses one of the most critical challenges related to PAI: the quantification of the optical absorption coefficient. In contrast to other approaches, the CE-qPAI method uses machine learning to estimate the light fluence in a voxel to deduce the corresponding optical absorption. Comprehensive *in silico* experiments demonstrate the potential of this approach to estimate local light fluence and subsequently, the optical absorption coefficients as well as derived functional properties, even in the presence of high additive and multiplicative Gaussian noise.

#### Discussion of results

The conducted experiments show that it is feasible to estimate local fluence from CIs with regression techniques such as random forest regression and deep learning. Even though the random forest outperformed the deep learning network on the fluence estimation tasks by a large margin, there was only a negligible difference between random forest and deep learning regression for functional CE-qPAI. The comparatively weak deep learning results have much room for improvement. For this feasibility examination, only a relatively small fixed computational budget in terms of epochs, batches size, and batches per epoch was chosen. Also, a static learning rate might be non-optimal in this case, and the results might be better when considering different loss functions. When fine-tuning all adjustable hyperparameters, it should be possible to get deep learning results on the same scale if not better compared to using random forests, but this optimization was not the aim of this work.

While the experimental results of CE-qPAI demonstrate the general feasibility of the method, the conclusions that can be drawn are limited. Firstly, a very low scattering coefficient was chosen, which enables a deeper penetration depth of the light, but is not likely to occur in any clinically relevant scenarios, where the scattering is usually an order of magnitude higher. Furthermore, the method has only been tested on *in silico* data, and also exclusively been tested on data simulated by a single Monte Carlo framework. The robustness of the CE-qPAI method remains an open question independent of the regression algorithm when regarding the imprecise reconstruction of the initial pressure distribution, for example, due to limited view problems, inaccurate calibration of the PA device, systematic noise or imaging artifacts, or other influences. Unfortunately, the CE-qPAI method, as presented in its current state, is not applicable *in vivo* or even *in vitro*, as a near-perfect reconstruction of the initial pressure distribution is needed that is only subject to Gaussian noise.

#### Discussion of methodology

The introduction of *a priori* knowledge in the form of the FCMs was the main contribution of this work, as they allow to “disassemble” the initial pressure distribution to encode the 3D  $p_0$  context with respect to the voxel-specific fluence contribution. However, the calculation of the FCMs imposes practical constraints. Firstly, the FCMs have to be calculated for each voxel of the

imaging plane in the medium (potentially reducible to less when symmetries in the illumination geometry are present), and these simulations have to be conducted using many photons because of the relatively small amount of photons traversing the respective detector voxel. Secondly, due to the nature of the Monte Carlo simulation, the resulting FCMs are calculated on a fixed computational grid that is of a specific size with a specific voxel spacing. This leads to the limitation that the FCMs can only be used for calculations with  $p_0$  images of the same spacing. For different settings, the FCMs would have to be calculated anew. The settings chosen in this work already represented the feasibility limitations of the currently available hardware in terms of computation time and storage needed for the FCMs.

One of the biggest strengths of machine learning-based approaches is their universally high computational efficiency after the training process. A trained CE-qPAI regressor can provide estimates for the fluence in a few milliseconds on powerful hardware. However, online estimation of the light fluence for the application with clinical handheld 2D PAI probes would not be achievable due to the high computational cost of compounding several 2D slices into a 3D volume and subsequently calculating the histograms for the context images. In this case, the fluence could only be estimated on a sweep-by-sweep basis. However, this drawback could potentially be remedied with the use of 3D PAI probes.

The CE-qPAI results demonstrate high robustness to noise when estimating the fluence, but the direct estimation of the absorption coefficient would be preferable. This is because the effects of noisy data can be mitigated when not using it for subsequent calculations. This becomes even more relevant for low SNR signals because performing fluence correction in these cases does not only enhance the signal but also amplifies the noise, which can have detrimental effects, for example, on the performance of classical multispectral unmixing algorithms for  $\text{sO}_2$  estimation. Very promising in this direction was the applicability of the CE-qPAI method to functional PAI, which yielded highly accurate and very smooth results. It is to be expected that the direct functional CE-qPAI approach does not significantly lose accuracy when having to face noise in the input data, whereas unmixing based on fluence corrected images is heavily impeded when there is an increase in the overall noise level.

### **Transfer to reality**

Despite the promising results *in silico*, the method could not successfully be applied to real measured PA data. The main reason for this is the fact that perfect deposited energy distributions were used for the creation of the CIs for the training of the machine learning algorithm. With clinical handheld transducers, due to limited view artifacts and limited sensitivity, only very rough approximations of the initial pressure distribution are possible. Furthermore, factors such as the temperature-dependent Grüneisen parameter or a rigorous calibration of the detection device are usually neglected. The most important step towards the applicability of the CE-qPAI method to real data is the generation of realistic PA signals in a well-validated simulation pipeline. Such a realistic simulation might incorporate heterogeneous scattering coefficients and a heterogeneous Grüneisen parameter map and should also consider acoustic forward modeling and suitable image reconstruction, in order to obtain rough approximations of the initial pressure similar to realistic scenarios.

### Related work

While state-of-the-art techniques for the estimation of the light fluence exist in the form of iterative model-based approaches [Cox et al., 2005], the CE-qPAI method was not compared to these in this thesis. The main reason for this is the long computation times that these methods are associated with in complex 3D scenarios. However, future work on a systematic comparison of many different fluence estimation methods on a well-designed data set would be highly relevant and exciting.

### 3.1.5 Conclusion

In this work, two questions were derived from the overarching hypothesis  $H_1$ , whether the application of data-driven approaches to the inverse problems of qPAI is feasible:

1) *Is it possible to encode the voxel-specific 3D context into a single low-dimensional feature vector to derive a voxel-wise estimate of the local fluence?*

In regard to this question, the concept of the *context image* was developed that combines the information of a measured 3D map of initial pressure with *a priori* information of the illumination geometry. To this end, a 2D histogram is computed that contains correspondences of initial pressure and voxel-specific fluence contribution. The conducted *in silico* experiments suggest that this hand-crafted feature vector indeed contains enough information to enable the inference of voxel-specific fluence, which - in principle - allows for the correction of the initial pressure distribution to obtain estimates of the optical absorption coefficient.

2) *Is the calculation of voxel-wise feature vectors from only a limited number of simulated volumes sufficient for generalization given the large possible parameter space of the distribution of optical properties?*

Even with the use of modern hardware infrastructure, the computational time for the forward simulation of a few thousand tissue samples is in the order of magnitude of days. As such, the number of samples available for training is limited. In this work, it was shown that 150 composed volumes (created from a total of 4800 simulations) were sufficient in order to generalize well enough to predict values for previously unseen situations accurately. This was investigated using both random forests as well as a neural network for inference. Furthermore, the analysis showed that the method is very robust to noise as long as the noise is accurately modeled in the training data set.

While the results *in silico* are promising, further work needs to be conducted to establish the potential of the methods in realistic *in vitro* scenarios or *in vivo* clinical applications. This future work should include the investigation on how to generate realistic training data and might also include a rigorous analysis of the upper bound of the achievable accuracy of the method in different scenarios, for example, by gradually increasing the amount of training data until a saturation limit can be detected. The presented approach might be applicable in reality when accurate reconstructions of  $p_0$  is possible. With this in mind, the development and of very broadband detectors, such as the Fabry-Pérot interferometer, to PAI is an exciting trend (cf. e.g. [Buchmann et al., 2017a]).



When training a neural network with **synthetic or simulated data**, it is expected that performance will increase when the simulated data **closely resembles the real environment**.

— [Allman et al., 2018b]

## 3.2 | Quantitative photoacoustic imaging with end-to-end deep learning

### Disclosures to this work:

Lena Maier-Hein supervised this work and was - along with Thomas Kirchner, Tim Adler, and Dominik Waibel - part of the development of the methodology and involved in the writing process of related publications. Parts of this work have been published in the *Journal of Imaging* [Gröhl et al., 2018a] and in the proceedings of the BiOS *Photons Plus Ultrasound* meeting of the SPIE Photonics West conference in 2018 [Waibel et al., 2018].

### 3.2.1 Introduction

The identification of fast, accurate, and reliable quantification methods to obtain optical tissue property distributions from photoacoustic imaging is one of the primary goals for photoacoustic research. Model-based iterative methods for the optical inverse problem in quantitative photoacoustic imaging suffer from long computation times, high sensitivity to noise, and the potential uncertainty of necessary *a priori* assumptions. Methods to tackle the acoustic inverse problem have to deal with imaging artifacts, the limited view problem, and detector noise in order to be able to obtain accurate estimates of the initial pressure distribution.

While the proposed iterative methods have been demonstrated to produce accurate results on simulated data sets, an application of the methods in real settings has only been achieved once with careful modeling of the expected values for the unknowns as well as additional *ad hoc* corrections [Buchmann et al., 2019a]. The main reasons for this apparent gap are the practical limitations of the model-based method that make the inverse problem - which is complicated as it is - even more challenging to solve. As machine learning methods are, in principle, capable of mapping arbitrary functions, the main objective of the experiments outlined in this chapter is to determine whether it is feasible to use deep learning for the optical and acoustic inverse problems of qPAI. The final goal of these techniques is to lay the foundation to obtain quantitative estimates for the optical absorption distribution.

In the method introduced in the previous chapter - Context Encoding Quantitative Photoacoustic Imaging (CE-qPAI) - *a priori* knowledge was used in the form of the Fluence Contribution Maps (FCMs) to create a hand-crafted feature vector (Context Image (CI)) for each image pixel and

then a machine learning regressor was used to estimate the target property based on this feature vector. In the field of theoretical machine learning, it has long been shown that hand-crafted features are usually inferior to features that are chosen by the machine learning algorithm and specially tailored to the data set (cf. e.g. [Liao et al., 2013]).

In the experiments demonstrated in this chapter, it is investigated how deep learning methods trained in an end-to-end manner can cope with the optical and acoustic inverse problems based on data simulated on the basis of the DiPhAs handheld photoacoustic probe design (see section 2.3.1). To this end, to investigate the optical inverse problem, a convolutional neural network - specifically the widely used U-Net - was trained to infer optical absorption coefficients from 2D images of initial pressure data and 2D raw time series data corresponding to a linear array geometry. Towards the acoustic inverse problem, the same network architecture was used to reconstruct initial pressure distributions from raw time series data.

### Hypothesis investigated in this chapter

*H1: Data-driven methods can be used to solve the optical and acoustic inverse problem.*

Similar to chapter 3.1, this chapter deals with the first research hypothesis. In contrast to the other work, however, the experiments presented in this chapter try to determine whether it is feasible to employ end-to-end deep learning models for the inverse problems of qPAI. Following the principle ideas of the prior work in literature, three research questions are examined regarding the hypothesis:

- 1) Is it feasible to solve the acoustic inverse problem with end-to-end deep learning?
- 2) Is it feasible to solve the optical inverse problem with end-to-end deep learning?
- 3) Is it feasible to solve the combined inverse problem with a one-stage end-to-end deep learning method?

### 3.2.2 Methods

Regarding the *H1* hypothesis, it is attempted to solve the acoustic and optical problems separately, as well as in a combined approach. Each of the approaches is then validated in a set of experiments. In order to formulate these approaches in the framework of end-to-end deep learning, it is crucial to correctly define the input data (*features*), as well as the output data (*labels*) for the respective approach.

In the context of the qPAI inverse problems, the label domain of the optical inverse and the combined approach is the distribution of optical absorption coefficients  $\mu_a$ , and the label domain



of the acoustic inverse approach is the initial pressure distribution  $p_0$ . The feature domain of the optical inverse approach is  $p_0$ , and the feature domain of the acoustic inverse approach and the combined approach is the raw time series data  $p(t)$  (see table 3.5 and figure 3.10).

Target approach	Feature domain	Label domain
Acoustic inverse approach	$p(t)$	$p_0$
Optical inverse approach	$p_0$	$\mu_a$
Combined approach	$p(t)$	$\mu_a$

Table 3.5: Feature and label domains for the end-to-end deep learning approaches.

A regression implementation of the U-Net [Ronneberger et al., 2015] is used for all of these experiments (cf. section 2.4.3). While this model architecture had already been proposed in 2015, it is still regarded as the best architecture for medical image segmentation problems, frequently winning international benchmark challenges [Isensee et al., 2018].

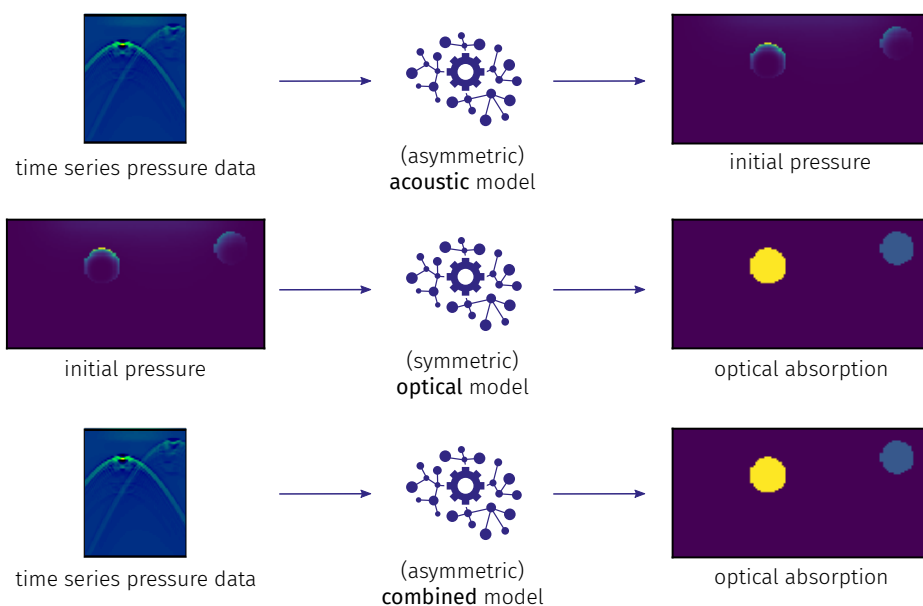


Figure 3.10: Overview figure of the acoustic inverse approach, the optical inverse approach, and the combined approach. When instructing the network to learn a mapping from raw time series data to a spatial distribution of interest, asymmetric U-Net derivatives have to be used, as the time-domain axis needs to be resampled to match the spatial domain.

When applying this architecture to the acoustic inverse approach and the combined approach, it has to be considered that the "depth-axis" of the feature domain is of a different shape than the label domain, as there are way more time-step samples in  $p(t)$  than spatial pixels in the discretization of  $\mu_a$  or  $p_0$ . An asymmetric U-Net derivative has to be used to remedy this, as the

time-domain axis needs to be resampled to match the spatial domain. This was implemented with the introduction of asymmetric skip connections in the network that can scale the input by a fixed factor. Due to this, a sampling rate needs to be chosen for the experimental design, which allows the time samples to be an integer multiple of the number of spatial pixels when covering the same imaging depth.

### Experimental design

A total of four different data sets were created for the experimental validation of the proposed end-to-end deep learning methodology. These are described in detail in the next section. They were simulated using well-known frameworks that implement the respective forward models. The optical forward process was simulated with the Monte Carlo method using the Monte Carlo eXtreme (MCX) framework [Fang and Boas, 2009], which enables running the simulation on a GPU, effectively increasing the simulation speed by an order of magnitude. For the acoustic forward simulation of the wave equations, the k-Wave Matlab toolkit was chosen, which is “based on a k-space pseudo-spectral time domain solution to coupled first-order acoustic equations for homogeneous or heterogeneous media [Treeby and Cox, 2010]”. The k-Wave toolkit also features GPU support and can provide solutions in a reasonable amount of time. In both of these frameworks, the DiPhAs photoacoustic probe (see section 2.3.1) is modeled, and the parameters listed in tables 3.6 and 3.7 are chosen for the frameworks. As the illumination geometry of the probe is already designed to account for the refractive index mismatch from air to a fluid medium, a refractive index is not modeled in the optical forward simulation.

#### Optical forward model

Parameter	Value [Units]
Pixel spacing	0.3 [mm]
Number of photons	$10^7$
Laser pulse energy	50 [mJ]
Tissue temperature	37° [C]
Grüneisen parameter	0.2

Table 3.6: Optical forward model parameters. The value for the Grüneisen parameter was calculated from the formula given in [Wang and Wu, 2012] for a temperature of 37° C.

The optical tissue parameters (listed in table 3.7 and detailed for the respective data sets) were chosen with great care to resemble close approximations of reality regarding the different setups. In this regard, single-voxel resolved distributions for the optical absorption and scattering were used. The acoustic properties (listed in table 3.7), on the other hand, have been assumed to be homogeneous throughout the medium. The main reason for this was that acoustic scattering and the effect of acoustic attenuation is orders of magnitude smaller than that of the optical counterparts.

The relative estimation error  $e_{\mu}^r$  is calculated to evaluate the performance of each of the approaches and reported in the results section:

**Acoustic forward model**

Parameter	Value [Units]
Speed of sound	1540 [m/s]
Alpha coefficient $\alpha$	0.1 [dB/(MHz $^\gamma$ cm)]
Alpha power $\gamma$	1.5
Medium density	1000 [kg/m $^3$ ]
Sampling rate	66.6 [MHz]
Detector central frequency	7.5 [MHz]
Detector bandwidth	80 [%]

Table 3.7: Acoustic forward model parameters.

$$e_\mu^r = \frac{|\hat{\mu} - \mu_{GT}|}{\mu_{GT}} \times 100[\%], \quad (3.1)$$

where  $\mu$  refers to the property in question (in this case, either  $\mu_a$  or  $p_0$ ),  $\hat{\mu}$  refers to its estimate, and  $\mu_{GT}$  refers to its ground truth value. The median value is reported as a metric that is robust to outliers and does not need to be derived from a Gaussian distribution to be meaningful. The error is calculated only for those parts of the image, where there were vascular structures in the a priori tissue geometry, and where the CNR of the initial pressure distribution is greater than 2. For this case, the CNR is calculated as  $CNR = (I - \text{mean}(I_{\text{noise}})) / \text{std}(I_{\text{noise}})$ , where  $I$  is the pixel intensity, and  $I_{\text{noise}}$  is the distribution of pixel intensities in the background.

**Data sets**

The proposed methods are validated on a total of four different data sets. Two of these are representative of blood flow phantoms in which a vessel structure is submerged in a milk bath, and the other two are created to resemble tissue-realistic setups. Of all data sets, 400 variations of the contained structures are created and imaged at five different wavelengths that are in the imaging range of all conventional photoacoustic devices: 700nm, 750nm, 800nm, 850nm, and 900nm. In order to circumvent off-plane fluence effects, the simulated volumes are constant along the y-axis, which enables the end-to-end estimation on 2D images recorded with a handheld linear array transducer, without introducing an unknown error term.

The two milk bath phantoms are modeled with a wavelength-independent background absorption and scattering, which were set to  $\mu'_s = 10 \text{ cm}^{-1}$  ( $\mu_s = 100 \text{ cm}^{-1}$ , and  $g = 0.9$ ), and  $\mu_a = 0.1 \text{ cm}^{-1}$ . The first phantom contains one tubular structure that is always placed at a depth of around 0.5 cm in the center of the phantom with a variation of ( $\pm 2\text{mm}$ ) in the x-axis and ( $\pm 1\text{mm}$ ) in the z-axis. The second setup contains one to two tubular structures that are located freely within a spawn area in the phantom in a depth of 4 to 10 mm. In both cases, the vessel structures are modeled to have a hemoglobin concentration of 150g/l and are each assigned a random oxygenation between 0% and 100%. There is a margin of 4 mm to either side of the phantom to avoid the unrealistic clipping of structures. A graphical representation of the composition of

both phantoms is shown in figure 3.11.

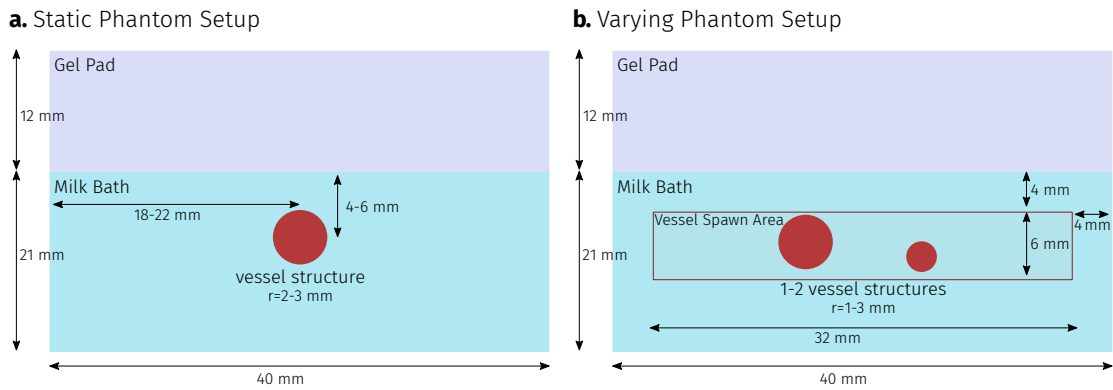


Figure 3.11: Graphical representation of the simple static phantom setup (a.) and the phantom setup with varying target structures (b.). Except for the shown distance measures, this representation is not to scale.

One of the tissue-realistic setups resembles a tissue model containing 68% water and 0.5% blood in the background with a mean blood oxygenation of 50%. It contains 3 to 12 tubular vessel structures which are freely distributed in the tissue. The final tissue realistic setup was created to mimic the human forearm, including models of epidermis, dermis, muscular homogeneous background, arteries, veins, and bones. In order to increase the variation in the volumes, zero to three random vascular structures were added to the realistic forearm model in the same spawn area as in the previous models. A graphical representation of the composition of both phantoms is shown in figure 3.12.

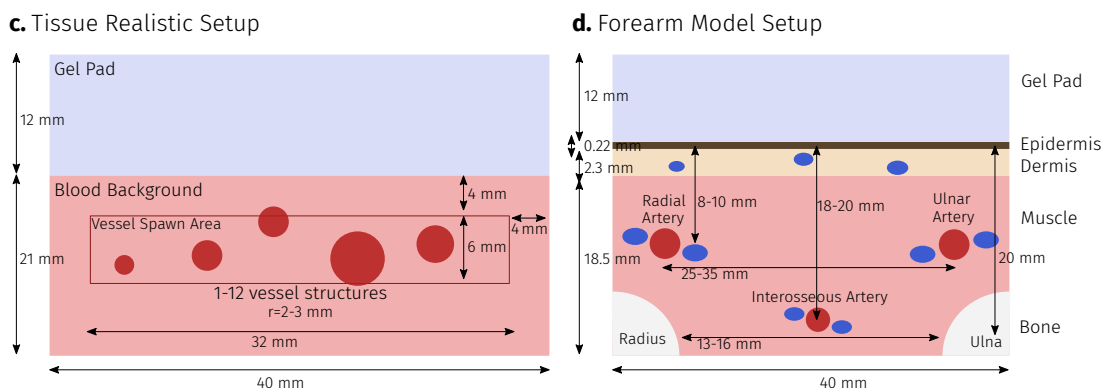


Figure 3.12: Graphical representation of the complex realistic tissue setup (c.) and the forearm model setup (d.). Except for the shown distance measures, this representation is not to scale.

The optical properties of the different structures are modeled as reviewed by Jacques [Jacques, 2013] and the multispectral absorption coefficient curves are taken from the paper, and the

omlc website of Scott Prahl <https://omlc.org/spectra/index.html> (as visited on the 8th of November 2019) and a constant anisotropy of  $g = 0.9$  is assumed. The homogeneous muscle tissue is assumed to have 0.5% blood volume fraction and water content of 68% [Forbes et al., 1953]. The dermis is modeled to be  $2.3 \pm 1.2$  mm thick and to have a water volume fraction of 58% [Forbes et al., 1953], and the epidermis is assumed to be  $0.22 \pm 0.1$  mm thick [Oltulu et al., 2018], to not contain water, and to have a mean melanin concentration of  $2.2\% \pm 1\%$  [Alaluf et al., 2002]. The radii of the radial and ulnar artery are simulated to be  $1.65 \text{ mm} \pm 0.2 \text{ mm}$  [Ashraf et al., 2010], and their accompanying veins are modeled to have a radius of  $1 \pm 0.2$  mm [Yang and Chung, 2018]. The radius of the interosseous artery is set to  $0.3 \pm 0.125$  mm [Hubmer et al., 2004] and the accompanying veins are modeled to be of half the size as the radial and accompanying ulnar veins. Several subcutaneous veins with a pitch not smaller than 5 mm are simulated at a depth of  $1.5 \pm 0.7$  mm and a radius of  $0.4 \pm 0.3$  mm. The oxygenation of an arterial vessel is modeled to be 90% - 100% [Merrick and Hayes, 1976], and the oxygenation of venous oxygenation is assumed to be around  $70\% \pm 10\%$  [Molnar and Nemeth, 2018]. The two bones that are present in the model (radius and ulnar) are modeled to have a water content of  $19\% \pm 1\%$  [Timmins and Wall, 1977] and to have a mean separation of 32 mm [Christensen et al., 1968]. Other parameters, such as the depth of the arteries and sizes of the bones, were experimentally determined based on photoacoustic (of healthy human volunteers) and CT images of forearms. For all data sets, a gel pad layer was added to the simulation, which is usually used in conjunction with the DiPhAS PAI device design in order to reduce the amplitude of the transducer artifact caused by light absorption of the US transducer membrane.

### Deep learning hyperparameters

The simulated input and output images were pre-processed to be of the size of a power of two. For this, a center crop operation was performed, and the  $p_0$  and  $\mu_a$  images were cropped to a width of 128 pixels and a height of 64 pixels (after removing the air and gel layer), corresponding to 3.84 and 1.92 cm with a spacing of 0.3 mm. With a speed of sound of 1540 m/s and a sampling rate of 66.6 MHz, a depth of 1.92 cm corresponds to roughly 831 samples. As such, the  $p(t)$  data was cropped to  $832 = 13 \times 64$  samples. An integer multiple is necessary in order to be able to use the asymmetric U-Net implementation - in this case, with a scaling factor of 13 in the skip connections. The training was performed for 500 epochs, using the Adam optimizer [Kingma and Ba, 2014], a mean squared error loss function, an initial learning rate of  $10^{-3}$ , and a batch size of 50. For the time series input data experiments, the batch size had to be reduced to 20 due to memory limitations, and for the combined approach, the number of epochs was increased to 1000, due to slower convergence. The learning rate was reduced to 95% of the previous value every 20 epochs, leading to a steady decrease in the learning rate throughout the experiment. For each new training batch, the input features were noised with a Gaussian multiplicative white noise of 5% standard deviation. All training was performed on a Ubuntu 18.04 workstation with a Ryzen 5 processor, 32 GigaBytes (GB) Random Access Memory (RAM), and an Nvidia GTX 1080 Ti graphics card with 11GB RAM.

### 3.2.3 Results

This section will state the quantitative results of the experiments evaluated in the ROI and will show sample images of the estimation results on all four data sets that are representative of the median estimation error.

### Quantitative description of results

Table 3.8 shows the descriptive statistics of the recovery of the optical absorption coefficient (and the initial pressure distribution for the acoustic inverse model). While the results are not comparable due to their different complexities, the optical inverse models achieved the lowest errors all data sets, achieving a median estimation error of up to 8.1% when trained on and applied to the tissue model data set. The performances of the acoustic inverse models were roughly between that of the combined approach and the optical inverse approach, except for the forearm model data set, where the acoustic inverse model performed the worst with a median relative error of 34.3%. The performance of the combined models was generally the worst on the data sets, averaging on a median relative error roughly twice as high as that of the optical inverse models.

Method	Static phantom		Varying phantom		Tissue model		Forearm model	
	Med.	IQR	Med.	IQR	Med.	IQR	Med.	IQR
AIP	13.0	(6.1, 22.8)	17.1	(8.1, 31.0)	15.2	(7.0, 26.7)	34.3	(16.4, 58.1)
OIP	10.6	(4.8, 19.7)	13.0	(5.9, 22.2)	8.1	(3.8, 15.0)	13.3	(6.2, 23.4)
Combined	17.0	(7.7, 31.8)	20.9	(9.6, 38.8)	19.8	(9.5, 35.0)	27.9	(13.3, 47.2)

Table 3.8: Descriptive results of the three different deep learning inversion methods. The median error (Med.), as well as the interquartile ranges (IQR), are shown in the table for each of the methods on each of the data sets.

The distribution of the errors regarding all approach/data set combinations are shown in figure 3.13, which shows violin plots and box plots of the error distributions.

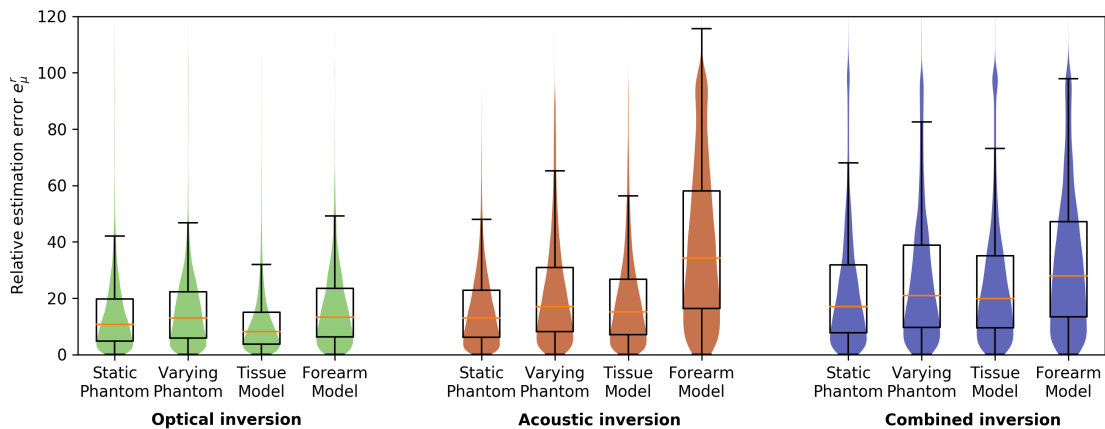


Figure 3.13: Violin plots and box plots visualizing the distribution of the relative estimation errors of the approach and data set combinations.

One question that now arises is how the accuracy of the direct approach compares to a sequential

application of the acoustic inverse model and the optical inverse model. Here, it would be expected that a sequential application would perform much worse, due to the propagation of the errors of the first model to the second. Figure 3.14 shows the results of this sequential application of the inverse models compared to the combined inverse model. The results reveal an explosion of the relative  $\mu_a$  estimation error when the two inverse models are being applied sequentially.

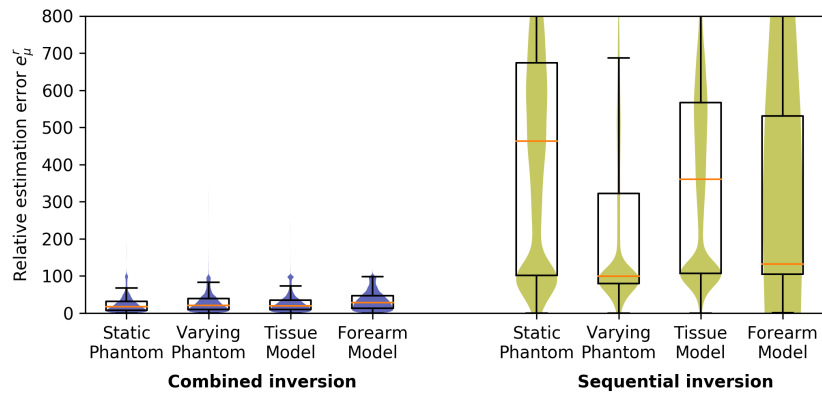


Figure 3.14: Violin plots of the direct approach versus a sequential application of the acoustic and optical inverse models. The combined model is displayed in the same color as in figure 3.13.

### Qualitative visualization of results

In the following pages, sample results for the combinations of the three examined methods and data set are shown. In order to show a representative example, the datum that produced the median performance was chosen in each case. Because of this way of choosing the representative image, different representative images from the test data set might be chosen for each method.

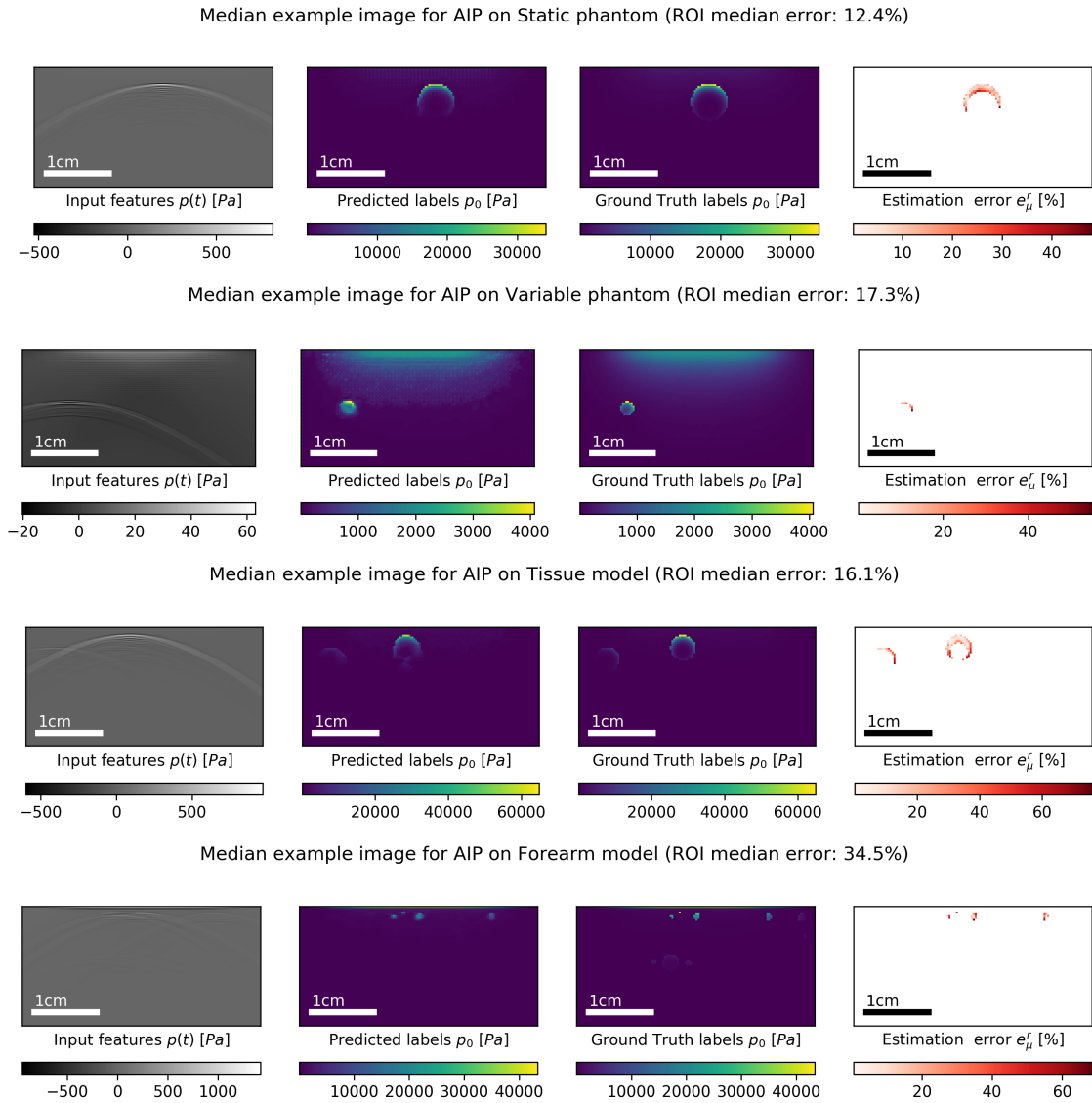
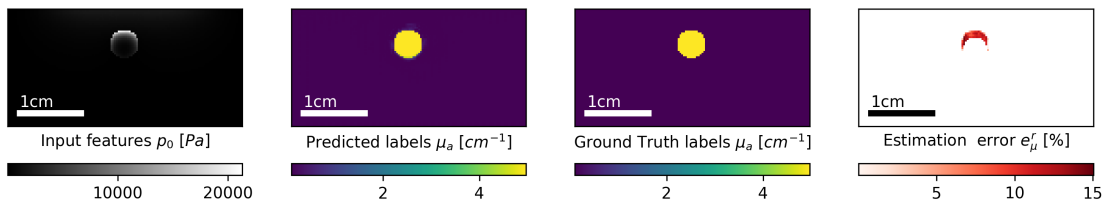
**Acoustic inverse problem:**

Figure 3.15: Representative results for the acoustic inverse model for the static phantom, varying phantom, tissue model, and forearm model data sets (from top to bottom). The images are the input data, the estimated output labels, the ground truth output labels, and the ROI with error estimates (from left to right). The caption contains the median error of the estimates in the ROI.

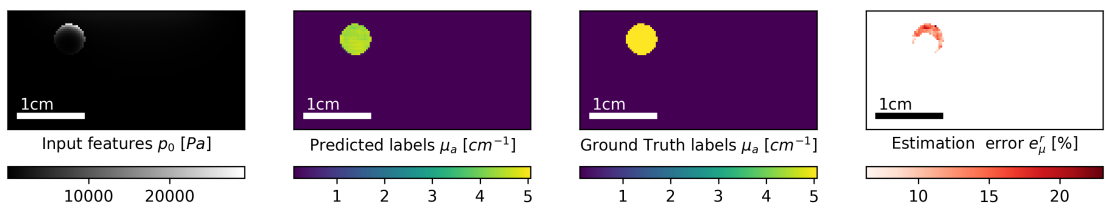


**Optical inverse problem:**

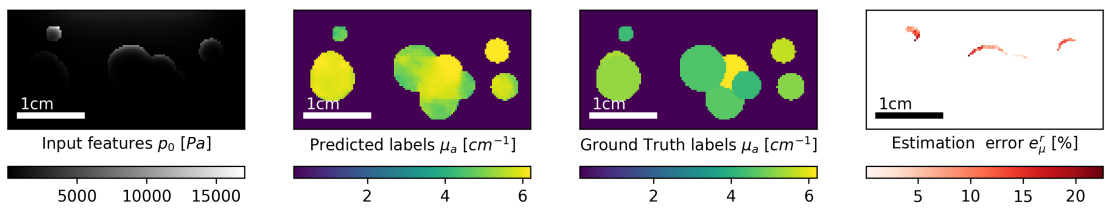
Median example image for OIP on Static phantom (ROI median error: 10.6%)



Median example image for OIP on Variable phantom (ROI median error: 12.7%)



Median example image for OIP on Tissue model (ROI median error: 8.4%)



Median example image for OIP on Forearm model (ROI median error: 14.2%)

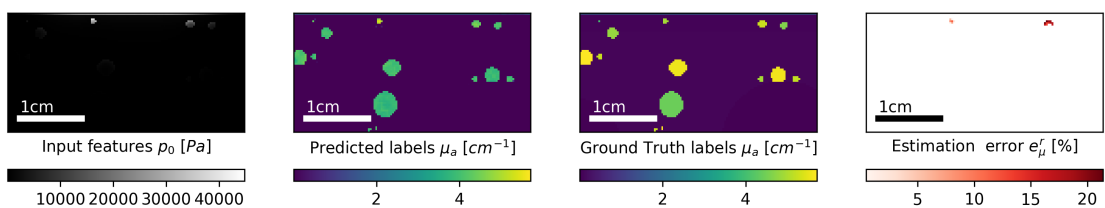
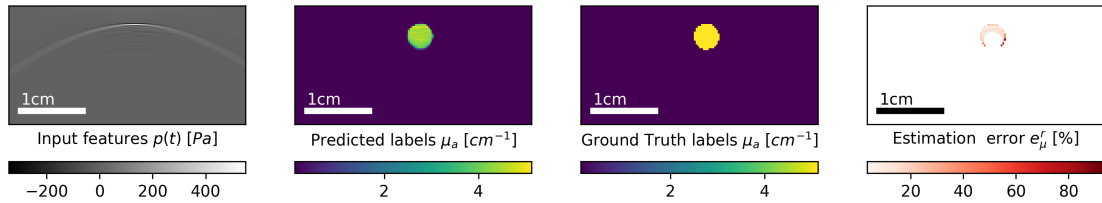


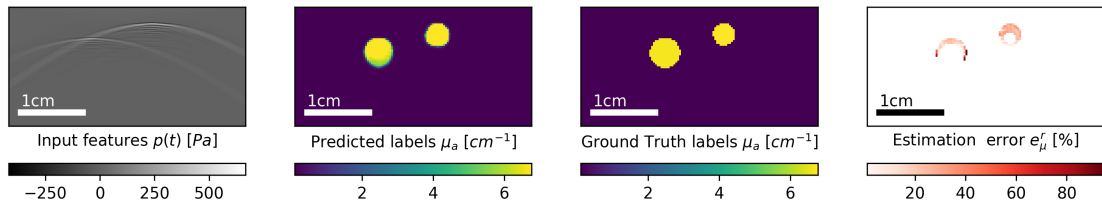
Figure 3.16: Representative results for the optical inverse model for the static phantom, varying phantom, tissue model, and forearm model data sets (from top to bottom). The images are the input data, the estimated output labels, the ground truth output labels, and the ROI with error estimates (from left to right). The caption contains the median error of the estimates in the ROI.

**Combined inverse problem:**

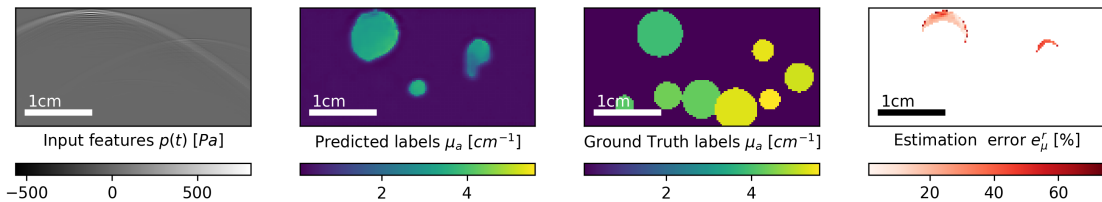
Median example image for Combined on Static phantom (ROI median error: 15.1%)



Median example image for Combined on Variable phantom (ROI median error: 22.4%)



Median example image for Combined on Tissue model (ROI median error: 19.8%)



Median example image for Combined on Forearm model (ROI median error: 27.6%)

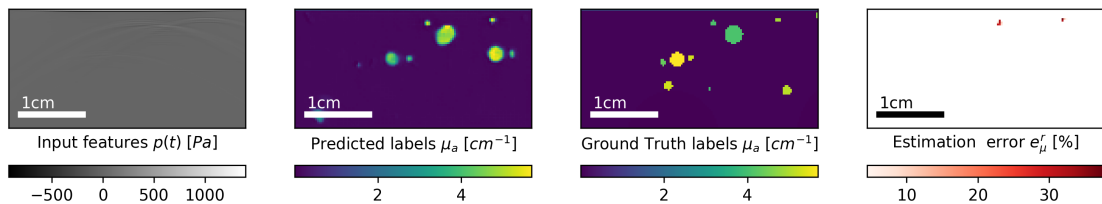


Figure 3.17: Representative results for the combined inverse model for the static phantom, varying phantom, tissue model, and forearm model data sets (from top to bottom). The images are the input data, the estimated output labels, the ground truth output labels, and the ROI with error estimates (from left to right). The caption contains the median error of the estimates in the ROI.

### 3.2.4 Discussion

Based on the conducted experiments and given the constraints of the used data sets, it appears that the creation of models that are capable of tackling the optical inverse problem in an end-to-end manner is indeed feasible. However, the experiments also revealed that the integration of the acoustic inverse problem introduced large errors in the obtained estimates.

#### Discussion of results

The experiments contained three inversion methods: (1) acoustic inversion, (2) optical inversion, and (3) combined inversion, which were evaluated on four different data sets.

(1) *acoustic inversion*: The acoustic inversion resulted in median estimation errors as low as 13 % on the most simple data set, but increased to 34 % on the forearm data set. Qualitatively the estimation results look adequate. However, there seems to be a systematic error in the reconstructions in terms of the amplitude of the reconstructed signal. It is also evident that deeper structures could not be reconstructed as well as more superficial structures. This is especially apparent in the sample image taken from the forearm data set. A possible reason for this might be the loss of information due to the depth-dependent acoustic attenuation. However, such systematic errors can usually be tackled by increasing the amount of training data [Kendall and Gal, 2017]. Further modifications of the architecture for the acoustic inversion should be considered as well and are discussed at a later stage.

(2) *optical inversion*: The optical inversion resulted in median estimation errors as low as 8 % on the generic tissue data set, but increased to 13 % on the forearm data set. Qualitatively the results appear to be very promising, only really struggling when different structures were overlapping as apparent in the sample image from the generic tissue data set. It can be expected that an increase in the amount of training data can drastically improve the performance of the optical inversion model.

(3) *combined inversion*: The combined inversion resulted in median estimation errors as low as 17% on the most simple data set, but increased to 28% on the forearm data set. One thing to note in the qualitative samples from the data set is that the combined approach seemed to struggle with accurately estimating the sizes of the chromophore structures. Furthermore, after a certain depth, the combined inversion model was not able to correctly predict the presence of absorbers, but instead predicted the presence of absorbers in places where there were none. One of the reasons this might have happened is that only slight noise augmentation was done during training, which might have caused the inversion model to become sensitive to slight variations to the time series data.

#### Single-stage versus two-stage inversion

In addition to the isolated analysis of the three inversion methods, it has also been investigated how the sequential application of the acoustic inverse model and the optical inverse model compare to solving the combined inverse problem directly. Here, it has to be concluded that the error introduced by the acoustic inverse model was amplified by the optical inverse model leading to an explosion of the error. Figure 3.18 shows the median performance of this two-stage approach on the *generic tissue data set*.

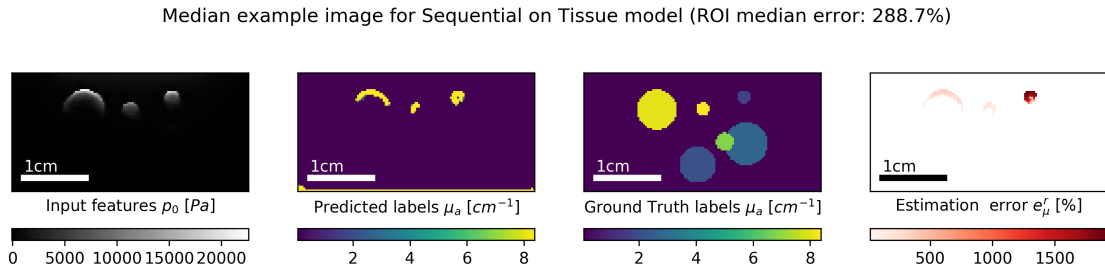


Figure 3.18: Performance of the two-stage approach regarding the generic tissue data set. The input of the optical inverse model is the output of the acoustic inverse model.

Qualitatively looking at these results, it is evident how reliant the optical inversion model is on an accurate reconstruction of the initial pressure distribution as an input. In this case, even though the estimated initial pressure distributions look reasonable, the optical inversion model cannot cope with the present deviations from the "perfect"  $p_0$  distributions it was trained on. As a result, the estimated  $\mu_a$  values are way too high and do not reflect the shape of the underlying structures. However, this crude sequential application of the two inverse models does not reflect the true potential of data-driven models to solve the inverse tasks (as suggested by the performance of the combined approach). Several measures might be taken to improve the accuracy of these results. One of these could be to train the two models in tandem, where the optical inversion algorithm is also trained to be capable of inverting the estimates of the acoustic inverse model. Furthermore, a sophisticated augmentation of the training data might also be a consideration to improve the robustness of the optical inversion model. Modifications to the architecture of the acoustic inversion model should also be considered and are discussed at a later stage.

#### *Depth dependence of the results*

For the acoustic inversion and the combined inversion (when working with time series data as the input features) structures that were deeper than about 0.5 - 1 cm could not be properly be predicted by the inversion algorithms and no quantification of the underlying tissue properties was possible. Generally, the estimation error increased drastically for deeper structures, as shown in figure 3.19, especially in the approaches that had to work with raw time series data. However, surprisingly, also the optical inverse approaches show a trend where the median estimation error increases with increasing distance.

#### *The acoustic inverse problem*

The acoustic inverse problem has been studied extensively compared to the optical inverse problem [Cox et al., 2012]. As such, it might be obvious to assume that the acoustic inverse problem is more straightforward to tackle. However, the results obtained in the presented experiments suggest that including the acoustic inverse problem and trying to tackle it in an end-to-end manner with an architecture as presented here is a lot more difficult than anticipated. While the acoustic inverse problem is well-posed for certain detection geometries, in the case of

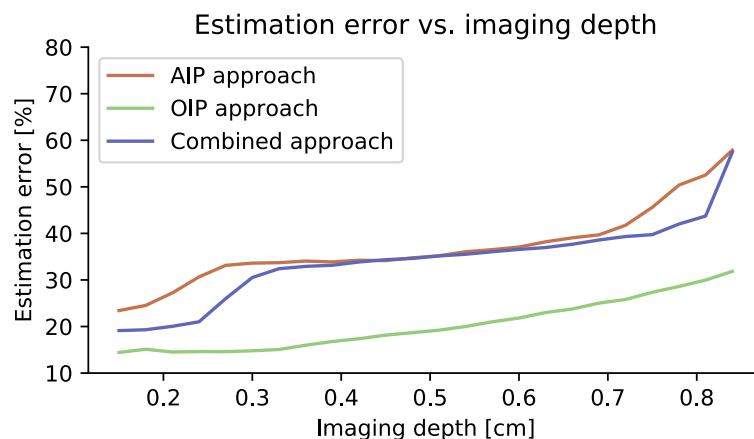


Figure 3.19: These graphs show the development of the median estimation error when only evaluating the pixels at a certain depth over all data sets for the three models. The optical inverse model is shown in green, the acoustic inverse model in red, and the combined inverse model in blue. It can be seen that the estimation error generally increases with increasing imaging depth.

linear detector arrays, solving it is much more difficult [Paltauf et al., 2007].

Some of the possible reasons for the difficulties of end-to-end methods when facing the acoustic inverse problem might be the drastic loss of information due to the limited-view detection geometry, due to errors propagated by up-sampling and smoothing operations in the transition from the optical to the acoustic forward simulation, and the presence of acoustic attenuation, which introduces a depth and frequency-dependent degradation of information. Presumably, increasing the amount of training data might lead to a significant improvement of the reconstruction results. This may be true, especially when including the acoustic inverse problem, as here, the complexity of the problem is higher. Another possible reason might be the chosen model architecture, which was an asymmetric modification of the U-Net. Maybe, a fully convolutional architecture without a decoding and an encoding part would have been a better choice. Facing these problems, many promising machine learning-based approaches to the acoustic inverse problem have been published lately (e.g. [Haltmeier and Nguyen, 2019, Schwab et al., 2019b, Antholzer et al., 2019a, Allman et al., 2019]) and in future work, it should be investigated how the insights from these studies could be applied to the problems that were encountered in this work.

## Discussion of methodology

### *Data sets*

The data sets were designed to be of increasing difficulty for the inverse models. The first two data sets had the purpose of resembling two milk bath phantoms, where the simplest had only one vessel structure in a confined space, and the second had multiple vessels in random positions. The third and fourth data sets were designed to more closely resemble physiological tissue. Here, the third dataset was designed to contain many vessel structures in a muscle background to represent generic tissue, and the fourth data set was designed to contain all structures that are present in the human forearm.

It was expected that the estimation performance of the models would reflect this steady increase in complexity by yielding a steady decrease in median accuracy on the respective test sets. However, when the methods were trained on and applied to the generic tissue model, they generally yielded better performance than models trained on some of the simpler phantom setups. A possible reason for this could be the lower average background absorption that results from the optical parameter assumptions that were made for muscle tissue. Furthermore, the performances of the models on the data sets are not really comparable, as the models are also trained on a specialized training set. Due to this, an increase in complexity in the data set might cause the model to learn a more generalized representation of the physical model without relying as much on specific constraints present in the data set. This increase in complexity and variation within the data sets manifested mainly by an increased number of vessel structures and multispectral changes of background scattering and absorption, which were fixed in the phantom setups.

A multitude of informed assumptions for the morphological, optical, and acoustic tissue properties was made in the optical and acoustic forward simulation. While these were chosen carefully with respect to literature references that were deemed as reliable as possible, these assumptions might still not be representative of reality and, therefore, ill-chosen. Nevertheless, this does not significantly impede the results of this work towards evaluating the feasibility of data-driven approaches to the inverse problems of qPAI. This is because the methods were trained and tested on subsets of the same data set containing the same assumptions. It would be expected that changes in the assumptions would not impede the inversion accuracy to the extent that different conclusions would be drawn.

### *Experimental design*

All chromophore structures in the data sets were modeled to have a roughly circular shape. This was done because it was computationally very efficient to approximate vessels as tubular structures. Due to this, the trained models learn to rely on the fact that all structures are of circular shape. For the application of an end-to-end method in more diverse or realistic applications, constraints need to be employed that circumvent any (unrealistic) bias towards the shape of chromophores.

During the evaluation of the test data, no noise was applied to the input features, and only slight noise augmentation was performed during training. This was done because prior work towards deep learning techniques applied to the PA inverse problems showed that these algorithms are

generally very robust to noise [Kirchner et al., 2018a]. However, in retrospect, it would have been interesting to see the change of accuracy in relation to the imaging depth with the presence of noisy data, as the SNR would deteriorate much faster in a noisy setting.

Due to hardware limitations, a grid size of  $128 \times 128$  pixels was chosen for the optical forward model. This necessitated an up-sampling of the resulting pressure distribution for the acoustic forward model and led to a maximum possible simulated frequency of about 11MHz. These constraints and additional processing steps could also lead to inaccuracies and the loss of important information.

#### *Deep learning approaches*

In principle, convolutional deep learning architectures have been shown to be capable of efficiently processing large 3D input data [Çiçek et al., 2016]. However, for the sake of computational efficiency and feasibility of data storage, the acoustic forward simulation has been performed on 2D image slices extracted from 3D optical forward simulations of the symmetric data sets. This simplification makes it feasible to simulate the large amounts ( $4 \times 2.000$  volumes) of data needed for the training of the algorithms in a feasible amount of time. However, it also introduces inaccuracies in the acoustic forward model and necessitates a symmetry constraint along the y-axis of the data set that not present in realistic scenarios. Algorithms will have to be able to compute 3D inversions that are capable of capturing complex 3D geometries and out of plane light absorption and changes to the fluence. From the results seen in this work, it is to be expected, that algorithms could be capable of computing the 3D inverse problems, but it is highly likely that a lot more training data would be needed for the algorithms to achieve comparable accuracy.

Deep learning methods have been applied to a variety of reconstruction problems [Zhang and Dong, 2019], e.g. US image reconstruction [Nair et al., 2018, Yoon et al., 2018], CT image reconstruction [Kang et al., 2017], or MRI image reconstruction [Wang et al., 2016]. The biggest criticism towards these applications is the lack of robustness to unpredictable variations to the input data or the question, whether the algorithms learn a sensible reconstruction (cf. e.g. [Wang et al., 2018a, Zhang and Dong, 2019]). Approaches need to be established that can robustly handle such anomalies. This might be achieved using pre-processing steps such as out of distribution detection [Adler et al., 2019b] or by thinking about combined model-based and data-driven approaches. Even if it can be assumed that the employed algorithms for the acoustic inverse problem are capable of finding a generalized model to solve the inverse problem of the wave equations, they will find an approximation that is specifically tailored to the data set and might not generalize to different distributions and data sets. Methods would need to be developed that ensure a generalization of the machine learning method to arbitrary solutions of these equations. This could be achieved, for instance, by creating algorithms that are informed by the physical models, as, for example, demonstrated by [Lan et al., 2019].

#### *Transfer to reality*

A transfer of this method to real data was not demonstrated, as the neural network trained on the generated synthetic data was not capable of generating any plausible results for real recorded raw time series data, and no data with adequate  $p_0$  reconstructions were available to test the optical inverse method. In order to achieve a transfer of data-driven qPAI approaches to

real data scenarios, methods have to be investigated that narrow the apparent domain gap.

The two next big steps that can be taken towards a transfer of these methods to realistic scenarios are (1) the creation of more and more realistic forward simulations that are explicitly tailored to the imaging device that is being used and (2) the development of hybrid inversion models, where data-driven method is informed by physical models and vice versa.

#### *Comparison to related techniques*

In [Waibel et al., 2018], the results of the learned initial pressure reconstruction from time series data were compared to that of Delay-And-Sum (DAS) beamforming. The results revealed an insurmountable performance gap between these two methods. The main reason for this is that the main objective of the DAS algorithm is to localize the source of the pressure wave and does not include any compensation for the loss of information caused by the limited view problem and the limited frequency response bandwidth of the US transducer. This is in contrast to the learned algorithm, that is specifically trained to account for this loss of information and attempts to reconstruct the full initial pressure distribution. Because of these reasons, a comparison of the proposed method to the commonly used DAS algorithm was not considered for the evaluation presented in this section.



### 3.2.5 Conclusion

At the beginning of this chapter, three important research questions regarding the overarching hypothesis  $H_2$  were formulated:

1) *Is it feasible to solve the acoustic inverse problem with end-to-end deep learning?*

To investigate this objective, acoustic simulations for four data sets were created, and it was attempted to estimate the initial pressure distribution from the raw time series data. Some results look promising, for example, on the simple phantom with a median estimation error of 13% in the region of interest, which could indicate the feasibility of applying data-driven approaches for quantitative PA image reconstruction. However, on the forearm data set, which was created to more closely resemble structures present in real tissue, the performance was much worse, only achieving a median estimation error of 34% in the region of interest. While these errors seem very high, it is difficult to estimate what magnitude of error would still be sufficient in order for post-processing steps to work correctly. In conclusion, some of the reconstructed images qualitatively look very promising, but the general feasibility of end-to-end learning for accurate image reconstruction could not be sufficiently verified with the conducted experiments.

2) *Is it feasible to solve the optical inverse problem with end-to-end deep learning?*

The models tasked to learn the optical inverse problem, on the other hand, achieved much better results. Here, the relative median estimation errors were never above 15%, even for the forearm data set, and the qualitative results also look very promising. As such, it can be concluded that - within the constraints of the used data sets and the model architecture - the estimation of the optical absorption coefficient distribution from initial pressure data is feasible. Of course, one of the primary constraints for this method to be applied is the presence of accurate measurements of the initial pressure distribution  $p_0$ , which raises further questions towards the practical applicability of such an approach.

3) *Is it feasible to solve the combined inverse problem with a one-stage end-to-end deep learning method?*

The model trained to estimate  $\mu_a$  from raw time series data  $p(t)$  directly, generally performed the worst out of all approaches. Here, the best achieved median relative estimation error was 18.1% on the most simple data set, and the approach performed the worst on the forearm data set with a median relative estimation error of 20% in the ROI. The qualitative result images reveal that the single-stage approach especially struggled in deeper tissue regions and also had difficulties in identifying the size of the absorbing structures correctly. With these results, the general feasibility of end-to-end learning for a single-stage approach, accurate image reconstruction, and absorption estimation could not be verified with the conducted experiments. However, it was also shown that a single-stage approach can be much more accurate, as the errors of the individual inversion algorithms are not propagated to the following steps.

It is apparent from the results, that it is very difficult to learn the acoustic inverse problem in an end-to-end manner subject to the chosen model architecture and the nature of the data sets. There are many possible reasons why this might be the case, such as the network architecture, ill-chosen parameters in the forward model, or general difficulty of the problem. Given

these problems, one interesting future direction is the investigation of *physically constrained* data-driven approaches, where the reconstruction algorithm is informed by model-based reconstruction algorithms that are based on the underlying physical principles of sound propagation in tissue. This could be imagined to work in a similar fashion, as demonstrated by [Lan et al., 2019].

While this work has demonstrated that the application of data-driven approaches to the photoacoustic inverse problems is not straightforward, the initial results look very promising, especially considering the results regarding the optical inverse problem. Data-driven approaches promise fast inference times and can be tailored specifically for specific use cases by intelligently choosing the training data for the model. However, it has to be concluded that more work has to be conducted towards the application of data-driven methods for the photoacoustic problems in order to be able to exploit their full potential on clinically acquired data.

*Structural distortion and **spectral coloring** are two manifestations of the same phenomenon, namely the **effect of the fluence** on the photoacoustic image.*

– [Cox et al., 2012]

## 3.3 | Functional photoacoustic imaging with learned spectral decoloring

### Disclosures to this work:

Lena Maier-Hein supervised this work. Herself, Thomas Kirchner, and Tim Adler were involved in various stages of the research process and contributed valuable ideas and suggestions. Parts of this work have been published in [Gröhl et al., 2019]. The oxygenation flow phantom data was contributed by L. Hacker and S. Bohndiek of the VisionLab of Cambridge University; the porcine brain data were imaged by E. Santos, M. Herrera, and A. Hernández-Aguilera of the University Hospital of Heidelberg University; and the forearm data was recorded by my colleagues T. Kirchner, M. Schellenberg, and N. Holzwarth.

### 3.3.1 Introduction

In healthcare, knowledge of blood oxygen saturation ( $sO_2$ ) is a critical indicator of the health status of a patient. However, state-of-the-art methods to obtain this value are limited as they are either invasive (for example, blood gas analysis especially when drawing arterial samples [Kelly, 2010]), lack practicability and accuracy (for example, Blood Oxygen Level Dependent MRI (BOLD MRI) [Haacke et al., 1995] or functional Near-Infrared Spectroscopy (fNIRS) [Tak and Ye, 2014]), or only yields a rough estimate from peripheral vasculature (for example, pulse-oximetry [Severinghaus and Honda, 1987]). None of these techniques can give real-time, spatially-resolved  $sO_2$  estimates. Due to its non-invasive nature, the live and spatially-resolved monitoring of  $sO_2$  is considered to be one of the main applications of Photoacoustic Imaging (PAI). Typically, this is achieved by using Linear Unmixing (LU) (see section 2.2.5) of the optical absorption spectra for the two relevant chromophores (oxyhemoglobin  $HbO_2$  and hemoglobin  $Hb$ ) from multispectral PA images. The core assumption for these algorithms is that the signal intensities of the reconstructed PA image  $S$  (which is an approximation of the underlying initial pressure distribution  $p_0$ ) are only proportional to the optical absorption coefficients  $\mu_a$  of the chromophore distribution:

$$S \approx p_0 \propto \mu_a. \quad (3.2)$$

However, as explained in detail in section 2.2.1, this assumption is generally speaking not true, because (even when assuming the Grüneisen parameter to be constant) the reconstructed image is proportional to both the optical absorption coefficients  $\mu_a$  and the light fluence  $\phi$ :

$$S \approx p_0 \propto \mu_a \cdot \phi. \quad (3.3)$$

Due to this non-linear relationship, the fluence has a considerable influence on the recorded multispectral signal. As the optical absorption and scattering coefficients change with wavelength, so does the fluence, which leads to changes in the multispectral images that are very hard to predict. More specifically, this effect occurs due to wavelength-dependent absorption in the surrounding tissue and is generally referred to as *spectral corruption* [Tzoumas et al., 2016] or *spectral coloring* [Cox et al., 2012]. For example, even small absorption coefficients in the background medium can lead to more or less pronounced coloring effects depending on the depth in the medium. Figure 3.20 shows the results of simulation experiments to demonstrate the non-linear changes in recorded  $p_0$  spectra at different locations in a vessel structure. All parameters except the vessel radius are kept constant for the demonstration.

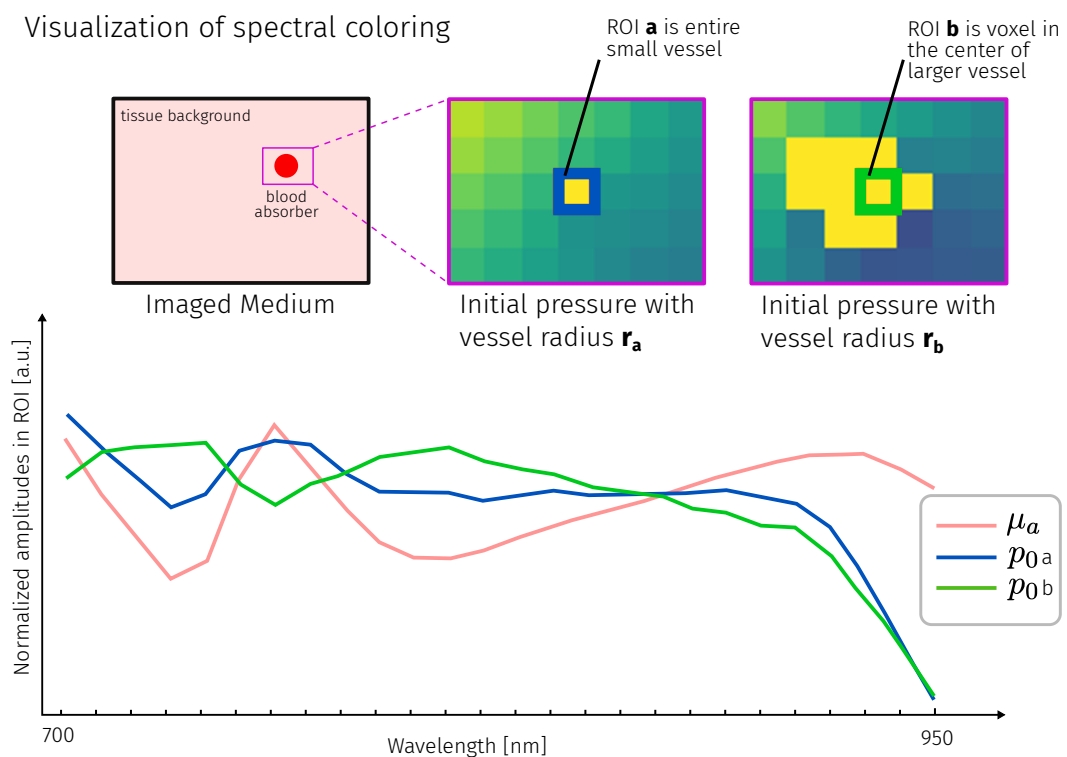


Figure 3.20: Illustration of spectral coloring in a simulation experiment. The green and blue lines show the differences in  $p_0$  spectra in two different geometric scenarios **a** and **b**, where all optical parameters are constant, and only the vessel radius is changed. The colored boxes highlight the respective pixel that the spectra were extracted from. The red line shows the used absorption spectrum of the blood within the tube.

The work presented in this chapter introduces a data-driven approach to tackle the problem of  $sO_2$  estimation by introducing Learned Spectral Decoloring for quantitative Photoacoustic Imaging (LSD-qPAI). LSD-qPAI is based on the assumption that a data-driven algorithm can learn how the  $p_0$  spectrum is colored at different depths in tissue when presented enough data to analyze. While other approaches try to quantify the underlying optical absorption coefficients

first and then perform LU on the estimation results, LSD-qPAI directly estimates  $sO_2$  from pixel-wise  $p_0$  spectra. It is trained on *in silico* training data and applied to various *in silico*, *in vitro*, and *in vivo* data sets. According to the results,  $sO_2$  estimates using LSD-qPAI exhibit a higher dynamic range and seem more plausible than those of LU techniques.

### Hypothesis investigated in this chapter

*H2: Data-driven methods can be used for spectral unmixing in a realistic context.*

The two previous chapters investigated the principle feasibility of using data-driven methods for the inverse problems of qPAI. In contrast to that, this chapter focuses on two research questions derived from hypothesis **H2**:

- 1) Can a data-driven method be used to obtain quantitative  $sO_2$  estimates in the presence of spectrally colored data?
- 2) Can such a model trained solely on simulated data be applied in a realistic scenario and still yield plausible results?

## 3.3.2 Methods

### Concept overview

LSD-qPAI is a data-driven regressor that approximates a function  $f_{\text{LSD-qPAI}}$  that maps initial pressure  $p_0$  spectra  $S_{p_0}$  to blood oxygenation saturation  $sO_2$  values:

$$f_{\text{LSD-qPAI}} : S_{p_0} = \begin{pmatrix} p_{0\lambda_1} \\ \dots \\ p_{0\lambda_n} \end{pmatrix} \in \mathbb{R}^n \rightarrow sO_2 \in \mathbb{R} \quad (3.4)$$

with  $n$  being the number of recorded wavelengths.  $f_{\text{LSD-qPAI}}$  is a neural network that is trained to compensate for different extents of spectral coloring and learns a many-to-one mapping, where many differently colored  $p_0$  spectra correspond to the same  $sO_2$  value. As there is an inherent lack of ground truth  $sO_2$  values for real  $p_0$  measurements, the method is trained on simulated data sets that can be optimized for the specific applications and wavelengths (see section 3.3.2). When sampling single-pixel spectra from different spatial locations within the same image, various degrees of spectral coloring are simulated. Examples for this are shallow or deep samples within the homogeneous background. Here, the absorption of water is more expressed for deep samples, as there has been a greater amount of interaction of the light with the respective chromophore. A graphical representation of the spectra extraction from simulated  $p_0$  data is shown in figure 3.21.

During training, the algorithm is presented tuples  $(S_{p_0}, sO_2)$ , with  $S_{p_0} \in \mathbb{R}^n$  and  $sO_2 \in \mathbb{R}$ . Each spectrum  $S_{p_0}$  is normalized such that

$$\sum_{i=1}^n S_{p_{0\lambda_i}} = 1 \quad (3.5)$$

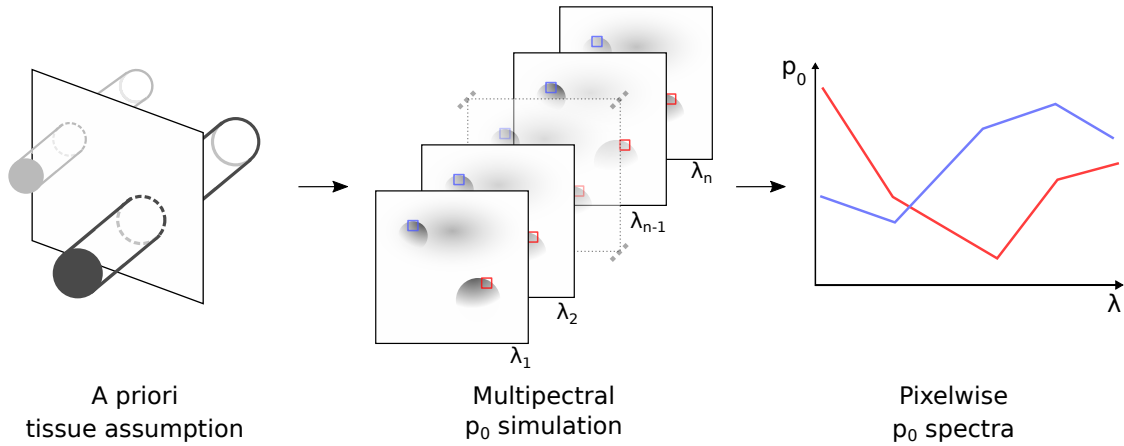


Figure 3.21: Overview of the core idea of extracting differently colored spectra from the same *a priori* simulation in LSD-qPAI. Various geometrical *a priori* tissue assumptions are generated, and light delivery is simulated at multiple wavelengths. Pixel-wise spectra are extracted from the multi-wavelength simulation by following the  $p_0$  intensity at a fixed pixel location as a function of wavelength.

Because the extracted training  $p_0$  spectra are normalized, the spectra acquired from *in vitro* or *in vivo* settings need to be normalized as well. This sacrifices the valuable amplitude information of the spectra to eliminate the need to calibrate the *in silico* training data to the acquisition device and the specific target domain. For quantitative recovery of not just the ratios but absolute concentrations, an accurate calibration would be preferred, but in this case, the calculation of the  $sO_2$  ratio is in focus.

### Data sets

Several data sets were used in this work to provide means of training and evaluation of the LSD-qPAI method. These evaluation data comprise (a) *in silico* simulated data, (b) *in vitro* recorded phantom data from a blood flow phantom, (c) *in vivo* animal data of an open porcine brain, and (d) *in vivo* data from forearms of healthy human volunteers.

**(a) Synthetic data.** A total of three simulated data sets were generated to train the models and test the method in different scenarios.

1) The first was a *generalized data set* of similar composition as the tissue model presented in section 3.2.2. It contained several randomly distributed vessel structures and a homogeneous background structure and were initialized with the same random blood oxygenation. All oxygenations were drawn from a uniform distribution from 0% to 100% oxygenation. The background scattering was set to  $\mu_s = 10 \text{ cm}^{-1}$ , while all other parameters were the same as reported in section 3.2.2. The vessel phantoms were simulated with 26 wavelengths equidistant from 700 nm to 950 nm (in 10 nm steps) using a multithreaded adaptation of the Monte Carlo framework mcxyz [Jacques, 2014] with  $10^7$  photons for each simulation. In this data set, only simulated spectra corresponding to vessel structures were included, effectively eliminating background

spectra from the data set.

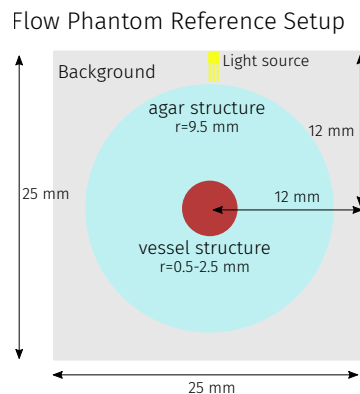


Figure 3.22: Schematic representation of the *in silico* flow phantom data set. A circular vessel structure with varying radius (0.5-2.5 mm) is embedded in a circular agar structure with a radius of 9.5 mm. A single point light source illuminates the phantom which is placed inside a background medium that is not optically active.

2) The *flow phantom data set* was simulated in order to closely resemble the geometrical setup of the oxygenation flow phantom depicted in figure 3.22. The phantom was simulated to be a tubular agar phantom with a reduced scattering coefficient  $\mu'_s$  of  $5 \text{ cm}^{-1}$  and a water content of 50% - 100%. The phantom contains a single tubular structure containing blood with a radius between 0.5 and 2.5 mm. It is submerged in a solution which is assumed to have negligible absorption and scattering. Even though the radius of the tubular structure was known a priori, the radius of the simulated tubular structure was varied in order to obtain more diverse samples and to reduce the effects from voxel discretization. These structures were simulated with the same wavelengths that were used for the flow phantom data acquisition {660 nm, 664 nm, 680 nm, 684 nm, 694 nm, 700 nm, 708 nm, 715 nm, 730 nm, 735 nm, 760 nm, 770 nm, 775 nm, 779 nm, 800 nm, 850 nm, 950 nm}. For this data set, the Monte Carlo eXtreme (MCX) simulation framework [Fang and Boas, 2009, Yu et al., 2018] was used due to its fast computational speed with  $10^7$  photons for each simulation. Spectra were extracted exclusively from an Region Of Interest (ROI) defined as vessel structures where the signal at the isosbestic point of 800 nm was higher than a noise equivalent level (determined by calculating the pixel-wise CNR and setting a threshold of  $\text{CNR} \geq 2$ ).

3) The *forearm model data set* was exactly created, as described in section 3.2.2. Here, phantoms were also simulated with 26 wavelengths equidistant from 700 nm to 950 nm in 10 nm steps, also using the MCX simulation framework with  $10^7$  photons for each simulation. Only spectra corresponding to vascular structures and expressing a  $\text{CNR} \geq 2$  were included in the data set. This is done because frequency responses of the detection device and detection noise make it impossible to extract any meaningful information that deep in a blood vessel within *in vivo* images.

**(b) Blood flow phantom data.** Three data sets of PA data from a blood flow phantom setup,

including reference blood oxygenation measurements provided by partial oxygen pressure ( $pO_2$ ) needle probes, were generously provided by Lina Hacker of the VisionLab led by Prof. Dr. Sarah Bohndiek at Cambridge University. These data sets contained measurements of two human blood samples and a rat blood sample. A schematic of the measurement setup is reprinted from the original publication [Gehring et al., 2019] from the lab in figure 3.23. The data was measured using an MSOT inVision 256-TF imaging system (iThera Medical GmbH, Munich, Germany). The blood samples were first chemically oxygenated and then chemically deoxygenated during the measurement process, with continuous reference measurements of the  $pO_2$  needle probes. These  $pO_2$  measurements were translated into  $sO_2$  estimates using the Severinghaus equation [Severinghaus, 1979, Collins et al., 2015]:

$$sO_2(\%) = \left( \left( \left( (pO_2^3 + 150 \cdot pO_2)^{-1} \times 23,400 \right) + 1 \right)^{-1} \right) \times 100 \quad (3.6)$$

A detailed description of the data acquisition process can be found in the original publication.

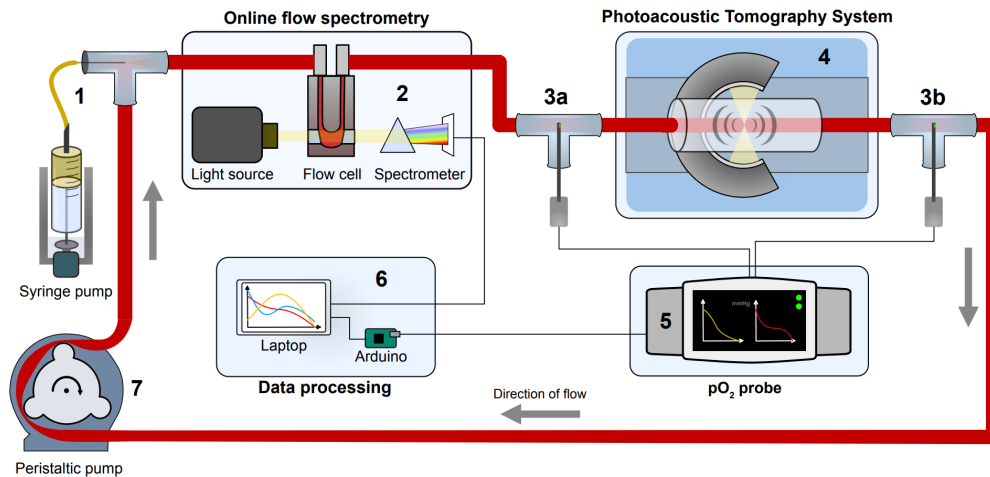


Figure 3.23: Overview of the flow system as presented in [Gehring et al., 2019]. (1) Injection site for introducing fluids into the flow system; (2) spectrometer setup for reference measurements; (3) needle probes measure the temperature and partial pressure of oxygen ( $pO_2$ ); (4) the PA imaging system; (5) touch-screen monitor to display temperature and oxygen data; (6) data interface via an Arduino; (7) peristaltic pump to induce blood circulation; This figure is reprinted from the original (CC BY 4.0) publication from [Gehring et al., 2019]. Figure description adapted.

For the evaluation of the method on this data set, the tube structure was automatically segmented by only taking into consideration pixels, where the signal at 800 nm was greater than  $2 \times 10^4$  MSOT signal units. This threshold was arbitrarily chosen and validated to yield a good fit of the vessel structure for each of the data sets. This step was necessary, as the tubular structure was subject to slight movements over the imaging duration, and, as such, using a constant hand-drawn ROI for all images was not feasible.

**(c) Porcine brain data.** A multispectral image series was taken from a previous animal experiment [Kirchner et al., 2019] in which a porcine brain was imaged by E. Santos, M. Herrera, and



A. Hernández-Aguilera during open brain surgery. These images were recorded at the same wavelengths as in the training data set (700-950 nm in 10 nm steps). The images were normalized by the recorded laser energy and reconstructed with the delay-and-sum algorithm using a hamming window. These images were recorded using the DiPhAs imaging system (see section 2.3.1). Further details on the general image acquisition can be found in the publication of the previous experiment [Kirchner et al., 2019]

**(d) Human forearm data.** Multispectral images of the forearms of 15 healthy human volunteers were acquired. The images were taken at two different distances from the radiocarpal joint: at the "watch location" of 2-3 cm and at a distance of 5-6 cm, leading to a total of  $N=30$  multispectral images of forearm structures. For each recording, the person operating the handheld PAI device tried to either capture the *arteria radialis* or the *arteria ulnaris* in the center of the imaging plane. These vessels could be identified by their pulsating motion induced by the heartbeat when slightly increasing the pressure of the probe. The data sets were recorded using the iThera MSOT Acuity Echo PAI device. The images were taken at a wavelength range from 660-1300 nm in 10 nm steps, and the matching wavelengths to the training data (700-950 nm in 10 nm steps) were extracted in a post-processing step. The last five images of the recording of each wavelength were averaged to account for laser intensity fluctuations and increase the robustness against motion artifacts.

## Implementation details

### Deep learning models

For each of the training data sets, a separate fully connected feed forward neural network (for the architecture details see 2.4.3) was trained. The number of input features was set to the number of wavelengths (26 for the forearm and generic data set and 17 for the flow phantom data set), and the size of the hidden layers was set to be twice the size of the input vector. The models were trained for 100 epochs, where one epoch was defined to contain 500 batches of size  $10^4$ . The initial learning rate was set to  $10^{-2}$  and was updated every two epochs to  $\text{new}_{lr} = 10^{-2} \times 0.9^{(\text{epoch}/2)}$ . Tracking of the validation losses over the number of epochs showed that the validation loss did not significantly decrease after as little as ten epochs using this training scheme.

### Linear unmixing

LU (cf. section 2.2.5) served as a reference method to compare the LSD-qPAI method to. It was performed using literature absorption spectra of pure Hb and HbO<sub>2</sub> as reviewed by Steven L. Jacques [Jacques, 2013], downloaded from the OMLC website and contained in the mcxyz package as *spectralLIB.dat* (<https://omlc.org/software/mc/mcxyz/index.html>, as visited on 4.12.2019). The implementation of the unmixing was done in Python 3.7., using the *minimize* function of the *scipy* python package that implements the SLSQP (Sequential Least Squares Programming) algorithm for finding the minimum. The unmixing was done for Hb and HbO<sub>2</sub> assuming initial values of 0.5 and imposing a non-negativity constraint on the result.

While LU is used as a reference method for the *in vitro* and *in vivo* results, the *in silico* data set is designed to contain a high degree of spectral coloring. Because applying LU techniques to systematically colored data is not feasible and visually only results in quasi-random sO<sub>2</sub> estimates, the results of the LU algorithm are omitted for the *in silico* results.

### 3.3.3 Results

This chapter shows the results of the LSD-qPAI method on the simulated *in silico* data sets, on the *in vitro* flow phantom data sets, and on the *in vivo* porcine brain and human forearm data.

#### *In silico* results

Figure 3.24 shows the performance of the respective deep learning model when tasked to predict  $sO_2$  values for a held-out test set from the generic data set, the flow phantom data set, and the forearm data set. Here, the relative  $sO_2$  estimation error is reported, which is calculated as  $e_{sO_2} = |sO_2^{EST} - sO_2^{GT}|/sO_2^{GT}$ , with  $sO_2^{EST}$  being the estimated oxygen saturation and  $sO_2^{GT}$  being the ground truth oxygen saturation.

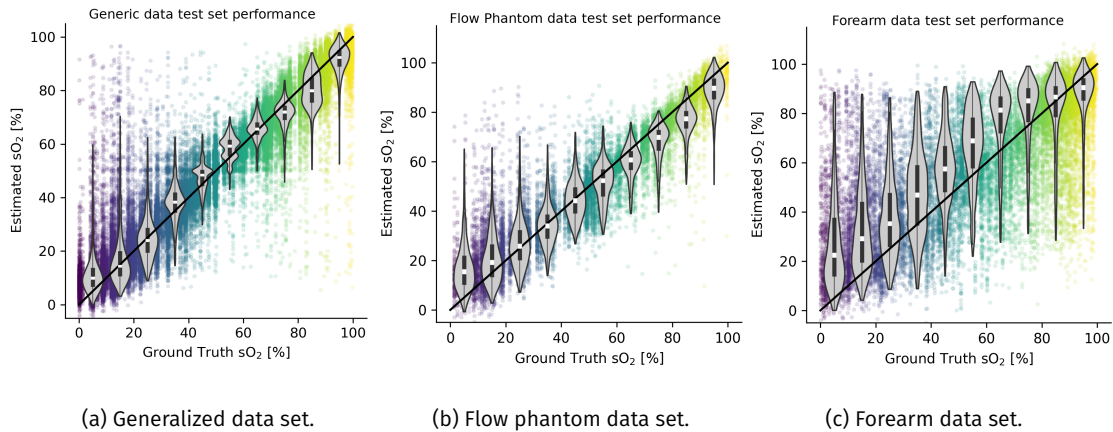


Figure 3.24: *In silico* estimation results for the generalized data set (a), the flow phantom data set (b), and the forearm data set (c). The scatter plot is colored with the ground truth oxygenation value. The violin plots show the estimated  $sO_2$  for ground truth  $sO_2$  intervals with a size of 10%. As such, in addition to the scatter plot, there is one violin plot for all ground truth  $sO_2$  values from 0% to 10%, one for 10% to 20%, and so forth.

The median relative  $sO_2$  estimation error for the model trained and tested on the generic tissue model data set was 6.1%, with an Interquartile Range (IQR) of (2.4%, 18.7%). The flow phantom data set achieved a median relative estimation error of 9.9%, with an IQR of (3.6%, 28.5%). The worst performing data set *in silico* was the forearm data set. The relative median quantification error was 16.6 %, with an IQR of (6.0%, 48.6%).

The median absolute  $sO_2$  quantification error on the test sets was in all cases well below 10%, with the model trained and tested on the forearm data set, achieving 8.8% absolute  $sO_2$  estimation error.

### In vitro results

The provided *in vitro* flow phantom data consisted out of two human blood samples and one rat blood sample, that were chemically oxygenated and then deoxygenated over a period of about 40 minutes each. Ten consecutive frames of the same wavelength were averaged to account for laser pulse energy fluctuations, and the averaged multi-wavelength stack was analyzed in the ROI, which was defined as: MSOT signal > 20,000. For the analysis, three sources of  $sO_2$  estimation are visualized in figure 3.25: spectral unmixing using LU, spectral unmixing using the proposed LSD-qPAI approach, and  $pO_2$  probe reference measurements. For both the LU and LSD-qPAI approach, the mean and standard deviation of the estimates are shown on the graphs.

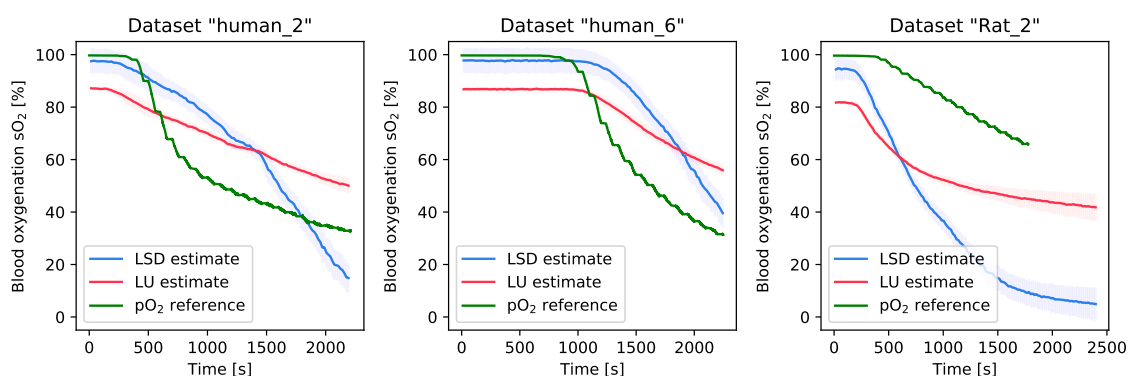


Figure 3.25: The mean oxygenation estimation results over time on three different blood samples with three different measurement methods: (1) LSD-qPAI spectral unmixing in blue, (2) LU in red, and (3)  $pO_2$  reference measurement in green. The standard deviation of the estimations within the ROI for LU and LSD-qPAI unmixing is visualized in the corresponding color around the mean estimate.

All three of these measurements showed a monotonous decrease of the blood oxygenation over the time frame of the experiment. The LU unmixing results exhibited a dynamic range of  $\approx 85\% - 40\%$ , whereas LSD-qPAI unmixing exhibited a dynamic range of  $\approx 95\% - 5\%$  on all three data sets. It has to be noted that the temporal calibration of the "human\_6" data set was manually adjusted by 1,000 seconds, as the curves did not seem to match initially and that the  $pO_2$  reference was prematurely stopped in the "Rat\_2" data set.

Figures 3.26 and 3.27 show qualitative oxygenation estimation results of the "human\_2" and "Rat\_2" data sets. At three points in time - at the beginning, the middle, and the end of the deoxygenation time - these figures show the spatially resolved  $sO_2$  estimations of the proposed method, as well as the LU result for comparison.

## Qualitative results of the "human\_2" data set

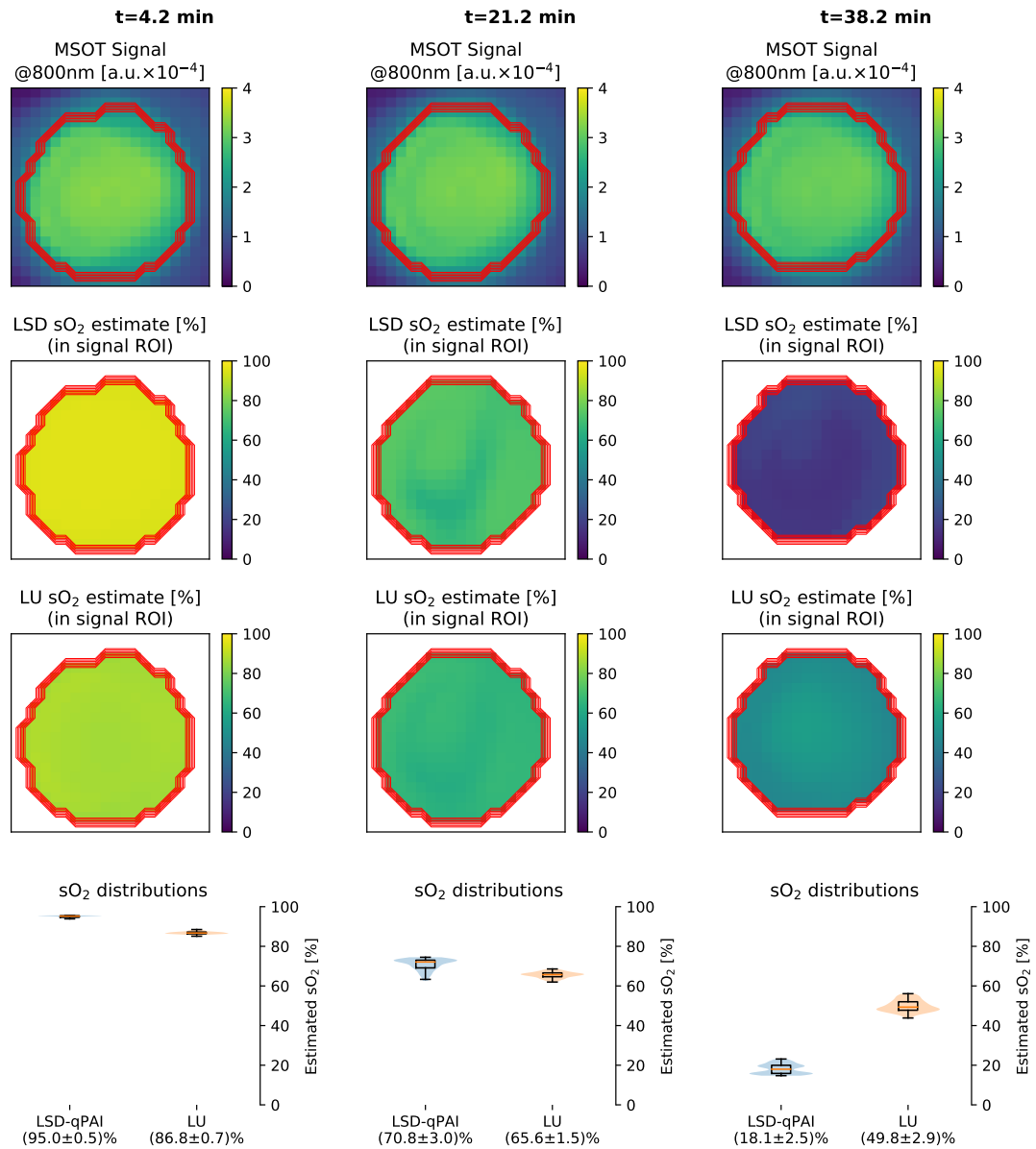


Figure 3.26: Qualitative spatially resolved images of estimated  $sO_2$  for the "human\_2" data set of the flow phantom data. From left to right, three points in time are shown, and from top to bottom the figure shows the MSOT signal at 800 nm, the LSD-qPAI  $sO_2$  estimate, the LU  $sO_2$  estimate, and violin/box plots of the distribution for LSD-qPAI (blue) and LU (orange).

Qualitative results of the "Rat\_2" data set

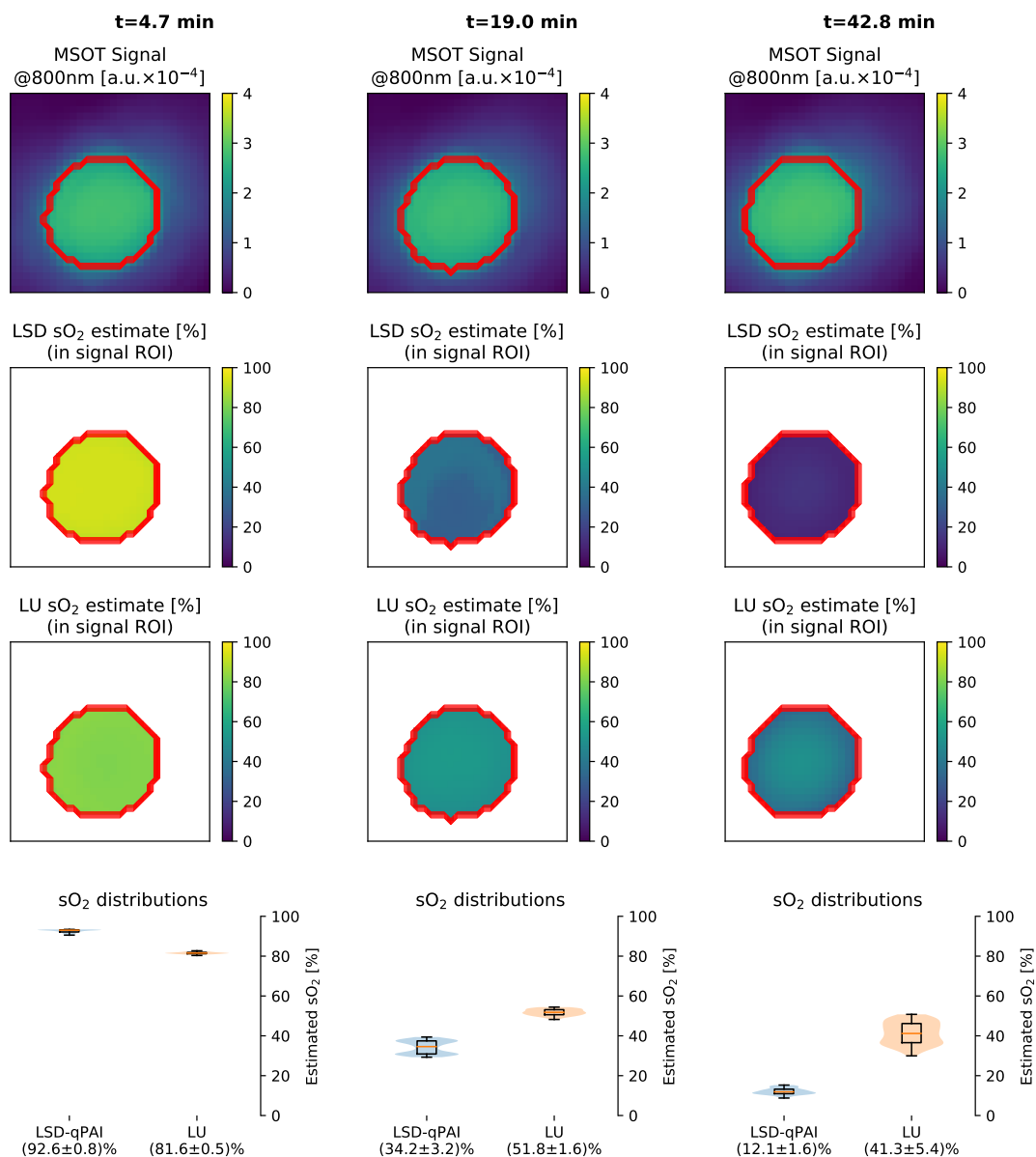


Figure 3.27: Qualitative spatially resolved images of estimated  $sO_2$  for the "Rat\_2" data set of the flow phantom data. From left to right, three points in time are shown, and from top to bottom the figure shows the MSOT signal at 800 nm, the LSD-qPAI  $sO_2$  estimate, the LU  $sO_2$  estimate, and violin/box plots of the distribution for LSD-qPAI (blue) and LU (orange).

### ***In vivo* results**

The proposed LSD-qPAI method was also applied to two *in vivo* data sets: a multispectral PA image of a porcine brain and PA images of the human forearm.

#### **Porcine brain results:**

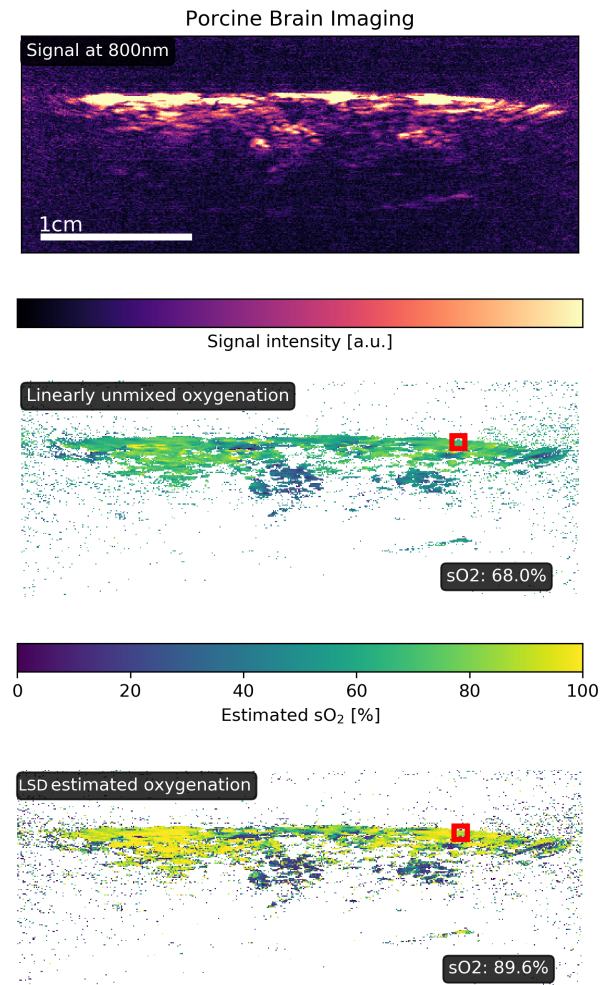


Figure 3.28: The results of LSD-qPAI *in vivo* on an open porcine brain with a deep learning model trained on the generic tissue data set. The LSD-qPAI results are compared to the LU results. The red rectangle identifies an ROI which corresponds to the mean  $sO_2$  estimate at the bottom of the two lower subfigures.

In the porcine brain image, LSD-qPAI estimated generally higher values for  $sO_2$  than compared to  $sO_2$  values obtained with LU. In general, LSD-qPAI was observed to increase the dynamic range of the predictions while maintaining the same tendency (high values were mapped to higher

values, and low values were mapped to lower values). For reference, a manually placed ROI was included in figure 3.28. The mean estimated  $sO_2$  in this region was 68.0% for LU and 89.6% for LSD-qPAI.

#### Forearm data results:

Figure 3.29 shows the  $sO_2$  distributions as estimated by LU and the proposed LSD-qPAI method. The results are analyzed on all spectra and over spectra that correspond to an ROI. This ROI was hand-selected in each image to contain one arterial vessel structure from each recorded image. The estimations on all spectra reveal a systematic positive shift of the median from  $\approx 30\%$  to  $\approx 45\%$  when using LSD-qPAI compared to LU. The same trend can be seen in the ROI spectra, where the median estimate is shifted from  $\approx 75\%$  to  $\approx 95\%$ . The LU results are exclusively between 0% and 100% because of the non-negativity constraint in the optimization, but the LSD-qPAI approach also predicts outside of these values.

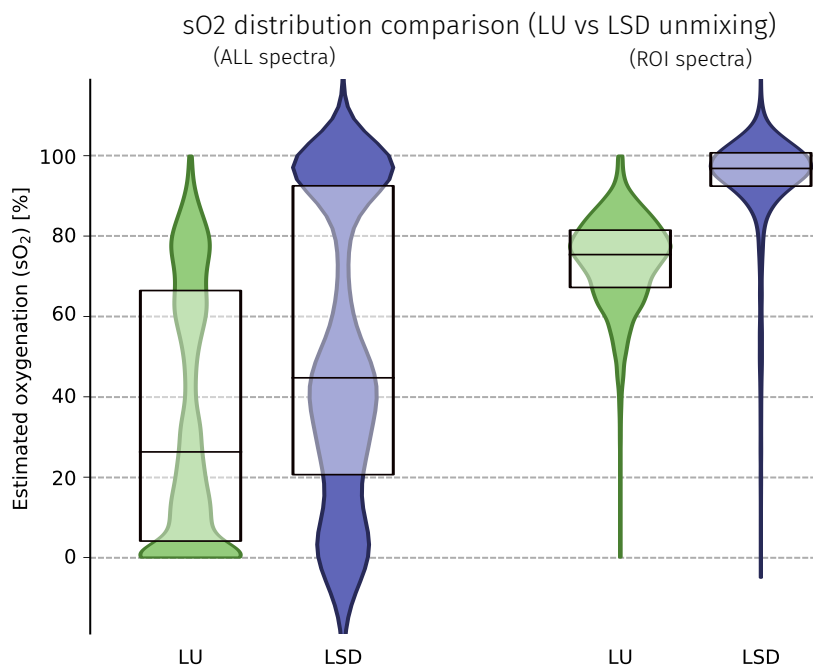


Figure 3.29: Comparison of the distributions of LU versus LSD-qPAI unmixed forearm spectra of  $N=30$  multispectral PA images. Green denotes LU estimates, and blue denotes LSD-qPAI estimates. On the left, the estimates of all spectra are shown, and on the right, only the spectra from a single arterial vessel structure (ROI) from each forearm image was taken into consideration.

Figure 3.30 shows two randomly chosen but representative examples of the *in vivo* forearm data set. The LSD-qPAI estimates were obtained from a deep learning model trained on the synthetic forearm data set. The results are compared to data analyzed with LU. On the left image, the selected vessel structure is estimated to have a blood oxygenation of 76.9% with LU and 97.5% with LSD. On the right image, LSD-qPAI estimates a blood oxygenation of 98.4%, while LU yields 70.7%. On the right image, one can clearly distinguish the artery from its accompanying vein.

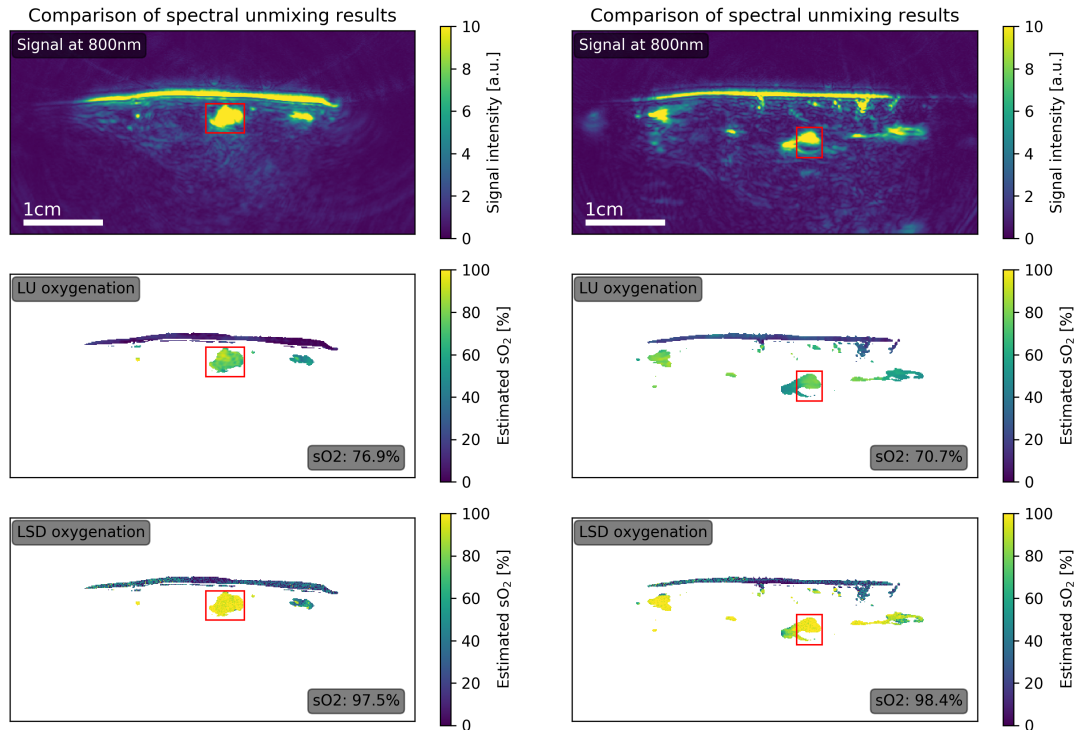


Figure 3.30: The results of LSD-qPAI *in vivo* on a human forearm, where the LSD-qPAI model specifically trained on the forearm data set. The results are compared to data analyzed with LU. The top row shows the PA signal at 800 nm, the middle row shows the LU results, and the bottom row shows the LSD-qPAI  $sO_2$  estimates.

### 3.3.4 Discussion

#### Discussion of results

The conducted experiments demonstrate the promising potential of the LSD-qPAI approach. The results of the synthetic data set suggest that accounting for various degrees of spectral coloring can be feasible. The more realistic variation with respect to the target structures was included in the respective simulated data set during forward simulation, the more inaccurate the  $sO_2$  inversion results became. This is most noticeable in the forearm data set. This data set was the only data set, where background and vessel structures were allowed to take different oxygenation values. This is expected to significantly increase the complexity of the inversion, as this might add more ambiguity to the problem.

The results of the *in vitro* flow phantom data show that the method can be applied to real multispectral PA images. Here, the  $sO_2$  estimates of the proposed method follow the same trend of a monotonous decrease in  $sO_2$  as reference  $pO_2$  and LU measurements. Furthermore, the LSD-qPAI  $sO_2$  estimates also exhibit a higher dynamic range on the spectra of the flow phantom data set than classical LU approaches.



The results of the *in vivo* data sets suggest that LSD-qPAI is capable of yielding physiologically plausible  $sO_2$  measurements *in vivo*. The forearm  $sO_2$  estimates included values higher than 100% and lower than 0%. This can be potentially explained by several different effects. On the one hand, the noise on the spectra might be different in the recorded images compared to the simple noise model that was applied to the simulated data. On the other hand, the approach does not exhibit the boundary exceeding values behavior in simulated data, which might indicate a domain mismatch between the simulation domain and the "real" domain. This is evident when inspecting the  $sO_2$  estimates of the skin signal in the forearm images. In the simulation, the epidermis was simulated not to contain blood, and therefore could not be assigned an oxygenation value. Because the inference of  $sO_2$  was the goal of the LSD-qPAI algorithm, the epidermis spectra were excluded from the training data set. Because of this, spectra dominated by melanin are not known by the inversion model, and they are often mapped to negative or very low  $sO_2$  values. This might potentially be remedied by using constraints on the solution space either during the training of the algorithm or during inference / as a post-processing step.

### Discussion of methodology

The proposed LSD-qPAI methodology suffers from two distinct drawbacks: (1) it works only on single-pixel spectra and (2) the inversion model is wavelength dependent.

(1) Single-pixel estimates might not be enough to guarantee an unambiguous and accurate inversion. This is because there is a trade-off regarding simplicity versus accuracy. A pixel-wise approach enables the use of a simple model and allows for a straightforward simulation of training data, whereas a spatially constrained approach would sacrifice these conveniences for the potential benefits of taking the signal context into account. Based on the results of the Context Encoding Quantitative Photoacoustic Imaging (CE-qPAI) method presented in section 3.1, it is to be expected that the use of spatial regularization, for example, by including more spatially co-located pixel spectra into the inversion can improve the accuracy of the inversion method. One of the reasons that only single-pixel spectra were taken into account is that the inverse problem is very easy to formulate and that training and application of a model is very straightforward. Furthermore, an approach that takes into consideration multiple estimates from a neighborhood also introduces further assumptions that have to be respected when trying to apply the method to new data. One of these assumptions is, for example, the pixel spacing. With further assumptions on the tissue geometry and the discretization grid, it would be crucial to simulate more realistic images, and the training process and live inversion capabilities would be more complicated and not as straightforward as with the proposed single-pixel approach.

(2) The proposed implementation of the LSD-qPAI method is wavelength dependent, which means that the  $sO_2$  estimation is only possible for specific wavelengths. This is a significant restriction compared to LU techniques that allow arbitrary combinations of wavelengths, as long as the absorption coefficients at these points are known, this is a big restriction. For each set of wavelengths that should be unmixed, a new model would need to be trained. This also makes the comparison of the method regarding more or less input wavelengths difficult, as the model complexity is dependent on the number of wavelengths. However, in clinical routine, it is most likely common practice to always use the same set of wavelengths for a specific use case. In such a scenario, this limitation might not be as impactful.

## Spectral coloring

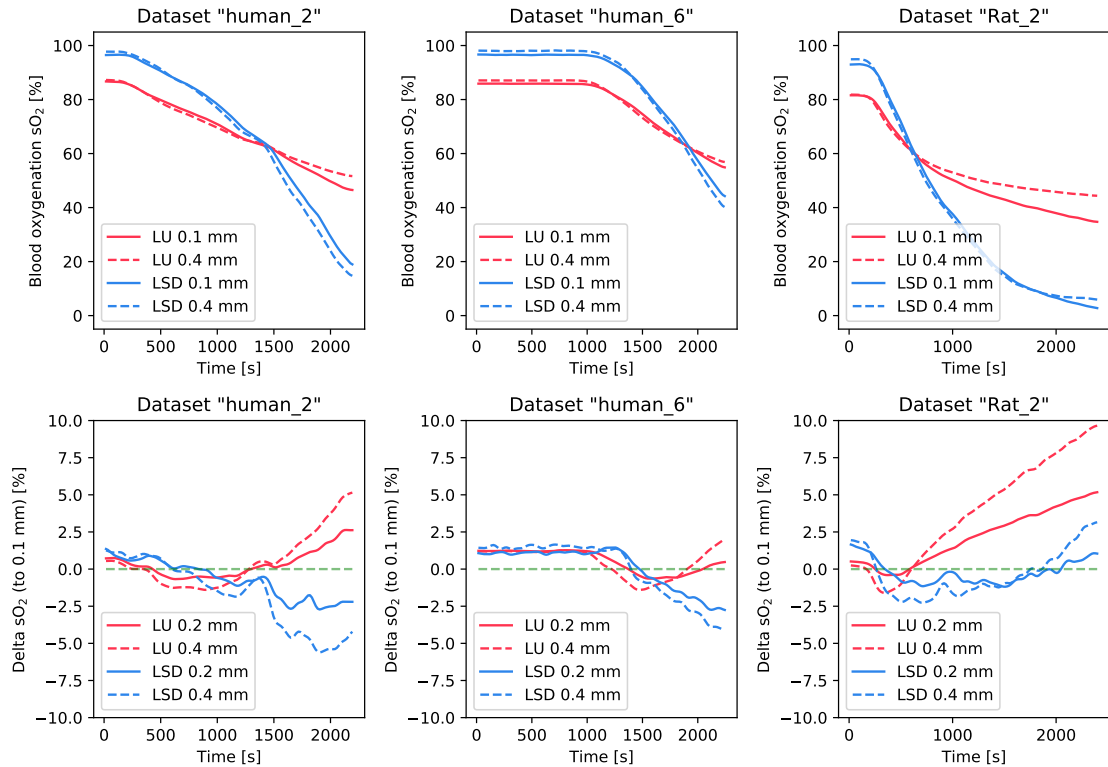


Figure 3.31: Overview visualization of the effect of spectral coloring in the flow phantom data sets. The three upper graphs show the absolute oxygenation estimates when only considering pixels of a certain depth within the tube. The three lower graphs show the absolute differences of the oxygenation estimates when comparing the estimated oxygenation at the surface of the tube (depth of 0.1 mm) with the estimated oxygenation in different depths (0.2 mm and 0.4 mm). The red lines denote the LU results, and the blue lines denote the LSD-qPAI unmixing results. The dashed green lines in the lower graphs denote a perfect result only achievable if no spatial dependence is present in the data.

The authors of the flow phantom publication [Gehring et al., 2019] included an analysis of the influence of spectral coloring into their work, by analyzing the oxygenation estimates over different depths of the tubular vessel structure. They observed that their  $sO_2$  estimations based on LU with hemoglobin spectra obtained from online spectrometer measurements had a tendency to decrease with increasing depth. LU based on literature spectra [Jacques, 2013] had a tendency to estimate lower oxygenation for deeper locations inside the tube when overall oxygenation was high and higher oxygenation deep inside the tube when overall oxygenation was low.

The bottom graphs of figure 3.31 demonstrate this behavior, where the red plots denoting the LU results exhibit the described tendency. In an ideal case, stable unmixing results would not

have a spatial dependence on the depth of the tube, and the differences of the estimates would lie around 0%, as depicted by the dashed green line in the graphs. The results of the LSD-qPAI method on the "Rat\_2" data set look very promising, while the LSD-qPAI performance seems to have the opposite effect as the LU results. In this context, one needs to be careful to attribute all changes in  $sO_2$  estimates to the depth of the vessel tube to spectral coloring. Blood is a viscous liquid, and Poiseuille's law shows that viscous liquids move in different circulation speeds at different locations in a tube [Pfitzner, 1976], which might also lead to a heterogeneous oxygenation distribution within the liquid that correlates to the tube depth.

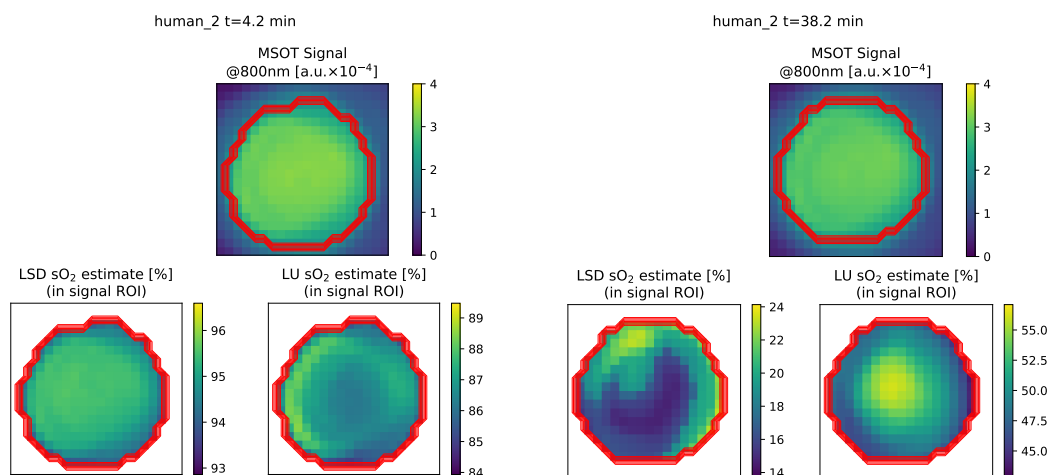


Figure 3.32: Visualization of spectral coloring on the flow phantom data set "human\_2" data set comparing LU and LSD-qPAI at two points in time: 4.2 min (left) and 38.2 min (right). For each of these time points, the upper images show the MSOT image and the signal intensity based ROI. The lower left image shows the LSD-qPAI  $sO_2$  estimates, and the lower right image shows the LU  $sO_2$  estimates. The dynamic range of the color mapping is adjusted for each image to emphasize the spatial changes of  $sO_2$  estimation.

Figures 3.32 and 3.33 show the same images as the qualitative result plots in figure 3.26 and figure 3.27, but with an adjusted dynamic range of the color mapping in order to emphasize the spatial variation of the  $sO_2$  estimations. In all of the LU visualizations, the depth-dependent effect of spectral coloring is visible, exhibiting a drop in  $sO_2$  for generally high  $sO_2$  values and a rise in  $sO_2$  for generally low  $sO_2$  values.

In contrast to this, the LSD-qPAI estimates corresponding to generally high oxygenation appear to be more homogeneous. For the low oxygenation case, the "Rat\_2" LSD-qPAI estimates also exhibit signs of spectral coloring - but of lower amplitude than with LU. Here, the "human\_2" estimates differ completely from the previously observed coloring pattern. There appears to be a structure in the estimates that is not only limited changes in the depth profile. This might be an indication, that this effect is not due to spectral coloring, but instead is caused by imaging artifacts or underlying physical phenomena, such as Poiseuille's law. However, due to the low number of data sets examined in this study, it is impossible to draw meaningful conclusions, and further investigation towards the meaning and reproducibility of these findings needs to be

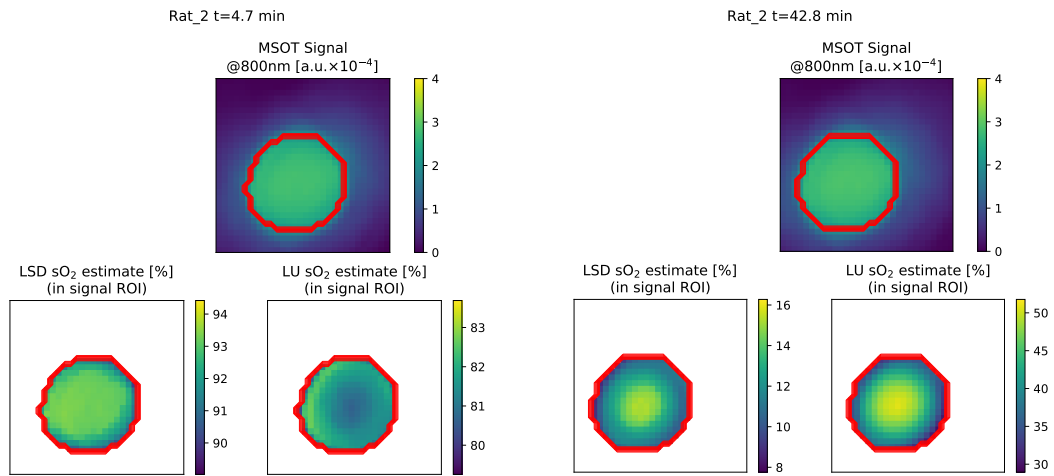


Figure 3.33: Visualization of spectral coloring on the flow phantom data set "Rat\_2" data set comparing LU and LSD-qPAI at two points in time: 4.7 min (left) and 42.8 min (right). For each of these time points, the upper images show the MSOT image and the signal intensity based ROI. The lower left image shows the LSD-qPAI sO<sub>2</sub> estimates, and the lower right image shows the LU sO<sub>2</sub> estimates. The dynamic range of the color mapping is adjusted for each image, to emphasize the spatial changes of sO<sub>2</sub> estimation.

conducted.

#### *Simulation gap to reality*

One of the main challenges when trying to apply data-driven methods to PAI is that simulations are only an approximation of the real underlying processes, where assumptions and approximations are made. As such, there is a large domain gap between simulated PA images and real PA images. Early iterations of this work have been solely conducted on a generalized data set, such as the generalized data set presented in this work. However, when switching the imaging domain, for example, from the brain to the forearm - where melanin plays a crucial role but is not modeled in the generic data set - the accuracy of the results degraded. In this regard, a specialized training data set is expected to yield more accurate sO<sub>2</sub> estimation results in realistic scenarios. In this case, this was done by using a forearm model data set for the forearm estimations that included an epidermis layer and had a fine-tuned range of the optical scattering and absorption coefficients.

One key aspect also seems to be a sophisticated extraction of the pixel spectra to train on. Background information is of very low frequency and can usually not be captured by most detection devices. For the structures, also only areas with sufficient CNR yield meaningful information to learn from, as most effects that happen in low CNR regions are not detectable in real case scenarios and - as such - do not contribute positively to the model's performance. This effect can be seen in figure 3.34, where the two bottom Principal Component (PC) distributions are calculated from the generic data set, in which the entire vessel structure is used for the

spectra extraction.

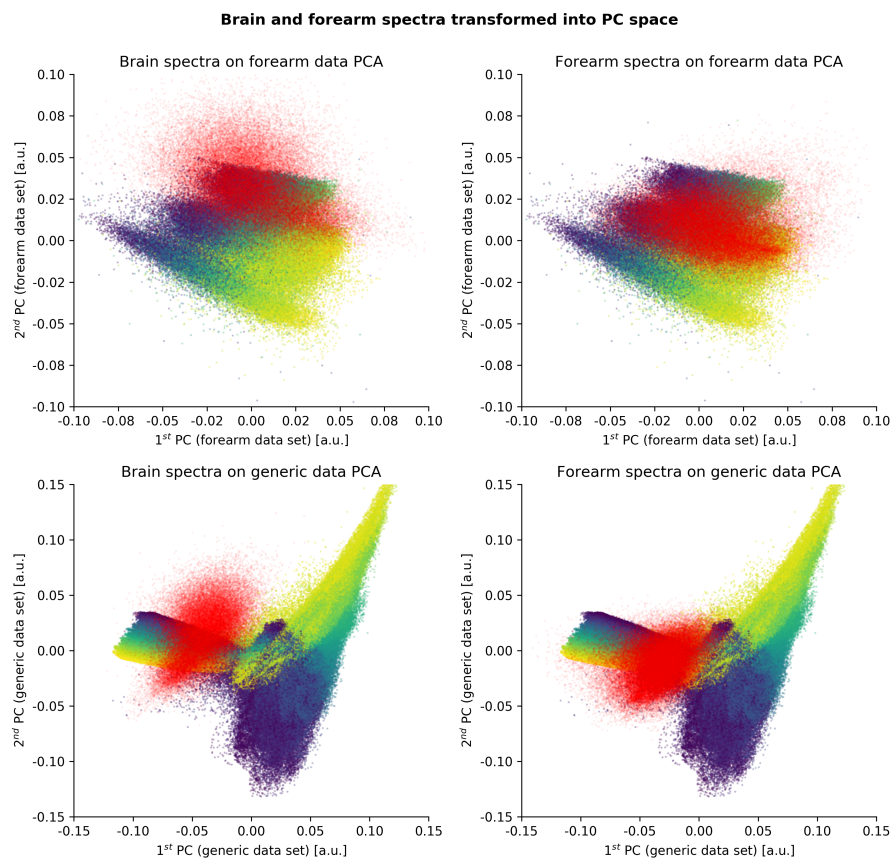


Figure 3.34: Principal component analysis of simulated and real spectra. The two upper scatter plots show how real recorded spectra are mapped into the space of the first two principal components of the forearm data set. The color of the plotted simulated spectra are annotated with their oxygenation values ranging from 0% (blue) to 100% (yellow) in a Viridis color mapping (cf. e.g. the color bar in figure 3.30). The two lower scatter plots show the same for the generic data set. Real data is plotted in red, where the two plots to the left show the mapping of brain spectra, and the two mappings to the right show the mapping of forearm spectra.

For the generic data set, the first PC is corresponding to the degree of spectral coloring, and one can see how the real recorded spectra (of both brain and forearm) are mapped to the left of this distribution. This might indicate the possibility that the spectral coloring inside the vessel structures is not captured in the PA data when using clinical ultrasound detection devices, because of limitations regarding, for example, the detection bandwidth and detectable SNR. In the two upper PC distributions, only those spectra were extracted from the forearm data set that corresponded to vascular structures and were above a CNR threshold. This data set did not

contain the same extreme spectral coloring effects as the previous data set, and the real recorded spectra (especially from the forearm data set) are mapped very nicely into the *in silico* distribution.

#### *Reference methods*

The accuracy of the measurements of the used  $pO_2$  probe and the accuracy of the Severinghouse equation to translate the measurements into  $sO_2$  values in this context is questionable. While the trend of a monotonous decrease is the same for all methods, both LU and LSD-qPAI show different rates of progression of the drop in  $sO_2$  over time. For absolute ground truth reference, a phantom setup would be needed in which, *a priori*, the absolute concentrations of the mixed chromophores are known very accurately.

The proposed LSD-qPAI method was only compared to LU, as it is the most commonly used method for  $sO_2$  estimation from multispectral PAI used in literature. A comparison to adjacent methods (such as eigenspectra Multispectral Optoacoustic Tomography (eMSOT) [Tzoumas et al., 2016]) would be very insightful, however, in this case, properly constraining eMSOT to handheld detection geometries has proven to be infeasible. In this regard, a thorough review and comparison of current spectral unmixing methods on well-designed data sets captured with various detection and illumination geometries would be highly interesting.

### 3.3.5 Conclusion

At the beginning of this section, two important research questions regarding the overarching hypothesis  $H_2$  were stated:

1) *Can a data-driven method be used to obtain quantitative  $sO_2$  estimates in the presence of spectrally colored data?*

Three highly spectrally colored data sets representing different use cases were simulated for the training of the LSD-qPAI models. In order to investigate the feasibility of accounting for spectral coloring, the models were evaluated on a held-out test set. This analysis revealed that the models were indeed capable of accounting for spectral coloring as encountered in the training data. It has to be noted that the more variation was included in the data sets, the less accurate the  $sO_2$  estimations became. The median absolute  $sO_2$  estimation error was well below 10% for all of these data sets.

2) *Can such a model trained solely on simulated data be applied in a realistic scenario and still yield plausible results?*

The results of this work demonstrate that it is indeed feasible to apply a data-driven method that is solely trained on *in silico* training data to realistic scenarios. Plausible results were demonstrated on phantom data, on porcine brain data, and on forearm data, with three different models that were all trained on different *in silico* data sets. The phantom results suggested that LSD-qPAI is capable of estimating a high dynamic range of  $sO_2$  values, whereas the results on the *in vivo* data sets suggest that LSD-qPAI is capable of recovering plausible  $sO_2$  estimates in arterial blood samples. In this case, it was possible because the proposed LSD-qPAI method does not see the entire image domain in which there is a huge domain gap to reality but instead a very condensed representation of the simulated data in the form of single-pixel spectra. Furthermore, a Principal Component Analysis (PCA) of the training spectra revealed that there is merit in carefully modeling the tissue geometry of the imaged region in order to obtain a better fit of the simulated spectra and the real recorded spectra.

It is to be expected that more careful consideration of imaging noise and artifacts, variations in laser pulse intensities, as well as spatial regularization techniques can lead to an improvement in the accuracy and robustness of LSD-qPAI or related techniques. In future work, it would be interesting to test the limitations of the technique, for example, in the minimum number of wavelengths needed to still be able to account for coloring artifacts as well as the influence of the choice of wavelengths for unmixing. Furthermore, an extension of the method of the estimation of different functional parameters or the recovery of absolute chromophore concentrations would be highly desirable.

Nevertheless, data-driven spectral unmixing techniques could be of great potential benefit for the clinical translation of PAI, as these methods are capable of accurate real-time estimation of functional tissue parameters and can easily be optimized for the respective clinical application by considering all relevant parameters in the training data.





Clinically, the use of [artificial intelligence] **in tandem with** PAI can potentially result in better disease management and patient outcomes by **augmenting** the qualitative assessments made by clinicians and introducing earlier interventions.

— [Attia et al., 2019]

## 3.4 | Uncertainty estimation for quantitative photoacoustic imaging

### Disclosures to this work:

Lena Maier-Hein has supervised this work. She, Tim Adler, and Thomas Kirchner have been part of various stages of the research process. L. Hacker and S. Bohndiek contributed the oxygenation flow phantom data. Parts of the methodological work of this chapter have been published in the *Journal of Imaging* [Gröhl et al., 2018a] and the proceedings of the *SPIE Photonics West: Photons plus Ultrasound 2018* conference [Gröhl et al., 2018b].

### 3.4.1 Introduction

So far, throughout this thesis, it has been demonstrated that data-driven methods can be suitable to tackle problems within the framework of photoacoustic imaging. However, it has also been discussed that many of these methods might not be usable in real-world scenarios, let alone in clinical practice. One of the reasons for this can be seen in the lack of robustness caused by systematic errors in *a priori* modeling assumptions, such as the illumination and detection geometry, calibrations, or the optical tissue properties. As such, developing methods to quantify the estimation uncertainty of these methods and to investigate the nature of error sources would be of great benefit for moving data-driven methods toward clinical applicability.

As introduced in section 2.3.4, sources of error in photoacoustic imaging can be differentiated into aleatoric and epistemic sources. Here, aleatoric uncertainty refers to stochastic uncertainty that is introduced, for example, by statistic fluctuations during the imaging process. Epistemic uncertainty, on the other hand, describes systematic errors in the model of the underlying process that can either be introduced by errors in the *a priori* assumptions or - especially for data-driven methods - it can be caused by a lack of representative data.

For clinical PAI applications, it is common to evaluate the photoacoustic signal or derived functional tissue parameters in an Region Of Interest (ROI) (for example, in a boxed region such as in [Hu et al., 2010]). This ROI is usually hand-drawn by the attending physician, and many times it is complemented by using an Signal-to-Noise Ratio (SNR)-based threshold in order to minimize the influence of aleatoric uncertainty in the derived results for the ROI. However,

this approach does not take into account the uncertainty accompanying the inferred functional parameter values. As uncertainty quantification for data-driven approaches is an essential research objective in computer sciences, many approaches to achieve this have been proposed to this date, such as Bayesian inference, latent space sampling, and ensemble techniques (for details and references see section 2.5.4). This work investigates the merit of applying such techniques to the inverse problems of photoacoustic imaging in order to improve on the accuracy of the parameter estimations by refining a purely SNR-based ROI with the use of uncertainty estimates.

---

**Hypothesis investigated in this chapter**

*H3: Confidence estimation methods be used to gain insight into qPAI methods.*

All previous sections have investigated the feasibility of the application of data-driven methods to obtain quantitative estimates of optical tissue properties or derived functional parameters. In order to successfully apply these in a clinical context, robust and reliable methods are needed that offer the attending surgeon an insight into the uncertainty of the given estimates. In this context, two research questions are derived from the overarching hypothesis:

- 1) Can a systematic exclusion of relatively uncertain estimates from a region of interest lead to a change of the aggregated estimate?
- 2) Are there differences between different state-of-the-art techniques for uncertainty estimation?

### 3.4.2 Methods

One of the goals of uncertainty estimation is to have the ability to calculate a confidence interval that the actual value is likely to reside in. In this work, the ability to estimate the standard error will be regarded as sufficient towards this goal, even though the ability to calculate the full posterior distributions would be preferential. The obtained estimates are then integrated into a multi-step approach, as demonstrated in figure 3.35.

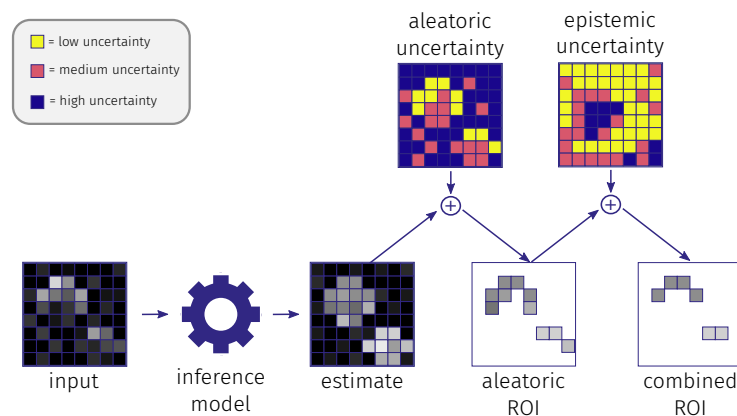


Figure 3.35: Overview figure for the uncertainty estimation workflow. In the first step, the properties of interest are calculated with an inference model from the input data. Secondly, the resulting estimate is masked with an SNR-based approximation of the aleatoric uncertainty, and finally, the SNR-based ROI is refined using a measure of the estimation uncertainty of the inference model to approximate the epistemic uncertainty.

Four different methods for retrieving estimation uncertainty are used and compared in this chapter: (1) dropout sampling, (2) the estimation of conditional probability densities instead of point estimates, (3) using an external model to estimate the expected error of the regressor, and (4) latent space sampling with a conditional INN:

#### (1) Dropout sampling

Dropout sampling was proposed as a method for network regularization to tackle the problem of overfitting [Srivastava et al., 2014]. The principle idea is to take away the possibility for the network to rely on individual nodes in the graph, by randomly ignoring (dropping out) a certain number of them in each forward pass. It was later also introduced as a method that attempts to approximate Bayesian processes in a neural network [Gal and Ghahramani, 2016]. Here, the input is evaluated multiple times, and one can compute the expected value as well as metrics for the variance from the samples (see figure 3.36). Even with the emergence of criticism towards the technical correctness of this approach [Hron et al., 2017] and with the emergence of more sophisticated methods for uncertainty estimation (such as variational approaches [Kohl et al., 2018]), it is still frequently being used as a well-understood benchmark and is, therefore, used for comparison in this thesis.

In this case, dropout sampling was implemented by adding dropout layers with a dropout probability of 20% into the network. During inference, the dropout layers remained active, and

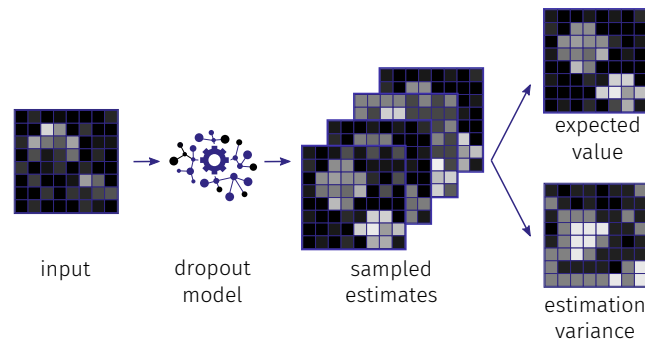


Figure 3.36: Visualization of dropout sampling for uncertainty estimation. With multiple forward passes, several estimations can be created that allow the computation of metrics such as the expected value and the estimation variance.

$N=100$  estimates were obtained by means of multiple forward passes through the network. The 15.87, 50.0, and 84.13 percentiles were then calculated from the obtained samples and used as estimates for the expected value and the surrounding error interval.

## (2) Conditional Probability Densities

In 2004, [Feindt, 2004] proposed the use of neural networks to predict probability densities for one-dimensional labels during inference instead of single-point estimates. The purpose of this is to approximate Bayesian inference by interpreting the network output as a posterior probability distribution. To achieve this, the label space of the target variable has to be pre-processed, and the network inference results have to be interpreted with care. These processes are briefly explained in the following steps. For a detailed description of the steps refer to the original publication [Feindt, 2004]:

1) Pre-processing: The one-dimensional label space  $y$  is transformed into a uniform distribution  $F: y \rightarrow s$  between zero and one, such that  $s(y_{\min}) = 0$  and  $s(y_{\max}) = 1$ . This transformation is done by sorting the label space according to  $y$  and assigning the  $i^{\text{th}}$  element in the list the value  $s(y_i) = i/N_{y\_samples}$ , where  $N_{y\_samples}$  is the number of samples in the label distribution. This mapping has to be stored, as it is used at a later stage for value reconstruction during inference.

2) Discretization: The newly constructed uniform probability distribution  $G(s)$  is now discretized into bins at  $N$  equidistant levels. A total of  $N + 1$  intervals are defined to ensure that  $G(0) = 0$  and  $G(1) = 1$ . In order to create the new training data set, each training label  $s(y_i)$  is transformed into these bins, by setting every bin value to -1, where  $s(y_i)$  is still smaller than the maximum value in the bin. Every bin value is set to 1, where  $s(y_i)$  is higher than the minimum bin value, and an interpolated value between -1 and 1, representing the position within the bin where the label value resides in (see figure 3.37).

3) Network training: The same base feed forward network architecture is used as explained in section 2.4.3, but in this case, for each discretization bin, a separate output node is added to the architecture. During training, an L2 loss function and the Adam optimizer are used, which essentially means that the nodes learn to regress values between -1 or 1 in the classification

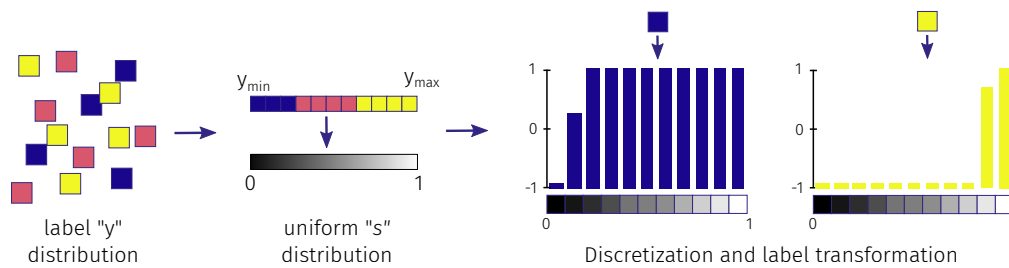


Figure 3.37: Visualization of the label space transformations to enable estimation of conditional probability densities. The label distribution is transformed into a uniform distribution. Afterward, the uniform distribution is discretized, and the labels are transformed into a new label space. Here, the bins are assigned -1, if the label value is lower than the bin value, and 1 if the label value is higher. The bin where the label value resides in is assigned a value that represents the precise position within the bin.

problem whether the actual value is above or below the threshold value represented by the bin.

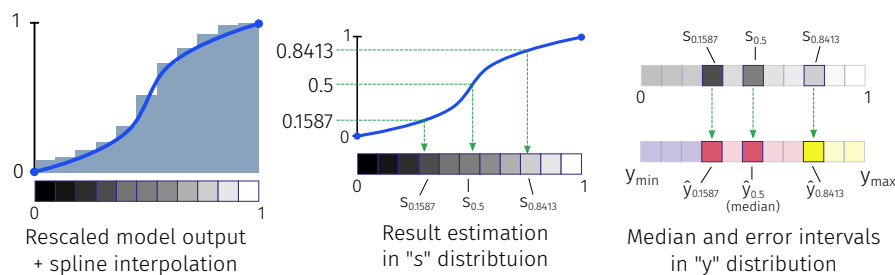


Figure 3.38: Visualization of how conditional probability density estimates are transformed back into the label distribution. The output is re-scaled to  $[0, 1]$ , and the bin values are fitted with a b-spline. This enables the extraction of percentiles in the uniform "s" distribution which can then be mapped back into the label distribution.

4) Inference: After the network converged during training, the output is re-scaled to  $[0, 1]$ . Then, the bin estimates are fitted with a b-spline that is forced through 0 and 1 at the extreme values. From this spline curve, relevant quantities of interest such as the median and the surrounding error intervals (in this case defined as the 0.1587, 0.5, and 0.8413 percentiles) can now be calculated in  $s$  space and then reconstructed into the original label space by performing a sorted search for the closest values in the original mapping list and by performing linear interpolation between these points (see figure 3.38).

### (3) External observing network

An external observing network was developed and presented in [Gröhl et al., 2018a]. The idea is to create a model that is capable of estimating the expected error completely independently of the algorithm doing the inference (see figure 3.39). The observing model observes the inference

model and, by doing that, learns to predict the expected error. The expected error can then be used to compute confidence intervals for the estimates of the inference model.

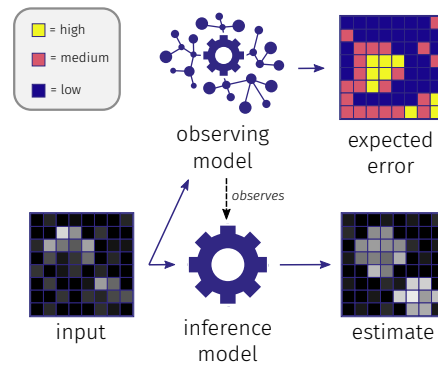


Figure 3.39: Visualization of using an observing model for uncertainty estimation. An inference model estimates the target properties given an input. The observing model is tasked to estimate the expected error of the inference model. It is being trained on observations of the inference model's performance on a validation set.

The observing network is trained on a validation data set, that the inference method is tasked to predict. The ground truth values for this data set are available and computed, and the observing model is tasked to estimate the expected error of the inference method on the validation data set. The external observing network has the advantage of being very flexible and independent of the inference method. However, it is limited to the fact that it cannot produce posterior estimation distributions.

#### (4) Conditional Invertible Neural Networks

Recently, [Ardizzone et al., 2019] have presented an extension to Invertible Neural Networks (INNs), in the form of a conditional INN. During training, the conditional INN learns to map the label distribution into a Gaussian latent distribution under the condition of the input features. This creates an easy-to-sample-from latent distribution, and during training, multiple latent samples are taken and inversely propagated through the network under the condition of the input features. With this approach, conditional INNs are capable of recovering full posterior label distributions given the input features as a condition. A more detailed description of the training process is given in section 2.4.3.

#### Experimental design

The primary hypothesis that should be investigated by the experiments is whether there is merit in excluding pixel estimates from a region of interest when they exhibit a high estimation uncertainty. With this in mind, the monitored thing of interest is the progression of the estimation error when excluding an increasing amount of estimates from the set of all estimates in a data set based on the coinciding uncertainty value.

Section 3.4.3 first briefly introduces the results from the previous publications where uncertainty estimation was investigated regarding the CE-qPAI method [Gröhl et al., 2018b] and end-to-end

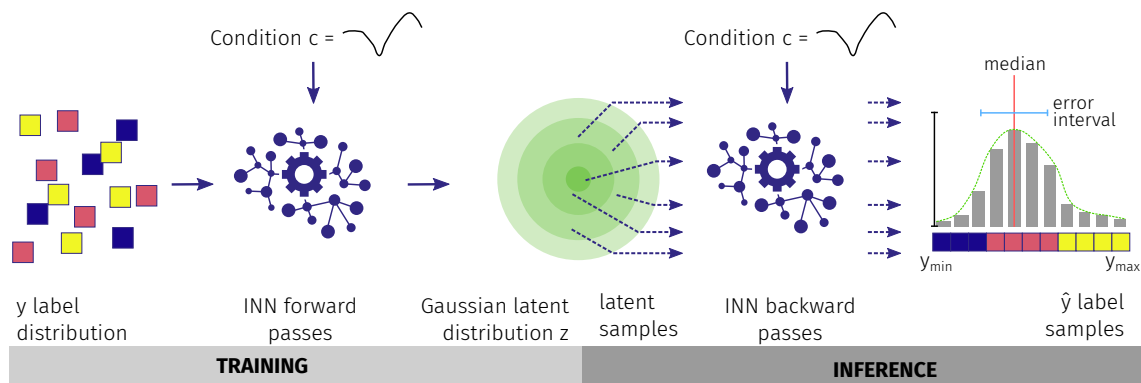


Figure 3.40: Visualization of using a conditional Invertible Neural Network (INN) for uncertainty estimation. During training, the INN learns to map the label distribution into a Gaussian distribution under the condition of the input features. During inference, one can sample from the Gaussian latent distribution, and together with the input feature conditions, the INN is capable of producing a full posterior distribution. The posterior distribution can then, for example, be used to calculate metrics for the median and the error interval.

deep learning for qPAI [Gröhl et al., 2018a]. Section 3.4.3 then reports detailed results in the context of  $\text{SO}_2$  estimation with the LSD-qPAI methodology. To this end, the four presented methods for uncertainty estimation were systematically investigated on the *in silico* flow phantom data set the LSD-qPAI method was trained on in order to be able to estimate  $\text{SO}_2$  results for the *in vitro* flow phantom data (see section 3.3). For each method, the absolute  $\text{SO}_2$  quantification error and a metric for the standard deviation were calculated on the test set. Then the median absolute  $\text{SO}_2$  estimation error was calculated with respect to a threshold in the maximum permissible standard deviations in the results, which was gradually decreased from 20% to 0%.

The quantitatively and qualitatively best performing method in terms of uncertainty estimation was applied to the *in vitro* recorded data. These results are qualitatively displayed on the "human\_2" and "Rat\_2" data sets and serve the purpose to qualitatively determine whether the envisioned application of uncertainty information has the expected effect also on real data.

### 3.4.3 Results

The results section of this chapter will present both previously published results using some of the presented methodologies in section 3.4.3, and the results of the comparison of all four presented uncertainty estimation methods in section 3.4.3.

#### Previously published results

This section will briefly introduce prior applications of some of the presented methods: (1) the estimation of conditional probability densities was applied to the Context Encoding Quantitative Photoacoustic Imaging (CE-qPAI) method [Gröhl et al., 2018b] and (2) an external observing network was applied to the end-to-end deep learning qPAI methodology [Gröhl et al., 2018a].

### (1) Confidence estimation for context encoding qPAI

In [Gröhl et al., 2018b] it was examined, whether the method of [Feindt, 2004] for uncertainty estimation could be utilized to improve the Context Encoding Quantitative Photoacoustic Imaging (CE-qPAI) results. To this end, it was investigated whether the fluence estimation error and the estimation uncertainty are correlated and if systematically excluding estimates with lower uncertainty would lead to an improvement in overall accuracy. Figure 3.41 shows the results of this study.

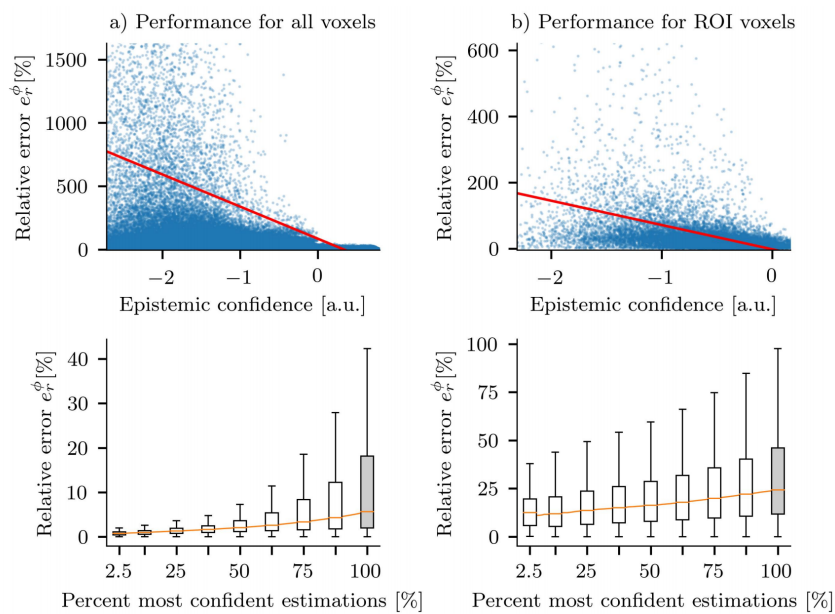


Figure 3.41: The results of applying the uncertainty method proposed by Feindt to the CE-qPAI method. The upper plots show the correlation between the relative fluence estimation error and the epistemic confidence estimate in a scatter plot. The lower plots show the behavior of the relative estimation error when only evaluating the  $n\%$  most confident estimates. The left plots show the results on all voxels, and the right-hand side plots show the results on the Region Of Interest (ROI). The figure was reprinted with permission from [Gröhl et al., 2018b].

The results in this work are calculated on the high-noise, multi-vessel data set, as presented in section 3.1. When evaluating the 2.5% most confident estimations exclusively, the median relative fluence estimation error calculated from Region Of Interest (ROI) estimates (within vessel structures with an SNR threshold above a noise equivalent level) dropped by up to 12 percentage points to 12%. Even more drastic was the increase in performance when considering all estimates, where the error dropped by up to 5 percentage points to 0.7%. It was concluded that the use of the proposed uncertainty estimation method from [Feindt, 2004] could potentially be of great benefit in the context of CE-qPAI.



## (2) Confidence estimation for end-to-end deep learning qPAI

In [Gröhl et al., 2018a], it was investigated how well an external observing neural network could estimate the uncertainty of the end-to-end 2D qPAI problem, as also presented in section 3.2. The performance of the observing network was evaluated with the same methodology as described above, where it was investigated whether systematically excluding estimates with lower uncertainty can lead to an improvement in overall accuracy. The approach was applied to different strategies of  $\mu_a$  estimation: (1) direct estimation of  $\mu_a$  and (2) estimation of the fluence  $\phi$  and correcting the input signal. These two methods were applied using either initial pressure or raw time series data as the input features. The main results are reproduced in figure 3.42. In addition to the four  $\mu_a$  estimation combinations, the observing model was also applied to the hypothetical case, that the image is corrected using a simulated homogeneous fluence map (referred to as the "naïve" approach in the figure).

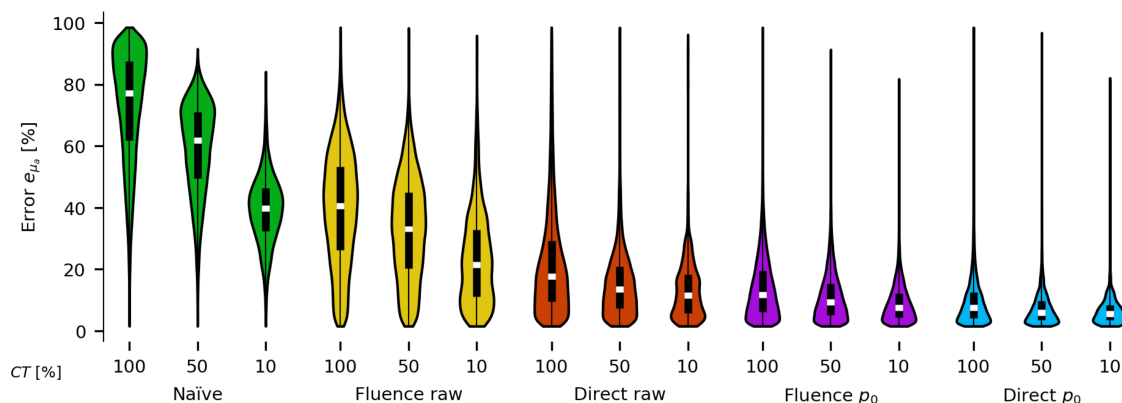


Figure 3.42: The results of using an external model to estimate the uncertainty of various 2D quantification methods. Each color shows the error distribution of one method when 100%, 50%, or 10% of the most confident estimates are being used for evaluation (referred to as a confidence threshold (CT)). Green shows the results for the naïve method of fluence compensation with a fluence map obtained from a homogeneous tissue assumption, yellow and purple show the results for fluence compensation, where the fluence is learned by a deep learning model, and red and blue show the results of direct  $\mu_a$  estimation - also learned by a deep learning model. Reprinted from the CC-BY licensed publication [Gröhl et al., 2018a].

The results were calculated for the ROI estimates, which was again defined as voxels within vessel structures that exhibited an SNR value above a noise equivalent threshold. On all methodologies, a steady decrease in the error when excluding more pixels with a higher estimated expected error could be seen. When only considering the top 50% most confident estimates, this approach led to a decrease in the absorption estimation error of up to  $\approx 30\%$  (increasing to up to a  $\approx 50\%$  improvement when evaluating only the top 10% of the most confident estimates). The violin plots visualize the changes in the distribution of the absorption estimation error when applying different confidence thresholds (referred to as CT in the figure). These visualizations reveal a meandering of the distribution toward lower error values.

### Results on the LSD-qPAI method

This chapter will first present the results of all four methods for uncertainty estimation applied to the *in silico* flow phantom data set, which was simulated with 17 different wavelengths using the mcx Monte Carlo implementation (for details on the data set and the Monte Carlo method see sections 3.3.2 and 2.3.2). For each of the methods, a regressor was trained to estimate  $sO_2$  based on a 17-dimensional single-pixel  $p_0$  spectrum.

### Performance of the methods on the *in silico* flow phantom data set

In this section, the results of the four presented methods of uncertainty estimation are presented in the same order as they were introduced: (1) dropout sampling, (2) the estimation of conditional probability densities instead of point estimates, (3) using an external model to estimate the expected error of the regressor, and (4) latent space sampling with a conditional INN.

Method	Dropout sampling	Cond. Probability densities	External model	Conditional INNs
Fractional change of median error	0.85	0.63	0.64	0.53

Table 3.9: This table shows the fraction of the original value that the median relative  $sO_2$  estimation error could be reduced to when excluding 50% of the most uncertain estimates as determined by the respective method (lower is better). A value of 0.6 would mean that the error, when excluding 50% of the most uncertain estimates, would be 60% of the error when evaluating on all estimates.

**(1) Results when using dropout sampling:**

The median absolute  $sO_2$  estimation error of the  $N=100$  dropout samples when computing the median on all samples was 3.4%, with an Interquartile Range (IQR) of (1.2%, 8.1%). When computing the results on half the data set, excluding the half with the highest estimated uncertainty, the error decreases to a median of 2.9%, with an IQR of (1.1%, 7.6%). Figure 3.43 shows which estimates were excluded by this process.

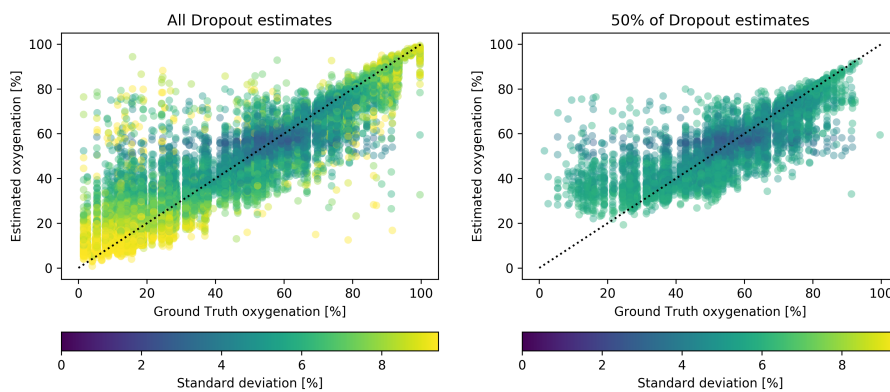


Figure 3.43: Scatter plots of  $sO_2$  estimates with annotated dropout uncertainty. The left plot shows all samples, and the right plot shows 50% of the samples with the highest estimation confidence.

Figure 3.44 shows a more detailed progression of the median error when it is computed while excluding estimates below a certain uncertainty threshold.

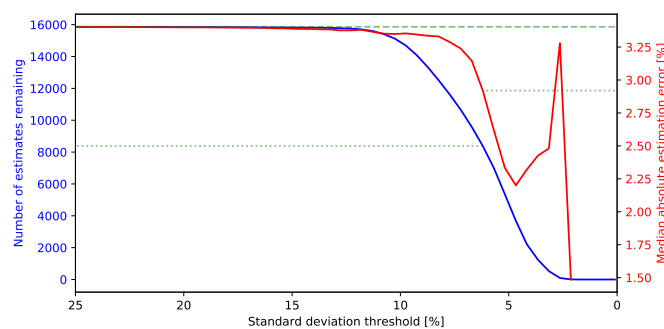


Figure 3.44: Detailed progression of the  $sO_2$  estimation error when excluding an increasing amount of estimates based on their dropout uncertainty. The x-axis shows the threshold of the uncertainty value that is permissible for the estimates. The red graph shows the progression of the median  $sO_2$  estimation error, and the blue graph shows the number of remaining estimates. The upper dashed green line shows the performance for 100% of estimates, and the lower green line shows the performance when considering 50% of estimates.

## (2) Results when using conditional probability densities:

The median absolute  $\text{sO}_2$  estimation error of the estimate reconstructions when computing the median on all samples was 5.4%, with an Interquartile Range (IQR) of (2.6%, 12.4%). When computing the results on half the data set, excluding the half with the highest estimated uncertainty, the error decreases to a median of 3.4%, with an IQR of (1.9%, 5.7%). Figure 3.45 shows which estimates were excluded by this process.

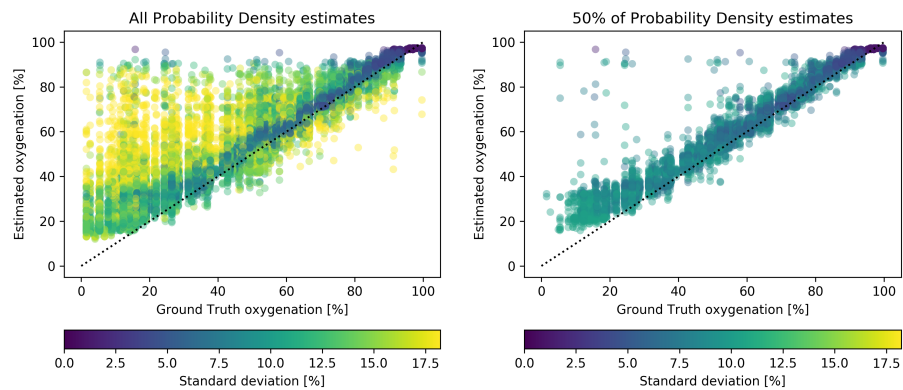


Figure 3.45: Scatter plots of  $\text{sO}_2$  estimates with annotated dropout uncertainty. The left plot shows all samples, and the right plot shows 50% of the samples with the highest estimation confidence.

Figure 3.46 shows a more detailed progression of the median error when it is computed while excluding estimates below a certain uncertainty threshold.

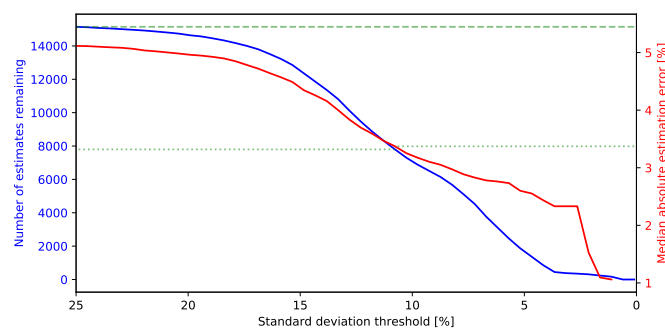


Figure 3.46: Detailed progression of the  $\text{sO}_2$  estimation error when excluding an increasing amount of estimates based on their reconstructed uncertainty. The x-axis shows the threshold of the uncertainty value that is permissible for the estimates. The red graph shows the progression of the median  $\text{sO}_2$  estimation error, and the blue graph shows the number of remaining estimates. The upper dashed green line shows the performance for 100% of estimates, and the lower green line shows the performance when considering 50% of estimates.

**(3) Results when using an external observing model:**

The median absolute  $\text{sO}_2$  estimation error of the inference model when computing the median on all samples was 4.4%, with an Interquartile Range (IQR) of (1.9%, 9.1%). When computing the results on half the data set, excluding the half with the highest estimated uncertainty according to the external observing network, the error decreases to a median of 2.8%, with an IQR of (1.3%, 5.1%). Figure 3.47 shows which estimates were excluded by this process.

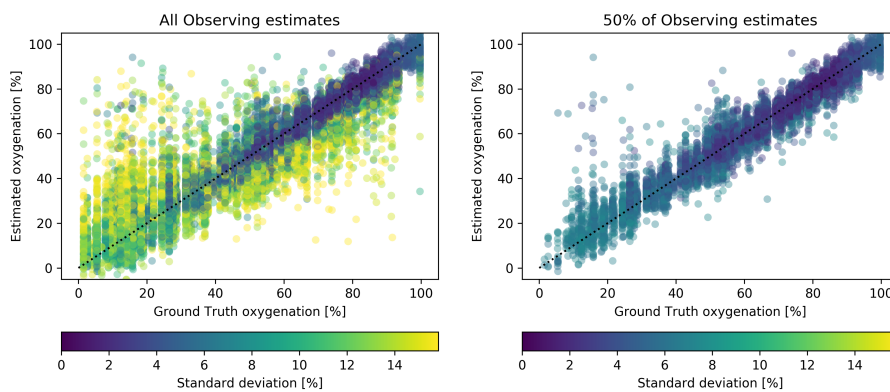


Figure 3.47: Scatter plots of  $\text{sO}_2$  estimates with annotated dropout uncertainty. The left plot shows all samples, and the right plot shows 50% of the samples with the highest estimation confidence.

Figure 3.48 shows a more detailed progression of the median error when it is computed while excluding estimates below a certain uncertainty threshold.

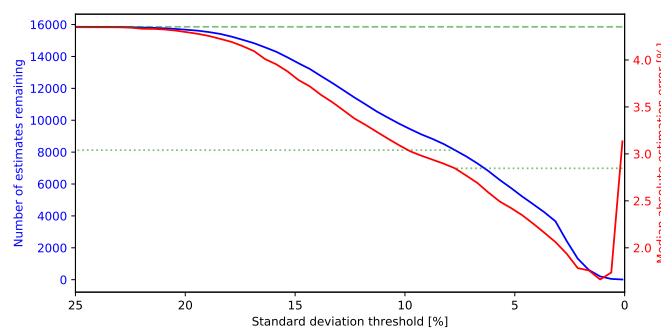


Figure 3.48: Detailed progression of the  $\text{sO}_2$  estimation error when excluding an increasing amount of estimates based on their dropout uncertainty. The x-axis shows the threshold of the uncertainty value that is permissible for the estimates. The red graph shows the progression of the median  $\text{sO}_2$  estimation error, and the blue graph shows the number of remaining estimates. The upper dashed green line shows the performance for 100% of estimates, and the lower green line shows the performance when considering 50% of estimates.

#### (4) Results when using Conditional Invertible Neural Networks:

The median absolute  $\text{sO}_2$  estimation error of the  $N=1000$  latent samples when computing the median on all samples was 5.7%, with an Interquartile Range (IQR) of (2.3%, 12.8%). When computing the results on half the data set, excluding the half with the highest estimated uncertainty, the error decreases to a median of 3.0%, with an IQR of (1.3%, 5.6%). Figure 3.49 shows which estimates were excluded by this process.

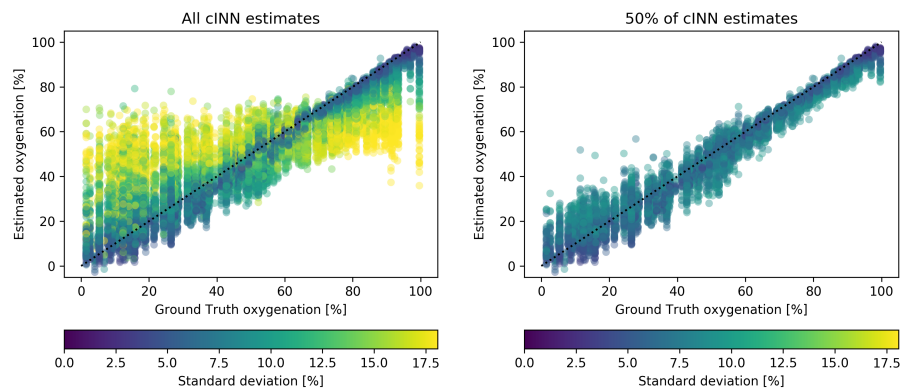


Figure 3.49: Scatter plots of  $\text{sO}_2$  estimates with annotated dropout uncertainty. The left plot shows all samples, and the right plot shows 50% of the samples with the highest estimation confidence.

Figure 3.50 shows a more detailed progression of the median error when it is computed while excluding estimates below a certain uncertainty threshold.

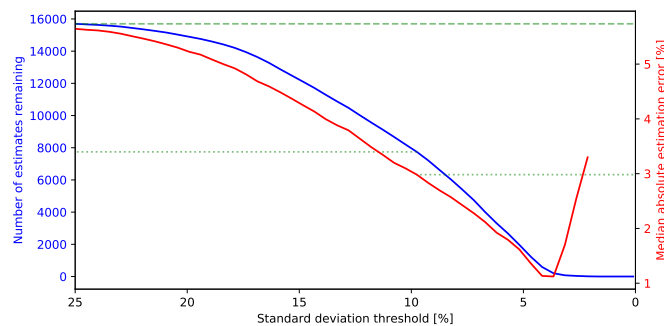


Figure 3.50: Detailed progression of the  $\text{sO}_2$  estimation error when excluding an increasing amount of estimates based on their dropout uncertainty. The x-axis shows the threshold of the uncertainty value that is permissible for the estimates. The red graph shows the progression of the median  $\text{sO}_2$  estimation error, and the blue graph shows the number of remaining estimates. The upper dashed green line shows the performance for 100% of estimates, and the lower green line shows the performance when considering 50% of estimates.

### Performance on the *in vitro* flow phantom data set

After evaluation of the methods on the *in silico* data set, the conditional INNs were applied to the recorded phantom data on ten equidistant time points between the start and the end of the experiment. The "human\_2" data set (figure 3.51) and the "Rat\_2" data set (figure 3.52) were evaluated for this purpose. The images that are shown here are representative of the entire data set and correspond to a rather early point in time, where the oxygenation in the vessel tube was still high (after approximately ten minutes), and a later point in time (after approximately 30 minutes), where the oxygenation had already decreased by a significant margin.

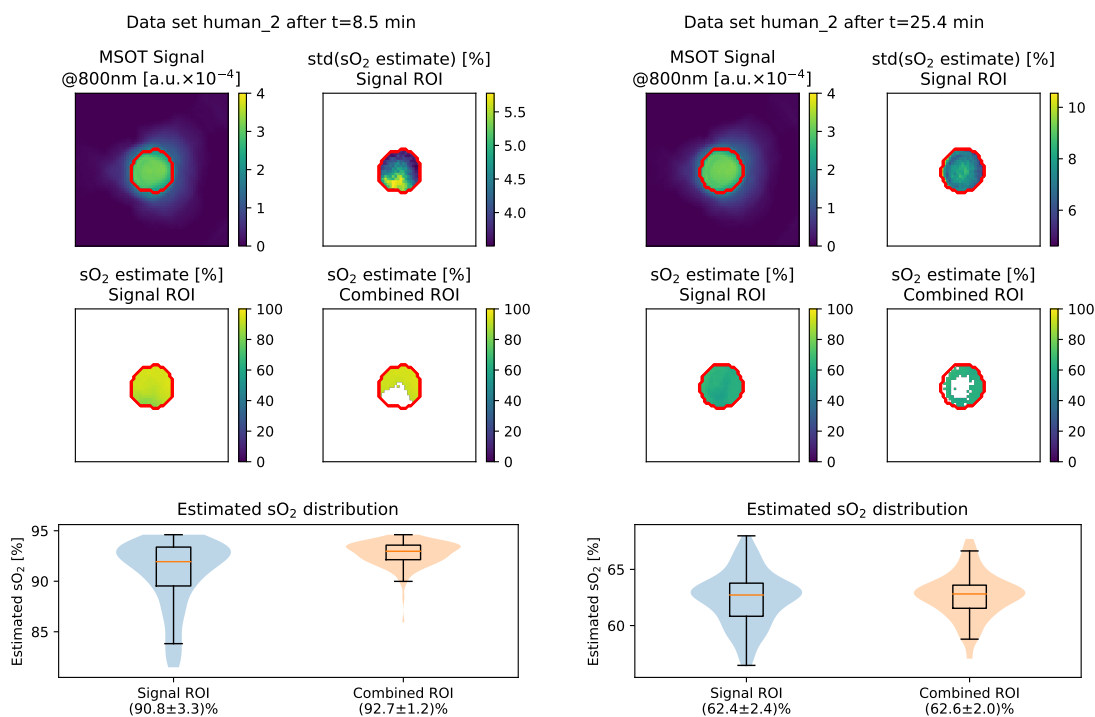


Figure 3.51: The result images of conditional INN implementation of LSD-qPAI and the uncertainty method on the flow phantom "human\_2" data set at two points in time (5.5 min (left) and 25.4 min (right)). The red circle corresponds to the ROI that was identified by an SNR metric. The top-left plot shows the MSOT signal; the lower-left plot shows the estimated oxygenation in the ROI; the top-right plot shows the error interval of the  $sO_2$  estimations; and the lower-right plot shows the  $sO_2$  estimates in the refined uncertainty-aware ROI. The bottom violin plots show the distribution of  $sO_2$  estimates in the entire ROI (left) and the uncertainty refined ROI (right).

The proposed method for incorporating uncertainty estimates into the estimations was implemented in the way that first, a purely SNR-based ROI was created, and afterward, the 50% estimates with the highest corresponding error interval were removed from it to obtain the refined ROI.

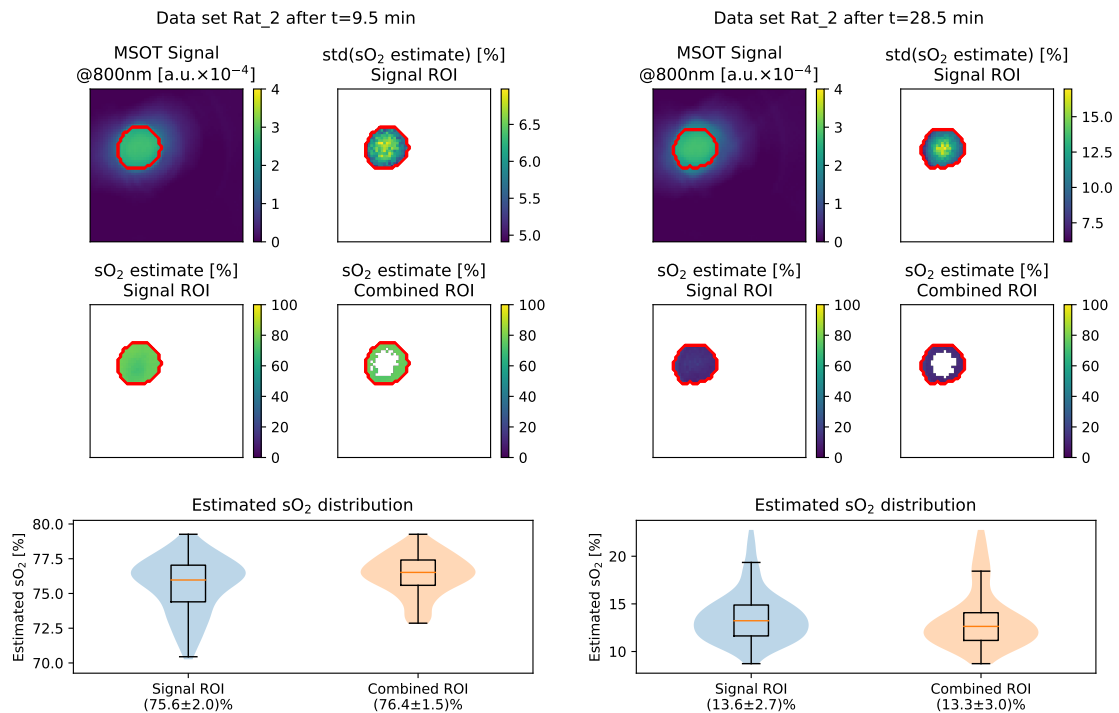


Figure 3.52: The result images of conditional INN implementation of LSD-qPAI and the uncertainty method on the flow phantom "Rat\_2" data set at two points in time (5.5 min (left) and 25.4 min (right)). The red circle corresponds to the ROI that was identified by an SNR metric. The top-left plot shows the MSOT signal; the lower-left plot shows the estimated oxygenation in the ROI; the top-right plot shows the error interval of the  $sO_2$  estimations; and the lower-right plot shows the  $sO_2$  estimates in the refined uncertainty-aware ROI. The bottom violin plots show the distribution of  $sO_2$  estimates in the entire ROI (left) and the uncertainty refined ROI (right).

These qualitative results show that in most images, the estimation uncertainty increases with depth inside the tube. Only in the "human\_2" data set, at some points in time, there were other variations as well. In most instances, the refinement of the ROI led to a slight shift in the mean estimated  $sO_2$  value and had a tendency to decrease the standard deviation of these slightly.

### Model calibrations

Modern neural network architectures have been demonstrated to produce wrongly calibrated uncertainty estimates if not properly trained [Guo et al., 2017]. This is, because of the large number of parameters in the networks and because of the sensitivity of the calibration to the choice of hyperparameters, such as the batch size or the loss functions. As evident from the blue graphs in the figure, none of the proposed methods produced calibrated uncertainty estimates directly after training. The conditional INN seemed to overestimate the uncertainty, whereas the other three methods (dropout sampling, the estimation of probability densities, and the use of



an external model) seemed to underestimate the uncertainty of their estimates.

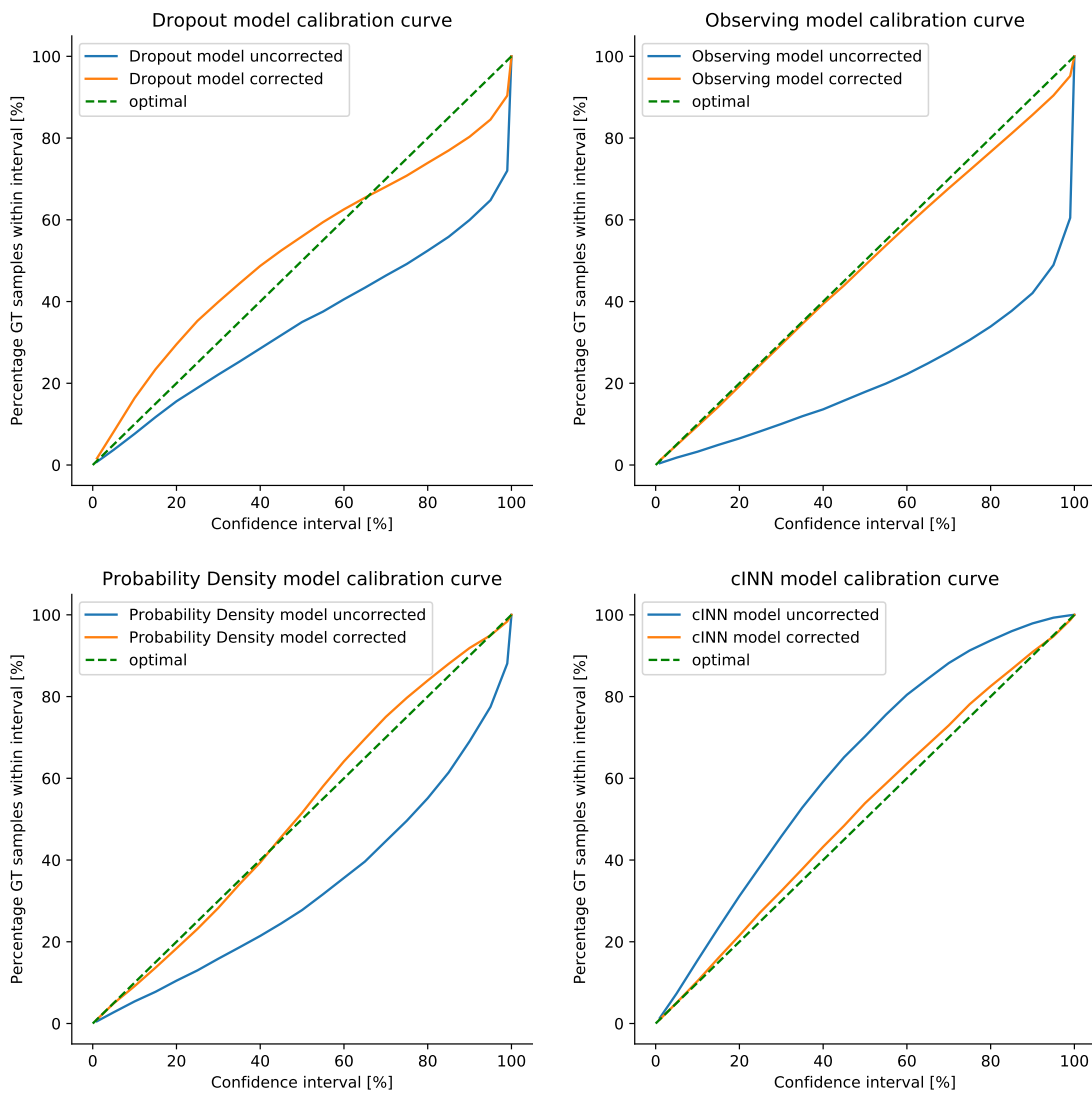


Figure 3.53: Uncertainty calibration curves for (left to right, top to bottom) dropout sampling, estimating the standard error with an external model, estimating probability densities, or sampling from a conditional INN model. Shown are the vanilla calibration curves (blue) and the corrected calibration curves (orange), where the correction factor is estimated from the validation set. The green dashed line shows the curve for optimal calibration.

One tool to investigate the calibration of uncertainty estimation methods is the calculation of calibration curves. In these, it is systematically examined if the uncertainties reflect the actual accuracy of the estimated values.

In order to re-calibrate the model uncertainty, a correction factor was introduced that was optimized on the validation set and then applied on the test set. This simple calibration method seemed to have a great effect in this case and produced results very similar to the optimal calibration on the test set, with dropout sampling being the only outlier. Figure 3.53 shows the calibration curves of the examined methods on the *in silico* data set before and after the application of a simple *post-hoc* correction method.

### 3.4.4 Discussion

#### Discussion of results

All presented results revealed that there is a benefit in excluding estimates with a relatively high corresponding uncertainty estimate. However, the *in silico* results revealed that there are distinct differences between the four approaches:

(1) Dropout sampling: In the case of dropout sampling, there is a clear tendency, that lower and higher ground truth  $\text{sO}_2$  values were estimated with higher uncertainty. The error progression analysis showed that this behavior is correct to a certain extent. However, after a certain threshold, the estimation error started to increase again.

(2) Conditional probability densities: The model trained on predicting probability densities showed a tendency to overestimate the ground truth  $\text{sO}_2$  value. Here, there was a slight tendency that lower  $\text{sO}_2$  values were assigned a higher estimation uncertainty. The chosen implementation of the method of Feindt led to artifacts at the end of the distributions: there were no estimates smaller than  $\approx 15\%$ , and very high  $\text{sO}_2$  values were generally predicted with little variance. This might have been caused by the small number of bins (25) that were used for the discretization of the uniform "s" space when creating the labels.

(3) External observing model: This method also resulted in a monotonous improvement of the median estimation error when excluding more and more uncertain estimates from the evaluation. The uncertainty estimates also showed a tendency where lower ground truth  $\text{sO}_2$  values were attached with higher uncertainty than higher  $\text{sO}_2$  values. Nevertheless, in contrast to dropout sampling, this did not seem to be the only criterion, as also estimates that strayed further from the bisector seem to correspond to higher uncertainty values.

(4) Conditional invertible neural networks: The conditional INNs showed the best correspondences between distance from the bisector and the corresponding uncertainty estimate. The median estimation error decreased monotonously with a decreasing number of estimates, until the point where only a very small number of estimates was left. For the conditional INN, there was no obvious correlation between the ground truth  $\text{sO}_2$  value and uncertainty. However, the plot with only 50% of samples remaining shows a tendency, where estimates below 60%  $\text{sO}_2$  show higher variation in the estimates than above 60%.

The scatter plots of three of the four methods suggested a correlation between the ground truth  $\text{sO}_2$  value and the corresponding uncertainty estimate, where lower  $\text{sO}_2$  values were generally estimated with higher uncertainty. Also, the remaining method revealed a difference in the variation of the estimates depending on the ground truth  $\text{sO}_2$  value. It remains to be investigated in future work whether this tendency can be reproduced on other data sets or confirmed analytically.

When analyzing the *in silico* results, one has to take into consideration the number of remaining estimates after applying the uncertainty threshold. This is because the median error estimate cannot be representative if only very few estimates remain. This can be seen in the very sharp rise or drop in the error towards the end of many of the graphs.

Furthermore, comparing the different uncertainty methods showed that there were noticeable

differences in the estimation accuracy (the median absolute estimation error when evaluating all samples) of the different models. One reason for this may be that different neural network architectures were used due to the differences in the number of input and output features, which lead to different total amounts of adjustable parameters in the networks. Furthermore, there were differences in the chosen loss functions (for example, a mean squared error loss for the simple feed forward architecture versus a maximum likelihood loss for the conditional INN). There were even differences when evaluating the same network in the dropout sampling and external model case. Here, exactly the same trained network had been used; however, for dropout sampling, the median over one hundred estimates was computed, which seemed to marginally improve the accuracy when compared to doing a single forward pass in evaluation mode.

### **Discussion of methodology**

In this work, uncertainty estimation was used in order to calculate a confidence interval, in which the actual value is likely to reside in. In this case, the standard deviation was used, which means that in  $\approx 68\%$  of cases, the true value is inside the reported  $\mu \pm \sigma$  interval. However, this approach offers only a narrow view of uncertainty estimation and does not exploit the full impact that the computation of a full posterior distribution could have had. This is true, especially with regards to the recovery of ambiguities in multi-modal distributions, which would get lost when not being able to produce the full correct posterior distribution [Ardizzone et al., 2018]. The post-processing that is done in the proposed method to evaluate the width of the posterior distributions (for example, represented by the standard deviation or the IQR) or the estimation value (for example, represented by the median or the mean) are not suitable in a multi-modal scenario. The mean would estimate the average of all modes, and the median might fall into one of the modes, depending on the number of samples for the modes.

In this context, the proposed observing neural network by design cannot produce the full posterior distribution and is only capable of estimating the standard error of the estimate. Dropout sampling has been shown not to be able to produce accurate posterior distributions as well [Ardizzone et al., 2018, Hron et al., 2017]. INNs, on the other hand, have been demonstrated to be capable of recovering these full posterior distributions, if adequately trained.

Out of distribution detection as presented by [Adler et al., 2019b] for optical imaging might be crucial. This is because data-driven approaches might not know what they do not know and, as such, might confidently estimate systematically wrong values when tasked to do inference on out of distribution input samples.

### 3.4.5 Conclusion

Two research questions were derived from the overarching hypothesis  $H_3$  that was addressed in this chapter:

1) *Can a systematic exclusion of relatively uncertain estimates from a region of interest lead to a change of the aggregated estimate?*

In order to investigate this question, the progression of the relative median estimation error when excluding more and more uncertain estimates from the aggregated result was examined. This was done for all three quantification methods presented in this thesis: on one of the CE-qPAI data sets, on the direct inversion of entire PA images, and on the LSD-qPAI *in silico* flow phantom data set. The results achieved in all of these scenarios suggest that there can be great merit in excluding uncertain estimates from the result calculation. In all cases, the median estimation error could be reduced (by up to a factor of two in the best case) when excluding 50% of the estimates from the evaluation. One of the uncertainty estimation methods was also applied *in vitro* and generally showed an increase in uncertainty for increasing depth in the vessel structures. Excluding 50% of the most uncertain estimates in the ROI usually led to a slight change of the median estimated  $sO_2$  and to a decrease of the standard deviation of the estimates.

2) *Are there differences between different state-of-the-art techniques for uncertainty estimation?*

Four different methods were compared in their performance using the LSD-qPAI method on the *in silico* flow phantom data set. This analysis showed that there are differences in the quantitative performance, such as the fractional change of median error when evaluating on 50% of estimates, where the conditional INN performed the best with a fractional change of the error of 0.53. A qualitative performance assessment showed that different types of estimates were identified as being uncertain by the different algorithms. Here, the conditional INN performed the most as expected (estimates furthest from the bisector were estimated as the most uncertain), whereas, with dropout sampling, high and low  $sO_2$  estimates were systematically assigned a high uncertainty. Furthermore, dropout sampling has been shown not to be able to produce correct posterior distributions, and the observing neural network only estimates the standard error by design. Methods that are capable of estimating the full posterior distribution, such as the conditional INNs, should be preferred for application in a clinical context, as they enable the identification of multi-modal distributions.

The results suggest that the reduction of the estimation error within an ROI is feasibly by discarding those estimates that are labeled with high uncertainty as estimated with a suitable method. Here, *in silico* experiments show a consistent improvement in the estimation accuracy of a parameter in various different settings and experiments. Initial results on real data show that these methods yield practical and interpretable results. As the development and use of reliable methods for uncertainty quantification are mandatory for the successful translation of qPAI methods to clinical applications, future work should investigate the potential of such methods in clinically relevant contexts.

Such future work might include the integration of one or several of the compared methods into a toolkit that is easily usable by physicians and allows for usability and impact studies that determine the value of uncertainty estimation methods in combination with the quantification

algorithms. Such a toolkit should also contain a method for detecting out-of-distribution samples, and it should be analyzed how well the uncertainty estimation methods would be capable of handling fictitious scenarios not sufficiently contained in the training data set as well as the robustness to noise and imaging artifacts.

All these technological advancements can potentially enable not only the **oxygenation quantification** but also other functional and molecular applications. Although **challenges remain**, none of them are beyond reach.

[Li et al., 2018]

# 4 | Discussion

A key challenge for the successful translation of photoacoustic imaging into the clinical routine is the lack of robust and accurate methods for quantifying the measured signals. State-of-the-art techniques in the field of qPAI tackle the associated inverse problems by using iterative model-based inversion algorithms. However, due to practical limitations, these have not successfully been applied in realistic scenarios. The experiments and methods presented in this thesis face this challenge and pioneer the application of data-driven methods to the inverse problems of qPAI. Three overarching hypotheses were investigated that postulated the *feasibility* of the application of data-driven methods towards (1) the optical and acoustic inverse problem of qPAI, (2) the accurate recovery of functional tissue parameters from multispectral PA images, and (3) the estimation of uncertainty for the inverse qPAI models.

## 4.1 Summary of Contributions

Several approaches towards quantitative Photoacoustic Imaging (qPAI) using clinical handheld PA devices were presented in this thesis. The specific contributions to the field of qPAI presented in this thesis can be summarized using the aforementioned overarching hypotheses:

*H1: Data-driven methods can be used to solve the optical and acoustic inverse problem.*

The contributions to this hypothesis were two-fold. It was investigated if the light fluence can be estimated from distributions of initial pressure in the case of 3D PAI. For this, a method was developed that allows for the calculation of voxel-wise feature vectors that can be used for voxel-wise fluence estimation in a machine learning framework. Extensive *in silico* validation demonstrated the robust and versatile usability of this approach, and the results were published in the *Journal of Biomedical Optics* [Kirchner et al., 2018a].

In further experiments, it was investigated if a direct estimation of optical absorption  $\mu_a$  from 2D

PA initial pressure distributions  $p_0$  and raw time series data  $p(t)$  is potentially feasible as well. To this end, state-of-the-art deep learning models were applied to these problems on specially designed *in silico* data sets. The results of these analyses on multiple different data sets suggest that an estimation of  $\mu_a$  directly from 2D images is indeed principally feasible. However, the integration of the acoustic inverse problem into this framework has proven difficult and requires careful investigation in future work. The methods used towards this end were published in the *Journal of Imaging* [Gröhl et al., 2018a]. Initial work towards the acoustic inverse problem was published in the proceedings *Photons Plus Ultrasound: Imaging and Sensing* of the SPIE Photonics West conference in 2018 [Waibel et al., 2018].

*H2: Data-driven methods can be used for spectral unmixing in a realistic context.*

A method has been developed that allows pixel-wise spectral unmixing of multispectral PA images. The method is completely trained on *in silico* data. Because linear unmixing techniques break under the non-linear influence of the light fluence, the primary purpose of the method is to have the ability to map spectrally colored multispectral signals to blood oxygenation  $sO_2$ . The approach has shown to be able to account for an extensive range of degrees of spectral coloring within multiple *in silico* experiments. Additionally, it has successfully been applied to real *in vitro* data acquired from a blood flow phantom and demonstrated a better dynamic range than linear unmixing techniques.

Furthermore, the use of the method for *in vivo* imaging applications of  $sO_2$  estimation using a porcine brain recording and human forearm images showed that also in these situations, it exhibited a better dynamic range and was capable of yielding physiological results. Here, in thirty images from fifteen healthy volunteers, the measured arterial blood oxygenation was very close to the expected value of 98-100%. Early results of this work have already been published on *arXiv* [Gröhl et al., 2019].

*H3: Confidence estimation methods can be used to gain insight into qPAI methods.*

The capability of several state-of-the-art methods to predict the uncertainty of parameter estimates was investigated, and a new method was proposed that can potentially estimate the expected error of any method. In rigorous *in silico* experiments, it was tested how well these uncertainty estimates correlate to the actual estimation error. The results of these experiments showed that there is great potential value in systematically excluding parameter estimates with corresponding high uncertainty. For nearly all uncertainty estimation approaches, the accuracy continuously improved with the exclusion of more and more estimates corresponding to high uncertainty. Applications of the uncertainty estimation techniques shown in this thesis were published in the proceedings *Photons Plus Ultrasound: Imaging and Sensing* of the SPIE Photonics West conference in 2018 [Gröhl et al., 2018b] and in the *Journal of Imaging* [Gröhl et al., 2018a].

## 4.2 Discussion of results

While a detailed discussion of the results of the experiments is included in each results section, this section more generally discusses the resilience of the obtained results of the experimental validation.



### Acoustic inverse problem

Many research groups have been working with the reconstruction of PA images from raw time series data in the past decades, and the acoustic inverse problem is generally considered to be well-understood. Because of this, one of the more astounding findings in the results of the conducted experiments was that the achieved parameter estimation accuracy drastically deteriorated when including the acoustic inverse problem. However, one needs to consider that PA image reconstruction algorithms are generally not optimized to *quantitatively* reconstruct the initial pressure distribution  $p_0$  from the raw time series data  $p(t)$ . Their primary focus is instead to optimize related image quality metrics such as the peak SNR. Furthermore, the accuracy of the inversion can be expected to be much lower, as  $p(t)$  does not contain all information present in  $p_0$ . This is mainly because of two main limitations: limited view and limited bandwidth. The limited view of clinical probe designs is due to the nature of the geometry of the detection elements, which are usually arranged in a linear array or an array that only has a slight curvature. With such arrays, a large amount of the wavefronts is not detectable, and this information is lost. The limited bandwidth is another limiting factor, which arises because the frequency responses of the detection elements might not ideally overlap with the frequencies present in the PA waves. For example, an ultrasound transducer with a central frequency of 4 MHz and a bandwidth of 80% would neither be capable of detecting the low-frequency main bulk of the signal in the Kilohertz regime nor the high-frequency tail of the signal (cf. fig. 2.7).

Section 3.2 introduced an asymmetric derivation of the U-Net to tackle the acoustic inverse problem. Based on the results, it has to be concluded that this type of architecture might not be ideal for solving the acoustic inverse problem. This might be the case, due to the loss of information caused by the introduction of convolutional layers in the skip connections. It can be expected that changes in the network architecture and a compilation of better and more training data might significantly improve the accuracy of end-to-end data-driven approaches towards the acoustic inverse problem. However, many of the problems could be averted using careful design of the detection geometry and using very broadband detection elements.

### Optical inverse problem

The results of the work presented in sections 3.1 and 3.2 demonstrate that data-driven approaches show great promise towards the optical inverse problem. With the usage of pixel-wise calculated feature vectors, as well as with direct usage of entire 2D images, it was possible to obtain accurate estimations of the underlying fluence or the optical absorption parameters. Even under the influence of multiple different noise levels, the results remained stable *in silico*. When facing high noise scenarios, direct  $\mu_a$  estimation methods are expected to outperform fluence correction approaches. This is because the fluence correction approaches are more directly influenced by the SNR of the PA image.

While the methods were extensively validated *in silico*, the biggest question that remains is how these methods could be translated into realistic scenarios and how they would perform in such a context.

### Transfer to reality

A straightforward application of the proposed methods for the optical or acoustic inverse problem to realistic scenarios is unfortunately not possible. The main reason for this is the

large domain gap between the simulation domain and the real imaging domain. Simulated PA images, while accurately representing the physical phenomena subject to the constraints of the underlying physical models, show distinct differences to real recorded images, for example, in the noise and artifact patterns or the morphological variety of the chromophores. As such, there are two approaches at hand that could be pursued towards a transfer of these methods to reality.

The first could be to try and bridge the simulation gap with more rigorous and accurate measurements of the underlying initial pressure distribution. The more accurate these become, the more likely it is to be able to model them accurately. When such measurements are performed on well-known phantoms, some of these phantom measurements could also potentially be used to extend the training data set.

The second approach could be to extend the simulation framework and bridge the simulation gap. For example, one could add hand-crafted artifact patterns or sophisticated noise models. However, maybe the use of neural networks to learn the difference and to augment the simulated data would also potentially be beneficial. Bridging of similar domain gaps in different fields has been proposed, for example, using generative adversarial networks (cf. e.g. [Hong et al., 2018, Bousmalis et al., 2018, Frangi et al., 2018, Liu et al., 2018]).

### **Functional tissue parameter estimation**

Despite the encountered difficulties of the application of data-driven approaches to real data due to the domain gap, a method was proposed in this thesis that was able to obtain plausible parameter estimates *in vitro* and *in vivo*. This method was tasked with the estimation of a functional tissue parameter that can be derived from the relative chromophore concentrations of hemoglobin Hb and oxyhemoglobin HbO<sub>2</sub>, namely, the oxygen saturation sO<sub>2</sub>. The method analyses the normalized multispectral behavior of the PA signal at a single pixel in order to derive oxygenation estimates. Because of the normalization and the isolated analysis of single pixels, the domain gap can be reduced significantly. This is because the main property of interest here is the relative change of the signal with respect to the wavelength.

Considering that the LSD-qPAI approach is one of the first machine learning-based approach towards accurate sO<sub>2</sub> estimation, the *in vitro* and *in vivo* results were very convincing. Not only did the algorithm exhibit a higher dynamic range than linear unmixing, but it also produced plausible and physiological results in a multitude of different real imaging settings. Together with the general advantages of machine learning-based methods, being the fast inference speed and high adaptability, there might be a lot of potential applications of the method.

It can be expected that adding constraints to the inversion will result in an increase in overall accuracy. Such constraints could be included in the form of value constraints (for example, that the sO<sub>2</sub> estimate cannot be below 0% or above 100%), or in the form of *spatial regularization* where also neighboring pixels are used during inference. Such spatial regularization schemes might, for example, assume that sudden changes in sO<sub>2</sub> for neighboring pixels are unlikely. However, adding such constraints might come at the cost of a decreased ease of use or longer computation times, and the advantages and disadvantages will have to be considered carefully. Further improvements might be achieved by more closely approximating the tissue to be imaged, by improving on the simulation pipeline, or by adding more information to the inversion, such as

measurements obtained with other optical sensing techniques or the co-registered ultrasound signal.

### Uncertainty estimation

In this work, four different methods for the recovery of uncertainty estimates have been examined. These methods have been applied in multiple different scenarios and have all demonstrated the ability to yield meaningful uncertainty estimates. In this context, it was investigated if it is possible to gain a boost in estimation accuracy by excluding less certain estimates from the evaluation and subsequent calculation of derived parameters. The results showed that systematically excluding estimates with corresponding high uncertainty can improve the estimation accuracy by a large margin. In the conducted experiments, the exclusion of 50% of the most uncertain estimates was able to decrease the median estimation error by up to 50% in some cases.

However, analysis of the results showed that the estimates need to be interpreted with care, as none of the techniques was capable of initially yielding well-calibrated uncertainties. The calibration would need to be improved using *post hoc* calibration techniques or a rigorous modification of the hyper-parameters. It also has to be noted that dropout sampling led to unstable results, where when applied to functional parameter estimation, high and low estimates were systematically correlated to high uncertainty, which was not sufficiently correlated with the estimation error.

In future work, it should be investigated how the additional information of estimation uncertainty can be integrated into clinical use cases to be of benefit for the physician.

## 4.3 Solving inverse problems with data-driven methods

This thesis has focused on the application of data-driven techniques to the inverse problems of quantitative PA imaging. Machine learning-based methods have shown great success in various applications, and research is continuously conducted in the field, leading to more and more advanced methods and techniques. This section will explore the general advantages and opportunities as well as threats of using machine learning techniques and will then discuss their application to the modeling of physical processes and the physical phenomena of PAI.

### Advantages and opportunities for data-driven methods

One of the primary advantages of machine learning-based methods is their ability to create a complex non-linear model to represent the training data. This has the main advantage that they become applicable to problems, where finding an analytic solution is very complex, for example, when solving partial differential equations, or to problems where no connection can be found at first glance (cf. e.g. [Sun et al., 2014, Han et al., 2018]). This is especially useful when dealing with highly dimensional and complex data, for example in the field of radiomics (c.f. e.g. [Larue et al., 2017, Vial et al., 2018]) or genomics (c.f. e.g. [Park and Kellis, 2015, Zou et al., 2019]), where data-driven approaches have been applied with great success. Machine-learned models have shown to outperform many classical data and image processing techniques, for example, in the field of medical image computing [Esteva et al., 2019].

Another advantage of machine learning models is their fast computation during model inference. Once trained, depending on the underlying algorithm and model complexity, a model is usually able to yield estimations on the order of magnitude of milliseconds. This is critical, for example for systems that need to react in time to their environment such as in autonomous driving [Trembl et al., 2016, Sallab et al., 2017] or in robotics [Pierson and Gashler, 2017, Levine et al., 2018], in which inference speed is of key importance. One of the most exciting directions in this regard is the development of *all-optical neural networks* that potentially allow for inference at the speed of light [Lin et al., 2018].

A great opportunity that data-driven approaches offer is the possibility for continuous learning. This is commonly tackled with strategies to keep improving with the availability of new data or by also solving related tasks. Fields of research where machine learning methods learn to improve themselves are commonly classified as *meta learning*, which is the umbrella term for techniques that enable *learning to learn*. Specialized fields of research derived from this broader area are, for example, *continuous learning* [Liu, 2017], *multi-task learning* [Caruana, 1997], or meta networks [Munkhdalai and Yu, 2017].

### **Disadvantages and risks of data-driven methods**

While the fast inference time of machine learning algorithms is a great benefit of machine learning algorithms, it comes at the cost of immense computational resources that are needed for model training. Especially in the area of deep learning, without the use of specialized hardware, training times can be as high as hours to days, depending on the data set size and model complexity.

Another disadvantage of machine learning techniques is their dependence on the presence of well structured and curated data sets and the subsequently high error susceptibility of these techniques. The machine learning algorithm can generally only be as good as the quality of the training data set allows it to be, and any biases in the data set can lead to terrible inference results on previously unseen input features. In the domain of medical imaging, the availability of such data sets is very sparse, and as such, methods have to be developed that can learn from synthetic data [Mahmood et al., 2018] or have the ability to learn from only a few examples [Ross et al., 2018].

Finally, a commonly seen criticism for machine learning methods is the lack of interpretability of the results. Machine learning algorithms are often referred to as a "black box", in which an input is transformed to an output. It is inherently difficult to reconstruct the exact steps to the decision. Because of this, entire fields of machine learning research are dedicated to the goal of making machine learning methods more interpretable. These are, for example, concerned with *explainability* [Holzinger et al., 2017] and latent space interpretability [Donahue et al., 2017] such that the results can be understood by human experts.

### **Learning physical processes**

There are many examples, where machine learning methods are applied to the forward simulation of physical processes, such as the motion of smoke [Chu and Thuerey, 2017], growth of plants [Hädrich et al., 2017], fluid dynamics [Kim et al., 2019], multiple light scattering in clouds [Kallweit et al., 2017], and subsurface scattering of light in 3D rendering applications.

In these papers, the assessment of feasibility when using machine learning algorithms often encompasses a critical discussion on the speed versus accuracy trade-off. While the unparalleled inference speed of machine learning methods is a great benefit, they usually are not as accurate as their model-based counterparts. These tend to be rather slow as they usually have to approximate the solution to a partial differential equation numerically. Depending on the application, these inaccuracies might be tolerable, for example, in 3D modeling, where the algorithm only needs to fool a human observer. However, this might be different in physical simulations where the results are used as an intermediate step for other computations, and estimation errors might unfavorably propagate.

In the case of PAI, the propagation of light and sound through turbid media are the physical phenomena that need to be modeled. Here, researchers are interested in fast forward models, for example, to speed up the iterative qPAI algorithms or for the fast generation of data to train data-driven models. In addition, they are also interested in creating accurate models to solve the inverse problems of reconstructing the physical process when only being able to observe the result. In this regard, the accuracy of the models is of great importance, as the quantitative estimation of the optical tissue properties is an important step to derive clinically relevant functional tissue parameters.

## 4.4 Data-driven quantitative photoacoustic imaging

This section will discuss the principle limitations and opportunities and general considerations that should be regarded when applying data-driven approaches to qPAI.

### The data famine

The primary bottleneck for the application of data-driven approaches to the field of qPAI is the lack of training data. This is because obtaining ground truth values for the optical tissue properties is inherently difficult. *In vivo*, the main limiting factor is the penetration depth of optical imaging techniques, and when analyzing *ex vivo* samples, the optical properties of the sample are bound to change, for example, due to the change of blood perfusion. There are two potential strategies to circumvent this inherent problem.

The first strategy was also pursued in this thesis and is the simulation of training data. When attempting to do this, accurate and fast forward models are needed, and these need to be integrated into a framework that is capable of producing simulation results that are as close to reality as possible. This strategy has the advantage of high flexibility and the possibility to easily create massive data sets, however, bears the risk of learning from a biased data set, in which assumptions might be ill-chosen or which has a vast domain gap to real images, making it extremely hard to be able to infer plausible values on real data.

The second strategy could be to create well-understood tissue phantoms and image these. This enables a generation of a data set with an excellent reference for the underlying optical and acoustic medium properties. A model adequately trained on such a data set might be capable of inferring the optical properties. However, this approach has the disadvantage of being limited to the structures and parameter ranges that can be created in a solid phantom *in vitro*.

### **Inversion strategies**

To approach a problem within the framework of machine learning techniques, one needs to carefully design the input and output domains in order to be able to learn the mapping between the two. For this, two principle strategies were pursued in this thesis: (1) single-pixel estimation and (2) whole image estimation.

For single-pixel estimation, data corresponding to a single pixel location in the image is used as the input feature vector, and the label for that one pixel is estimated. Usually, there is not enough information in single-pixel values, especially for the complex non-linear optical inverse problem. Because of this, a feature vector that contains relevant information for the inversion is needed for inference. In the CE-qPAI method (cf. section 3.1), this feature vector was a context image that comprised the 3D signal context relative to the light fluence in a homogeneous medium, and the LSD-qPAI method (cf. section 3.3) exploited multispectral information for each pixel. Such a pixel-wise approach allows for the generation of thousands of training samples from a single simulated image, which is very handy when facing modern data-hungry machine learning techniques. Furthermore, when using simulated images, a single-pixel approach yields a greater abstraction from the simulation domain, and, as such, the domain gap might not be as evident.

On the other hand, the extraction of hand-crafted features for the machine learning algorithm might not be optimal to represent the problem. It might be of benefit to let the algorithm *learn* which features to extract from the whole image information. This is done, for example, by learning the filter weights when using convolutional neural networks. However, when directly working with 2D or 3D images, careful creation and pre-processing of the image data is needed, as it is easy to have a huge domain gap in the training data domain versus the real data domain.

### **Multispectral information**

Another strategic decision that needs to be made is the consideration of whether to use monospectral image data for the inversion or whether to use multispectral image data. While monospectral information should, in theory, be sufficient for the estimation of the optical parameters, the multispectral behavior of the signal yields additional information regarding the underlying chromophore distribution, which could - in turn - help to more accurately estimate the optical absorption at a specific wavelength. Furthermore, only a multispectral signal can be used to derive functional tissue properties, as these are calculated based on the relative ratios of the chromophore distributions. When considering the application with data-driven approaches, the multispectral estimation of optical absorption coefficients could be formulated as a multi-task approach. Here, multiple labels would be defined, for example, each label corresponding to the absorption coefficients at a particular wavelength. Additional labels could be defined by also including the derived functional parameters into the estimation. Such an approach might be a good idea because multi-task approaches have been shown to have the potential to improve the performance of the individual tasks.

## **4.5 Uncertainty-aware parameter estimation**

In clinical applications, the quantification of uncertainty for data-driven estimates is of crucial importance. When the parameter estimates are used by a physician to determine the next steps

of the patient's intervention, the estimation confidence is a crucial piece of information for the physician to consider. In this thesis, several approaches for such uncertainty quantification techniques have been demonstrated.

There is a principal distinction that needs to be made regarding these methods, though, as some of these methods are only capable of yielding single point estimates for the expected error, while other techniques are capable of reconstructing the full posterior parameter distribution. Having access to the full posterior distribution is desirable, as it enables getting an unbiased view without artificially imposing a Gaussian distribution. Furthermore, it allows the identification of multiple modes in the posterior, which would not be detectable when only having access to, for example, the standard deviation.

Another exciting direction of research in regard to uncertainty estimation is the ability to automatically estimate out-of-distribution samples [Adler et al., 2019b]. When being tasked to predict data that is not within the distribution of the training data, a data-driven method might very confidently predict a completely wrong value, which could potentially lead to dire consequences for the patient. In future work, it should be examined if the computation of the techniques mentioned above for confidence estimation and out-of-distribution detection could be used for a robust derivation of clinically relevant parameters for use in clinical routine.

## 4.6 Conclusion

This thesis shows that data-driven approaches have the potential to solve many of the challenges of achieving quantitative photoacoustic imaging. To this end, three hypotheses were formulated and investigated in this thesis:

**(H1)** *Data-driven methods can be used to solve the optical and acoustic inverse problem:*

The capability of data-driven methods to estimate the light fluence or the optical absorption coefficients was examined in two rigorous *in silico* studies. Each of these was conducted on several different data sets of different difficulty levels. To this end, an algorithm was developed that can encode the entire three-dimensional signal context in a voxel-specific low-dimensional feature vector. Furthermore, state-of-the-art deep learning algorithms were used to estimate the optical absorption distribution directly on two-dimensional photoacoustic images. The results of these experiments show the general feasibility of data-driven approaches in the photoacoustic imaging context and also reveal the current limitations of these methods.

**(H2)** *Data-driven methods can be used for spectral unmixing in a realistic context:*

A data-driven approach was developed that was trained on simulated data and enabled the estimation of functional tissue properties in both *in vitro* and *in vivo* settings. Here, the capability of the method to predict plausible results for blood oxygenation in various contexts was demonstrated. The method continuously outperformed linear unmixing techniques in terms of the estimated dynamic range and overall agreement of the estimates with the expected values.

**(H3)** *Confidence estimation methods can be used to gain insight into qPAI methods:*

Four different techniques for uncertainty estimation were examined towards their applicability to photoacoustic imaging. The conducted experiments revealed that the integration of uncertainty estimates during result aggregation could potentially be of great benefit.

In conclusion, for sure, more work is required to achieve a successful and robust application of data-driven quantification of photoacoustic signals. This work, however, revealed the great potential of data-driven methods in this context and outlined several possible ways of applying them to the associated inverse problems. By now, other research groups have also applied data-driven approaches to the inverse problems in photoacoustic imaging, both to the acoustic inverse problem (cf. section 2.3.5) and the optical inverse problem (cf. section 2.5.2), further substantiating this conclusion. As such, in conjunction with the promising results of this Ph.D. thesis, it can confidently be assumed that the combination of model-based and data-driven approaches will form the foundation for a successful translation of quantitative photoacoustic imaging into clinical practice.

### Outlook

In prior work (e.g. [Cox et al., 2009b]) as well as in this thesis, it has been hypothesized and demonstrated *in silico* that the estimation of optical tissue properties from PA signals is principally possible. Based on the experiments conducted during this thesis, I believe that the quantification of the optical tissue properties of photon absorption and photon scattering is also feasible in practice. To this end, I hypothesize that a successful qPAI algorithm will have to be subject to at least several of the following constraints:

1. It is a **hybrid** algorithm and combines the advantages of data-driven approaches while being informed or regularized by rigorous models of the underlying physical principles.
2. It performs the inversion on **3D images** that are recorded in real-time with a probe specially designed for the target application.
3. It directly works with the **raw time series data** in order not to unnecessarily discard valuable information and in order to be able to include the most accurate recovery of  $p_0$  possible.
4. It performs the inversion on **multispectral PA images** in order to exploit the wavelength-dependent behavioral characteristics of the PA signal.
5. It is **specifically targeted** to a particular clinical use case. To this end, it might utilize a specially designed data set that contains a combination of simulated, as well as real recorded data.
6. It utilizes **US imaging** in order to exploit the naturally obtainable structural information of the imaged tissue to make educated *a priori* guesses on the tissue composition.
7. It uses more **experimentally determined information** on the imaged tissue using other imaging or sensing techniques. In an ideal case, these techniques are integrated into the PA device, such that they are instantly available as *a priori* knowledge for the inversion algorithm.



# 5 | Summary

This thesis shows that data-driven approaches have the potential to solve many of the challenges of achieving quantitative photoacoustic imaging. To this end, rigorous *in silico* evaluation of machine learning algorithms for the inverse problems associated with photoacoustic imaging is conducted, a data-driven approach for blood oxygenation estimation from multispectral photoacoustic measurements is applied *in vitro* and *in vivo*, and methods for uncertainty estimation for the developed algorithms are analyzed.

Photoacoustic imaging is an emerging imaging modality in healthcare. It promises noninvasive and radiation-free imaging of optical tissue properties. In contrast to commonly used optical imaging techniques, it can visualize optical tissue properties up to several centimeters deep in tissue. Photoacoustic imaging is based on the *photoacoustic effect*, which enables spatially resolved imaging of optically absorbing *chromophores*. When pulsed laser light is sent into tissue and is absorbed by chromophores, sound waves emerge at the location of the absorption event. These can be measured with ultrasound transducers and are reconstructed into a spatial image of absorbed energy. When using multiple measurements of the absorbed energy at different wavelengths of light, knowledge on clinically relevant functional tissue parameters, such as blood oxygenation, can be derived.

However, one of the critical challenges in photoacoustic imaging remains unsolved. This challenge is accurate and reliable quantification of the underlying optical tissue properties. Especially the estimation of the optical absorption coefficient of the tissue is essential for the derivation of functional tissue parameters. However, the absorbed energy is not only proportional to the optical absorption coefficients, but instead, it is also proportional to the light *fluence*. The fluence describes the distribution of light in tissue, which is predominantly determined by the optical absorption and scattering properties. Due to this, there is a non-linear interaction of optical absorption and fluence with respect to the absorbed energy and quantification of the signal is an ill-posed inverse problem, to which no general and easy-to-compute solution yet exists.

In the field, iterative model-based approaches have been proposed and thoroughly investigated. In numerous *in silico* investigations, these techniques have shown great theoretical potential. However, they have not successfully been applied to real images acquired in clinically relevant freehand imaging settings.

In this thesis, the feasibility of developing and applying data-driven methods to fill this gap is investigated. *Data-driven methods* refer to machine learning algorithms that learn an optimal inference model for a particular application based on training data that are relevant to the problem. For optical imaging applications generally, no ground truth information on the underlying tissue properties is available, and in this thesis, it is attempted to train the data-driven algorithms on computer-simulated data with the hope to gain the ability to infer tissue properties also in real scenarios.

The capability of data-driven methods to estimate the light fluence or the optical absorption coefficients was examined in two rigorous *in silico* studies. Each of these was conducted on several different data sets of different difficulty levels. To this end, an algorithm was developed that can encode the entire three-dimensional signal context in a voxel-specific low-dimensional feature vector. Furthermore, state-of-the-art deep learning algorithms were used to estimate the optical absorption distribution directly on two-dimensional photoacoustic images. The results of these experiments show the general feasibility of data-driven approaches in the photoacoustic imaging context and also reveal the current limitations of these methods.

Another data-driven approach was developed that was trained on simulated data and enabled the estimation of functional tissue properties in both *in vitro* and *in vivo* settings. Here, the capability of the method to predict plausible results for blood oxygenation in various contexts was demonstrated. The method continuously outperformed linear unmixing techniques in terms of the estimated dynamic range and agreement of the estimates with the expected values.

Finally, four different techniques for uncertainty estimation were examined towards their applicability to photoacoustic imaging. The conducted experiments revealed that the integration of uncertainty estimates during the calculation of results on a bigger region of interest could potentially be of great benefit.

For sure, more work is required to achieve a successful and robust application of data-driven quantification of photoacoustic signals. This work, however, revealed the potential of data-driven methods in this context and outlined several possible ways of applying them to the associated inverse problems. It can be assumed that the combination of model-based and data-driven approaches will form the foundation for a successful clinical translation of quantitative photoacoustic imaging into clinical practice.

# 6

## Zusammenfassung

Diese Dissertation zeigt, dass datengetriebenen Methoden das Potential haben, viele der Herausforderungen der quantitativen Photoakustik angehen zu können. Um dies zu untersuchen, werden rigorose *in silico* Evaluationen zur Applikation von maschinellen Lernverfahren auf die inversen Probleme vorgestellt, die mit der quantitativen Photoakustik verbunden sind. Außerdem wird ein datengetriebenes Modell benutzt, um die Blutsauerstoffsättigung aus multispektralen Photoakustikmessungen schätzen zu können. Es werden zusätzlich Methoden zur Unsicherheits-schätzung der entwickelten Algorithmen analysiert.

Die photoakustische Bildgebung ist eine aufstrebende Bildgebungsmodalität im Gesundheitswesen. Im Gegensatz zu herkömmlichen optischen Bildgebungsverfahren, können optische Gewebeeigenschaften bis zu mehrere Zentimeter tief im Gewebe sichtbar gemacht werden. Die Photoakustik basiert auf dem *photoakustischen Effekt*, der die orts aufgelöste Abbildung von optisch absorbierenden Molekülen ermöglicht. Wenn gepulstes Laserlicht in das Gewebe gesendet und absorbiert wird, treten am Ort des Absorptionereignisses Schallwellen auf. Diese können mit herkömmlichen Ultraschallgeräten gemessen und in ein räumliches Bild der absorbierten Energie rekonstruiert werden. Bei Verwendung mehrerer Messungen der absorbierten Energie bei verschiedenen Wellenlängen des Lichtes, kann Wissen über funktionale Gewebeparameter wie der Sauerstoffversorgung des Blutes abgeleitet werden.

Eine der entscheidenden Herausforderungen bei der photoakustischen Bildgebung bleibt jedoch ungelöst: die zuverlässige Quantifizierung der zugrundeliegenden optischen Gewebeeigenschaften. Insbesondere die Abschätzung des optischen Absorptionskoeffizienten des Gewebes ist für die Ableitung von funktionalen Gewebeparametern essentiell. Die absorbierte Energie ist allerdings nicht nur proportional zu den optischen Absorptionskoeffizienten, sondern auch proportional zu der Fluenz. Die Fluenz beschreibt die Lichtverteilung im Gewebe, die vorwiegend durch die optischen Absorptions- und Streuungseigenschaften bestimmt wird. Aufgrund dessen gibt es eine nichtlineare Wechselwirkung von optischer Absorption und Fluenz in Bezug auf die absorbierte Energie. Folglich ist die Quantifizierung des Signals ein schlecht gestelltes inverses

Problem, für das es noch keine allgemeine und leicht zu berechnende Lösung gibt.

Im Forschungsfeld wurden iterative modellbasierte Ansätze vorgeschlagen und eingehend untersucht. In zahlreichen Analysen haben diese Techniken ein großes theoretisches Potenzial gezeigt. Sie wurden jedoch nicht erfolgreich auf reale Bilder angewendet, die mit Hilfe von herkömmlichen Ultraschallgeräten aufgenommen wurden.

In dieser Arbeit wird untersucht, ob es generell machbar ist, datengetriebene Methoden zu entwickeln und anzuwenden, um diese Lücke zu schließen. *Datengetriebene Methoden* beziehen sich hierbei auf maschinelle Lernalgorithmen, die ein optimales Inferenzmodell für eine bestimmte Anwendung auf Grundlage der für das Problem relevanten Trainingsdaten erlernen. Für die optische Bildgebung sind im Allgemeinen keine "Ground Truth"-Informationen zu den zugrundeliegenden Gewebeeigenschaften verfügbar. Deshalb wird in dieser Arbeit versucht, die datengetriebenen Algorithmen auf computer-simulierten Daten zu trainieren, in der Hoffnung, auch in realen Szenarien auf die Gewebeeigenschaften schließen zu können.

Die Fähigkeit datengetriebener Methoden, die Fluenz und die optischen Absorptionskoeffizienten zu berechnen, wurde in zwei *in silico*-Studien untersucht. Jede dieser Analysen wurde mit mehreren Datensätzen mit unterschiedlichen Schwierigkeitsgraden durchgeführt. Zu diesem Zweck wurde ein Algorithmus entwickelt, der den gesamten dreidimensionalen Signalkontext in einem voxelspezifischen und niedrigdimensionalen Featurevektor codieren kann. Darüber hinaus wurden moderne Deep Learning-Algorithmen verwendet, um die optische Absorptionsverteilung direkt auf zweidimensionalen photoakustischen Bildern zu berechnen. Die Ergebnisse dieser Experimente zeigen die generelle Machbarkeit datengetriebener Ansätze im Kontext der photoakustischen Bildgebung und offenbaren aber auch die aktuellen Einschränkungen dieser Methoden.

Ein weiterer datengetriebener Ansatz wurde entwickelt, der auf simulierten Daten trainiert wurde und die Abschätzung der funktionalen Gewebeeigenschaften sowohl in *in vitro* als auch in *in vivo* ermöglicht. Hier wurde die Fähigkeit der Methode demonstriert, plausible Ergebnisse für die Blutsauerstoffsättigung in verschiedenen Kontexten vorherzusagen. Das Verfahren übertraf kontinuierlich die linearen Entmischungstechniken hinsichtlich des schätzbaren Wertebereichs und der Übereinstimmung der Berechnungen mit den erwarteten Werten.

Schließlich wurden vier verschiedene Techniken zur Schätzung der Unsicherheit auf ihre Anwendbarkeit für die photoakustische Bildgebung untersucht. Die durchgeführten Experimente haben gezeigt, dass die Integration von Unsicherheitsschätzungen während der Berechnung von Ergebnissen über einen größeren Messbereich möglicherweise von großem Nutzen sein kann.

Sicherlich sind weitere Arbeiten erforderlich, um eine erfolgreiche und robuste Anwendung der datengesteuerten Quantifizierung von photoakustischen Signalen zu erreichen. Diese Arbeit zeigte jedoch das Potenzial datengetriebener Methoden in diesem Zusammenhang auf. Sie skizzierte verschiedene Möglichkeiten, sie auf die inversen Probleme anzuwenden, die mit der Signalquantifizierung verbunden sind. Es ist davon auszugehen, dass die Kombination von modellbasierten und datengetriebenen Ansätzen die Grundlage für eine erfolgreiche klinische Umsetzung der quantitativen photoakustischen Bildgebung in die klinische Praxis bilden wird.

# Bibliography

- [Adler et al., 2019a] Tim J Adler, Lynton Ardizzone, Anant Vemuri, Leonardo Ayala, Janek Gröhl, Thomas Kirchner, Sebastian Wirkert, Jakob Kruse, Carsten Rother, Ullrich Köthe, et al. **Uncertainty-aware performance assessment of optical imaging modalities with invertible neural networks.** *International journal of computer assisted radiology and surgery*, pages 1–11, 2019a. doi: 10.1007/s11548-019-01939-9.
- [Adler et al., 2019b] Tim J Adler, Leonardo Ayala, Lynton Ardizzone, Hannes G Kenngott, Anant Vemuri, Beat P Müller-Stich, Carsten Rother, Ullrich Köthe, and Lena Maier-Hein. **Out of distribution detection for intra-operative functional imaging.** In *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging and Clinical Image-Based Procedures*, pages 75–82. Springer, 2019b. doi: 10.1007/978-3-030-32689-0\_8.
- [Aguirre et al., 2011] Andres Aguirre, Yasaman Ardeshirpour, Mary M Sanders, Molly Brewer, and Quing Zhu. **Potential role of coregistered photoacoustic and ultrasound imaging in ovarian cancer detection and characterization.** *Translational oncology*, 4(1):29, 2011. doi: 10.1593/tlo.10187.
- [Akhlaghi et al., 2019] Nima Akhlaghi, T Joshua Pfefer, Keith A Wear, Brian S Garra, and William C Vogt. **Multidomain computational modeling of photoacoustic imaging: verification, validation, and image quality prediction.** *Journal of biomedical optics*, 24(12):121910, 2019. doi: 10.1117/1.JBO.24.12.121910.
- [Alaluf et al., 2002] Simon Alaluf, Derek Atkins, Karen Barrett, Margaret Blount, Nik Carter, and Alan Heath. **Ethnic variation in melanin content and composition in photoexposed and photoprotected human skin.** *Pigment Cell Research*, 15(2):112–118, 2002. doi: 10.1034/j.1600-0749.2002.1o071.x.
- [Alessandrini et al., 2017] Giovanni Alessandrini, Michele Di Cristo, Elisa Francini, and Sergio Vessella. **Stability for quantitative photoacoustic tomography with well-chosen illuminations.** *Annali di Matematica Pura ed Applicata (1923-)*, 196(2):395–406, 2017. doi: 10.1007/s10231-016-0577-4.
- [Allman et al., 2018a] Derek Allman, Fabrizio Assis, Jonathan Chrispin, and Muyinatu A Lediju Bell. **Deep neural networks to remove photoacoustic reflection artifacts in ex vivo and in vivo tissue.** In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2018a. doi: 10.1109/ULTSYM.2018.8579723.
- [Allman et al., 2018b] Derek Allman, Austin Reiter, and Muyinatu Bell. **Exploring the effects of transducer models when training convolutional neural networks to eliminate reflection artifacts in experimental photoacoustic images.** In *Photons Plus Ultrasound: Imaging and Sensing 2018*, volume 10494, page 104945H. International Society for Optics and Photonics, 2018b. doi: 10.1117/12.2291279.
- [Allman et al., 2018c] Derek Allman, Austin Reiter, and Muyinatu A Lediju Bell. **Photoacoustic source detection and reflection artifact removal enabled by deep learning.** *IEEE transactions on medical imaging*, 37(6):1464–1477, 2018c. doi: 10.1109/TMI.2018.2829662.
- [Allman et al., 2019] Derek Allman, Fabrizio Assis, Jonathan Chrispin, and Muyinatu A Lediju Bell. **A deep learning-based approach to identify in vivo catheter tips during photoacoustic-guided cardiac interventions.** In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108785E. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2510993.
- [Alterovitz et al., 2008] Ron Alterovitz, Michael Branicky, and Ken Goldberg. **Motion planning under uncertainty for image-guided medical needle steering.** *The International journal of robotics research*, 27(11-12):1361–1374, 2008. doi: 10.1177/0278364908097661.
- [An et al., 2017] Lu An, Teedah Saratoon, Martina Fonseca, Robert Ellwood, and Ben Cox. **Statistical independence in nonlinear model-based inversion for quantitative photoacoustic tomography.** *Biomedical optics express*, 8(11):5297–5310, 2017. doi: 10.1364/BOE.8.005297.

- [Antholzer et al., 2018] Stephan Antholzer, Johannes Schwab, and Markus Haltmeier. **Deep Learning Versus  $l^1$ -Minimization for Compressed Sensing Photoacoustic Tomography**. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 206–212. IEEE, 2018. doi: 10.1109/ULTSYM.2018.8579737.
- [Antholzer et al., 2019a] Stephan Antholzer, Markus Haltmeier, and Johannes Schwab. **Deep learning for photoacoustic tomography from sparse data**. *Inverse problems in science and engineering*, 27(7):987–1005, 2019a. doi: 10.1080/17415977.2018.1518444.
- [Antholzer et al., 2019b] Stephan Antholzer, Johannes Schwab, Johannes Bauer-Marschallinger, Peter Burgholzer, and Markus Haltmeier. **NETT regularization for compressed sensing photoacoustic tomography**. In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108783B. International Society for Optics and Photonics, 2019b. doi: 10.1117/12.2508486.
- [Ardizzone et al., 2018] Lynton Ardizzone, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W Pellegrini, Ralf S Klessen, Lena Maier-Hein, Carsten Rother, and Ullrich Köthe. **Analyzing inverse problems with invertible neural networks**. *arXiv preprint arXiv:1808.04730*, 2018.
- [Ardizzone et al., 2019] Lynton Ardizzone, Carsten Lüth, Jakob Kruse, Carsten Rother, and Ullrich Köthe. **Guided image generation with conditional invertible neural networks**. *arXiv preprint arXiv:1907.02392*, 2019.
- [Ashraf et al., 2010] Tariq Ashraf, Ziauddin Panhwar, Sultana Habib, Muhammad Anis Memon, Fahad Shamsi, and Javed Arif. **Size of radial and ulnar artery in local population**. *JPMA-Journal of the Pakistan Medical Association*, 60(10): 817, 2010.
- [Attia et al., 2019] Amalina Binte Ebrahim Attia, Ghayathri Balasundaram, Mohesh Moothanchery, US Dinish, Renzhe Bi, Vasilis Ntziachristos, and Malini Olivo. **A review of clinical photoacoustic imaging: Current and future trends**. *Photoacoustics*, page 100144, 2019. doi: 10.1016/j.pacs.2019.100144.
- [Banerjee et al., 2008] Biswanath Banerjee, Srijeeta Bagchi, Ram Mohan Vasu, and Debasish Roy. **Quantitative photoacoustic tomography from boundary pressure measurements: noniterative recovery of optical absorption coefficient from the reconstructed absorbed energy map**. *JOSA A*, 25(9):2347–2356, 2008. doi: 10.1364/JOSAA.25.002347.
- [Bauer et al., 2011] Adam Q Bauer, Ralph E Nothdurft, Joseph P Culver, Todd N Erpelding, and Lihong V Wang. **Quantitative photoacoustic imaging: correcting for heterogeneous light fluence distributions using diffuse optical tomography**. *Journal of Biomedical Optics*, 16(9):096016, 2011. doi: 10.1117/1.3626212.
- [Bell, 1880] Alexander Graham Bell. **Upon the production and reproduction of sound by light**. *Journal of the Society of Telegraph Engineers*, 9(34):404–426, 1880. doi: 10.1049/jste-1.1880.0046.
- [Bell, 1881] Alexander Graham Bell. **The spectrophone**. Smithsonian Inst., 1881.
- [Beretta et al., 2015] Elena Beretta, Monika Muszkieta, Wolf Naetar, and Otmar Scherzer. **A variational method for quantitative photoacoustic tomography with piecewise constant coefficients**. *Variational Methods*, pages 202–224, 2015. doi: 10.1515/9783110430394-006.
- [Boink et al., 2018] Yoeri E Boink, Stephan A Van Gils, Srirang Manohar, and Christoph Brune. **Sensitivity of a partially learned model-based reconstruction algorithm**. *PAMM*, 18(1):e201800222, 2018. doi: 10.1002/pamm.201800222.
- [Bojarski, 1982] Norbert N Bojarski. **The k-space formulation of the scattering problem in the time domain**. *The Journal of the Acoustical Society of America*, 72(2):570–584, 1982. doi: 10.1121/1.2019394.
- [Bojarski, 1985] Norbert N Bojarski. **The k-space formulation of the scattering problem in the time domain: An improved single propagator formulation**. *The Journal of the Acoustical Society of America*, 77(3):826–831, 1985. doi: 10.1121/1.392051.
- [Bonnetier et al., 2019] Eric Bonnetier, Mourad Choulli, and Faouzi Triki. **Stability for quantitative photoacoustic tomography revisited**. *arXiv preprint arXiv:1905.07914*, 2019.
- [Bousmalis et al., 2018] Konstantinos Bousmalis, Alex Irpan, Paul Wohlhart, Yunfei Bai, Matthew Kelcey, Mrinal Kalakrishnan, Laura Downs, Julian Ibarz, Peter Pastor, Kurt Konolige, et al. **Using simulation and domain adaptation to improve efficiency of deep robotic grasping**. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4243–4250. IEEE, 2018. doi: 10.1109/ICRA.2018.8460875.
- [Bowen, 1983] Theodore Bowen. **Radiation-induced thermoacoustic imaging**, May 1983. US Patent 4,385,634.
- [Breiman, 2001] Leo Breiman. **Random forests**. *Machine learning*, 45(1):5–32, 2001. doi: 10.1023/A:1010933404324.
- [Brochu et al., 2016] Frederic M Brochu, Joanna Bruncker, James Joseph, Michal R Tomaszewski, Stefan Morscher, and Sarah E Bohndiek. **Towards quantitative evaluation of tissue absorption coefficients using light fluence correction in photoacoustic tomography**. *IEEE transactions on medical imaging*, 36(1):322–331, 2016. doi: 10.1109/TMI.2016.2607199.

- [Brunker et al., 2017] Joanna Brunker, Junjie Yao, Jan Laufer, and Sarah E Bohndiek. **Photoacoustic imaging using genetically encoded reporters: a review.** *Journal of biomedical optics*, 22(7):070901, 2017. doi: 10.1117/1.JBO.22.7.070901.
- [Bu et al., 2012] Shuhui Bu, Zhenbao Liu, Tsuyoshi Shiina, Kengo Kondo, Makoto Yamakawa, Kazuhiko Fukutani, Yasuhiro Sameda, and Yasufumi Asao. **Model-based reconstruction integrated with fluence compensation for photoacoustic tomography.** *IEEE Transactions on Biomedical Engineering*, 59(5):1354–1363, 2012. doi: 10.1109/TBME.2012.2187649.
- [Buchmann et al., 2017a] Jens Buchmann, James Guggenheim, Edward Zhang, Chris Scharfenorth, Bastian Spannekrebs, Claus Villringer, and Jan Laufer. **Characterization and modeling of Fabry–Perot ultrasound sensors with hard dielectric mirrors for photoacoustic imaging.** *Applied optics*, 56(17):5039–5046, 2017a. doi: 10.1364/AO.56.005039.
- [Buchmann et al., 2017b] Jens Buchmann, Bernhard A Kaplan, Steffen Prohaska, and Jan Laufer. **Experimental validation of a Monte-Carlo-based inversion scheme for 3D quantitative photoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 1006416. International Society for Optics and Photonics, 2017b. doi: 10.1117/12.2252359.
- [Buchmann et al., 2019a] Jens Buchmann, Bernhard Kaplan, Samuel Powell, Steffen Prohaska, and Jan Laufer. **Quantitative PA tomography of high resolution 3-D images: experimental validation in tissue phantoms.** *Photoacoustics*, 100157, 2019a. doi: 10.1016/j.pacs.2019.100157.
- [Buchmann et al., 2019b] Jens Buchmann, Bernhard A Kaplan, Samuel Powell, Steffen Prohaska, and Jan Laufer. **Three-dimensional quantitative photoacoustic tomography using an adjoint radiance Monte Carlo model and gradient descent.** *Journal of biomedical optics*, 24(6):066001, 2019b. doi: 10.1117/1.JBO.24.6.066001.
- [Cai et al., 2018] Chuangjian Cai, Kexin Deng, Cheng Ma, and Jianwen Luo. **End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging.** *Optics letters*, 43(12):2752–2755, 2018. doi: 10.1364/OL.43.002752.
- [Cai et al., 2019] Chuangjian Cai, Xuanhao Wang, Ke Si, Jun Qian, Jianwen Luo, and Cheng Ma. **Streak artifact suppression in photoacoustic computed tomography using adaptive back projection.** *Biomedical Optics Express*, 10(9):4803–4814, 2019. doi: 10.1364/BOE.10.004803.
- [Caruana, 1997] Rich Caruana. **Multitask learning.** *Machine learning*, 28(1):41–75, 1997. doi: 10.1007/978-1-4615-5529-2\_5.
- [Caruana and Niculescu-Mizil, 2006] Rich Caruana and Alexandru Niculescu-Mizil. **An empirical comparison of supervised learning algorithms.** In *Proceedings of the 23rd international conference on Machine learning*, pages 161–168. ACM, 2006. doi: 10.1145/1143844.1143865.
- [Chen et al., 2019] Cameron Chen, Krishna Kumar Gadepalli, Bob MacDonald, Yun Liu, Shiro Kadowaki, Kunal Nagpal, Timo Kohlberger, Jeff Dean, Greg Corrado, Jason Hipp, Craig Mermel, and Martin Stumpe. **An Augmented Reality Microscope with Real-time Artificial Intelligence Integration for Cancer Diagnosis.** *Nature Medicine*, 2019. doi: 10.1038/s41591-019-0539-7. URL <https://rdcu.be/b0yXv>.
- [Chen et al., 2017] Matthew C Chen, Robyn L Ball, Lingyao Yang, Nathaniel Moradzadeh, Brian E Chapman, David B Larson, Curtis P Langlotz, Timothy J Amrhein, and Matthew P Lungren. **Deep learning to classify radiology free-text reports.** *Radiology*, 286(3):845–852, 2017. doi: 10.1148/radiol.2017171115.
- [Choo-Smith et al., 2002] L-P Choo-Smith, HGM Edwards, H Ph Endtz, JM Kros, Freerk Heule, Hugh Barr, JS Robinson Jr, HA Bruining, and GJ Puppels. **Medical applications of Raman spectroscopy: from proof of principle to clinical implementation.** *Biopolymers: Original Research on Biomolecules*, 67(1):1–9, 2002. doi: 10.1002/bip.10064.
- [Christensen et al., 1968] John B Christensen, John P Adams, KO Cho, and Lawrence Miller. **A study of the interosseous distance between the radius and ulna during rotation of the forearm.** *The Anatomical Record*, 160(2):261–271, 1968. doi: 10.1002/ar.1091600212.
- [Chu and Thuerey, 2017] Mengyu Chu and Nils Thuerey. **Data-driven synthesis of smoke flows with CNN-based feature descriptors.** *ACM Transactions on Graphics (TOG)*, 36(4):69, 2017. doi: 10.1145/3072959.3073643.
- [Çiçek et al., 2016] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. **3D U-Net: learning dense volumetric segmentation from sparse annotation.** In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016. doi: 10.1007/978-3-319-46723-8\_49.
- [Collins et al., 2015] Julie-Ann Collins, Aram Rudenski, John Gibson, Luke Howard, and Ronan O’Driscoll. **Relating oxygen partial pressure, saturation and content: the haemoglobin–oxygen dissociation curve.** *Breathe*, 11(3):194–201, 2015. doi: 10.1183/20734735.001415.
- [Cox et al., 2005] Benjamin T Cox, SR Arridge, KP Kostli, and Paul C Beard. **Quantitative photoacoustic imaging: fitting a model of light transport to the initial pressure distribution.** In *Photons Plus Ultrasound: Imaging and Sensing 2005: The Sixth Conference on Biomedical Thermoacoustics, Optoacoustics, and Acousto-optics*, volume 5697, pages 49–55. International Society for Optics and Photonics, 2005. doi: 10.1117/12.597190.

- [Cox et al., 2006] Benjamin T Cox, Simon R Arridge, Kornel P Köstli, and Paul C Beard. **Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method.** *Applied Optics*, 45(8):1866–1875, 2006. doi: 10.1364/AO.45.001866.
- [Cox et al., 2009a] Benjamin T Cox, Simon R Arridge, and Paul C Beard. **Estimating chromophore distributions from multiwavelength photoacoustic images.** *JOSA A*, 26(2):443–455, 2009a. doi: 10.1364/JOSAA.26.000443.
- [Cox et al., 2012] Benjamin T Cox, Jan G Laufer, Paul C Beard, and Simon R Arridge. **Quantitative spectroscopic photoacoustic imaging: a review.** *Journal of biomedical optics*, 17(6):061202, 2012. doi: 10.1117/1.JBO.17.6.061202.
- [Cox and Beard, 2005] BT Cox and PC Beard. **Fast calculation of pulsed photoacoustic fields in fluids using k-space methods.** *The Journal of the Acoustical Society of America*, 117(6):3616–3627, 2005. doi: 10.1121/1.1920227.
- [Cox and Treeby, 2010] BT Cox and BE Treeby. **Effect of sensor directionality on photoacoustic imaging: a study using the k-wave toolbox.** In *Photons Plus Ultrasound: Imaging and Sensing 2010*, volume 7564, page 75640I. International Society for Optics and Photonics, 2010. doi: 10.1117/12.841671.
- [Cox et al., 2009b] BT Cox, JG Laufer, and PC Beard. **The challenges for quantitative photoacoustic imaging.** In *Photons Plus Ultrasound: Imaging and Sensing 2009*, volume 7177, page 717713. International Society for Optics and Photonics, 2009b. doi: 10.1117/12.806788.
- [Dash et al., 1997] Manoranjan Dash, Hua Liu, and Jun Yao. **Dimensionality reduction of unsupervised data.** In *Proceedings ninth ieee international conference on tools with artificial intelligence*, pages 532–539. IEEE, 1997. doi: 10.1109/TAI.1997.632300.
- [Davoudi et al., 2019] Neda Davoudi, Xosé Luís Deán-Ben, and Daniel Razansky. **Deep learning optoacoustic tomography with sparse data.** *Nature Machine Intelligence*, pages 1–8, 2019. doi: 10.1038/s42256-019-0095-3.
- [Deng and Yu, 2014] Li Deng and Dong Yu. **Deep Learning: Methods and Applications.** Technical Report MSR-TR-2014-21, Microsoft Research, May 2014.
- [Ding et al., 2017a] Lu Ding, Xosé Luís Deán-Ben, Neal C Burton, Robert W Sobol, Vasilis Ntziachristos, and Daniel Razansky. **Non-negative constrained inversion approaches for unmixing chromophores in multispectral optoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100641B. International Society for Optics and Photonics, 2017a. doi: 10.1117/12.2252807.
- [Ding et al., 2017b] Lu Ding, Xosé Luís Deán-Ben, and Daniel Razansky. **20 frames per second model-based reconstruction in cross-sectional optoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100641A. International Society for Optics and Photonics, 2017b. doi: 10.1117/12.2252666.
- [Dinh et al., 2016] Laurent Dinh, Jascha Sohl-Dickstein, and Samy Bengio. **Density estimation using real nvp.** *arXiv preprint arXiv:1605.08803*, 2016.
- [Dogdas et al., 2007] Belma Dogdas, David Stout, Arion F Chatziioannou, and Richard M Leahy. **Digimouse: a 3D whole body mouse atlas from CT and cryosection data.** *Physics in Medicine & Biology*, 52(3):577, 2007. doi: 10.1088/0031-9155/52/3/003.
- [Donahue et al., 2017] Chris Donahue, Zachary C Lipton, Akshay Balsubramani, and Julian McAuley. **Semantically decomposing the latent spaces of generative adversarial networks.** *arXiv preprint arXiv:1705.07904*, 2017.
- [Elbau et al., 2017] Peter Elbau, Leonidas Mindrinos, and Otmar Scherzer. **Quantitative reconstructions in multi-modal photoacoustic and optical coherence tomography imaging.** *Inverse Problems*, 34(1):014006, 2017. doi: 10.1088/1361-6420/aa9ae7.
- [Erpelding et al., 2010] Todd N Erpelding, Chulhong Kim, Manojit Pramanik, Ladislav Jankovic, Konstantin Maslov, Zijian Guo, Julie A Margenthaler, Michael D Pashley, and Lihong V Wang. **Sentinel lymph nodes in the rat: noninvasive photoacoustic and US imaging with a clinical US system.** *Radiology*, 256(1):102–110, 2010. doi: 10.1148/radiol.10091772.
- [Escobar and West, 1995] Michael D Escobar and Mike West. **Bayesian density estimation and inference using mixtures.** *Journal of the american statistical association*, 90(430):577–588, 1995. doi: 10.1080/01621459.1995.10476550.
- [Esteva et al., 2019] Andre Esteva, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, and Jeff Dean. **A guide to deep learning in healthcare.** *Nature medicine*, 25(1):24, 2019. doi: 10.1038/s41591-018-0316-z.
- [Fang and Boas, 2009] Qianqian Fang and David A Boas. **Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units.** *Optics express*, 17(22):20178–20190, 2009. doi: 10.1364/OE.17.020178.
- [Feindt, 2004] Michael Feindt. **A neural Bayesian estimator for conditional probability densities.** *arXiv preprint physics/0402093*, 2004.



- [Fleming and Cottrell, 1990] Michael K Fleming and Garrison W Cottrell. **Categorization of faces using unsupervised feature extraction**. In *1990 IJCNN International Joint Conference on Neural Networks*, pages 65–70. IEEE, 1990. doi: 10.1109/IJCNN.1990.137696.
- [Fonseca et al., 2016] Martina Fonseca, Teedah Saratoon, Bajram Zeqiri, Paul Beard, and Ben Cox. **Sensitivity of quantitative photoacoustic tomography inversion schemes to experimental uncertainty**. In *Photons Plus Ultrasound: Imaging and Sensing 2016*, volume 9708, page 97084X. International Society for Optics and Photonics, 2016. doi: 10.1117/12.2210916.
- [Fonseca et al., 2017] Martina Fonseca, Emma Malone, Felix Lucka, Rob Ellwood, Lu An, Simon Arridge, Paul Beard, and Ben Cox. **Three-dimensional photoacoustic imaging and inversion for accurate quantification of chromophore distributions**. In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 1006415. International Society for Optics and Photonics, 2017. doi: 10.1117/12.2250964.
- [Forbes et al., 1953] RM Forbes, AR Cooper, HH Mitchell, et al. **The composition of the adult human body as determined by chemical analysis**. *J Biol Chem*, 203(1):359–366, 1953.
- [Fournelle and Bost, 2019] Marc Fournelle and Wolfgang Bost. **Wave front analysis for enhanced time-domain beam-forming of point-like targets in optoacoustic imaging using a linear array**. *Photoacoustics*, 14:67–76, 2019. doi: 10.1016/j.pacs.2019.04.002.
- [Frangi et al., 2018] Alejandro F Frangi, Sotirios A Tsaftaris, and Jerry L Prince. **Simulation and synthesis in medical imaging**. *IEEE transactions on medical imaging*, 37(3):673, 2018. doi: 10.1007/978-3-319-46630-9.
- [Freeman et al., 1997] Jenny Freeman, Frank Downs, Lisa Marcucci, E Neil Lewis, Brad Blume, and Jeff Rish. **Multispectral and hyperspectral imaging: applications for medical and surgical diagnostics**. In *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Magnificent Milestones and Emerging Opportunities in Medical Engineering* (Cat. No. 97CH36136), volume 2, pages 700–701. IEEE, 1997. doi: 10.1109/IEMBS.1997.757727.
- [Gal and Ghahramani, 2016] Yarin Gal and Zoubin Ghahramani. **Dropout as a bayesian approximation: Representing model uncertainty in deep learning**. In *international conference on machine learning*, pages 1050–1059, 2016.
- [Gamelin et al., 2008] John Kenneth Gamelin, Andres Aquirre, Anastasios Maurudis, Fei Huang, Diego Castillo, Lihong V Wang, and Quing Zhu. **Curved array photoacoustic tomographic system for small animal imaging**. *Journal of biomedical optics*, 13(2):024007, 2008. doi: 10.1117/1.2907157.
- [Gao et al., 2015] Fei Gao, Qiwen Peng, Xiaohua Feng, Bo Gao, and Yuanjin Zheng. **Single-wavelength blood oxygen saturation sensing with combined optical absorption and scattering**. *IEEE Sensors Journal*, 16(7):1943–1948, 2015. doi: 10.1109/JSEN.2015.2510744.
- [Garcia-Urbe et al., 2015] Alejandro Garcia-Urbe, Todd N Erpelding, Arie Krumholz, Haixin Ke, Konstantin Maslov, Catherine Appleton, Julie A Margenthaler, and Lihong V Wang. **Dual-modality photoacoustic and ultrasound imaging system for noninvasive sentinel lymph node detection in patients with breast cancer**. *Scientific reports*, 5:15748, 2015. doi: 10.1038/srep15748.
- [Ge et al., 2019] Yuying Ge, Ruimao Zhang, Xiaogang Wang, Xiaou Tang, and Ping Luo. **DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images**. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5337–5345, 2019. doi: 10.1109/CVPR.2019.00548.
- [Geethanath et al., 2013] Sairam Geethanath, Rashmi Reddy, Amaresha Shridhar Konar, Shaikh Imam, Rajagopalan Sundaresan, and Ramesh Venkatesan. **Compressed sensing MRI: a review**. *Critical Reviews™ in Biomedical Engineering*, 41(3), 2013. doi: 10.1615/CritRevBiomedEng.2014008058.
- [Gehring et al., 2019] Marcel Gehring, Sarah E. Bohndiek, and Joanna Brunker. **Development of a blood oxygenation phantom for photoacoustic tomography combined with online pO<sub>2</sub> detection and flow spectrometry**. *Journal of Biomedical Optics*, 24(12):1–11, 2019. doi: 10.1117/1.JBO.24.12.121908. URL <https://doi.org/10.1117/1.JBO.24.12.121908>.
- [Goertzel and Pennachin, 2007] Ben Goertzel and Cassio Pennachin. **Artificial general intelligence**, volume 2. Springer, 2007. doi: 10.1007/978-3-319-09274-4.
- [Graaff and Hoenders, 2005] Reindert Graaff and Bernhard J Hoenders. **Diffusion theory for light propagation in biological tissue: limitations and adaptations**. In *Saratov Fall Meeting 2004: Optical Technologies in Biophysics and Medicine VI*, volume 5771, pages 28–41. International Society for Optics and Photonics, 2005. doi: 10.1117/12.634647.
- [Gröhl, 2016] Janek Gröhl. **Machine learning-based quantitative photoacoustic tomography**. Master's thesis, Medical faculty of the University of Heidelberg, Heidelberg, Germany, 2016.

- [Gröhl et al., 2018a] Janek Gröhl, Thomas Kirchner, Tim Adler, and Lena Maier-Hein. **Confidence estimation for machine learning-based quantitative photoacoustics**. *Journal of Imaging*, 4(12):147, 2018a. doi: 10.3390/jimaging4120147.
- [Gröhl et al., 2018b] Janek Gröhl, Thomas Kirchner, and Lena Maier-Hein. **Confidence estimation for quantitative photoacoustic imaging**. In *Photons Plus Ultrasound: Imaging and Sensing 2018*, volume 10494, page 104941C. International Society for Optics and Photonics, 2018b. doi: 10.1117/12.2288362.
- [Gröhl et al., 2019] Janek Gröhl, Thomas Kirchner, Tim Adler, and Lena Maier-Hein. **Estimation of blood oxygenation with learned spectral decoloring for quantitative photoacoustic imaging (LSD-qPAI)**. *arXiv preprint arXiv:1902.05839*, 2019.
- [Güera and Delp, 2018] David Güera and Edward J Delp. **Deepfake video detection using recurrent neural networks**. In *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6. IEEE, 2018. doi: 10.1109/AVSS.2018.8639163.
- [Guo et al., 2017] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. **On calibration of modern neural networks**. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1321–1330. JMLR. org, 2017. doi: 10.5555/3305381.3305518.
- [Haacke et al., 1995] E Mark Haacke, Song Lai, Dmitriy A Yablonskiy, and Weili Lin. **In vivo validation of the BOLD mechanism: a review of signal changes in gradient echo functional MRI in the presence of flow**. *International Journal of Imaging Systems and Technology*, 6(2-3):153–163, 1995. doi: 10.1002/ima.1850060204.
- [Hädrich et al., 2017] Torsten Hädrich, Bedrich Benes, Oliver Deussen, and Sören Pirk. **Interactive modeling and authoring of climbing plants**. In *Computer Graphics Forum*, volume 36, pages 49–61. Wiley Online Library, 2017. doi: 10.1111/cgf.13106.
- [Haltmeier and Nguyen, 2019] Markus Haltmeier and Linh V Nguyen. **Reconstruction Algorithms for Photoacoustic Tomography in Heterogeneous Damping Media**. *Journal of Mathematical Imaging and Vision*, pages 1–15, 2019. doi: 10.1007/s10851-019-00879-y.
- [Haltmeier et al., 2015] Markus Haltmeier, Lukas Neumann, and Simon Rabanser. **Single-stage reconstruction algorithm for quantitative photoacoustic tomography**. *Inverse Problems*, 31(6):065005, 2015. doi: 10.1088/0266-5611/31/6/065005.
- [Haltmeier et al., 2019] Markus Haltmeier, Gerhard Zangerl, Linh V Nguyen, and Robert Nuster. **Photoacoustic image reconstruction from full field data in heterogeneous media**. In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108783D. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2508541.
- [Han et al., 2018] Jiequn Han, Arnulf Jentzen, and E Weinan. **Solving high-dimensional partial differential equations using deep learning**. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018. doi: 10.1073/pnas.1718942115.
- [Hanahan and Weinberg, 2011] Douglas Hanahan and Robert A Weinberg. **Hallmarks of cancer: the next generation**. *cell*, 144(5):646–674, 2011. doi: 10.1016/j.cell.2011.02.013.
- [Hänninen et al., 2018] Niko Hänninen, Aki Pulkkinen, and Tanja Tarvainen. **Image reconstruction with reliability assessment in quantitative photoacoustic tomography**. *Journal of Imaging*, 4(12):148, 2018. doi: 10.3390/jimaging4120148.
- [Havaei et al., 2017] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. **Brain tumor segmentation with deep neural networks**. *Medical image analysis*, 35:18–31, 2017. doi: 10.1016/j.media.2016.05.004.
- [Held et al., 2016] K Gerrit Held, Michael Jaeger, Jaro Rička, Martin Frenz, and H Günhan Akarçay. **Multiple irradiation sensing of the optical effective attenuation coefficient for spectral correction in handheld OA imaging**. *Photoacoustics*, 4(2):70–80, 2016. doi: 10.1016/j.pacs.2016.05.004.
- [Hochuli, 2016] Roman Hochuli. **Monte Carlo methods in quantitative photoacoustic tomography**. PhD thesis, UCL (University College London), 2016.
- [Hochuli et al., 2016] Roman Hochuli, Samuel Powell, Simon Arridge, and Ben Cox. **Quantitative photoacoustic tomography using forward and adjoint Monte Carlo models of radiance**. *Journal of biomedical optics*, 21(12):126004, 2016. doi: 10.1117/1.JBO.21.12.126004.
- [Hochuli et al., 2019] Roman Hochuli, Lu An, Paul C Beard, and Benjamin T Cox. **Estimating blood oxygenation from photoacoustic images: can a simple linear spectroscopic inversion ever work?** *Journal of Biomedical Optics*, 24(12):121914, 2019. doi: 10.1117/1.JBO.24.12.121914.
- [Hoffer-Hawlik and Luke, 2019] Kevin Hoffer-Hawlik and Geoffrey P Luke. **absO2luteU-Net: Tissue Oxygenation Calculation Using Photoacoustic Imaging and Convolutional Neural Networks**. Master’s thesis, Dartmouth College, 2019.

- [Holzinger et al., 2017] Andreas Holzinger, Chris Biemann, Constantinos S Pattichis, and Douglas B Kell. **What do we need to build explainable AI systems for the medical domain?** *arXiv preprint arXiv:1712.09923*, 2017.
- [Hong et al., 2018] Weixiang Hong, Zhenzhen Wang, Ming Yang, and Junsong Yuan. **Conditional generative adversarial network for structured domain adaptation.** In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1335–1344, 2018. doi: 10.1109/CVPR.2018.00145.
- [Hristova et al., 2008] Yulia Hristova, Peter Kuchment, and Linh Nguyen. **Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media.** *Inverse Problems*, 24(5):055006, 2008. doi: 10.1088/0266-5611/24/5/055006.
- [Hron et al., 2017] Jiri Hron, Alexander G de G Matthews, and Zoubin Ghahramani. **Variational Gaussian Dropout is not Bayesian.** *arXiv preprint arXiv:1711.02989*, 2017.
- [Hu et al., 2009] Song Hu, Konstantin I Maslov, Vassiliy Tsytarev, and Lihong V Wang. **Functional transcranial brain imaging by optical-resolution photoacoustic microscopy.** *Journal of biomedical optics*, 14(4):040503, 2009. doi: 10.1117/1.3194136.
- [Hu et al., 2010] Song Hu, Bin Rao, Konstantin Maslov, and Lihong V Wang. **Label-free photoacoustic ophthalmic angiography.** *Optics letters*, 35(1):1–3, 2010. doi: 10.1364/OL.35.000001.
- [Hubmer et al., 2004] Martin G Hubmer, Thomas Fasching, Franz Haas, Horst Koch, Franz Schwarzl, Andreas Weiglein, and Erwin Scharnagl. **The posterior interosseous artery in the distal part of the forearm. Is the term 'recurrent branch of the anterior interosseous artery' justified?** *British journal of plastic surgery*, 57(7):638–644, 2004. doi: 10.1016/j.bjps.2004.06.011.
- [Hussain et al., 2018] Altaf Hussain, Erwin Hondebrink, Jacob Staley, and Wiendelt Steenbergen. **Photoacoustic and acousto-optic tomography for quantitative and functional imaging.** *Optica*, 5(12):1579–1589, 2018. doi: 10.1364/OPTICA.5.001579.
- [Huynh et al., 2017] Nam Huynh, Felix Lucka, Edward Zhang, Marta Betcke, Simon Arridge, Paul Beard, and Ben Cox. **Sub-sampled Fabry-Perot photoacoustic scanner for fast 3D imaging.** In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100641Y. International Society for Optics and Photonics, 2017. doi: 10.1117/12.2250868.
- [Isensee et al., 2018] Fabian Isensee, Jens Petersen, Andre Klein, David Zimmerer, Paul F Jaeger, Simon Kohl, Jakob Wasserthal, Gregor Koehler, Tobias Norajitra, Sebastian Wirkert, et al. **nnu-net: Self-adapting framework for u-net-based medical image segmentation.** *arXiv preprint arXiv:1809.10486*, 2018. doi: 10.1007/978-3-658-25326-4\_7.
- [Jacques, 2013] Steven L Jacques. **Optical properties of biological tissues: a review.** *Physics in Medicine & Biology*, 58(11):R37, 2013. doi: 10.1088/0031-9155/58/11/R37.
- [Jacques, 2014] Steven L Jacques. **Coupling 3D Monte Carlo light transport in optically heterogeneous tissues to photoacoustic signal generation.** *Photoacoustics*, 2(4):137–142, 2014. doi: 10.1016/j.pacs.2014.09.001.
- [Javaherian and Holman, 2019] Ashkan Javaherian and Sean Holman. **Direct Quantitative Photoacoustic Tomography for realistic acoustic media.** *Inverse Problems*, 2019. doi: 10.1088/1361-6420/ab091e.
- [Jnawali et al., 2019] Kamal Jnawali, Bhargava Chinni, Vikram Dogra, Saugata Sinha, and Naval Gund Rao. **Deep 3D convolutional neural network for automatic cancer tissue detection using multispectral photoacoustic imaging.** In *Medical Imaging 2019: Ultrasonic Imaging and Tomography*, volume 10955, page 109551D. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2518686.
- [Jo et al., 2018] Junggun Jo, Guan Xu, Yunhao Zhu, Mary Burton, Jeffrey Sarazin, Elena Schiopu, Girish Gandikota, and Xueding Wang. **Detecting joint inflammation by an LED-based photoacoustic imaging system: a feasibility study.** *Journal of biomedical optics*, 23(11):110501, 2018. doi: 10.1117/1.JBO.23.11.110501.
- [Johnstonbaugh et al., 2019] Kerrick Johnstonbaugh, Sumit Agrawal, Deepit Abhishek, Matthew Homewood, Sri Phani Krisna Karri, and Sri-Rajasekhar Kothapalli. **Novel deep learning architecture for optical fluence dependent photoacoustic target localization.** In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108781L. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2511015.
- [Kallweit et al., 2017] Simon Kallweit, Thomas Müller, Brian McWilliams, Markus Gross, and Jan Novák. **Deep scattering: Rendering atmospheric clouds with radiance-predicting neural networks.** *ACM Transactions on Graphics (TOG)*, 36(6):231, 2017. doi: 10.1145/3130800.3130880.
- [Kang et al., 2017] Eunhee Kang, Junhong Min, and Jong Chul Ye. **A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction.** *Medical physics*, 44(10):e360–e375, 2017. doi: 10.1002/mp.12344.
- [Kaplan et al., 2017] Bernhard A Kaplan, Jens Buchmann, Steffen Prohaska, and Jan Laufer. **Monte-Carlo-based inversion scheme for 3D quantitative photoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100645J. International Society for Optics and Photonics, 2017. doi: 10.1117/12.2251945.

- [Kelly, 2010] Anne-Maree Kelly. **Can venous blood gas analysis replace arterial in emergency medical care.** *Emergency Medicine Australasia*, 22(6):493–498, 2010. doi: 10.1111/j.1742-6723.2010.01344.x.
- [Kendall and Gal, 2017] Alex Kendall and Yarin Gal. **What uncertainties do we need in bayesian deep learning for computer vision?** In *Advances in neural information processing systems*, pages 5574–5584, 2017.
- [Kerr and Atwood, 1968] Edwin L Kerr and John G Atwood. **The laser illuminated absorptivity spectrophone: a method for measurement of weak absorptivity in gases at laser wavelengths.** *Applied optics*, 7(5):915–921, 1968. doi: 10.1364/AO.7.000915.
- [Keshava, 2003] Nirmal Keshava. **A survey of spectral unmixing algorithms.** *Lincoln laboratory journal*, 14(1):55–78, 2003.
- [Keshava and Mustard, 2002] Nirmal Keshava and John F Mustard. **Spectral unmixing.** *IEEE signal processing magazine*, 19(1):44–57, 2002. doi: 10.1109/79.974727.
- [Kim et al., 2019] Byungsoo Kim, Vinicius C Azevedo, Nils Thuerey, Theodore Kim, Markus Gross, and Barbara Solenthaler. **Deep fluids: A generative network for parameterized fluid simulations.** In *Computer Graphics Forum*, volume 38, pages 59–70. Wiley Online Library, 2019. doi: 10.1111/cgf.13619.
- [Kim et al., 2010] Chulhong Kim, Todd N Erpelding, Konstantin I Maslov, Ladislav Jankovic, Walter J Akers, Liang Song, Samuel Achilefu, Julie A Margenthaler, Michael D Pashley, and Lihong V Wang. **Handheld array-based photoacoustic probe for guiding needle biopsy of sentinel lymph nodes.** *Journal of biomedical optics*, 15(4):046010, 2010. doi: 10.1117/1.3469829.
- [Kingma and Ba, 2014] Diederik P Kingma and Jimmy Ba. **Adam: A method for stochastic optimization.** *arXiv preprint arXiv:1412.6980*, 2014.
- [Kingma and Welling, 2013] Diederik P Kingma and Max Welling. **Auto-encoding variational bayes.** *arXiv preprint arXiv:1312.6114*, 2013.
- [Kingma and Dhariwal, 2018] Durk P Kingma and Prafulla Dhariwal. **Glow: Generative flow with invertible 1x1 convolutions.** In *Advances in Neural Information Processing Systems*, pages 10215–10224, 2018.
- [Kirchner et al., 2016] Thomas Kirchner, Esther Wild, Klaus H Maier-Hein, and Lena Maier-Hein. **Freehand photoacoustic tomography for 3D angiography using local gradient information.** In *Photons Plus Ultrasound: Imaging and Sensing 2016*, volume 9708, page 97083G. International Society for Optics and Photonics, 2016. doi: 10.1117/12.2209368.
- [Kirchner et al., 2018a] Thomas Kirchner, Janek Gröhl, and Lena Maier-Hein. **Context encoding enables machine learning-based quantitative photoacoustics.** *Journal of biomedical optics*, 23(5):056008, 2018a. doi: 10.1117/1.JBO.23.5.056008.
- [Kirchner et al., 2018b] Thomas Kirchner, Franz Sattler, Janek Gröhl, and Lena Maier-Hein. **Signed real-time delay multiply and sum beamforming for multispectral photoacoustic imaging.** *Journal of Imaging*, 4(10):121, 2018b. doi: 10.3390/jimaging4100121.
- [Kirchner et al., 2019] Thomas Kirchner, Janek Gröhl, Mildred A Herrera, Tim Adler, Adrián Hernández-Aguilera, Edgar Santos, and Lena Maier-Hein. **Photoacoustics can image spreading depolarization deep in gyrencephalic brain.** *Scientific reports*, 9(1):8661, 2019. doi: 10.1038/s41598-019-44935-4.
- [Klein et al., 2018] André Klein, Jan Warszawski, Jens Hillengaß, and Klaus Hermann Maier-Hein. **Towards whole-body CT bone segmentation.** In *Bildverarbeitung für die Medizin 2018*, pages 204–209. Springer, 2018. doi: 10.1007/978-3-662-56537-7\_59.
- [Klein et al., 2019] André Klein, Jan Warszawski, Jens Hillengaß, and Klaus H Maier-Hein. **Automatic bone segmentation in whole-body CT images.** *International journal of computer assisted radiology and surgery*, 14(1):21–29, 2019. doi: 10.1007/s11548-018-1883-7.
- [Knieling et al., 2017] Ferdinand Knieling, Clemens Neufert, Arndt Hartmann, Jing Claussen, Alexander Urich, Cornelia Egger, Marcel Vetter, Sarah Fischer, Lukas Pfeifer, Alexander Hagel, et al. **Multispectral optoacoustic tomography for assessment of Crohn’s disease activity.** *New England Journal of Medicine*, 376(13):1292–1294, 2017. doi: 10.1056/NEJMc1612455.
- [Kohl et al., 2018] Simon Kohl, Bernardino Romera-Paredes, Clemens Meyer, Jeffrey De Fauw, Joseph R Ledsam, Klaus Maier-Hein, SM Ali Eslami, Danilo Jimenez Rezende, and Olaf Ronneberger. **A probabilistic u-net for segmentation of ambiguous images.** In *Advances in Neural Information Processing Systems*, pages 6965–6975, 2018.
- [Kononenko, 2001] Igor Kononenko. **Machine learning for medical diagnosis: history, state of the art and perspective.** *Artificial Intelligence in medicine*, 23(1):89–109, 2001. doi: 10.1016/S0933-3657(01)00077-X.
- [Kremkau and Taylor, 1986] Frederick W Kremkau and KJ Taylor. **Artifacts in ultrasound imaging.** *Journal of ultrasound in medicine*, 5(4):227–237, 1986. doi: 10.7863/jum.1986.5.4.227.

- [Kruger et al., 2003] Robert A Kruger, William L Kiser, Daniel R Reinecke, Gabe A Kruger, and Kathy D Miller. **Thermoacoustic molecular imaging of small animals.** *Molecular Imaging*, 2(2):15353500200303109, 2003. doi: 10.1162/15353500200303109.
- [Ku et al., 2005] Geng Ku, Xueding Wang, Xueyi Xie, George Stoica, and Lihong V Wang. **Imaging of tumor angiogenesis in rat brains in vivo by photoacoustic tomography.** *Applied optics*, 44(5):770–775, 2005. doi: 10.1364/AO.44.000770.
- [Kuchment, 2013] Peter Kuchment. **The Radon transform and medical imaging.** SIAM, 2013. doi: 10.1137/1.9781611973297.
- [Lakshminarayanan et al., 2017] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. **Simple and scalable predictive uncertainty estimation using deep ensembles.** In *Advances in Neural Information Processing Systems*, pages 6402–6413, 2017.
- [Lan et al., 2019] Hengrong Lan, Kang Zhou, Changchun Yang, Jun Cheng, Jiang Liu, Shenghua Gao, and Fei Gao. **KiGAN: Knowledge Infusion Generative Adversarial Network for Photoacoustic Image Reconstruction In Vivo.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 273–281. Springer, 2019. doi: 10.1007/978-3-030-32239-7\_31.
- [Larue et al., 2017] Ruben THM Larue, Gilles Defraene, Dirk De Ruyscher, Philippe Lambin, and Wouter Van Elmpt. **Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures.** *The British journal of radiology*, 90(1070):20160665, 2017. doi: 10.1259/bjr.20160665.
- [Laufer et al., 2006] Jan Laufer, Dave Delpy, Clare Elwell, and Paul Beard. **Quantitative spatially resolved measurement of tissue chromophore concentrations using photoacoustic spectroscopy: application to the measurement of blood oxygenation and haemoglobin concentration.** *Physics in Medicine & Biology*, 52(1):141, 2006. doi: 10.1088/0031-9155/52/1/010.
- [Lee et al., 2017] Cecilia S Lee, Doug M Baughman, and Aaron Y Lee. **Deep learning is effective for classifying normal versus age-related macular degeneration OCT images.** *Ophthalmology Retina*, 1(4):322–327, 2017. doi: 10.1016/j.oret.2016.12.009.
- [Lei et al., 2019] Hao Lei, Laura A Johnson, Kathryn A Eaton, Shengchun Liu, Jun Ni, Xueding Wang, Peter DR Higgins, and Guan Xu. **Characterizing intestinal strictures of Crohn’s disease in vivo by endoscopic photoacoustic imaging.** *Biomedical optics express*, 10(5):2542–2555, 2019. doi: 10.1364/BOE.10.002542.
- [Leibig et al., 2017] Christian Leibig, Vaneeda Allken, Murat Seçkin Ayhan, Philipp Berens, and Siegfried Wahl. **Leveraging uncertainty information from deep neural networks for disease detection.** *Scientific reports*, 7(1):17816, 2017. doi: 10.1038/s41598-017-17876-z.
- [Levine et al., 2018] Sergey Levine, Peter Pastor, Alex Krizhevsky, Julian Ibarz, and Deirdre Quillen. **Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection.** *The International Journal of Robotics Research*, 37(4-5):421–436, 2018. doi: 10.1177/0278364917710318.
- [Li et al., 2019] Hua Li, Shasha Zhuang, Deng-ao Li, Jumin Zhao, and Yanyun Ma. **Benign and malignant classification of mammogram images based on deep learning.** *Biomedical Signal Processing and Control*, 51:347–354, 2019. doi: 10.1016/j.bspc.2019.02.017.
- [Li et al., 2018] Mucong Li, Yuqi Tang, and Junjie Yao. **Photoacoustic tomography of blood oxygenation: a mini review.** *Photoacoustics*, 10:65–73, 2018. doi: 10.1016/j.pacs.2018.05.001.
- [Li and Gal, 2017] Yingzhen Li and Yarin Gal. **Dropout inference in bayesian neural networks with alpha-divergences.** In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2052–2061. JMLR. org, 2017.
- [Li et al., 2017] Zheng Li, Dan-Qing Yang, Shi-Lei Liu, Si-Yuan Yu, Ming-Hui Lu, Jie Zhu, Shan-Tao Zhang, Ming-Wei Zhu, Xia-Sheng Guo, Hao-Dong Wu, et al. **Broadband gradient impedance matching using an acoustic metamaterial for ultrasonic transducers.** *Scientific reports*, 7:42863, 2017. doi: 10.1038/srep42863.
- [Liao et al., 2013] Shu Liao, Yaozong Gao, Aytakin Oto, and Dinggang Shen. **Representation learning: a unified deep learning framework for automatic prostate MR segmentation.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 254–261. Springer, 2013. doi: 10.1007/978-3-642-40763-5\_32.
- [Lin et al., 2018] Xing Lin, Yair Rivenson, Nezhir T Yardimci, Muhammed Veli, Yi Luo, Mona Jarrahi, and Aydogan Ozcan. **All-optical machine learning using diffractive deep neural networks.** *Science*, 361(6406):1004–1008, 2018. doi: 10.1126/science.aat8084.
- [Liu, 2017] Bing Liu. **Lifelong machine learning: a paradigm for continuous learning.** *Frontiers of Computer Science*, 11(3):359–361, 2017. doi: 10.1007/s11704-016-6903-6.
- [Liu et al., 2019a] Fangyan Liu, Xiaojing Gong, Lihong V Wang, Jingjing Guan, Liang Song, and Jing Meng. **Dictionary learning sparse-sampling reconstruction method for in-vivo 3D photoacoustic computed tomography.** *Biomedical Optics Express*, 10(4):1660–1677, 2019a. doi: 10.1364/BOE.10.001660.

- [Liu et al., 2018] Wenlong Liu, Yuejin Zhao, Ming Liu, Liquan Dong, Xiaohua Liu, and Mei Hui. **Generating simulated SAR images using Generative Adversarial Network**. In *Applications of Digital Image Processing XLI*, volume 10752, page 1075205. International Society for Optics and Photonics, 2018. doi: 10.1117/12.2320024.
- [Liu and Zhang, 2016] Wenzhong Liu and Hao F Zhang. **Photoacoustic imaging of the eye: a mini review**. *Photoacoustics*, 4(3):112–123, 2016. doi: 10.1016/j.pacs.2016.05.001.
- [Liu et al., 2019b] Yang Liu, Mingjian Sun, Ting Liu, Yiming Ma, Depeng Hu, Chao Li, and Naizhang Feng. **Quantitative Reconstruction of Absorption Coefficients for Photoacoustic Tomography**. *Applied Sciences*, 9(6):1187, 2019b. doi: 10.3390/app9061187.
- [Liu et al., 2016] Yubin Liu, Huabei Jiang, and Zhen Yuan. **Two schemes for quantitative photoacoustic tomography based on Monte Carlo simulation**. *Medical physics*, 43(7):3987–3997, 2016. doi: 10.1118/1.4953185.
- [Luke et al., 2012] Geoffrey P Luke, Doug Yeager, and Stanislav Y Emelianov. **Biomedical applications of photoacoustic imaging with exogenous contrast agents**. *Annals of biomedical engineering*, 40(2):422–437, 2012. doi: 10.1007/s10439-011-0449-4.
- [Luke et al., 2019] Geoffrey P Luke, Kevin Hoffer-Hawlik, Austin C Van Namen, and Ruibo Shang. **O-Net: A Convolutional Neural Network for Quantitative Photoacoustic Image Segmentation and Oximetry**. *arXiv preprint arXiv:1911.01935*, 2019.
- [Ma et al., 2014] Teng Ma, Xiangyang Zhang, Chi Tat Chiu, Ruimin Chen, Koping Kirk Shung, Qifa Zhou, and Shuliang Jiao. **Systematic study of high-frequency ultrasonic transducer design for laser-scanning photoacoustic ophthalmoscopy**. *Journal of biomedical optics*, 19(1):016015, 2014. doi: 10.1117/1.JBO.19.1.016015.
- [Maes et al., 1997] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens. **Multimodality image registration by maximization of mutual information**. *IEEE transactions on Medical Imaging*, 16(2):187–198, 1997. doi: 10.1109/42.563664.
- [Mahan et al., 1998] GD Mahan, WE Engler, JJ Tiemann, and E Uzgiris. **Ultrasonic tagging of light: theory**. *Proceedings of the National Academy of Sciences*, 95(24):14015–14019, 1998. doi: 10.1073/pnas.95.24.14015.
- [Mahmood et al., 2018] Faisal Mahmood, Richard Chen, and Nicholas J Durr. **Unsupervised reverse domain adaptation for synthetic medical images via adversarial training**. *IEEE transactions on medical imaging*, 37(12):2572–2581, 2018. doi: 10.1109/TMI.2018.2842767.
- [Maier-Hein et al., 2011] Lena Maier-Hein, Alfred Michael Franz, Thiago R Dos Santos, Mirko Schmidt, Markus Fangerau, Hans-Peter Meinzer, and J Michael Fitzpatrick. **Convergent iterative closest-point algorithm to accommodate anisotropic and inhomogeneous localization error**. *IEEE transactions on pattern analysis and machine intelligence*, 34(8):1520–1532, 2011. doi: 10.1109/TPAMI.2011.248.
- [Mallidi et al., 2011] Srivalleesha Mallidi, Geoffrey P Luke, and Stanislav Emelianov. **Photoacoustic imaging in cancer detection, diagnosis, and treatment guidance**. *Trends in biotechnology*, 29(5):213–221, 2011. doi: 10.1016/j.tibtech.2011.01.006.
- [Mandal et al., 2019] Subhamoy Mandal, Melanie Mueller, and Dorde Komljenovic. **Multimodal priors reduce acoustic and optical inaccuracies in photoacoustic imaging**. In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108781M. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2511662.
- [Manohar and Razansky, 2016] Srirang Manohar and Daniel Razansky. **Photoacoustics: a historical review**. *Advances in optics and photonics*, 8(4):586–617, 2016. doi: 10.1364/AOP.8.000586.
- [Masood et al., 2015] Ammara Masood, Adel Al-Jumaily, and Khairul Anam. **Self-supervised learning model for skin cancer diagnosis**. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1012–1015. IEEE, 2015. doi: 10.1109/NER.2015.7146798.
- [Mastanduno and Gambhir, 2016] Michael A Mastanduno and Sanjiv S Gambhir. **Quantitative photoacoustic image reconstruction improves accuracy in deep tissue structures**. *Biomedical optics express*, 7(10):3811–3825, 2016. doi: 10.1364/BOE.7.003811.
- [Matrone et al., 2014] Giulia Matrone, Alessandro Stuart Savoia, Giosuè Caliano, and Giovanni Magenes. **The delay multiply and sum beamforming algorithm in ultrasound B-mode medical imaging**. *IEEE transactions on medical imaging*, 34(4):940–949, 2014. doi: 10.1109/TMI.2014.2371235.
- [McDonald, 1986] F Alan McDonald. **Photoacoustic, photothermal, and related techniques: a review**. *Canadian Journal of Physics*, 64(9):1023–1029, 1986. doi: 10.1139/p86-174.
- [Merrick and Hayes, 1976] Edwin B Merrick and Thomas J Hayes. **Continuous, non-invasive measurements of arterial blood oxygen levels**. *Hewlett-packard J*, 28(2):2–9, 1976.

- [Mescheder et al., 2017] Lars Mescheder, Sebastian Nowozin, and Andreas Geiger. **Adversarial variational bayes: Unifying variational autoencoders and generative adversarial networks**. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2391–2400. JMLR. org, 2017.
- [Metropolis and Ulam, 1949] Nicholas Metropolis and Stanislaw Ulam. **The monte carlo method**. *Journal of the American statistical association*, 44(247):335–341, 1949. doi: 10.1080/01621459.1949.10483310.
- [Miller et al., 2018] Dimity Miller, Lachlan Nicholson, Feras Dayoub, and Niko Sünderhauf. **Dropout sampling for robust object detection in open-set conditions**. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–7. IEEE, 2018. doi: 10.1109/ICRA.2018.8460700.
- [Min et al., 2017] Seonwoo Min, Byunghan Lee, and Sungroh Yoon. **Deep learning in bioinformatics**. *Briefings in bioinformatics*, 18(5):851–869, 2017. doi: 10.1093/bib/bbw068.
- [Miscuglio et al., 2018] Mario Miscuglio, Armin Mehrabian, Zibo Hu, Shaimaa I Azzam, Jonathan George, Alexander V Kildishev, Matthew Pelton, and Volker J Sorger. **All-optical nonlinear activation function for photonic neural networks**. *Optical Materials Express*, 8(12):3851–3863, 2018. doi: 10.1364/OME.8.003851.
- [Mohammadi et al., 2019] Leila Mohammadi, Rayyan Manwar, Hamid Behnam, Jahan Tavakkoli, and Mohammad Reza Nasiri Avanaki. **Skull's aberration modeling: towards photoacoustic human brain imaging**. In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108785W. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2510549.
- [Molnar and Nemeth, 2018] Zsolt Molnar and Marton Nemeth. **Monitoring of tissue oxygenation: an everyday clinical challenge**. *Frontiers in medicine*, 4:247, 2018. doi: 10.3389/fmed.2017.00247.
- [Morse et al., 1969] Philip M Morse, K Uno Ingard, and RT Beyer. **Theoretical acoustics**, 1969.
- [Mozaffarzadeh et al., 2017] Moein Mozaffarzadeh, Ali Mahloojifar, Mahdi Orooji, Saba Adabi, and Mohammadreza Nasirivanaki. **Double-stage delay multiply and sum beamforming algorithm: Application to linear-array photoacoustic imaging**. *IEEE Transactions on Biomedical Engineering*, 65(1):31–42, 2017. doi: 10.1109/TBME.2017.2690959.
- [Mozaffarzadeh et al., 2018] Moein Mozaffarzadeh, Vijitha Periyasamy, Manojit Pramanik, and Bahador Makkiabadi. **Efficient nonlinear beamformer based on P'th root of detected signals for linear-array photoacoustic tomography: application to sentinel lymph node imaging**. *Journal of biomedical optics*, 23(12):121604, 2018. doi: 10.1117/1.JBO.23.12.121604.
- [Munkhdalai and Yu, 2017] Tsendsuren Munkhdalai and Hong Yu. **Meta networks**. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2554–2563. JMLR. org, 2017.
- [Naetar and Scherzer, 2014] Wolf Naetar and Otmar Scherzer. **Quantitative photoacoustic tomography with piecewise constant material parameters**. *SIAM Journal on Imaging Sciences*, 7(3):1755–1774, 2014. doi: 10.1137/140959705.
- [Nair et al., 2018] Arun Asokan Nair, Mardava Rajugopal Gubbi, Trac Duy Tran, Austin Reiter, and Muyinatu A Lediju Bell. **A fully convolutional neural network for beamforming ultrasound images**. In *2018 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2018. doi: 10.1109/ULTSYM.2018.8579960.
- [Nair et al., 2019] Tanya Nair, Doina Precup, Douglas L Arnold, and Tal Arbel. **Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation**. *Medical Image Analysis*, page 101557, 2019. doi: 10.1016/j.media.2019.101557.
- [Naser et al., 2018] Mohamed A Naser, Diego RT Sampaio, Nina M Muñoz, Cayla A Wood, Trevor M Mitcham, Wolfgang Stefan, Konstantin V Sokolov, Theo Z Pavan, Rony Avrithcher, and Richard R Bouchard. **Improved photoacoustic-based oxygen saturation estimation with SNR-regularized local fluence correction**. *IEEE transactions on medical imaging*, 38(2):561–571, 2018. doi: 10.1109/TMI.2018.2867602.
- [Needles et al., 2013] Andrew Needles, Andrew Heinmiller, John Sun, Catherine Theodoropoulos, David Bates, Desmond Hirson, Melissa Yin, and F Stuart Foster. **Development and initial application of a fully integrated photoacoustic micro-ultrasound system**. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 60(5):888–897, 2013. doi: 10.1109/TUFFC.2013.2646.
- [Ng, 2011] Andrew Ng. **Sparse autoencoder**, 2011.
- [Niederhauser et al., 2005] Joel J Niederhauser, Michael Jaeger, Robert Lemor, Peter Weber, and Martin Frenz. **Combined ultrasound and optoacoustic system for real-time high-contrast vascular imaging in vivo**. *IEEE transactions on medical imaging*, 24(4):436–440, 2005. doi: 10.1109/TMI.2004.843199.
- [Nuster et al., 2011] Robert Nuster, Hubert Gruen, Bernhard Reiteringer, Peter Burgholzer, Sibylle Gratt, Klaus Passler, and Guenther Paltauf. **Downstream Fabry-Perot interferometer for acoustic wave monitoring in photoacoustic tomography**. *Optics letters*, 36(6):981–983, 2011. doi: 10.1364/OL.36.000981.

- [Nykänen et al., 2017] Olli Nykänen, Aki Pulkkinen, and Tanja Tarvainen. **Quantitative photoacoustic tomography augmented with surface light measurements.** *Biomedical optics express*, 8(10):4380–4395, 2017. doi: 10.1364/BOE.8.004380.
- [Oh et al., 2006] Jung-Taek Oh, Meng-Lin Li, Hao F Zhang, Konstantin Maslov, and Lihong V Wang. **Three-dimensional imaging of skin melanoma in vivo by dual-wavelength photoacoustic microscopy.** *Journal of biomedical optics*, 11(3):034032, 2006. doi: 10.1117/1.2210907.
- [Olefir et al., 2018] Ivan Olefir, Stratis Tzoumas, Hong Yang, and Vasilis Ntziachristos. **A Bayesian Approach to Eigenspectra Optoacoustic Tomography.** *IEEE transactions on medical imaging*, 37(9):2070–2079, 2018. doi: 10.1109/TMI.2018.2815760.
- [Oltulu et al., 2018] Pembe Oltulu, Bilsev Ince, Naile Kokbudak, Sidika Findik, Fahriye Kilinc, et al. **Measurement of epidermis, dermis, and total skin thicknesses from six different body regions with a new ethical histometric technique.** *Turkish Journal of Plastic Surgery*, 26(2):56, 2018. doi: 10.4103/tjps.TJPS\_2\_17.
- [Omar et al., 2019] Murad Omar, Dominik Soliman, and Vasilis Ntziachristos. **Multimodal Optoacoustic Imaging.** In *Image Fusion in Preclinical Applications*, pages 69–99. Springer, 2019. doi: 10.1007/978-3-030-02973-9\_4.
- [Oraevsky et al., 1999] Alexander A Oraevsky, Valeri A Andreev, Alexander A Karabutov, and Rinat O Esenaliev. **Two-dimensional optoacoustic tomography: transducer array and image reconstruction algorithm.** In *Laser-Tissue Interaction X: Photochemical, Photothermal, and Photomechanical*, volume 3601, pages 256–267. International Society for Optics and Photonics, 1999. doi: 10.1117/12.350007.
- [Oraevsky et al., 2001] Alexander A Oraevsky, Alexander A Karabutov, Sergey V Solomatina, Elena V Savateeva, Valeri A Andreev, Zoran Gatalica, Harbans Singh, and R Declan Fleming. **Laser optoacoustic imaging of breast cancer in vivo.** In *Biomedical Optoacoustics II*, volume 4256, pages 6–15. International Society for Optics and Photonics, 2001. doi: 10.1117/12.429300.
- [Padole et al., 2015] Atul Padole, Ranish Deedar Ali Khawaja, Mannudeep K Kalra, and Sarabjeet Singh. **CT radiation dose and iterative reconstruction techniques.** *American Journal of Roentgenology*, 204(4):W384–W392, 2015. doi: 10.2214/AJR.14.13241.
- [Paltauf et al., 2007] Guenther Paltauf, Robert Nuster, Markus Haltmeier, and Peter Burgholzer. **Experimental evaluation of reconstruction algorithms for limited view photoacoustic tomography with line detectors.** *Inverse Problems*, 23(6):S81, 2007. doi: 10.1088/0266-5611/23/6/S07.
- [Paridar et al., 2019] Roya Paridar, Moein Mozaffarzadeh, Vijitha Periyasamy, Manojit Pramanik, Mohammad Mehrmohammadi, and Mahdi Orooji. **Sparsity-Based Beamforming to Enhance Two-Dimensional Linear-Array Photoacoustic Tomography.** *Ultrasonics*, 2019. doi: 10.1016/j.ultras.2019.03.010.
- [Park and Kellis, 2015] Yongjin Park and Manolis Kellis. **Deep learning for regulatory genomics.** *Nature biotechnology*, 33(8):825, 2015. doi: 10.1038/nbt.3313.
- [Paszke et al., 2017] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. **Automatic Differentiation in PyTorch.** In *NIPS Autodiff Workshop*, 2017.
- [Pedregosa et al., 2011] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. **Scikit-learn: Machine learning in Python.** *Journal of machine learning research*, 12(Oct):2825–2830, 2011.
- [Perekatova et al., 2017] VV Perekatova, PV Subochev, M Yu Kirillin, and IV Turchin. **Fluence compensated optoacoustic measurements of blood oxygen saturation in vivo at two optimal wavelengths.** In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100645K. International Society for Optics and Photonics, 2017. doi: 10.1117/12.2250851.
- [Pezzi et al., 2019] Luigia Pezzi, Alfredo Pane, Ferdinanda Annesi, Maria Adele Losso, Alexa Guglielmelli, Cesare Umetton, and Luciano De Sio. **Antimicrobial Effects of Chemically Functionalized and/or Photo-Heated Nanoparticles.** *Materials*, 12(7):1078, 2019. doi: 10.3390/ma12071078.
- [Pfitzner, 1976] J Pfitzner. **Poiseuille and his law.** *Anaesthesia*, 31(2):273–275, 1976. doi: 10.1111/j.1365-2044.1976.tb11804.x.
- [Pickering et al., 1993] John W Pickering, Scott A Prahl, Niek Van Wieringen, Johan F Beek, Henricus JCM Sterenborg, and Martin JC Van Gemert. **Double-integrating-sphere system for measuring the optical properties of tissue.** *Applied optics*, 32(4):399–410, 1993. doi: 10.1364/AO.32.000399.
- [Pierson and Gashler, 2017] Harry A Pierson and Michael S Gashler. **Deep learning in robotics: a review of recent research.** *Advanced Robotics*, 31(16):821–835, 2017. doi: 10.1080/01691864.2017.1365009.



- [Pluta and Maksymilian, 1988] Maksymilian Pluta and Pluta Maksymilian. **Advanced light microscopy**, volume 1. Elsevier Amsterdam, 1988.
- [Poudel et al., 2019] Joemini Poudel, Yang Lou, and Mark A Anastasio. **A survey of computational frameworks for solving the acoustic inverse problem in three-dimensional photoacoustic computed tomography**. *Physics in Medicine & Biology*, 2019. doi: 10.1088/1361-6560/ab2017.
- [Prakash et al., 2019] Jaya Prakash, Subhamoy Mandal, Daniel Razansky, and Vasilis Ntziachristos. **Maximum entropy based non-negative optoacoustic tomographic image reconstruction**. *IEEE Transactions on Biomedical Engineering*, 2019. doi: 10.1109/TBME.2019.2892842.
- [Pulkkinen et al., 2015] Aki Pulkkinen, Benjamin T Cox, Simon R Arridge, Jari P Kaipio, and Tanja Tarvainen. **Quantitative photoacoustic tomography using illuminations from a single direction**. *Journal of biomedical optics*, 20(3):036015, 2015. doi: 10.1117/1.JBO.20.3.036015.
- [Pulkkinen et al., 2016] Aki Pulkkinen, Ben T Cox, Simon R Arridge, Hwan Goh, Jari P Kaipio, and Tanja Tarvainen. **Direct estimation of optical parameters from photoacoustic time series in quantitative photoacoustic tomography**. *IEEE transactions on medical imaging*, 35(11):2497–2508, 2016. doi: 10.1109/TMI.2016.2581211.
- [Pulkkinen et al., 2017] Aki Pulkkinen, Ben T Cox, Simon R Arridge, Jari P Kaipio, and Tanja Tarvainen. **Estimation and uncertainty quantification of optical properties directly from the photoacoustic time series**. In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100643N. International Society for Optics and Photonics, 2017. doi: 10.1117/12.2248837.
- [Qi et al., 2017] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. **Pointnet: Deep learning on point sets for 3d classification and segmentation**. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 652–660, 2017.
- [Qin et al., 2019] Tian Qin, Zongjian Zheng, Rui Zhang, Chengxiang Wang, and Wei Yu.  **$l_0$  gradient minimization for limited-view photoacoustic tomography**. *Physics in Medicine & Biology*, 2019. doi: 10.1088/1361-6560/ab3704.
- [Quiros-Gonzalez et al., 2018] Isabel Quiros-Gonzalez, Michal R Tomaszewski, Sarah J Aitken, Laura Ansel-Bollepalli, Leigh-Ann McDuffus, Michael Gill, Lina Hacker, Joanna Bruncker, and Sarah E Bohndiek. **Optoacoustics delineates murine breast cancer models displaying angiogenesis and vascular mimicry**. *British journal of cancer*, 118(8):1098, 2018. doi: 10.1038/s41416-018-0033-x.
- [Radrich and Ntziachristos, 2016] Karin Radrich and Vasilis Ntziachristos. **Quantitative multi-spectral oxygen saturation measurements independent of tissue optical properties**. *Journal of biophotonics*, 9(1-2):83–99, 2016. doi: 10.1002/jbio.201400092.
- [Rajian et al., 2012] Justin R Rajian, Gandikota Girish, and Xueding Wang. **Photoacoustic tomography to identify inflammatory arthritis**. *Journal of biomedical optics*, 17(9):096013, 2012. doi: 10.1117/1.JBO.17.9.096013.
- [Rajpurkar et al., 2017] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, et al. **CheXnet: Radiologist-level pneumonia detection on chest x-rays with deep learning**. *arXiv preprint arXiv:1711.05225*, 2017.
- [Regensburger et al., 2019] Adrian P Regensburger, Lina M Fonteyne, Jörg Jüngert, Alexandra L Wagner, Teresa Gerhalter, Armin M Nagel, Rafael Heiss, Florian Flenkenthaler, Matthias Qurashi, Markus F Neurath, et al. **Detection of collagens by multispectral optoacoustic tomography as an imaging biomarker for Duchenne muscular dystrophy**. *Nature Medicine*, pages 1–11, 2019. doi: 10.1038/s41591-019-0669-y.
- [Reynolds, 2015] Douglas Reynolds. **Gaussian mixture models**. *Encyclopedia of biometrics*, pages 827–832, 2015. doi: 10.1007/978-1-4899-7488-4\_196.
- [Richardson et al., 2016] Elad Richardson, Matan Sela, and Ron Kimmel. **3D face reconstruction by learning from synthetic data**. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 460–469. IEEE, 2016. doi: 10.1109/3DV.2016.56.
- [Richling et al., 2001] Sabine Richling, Erik Meinköhn, N Kryzhevoi, and Guido Kanschat. **Radiative transfer with finite elements-I. Basic method and tests**. *Astronomy & Astrophysics*, 380(2):776–788, 2001. doi: 10.1051/0004-6361:20011411.
- [Risholm et al., 2013] Petter Risholm, Firdaus Janoos, Isaiah Norton, Alex J Golby, and William M Wells III. **Bayesian characterization of uncertainty in intra-subject non-rigid registration**. *Medical image analysis*, 17(5):538–555, 2013. doi: 10.1016/j.media.2013.03.002.
- [Ronneberger et al., 2015] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. **U-net: Convolutional networks for biomedical image segmentation**. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. doi: 10.1007/978-3-319-24574-4\_28.

- [Rosencwaig, 1975] Allan Rosencwaig. **Photoacoustic spectroscopy. New tool for investigation of solids.** *Analytical Chemistry*, 47(6):592A–604a, 1975. doi: 10.1021/ac60356a015.
- [Rosenthal et al., 2013] Amir Rosenthal, Vasilis Ntziachristos, and Daniel Razansky. **Acoustic inversion in optoacoustic tomography: A review.** *Current medical imaging reviews*, 9(4):318–336, 2013. doi: 10.2174/15734056113096660006.
- [Ross et al., 2018] Tobias Ross, David Zimmerer, Anant Vemuri, Fabian Isensee, Manuel Wiesenfarth, Sebastian Bodenstedt, Fabian Both, Philip Kessler, Martin Wagner, Beat Müller, et al. **Exploiting the potential of unlabeled endoscopic video data with self-supervised learning.** *International journal of computer assisted radiology and surgery*, 13(6): 925–933, 2018. doi: 10.1007/s11548-018-1772-0.
- [Rostami et al., 2019] Seyyed Reza Miri Rostami, Moein Mozaffarzadeh, Ali Hariri, Jesse V Jokerst, and Mohsen Ghaffari-Miab. **OpenACC GPU implementation of double-stage delay-multiply-and-sum algorithm: toward enhanced real-time linear-array photoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108785C. International Society for Optics and Photonics, 2019. doi: 10.1117/12.2511115.
- [Sallab et al., 2017] Ahmad EL Sallab, Mohammed Abdou, Etienne Perot, and Senthil Yogamani. **Deep reinforcement learning framework for autonomous driving.** *Electronic Imaging*, 2017(19):70–76, 2017. doi: 10.2352/ISSN.2470-1173.2017.19.AVM-023.
- [Schwab et al., 2019a] Johannes Schwab, Stephan Antholzer, and Markus Haltmeier. **Learned backprojection for sparse and limited view photoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 1087837. International Society for Optics and Photonics, 2019a. doi: 10.1117/12.2508438.
- [Schwab et al., 2019b] Johannes Schwab, Stephan Antholzer, Robert Nuster, Günther Paltauf, and Markus Haltmeier. **Deep Learning of truncated singular values for limited view photoacoustic tomography.** In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 1087836. International Society for Optics and Photonics, 2019b. doi: 10.1117/12.2508418.
- [Sethuraman et al., 2007] Shriram Sethuraman, Salavat R Aglyamov, James H Amirian, Richard W Smalling, and Stanislav Y Emelianov. **Intravascular photoacoustic imaging using an IVUS imaging catheter.** *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 54(5):978–986, 2007. doi: 10.1109/TUFFC.2007.343.
- [Severinghaus, 1979] John W Severinghaus. **Simple, accurate equations for human blood O<sub>2</sub> dissociation computations.** *Journal of Applied Physiology*, 46(3):599–602, 1979. doi: 10.1152/jappl.1979.46.3.599.
- [Severinghaus and Honda, 1987] John W Severinghaus and Yoshiyuki Honda. **History of blood gas analysis. VII. Pulse oximetry.** *Journal of clinical monitoring*, 3(2):135–138, 1987. doi: 10.1007/BF00858362.
- [Shan et al., 2019] Hongming Shan, Ge Wang, and Yang Yang. **Simultaneous reconstruction of the initial pressure and sound speed in photoacoustic tomography using a deep-learning approach.** *arXiv preprint arXiv:1907.09951*, 2019. doi: 10.1117/12.2529984.
- [Shao et al., 2011] Peng Shao, Ben Cox, and Roger J Zemp. **Estimating optical absorption, scattering, and Grueneisen distributions with multiple-illumination photoacoustic tomography.** *Applied optics*, 50(19):3145–3154, 2011. doi: 10.1364/AO.50.003145.
- [Singh et al., 2017] Mithun Kuniyil Ajith Singh, Michael Jaeger, Martin Frenz, and Wiendelt Steenbergen. **Photoacoustic reflection artifact reduction using photoacoustic-guided focused ultrasound: comparison between plane-wave and element-by-element synthetic backpropagation approach.** *Biomedical optics express*, 8(4):2245–2260, 2017. doi: 10.1364/BOE.8.002245.
- [Sivaramakrishnan et al., 2007] Mathangi Sivaramakrishnan, Konstantin Maslov, Hao F Zhang, George Stoica, and Lihong V Wang. **Limitations of quantitative photoacoustic measurements of blood oxygenation in small vessels.** *Physics in Medicine & Biology*, 52(5):1349, 2007. doi: 10.1088/0031-9155/52/5/010.
- [Skaar, 2006] Johannes Skaar. **Fresnel equations and the refractive index of active media.** *Physical Review E*, 73(2): 026605, 2006. doi: 10.1103/PhysRevE.73.026605.
- [Smith and Gal, 2018] Lewis Smith and Yarin Gal. **Understanding measures of uncertainty for adversarial example detection.** *arXiv preprint arXiv:1803.08533*, 2018.
- [Solberg et al., 2007] Ole Vegard Solberg, Frank Lindseth, Hans Torp, Richard E Blake, and Toril A Nagelhus Hernes. **Freehand 3D ultrasound reconstruction algorithms—a review.** *Ultrasound in medicine & biology*, 33(7):991–1009, 2007. doi: 10.1016/j.ultrasmedbio.2007.02.015.
- [Song et al., 2008] Kwang Hyun Song, Erich W Stein, Julie A Margenthaler, and Lihong V Wang. **Noninvasive photoacoustic identification of sentinel lymph nodes containing methylene blue in vivo in a rat model.** *Journal of biomedical optics*, 13(5):054033, 2008. doi: 10.1117/1.2976427.

- [Song et al., 2014] Ningning Song, Carole Deumié, and Anabela Da Silva. **Considering sources and detectors distributions for quantitative photoacoustic tomography**. *Biomedical optics express*, 5(11):3960–3974, 2014. doi: 10.1364/BOE.5.003960.
- [Srivastava et al., 2014] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. **Dropout: a simple way to prevent neural networks from overfitting**. *The journal of machine learning research*, 15(1):1929–1958, 2014.
- [Su et al., 2010] Jimmy L Su, Andrei B Karpouk, Bo Wang, and Stanislav Y Emelianov. **Photoacoustic imaging of clinical metal needles in tissue**. *Journal of biomedical optics*, 15(2):021309, 2010. doi: 10.1117/1.3368686.
- [Su et al., 2019] Ting Su, Shi Zhang, Dayu Li, Zhiwei Lv, and Shengwei Dong. **Low Complex Beam-Space Beamformer for Medical Ultrasound Imaging Based on Delay Multiply and Sum Method**. *Journal of Medical Imaging and Health Informatics*, 9(8):1630–1638, 2019. doi: 10.1166/jmih.2019.2776.
- [Sun et al., 2014] Yi Sun, Xiaogang Wang, and Xiaoou Tang. **Deep learning face representation from predicting 10,000 classes**. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1891–1898, 2014. doi: 10.1109/CVPR.2014.244.
- [Sykes et al., 2009] JR Sykes, DS Brettell, DR Magee, and DI Thwaites. **Investigation of uncertainties in image registration of cone beam CT to CT on an image-guided radiotherapy system**. *Physics in Medicine & Biology*, 54(24):7263, 2009. doi: 10.1088/0031-9155/54/24/002.
- [Tak and Ye, 2014] Sungho Tak and Jong Chul Ye. **Statistical analysis of fNIRS data: a comprehensive review**. *Neuroimage*, 85:72–91, 2014. doi: 10.1016/j.neuroimage.2013.06.016.
- [Taruttis and Ntziachristos, 2015] Adrian Taruttis and Vasilis Ntziachristos. **Advances in real-time multispectral photoacoustic imaging and its applications**. *Nature photonics*, 9(4):219, 2015. doi: 10.1038/nphoton.2015.29.
- [Tarvainen et al., 2013] Tanja Tarvainen, Aki Pulkkinen, Ben T Cox, Jari P Kaipio, and Simon R Arridge. **Bayesian image reconstruction in quantitative photoacoustic tomography**. *IEEE transactions on medical imaging*, 32(12):2287–2298, 2013. doi: 10.1109/TMI.2013.2280281.
- [Tarvainen et al., 2017] Tanja Tarvainen, Aki Pulkkinen, Ben T Cox, and Simon R Arridge. **Utilising the radiative transfer equation in quantitative photoacoustic tomography**. In *Photons Plus Ultrasound: Imaging and Sensing 2017*, volume 10064, page 100643E. International Society for Optics and Photonics, 2017. doi: 10.1117/12.2249310.
- [Thomenius, 1996] Kai E Thomenius. **Evolution of ultrasound beamformers**. In *1996 IEEE Ultrasonics Symposium. Proceedings*, volume 2, pages 1615–1622. IEEE, 1996. doi: 10.1109/ULTSYM.1996.584398.
- [Tick et al., 2016] Jenni Tick, Aki Pulkkinen, and Tanja Tarvainen. **Image reconstruction with uncertainty quantification in photoacoustic tomography**. *The Journal of the Acoustical Society of America*, 139(4):1951–1961, 2016. doi: 10.1364/MATH.2017.MTu1C.1.
- [Tick et al., 2018] Jenni Tick, Aki Pulkkinen, Felix Lucka, Robert Ellwood, Ben T Cox, Jari P Kaipio, Simon R Arridge, and Tanja Tarvainen. **Three dimensional photoacoustic tomography in Bayesian framework**. *The Journal of the Acoustical Society of America*, 144(4):2061–2071, 2018. doi: 10.1121/1.5057109.
- [Timmins and Wall, 1977] PA Timmins and JC Wall. **Bone water**. *Calcified tissue research*, 23(1):1–5, 1977. doi: 10.1007/BF02012759.
- [Toshev and Szegedy, 2014] Alexander Toshev and Christian Szegedy. **DeepPose: Human pose estimation via deep neural networks**. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1653–1660, 2014. doi: 10.1109/CVPR.2014.214.
- [Treeby and Cox, 2010] Bradley E Treeby and Benjamin T Cox. **k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields**. *Journal of biomedical optics*, 15(2):021314, 2010. doi: 10.1117/1.3360308.
- [Treeby et al., 2010] Bradley E Treeby, Edward Z Zhang, and Benjamin T Cox. **Photoacoustic tomography in absorbing acoustic media using time reversal**. *Inverse Problems*, 26(11):115003, 2010. doi: 10.1088/0266-5611/26/11/115003.
- [Tremel et al., 2016] Michael Tremel, José Arjona-Medina, Thomas Unterthiner, Rupesh Durgesh, Felix Friedmann, Peter Schuberth, Andreas Mayr, Martin Heusel, Markus Hofmarcher, Michael Widrich, et al. **Speeding up semantic segmentation for autonomous driving**. In *MLITS, NIPS Workshop*, volume 2, page 7, 2016.
- [Tzoumas et al., 2013] Stratis Tzoumas, Nikolaos C Deliolanis, Stefan Morscher, and Vasilis Ntziachristos. **Unmixing molecular agents from absorbing tissue in multispectral photoacoustic tomography**. *IEEE transactions on medical imaging*, 33(1):48–60, 2013. doi: 10.1109/TMI.2013.2279994.
- [Tzoumas et al., 2016] Stratis Tzoumas, Antonio Nunes, Ivan Olefir, Stefan Stangl, Panagiotis Symvoulidis, Sarah Glasl, Christine Bayer, Gabriele Multhoff, and Vasilis Ntziachristos. **Eigenspectra photoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues**. *Nature communications*, 7:12121, 2016. doi: 10.1038/ncomms12121.

- [Ulrich et al., 2019] Leonie Ulrich, Linda Ahnen, Hidayet Günhan Akarçay, Salvador Sánchez Majos, Michael Jaeger, Kai Gerrit Held, Martin Wolf, and Martin Frenz. **Spectral correction for handheld optoacoustic imaging by means of near-infrared optical tomography in reflection mode.** *Journal of biophotonics*, 12(1):e201800112, 2019. doi: 10.1002/jbio.201800112.
- [van Veen et al., 2004] Robert LP van Veen, HJCM Sterenborg, A Pifferi, A Torricelli, and R Cubeddu. **Determination of VIS-NIR absorption coefficients of mammalian fat, with time-and spatially resolved diffuse reflectance and transmission spectroscopy.** In *Biomedical Topical Meeting*, page SF4. Optical Society of America, 2004. doi: 10.1364/BIO.2004.SF4.
- [VanderPlas, 2016] Jake VanderPlas. **Python data science handbook: essential tools for working with data.** " O'Reilly Media, Inc.", 2016.
- [Vial et al., 2018] Alanna Vial, David Stirling, Matthew Field, Montserrat Ros, Christian Ritz, Martin Carolan, Lois Holloway, and Alexis A Miller. **The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review.** *Translational Cancer Research*, 7(3):803–816, 2018. doi: 10.21037/ter.2018.05.02.
- [Vogt et al., 2019] William C Vogt, Xuewen Zhou, Rudy Andriani, Keith A Wear, T Joshua Pfefer, and Brian S Garra. **Photoacoustic oximetry imaging performance evaluation using dynamic blood flow phantoms with tunable oxygen saturation.** *Biomedical optics express*, 10(2):449–464, 2019. doi: 10.1364/BOE.10.000449.
- [Wagstaff et al., 2001] Kiri Wagstaff, Claire Cardie, Seth Rogers, Stefan Schrödl, et al. **Constrained k-means clustering with background knowledge.** In *Icml*, volume 1, pages 577–584, 2001.
- [Waibel et al., 2018] Dominik Waibel, Janek Gröhl, Fabian Isensee, Thomas Kirchner, Klaus Maier-Hein, and Lena Maier-Hein. **Reconstruction of initial pressure from limited view photoacoustic images using deep learning.** In *Photons Plus Ultrasound: Imaging and Sensing 2018*, volume 10494, page 104942S. International Society for Optics and Photonics, 2018. doi: 10.1117/12.2288353.
- [Waldner et al., 2016] Maximilian J Waldner, Ferdinand Knieling, Cornelia Egger, Stefan Morscher, Jing Claussen, Marcel Vetter, Christian Kielisch, Sarah Fischer, Lukas Pfeifer, Alexander Hagel, et al. **Multispectral optoacoustic tomography in Crohn's disease: noninvasive imaging of disease activity.** *Gastroenterology*, 151(2):238–240, 2016. doi: 10.1053/j.gastro.2016.05.047.
- [Wang and Zhou, 2017] Chao Wang and Tie Zhou. **On iterative algorithms for quantitative photoacoustic tomography in the radiative transport regime.** *Inverse Problems*, 33(11):115006, 2017. doi: 10.1088/1361-6420/aa89c5.
- [Wang et al., 2018a] Ge Wang, Jong Chu Ye, Klaus Mueller, and Jeffrey A Fessler. **Image reconstruction is a new frontier of machine learning.** *IEEE transactions on medical imaging*, 37(6):1289–1296, 2018a. doi: 10.1109/TMI.2018.2833635.
- [Wang, 2012] Lihong V Wang. **Photoacoustic tomography: Ultrasonically breaking through the optical diffusion limit.** In *Biomedical Optics*, pages BSu1A–1. Optical Society of America, 2012. doi: 10.1364/BIOMED.2012.BSu1A.1.
- [Wang and Wu, 2012] Lihong V Wang and Hsin-i Wu. **Biomedical optics: principles and imaging.** John Wiley & Sons, 2012. doi: 10.1002/9780470177013.
- [Wang and Yao, 2016] Lihong V Wang and Junjie Yao. **A practical guide to photoacoustic tomography in the life sciences.** *Nature methods*, 13(8):627, 2016. doi: 10.1038/nmeth.3925.
- [Wang et al., 2016] Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang. **Accelerating magnetic resonance imaging via deep learning.** In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 514–517. IEEE, 2016. doi: 10.1109/ISBI.2016.7493320.
- [Wang et al., 2003a] Xueding Wang, Yongjiang Pang, Geng Ku, George Stoica, and Lihong V Wang. **Three-dimensional laser-induced photoacoustic tomography of mouse brain with the skin and skull intact.** *Optics letters*, 28(19):1739–1741, 2003a. doi: 10.1364/OL.28.001739.
- [Wang et al., 2003b] Xueding Wang, Yongjiang Pang, Geng Ku, Xueyi Xie, George Stoica, and Lihong V Wang. **Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain.** *Nature biotechnology*, 21(7):803, 2003b. doi: 10.1038/nbt839.
- [Wang et al., 2007] Xueding Wang, David L Chamberland, and David A Jamadar. **Noninvasive photoacoustic tomography of human peripheral joints toward diagnosis of inflammatory arthritis.** *Optics letters*, 32(20):3002–3004, 2007. doi: 10.1364/OL.32.003002.
- [Wang et al., 2018b] Yihan Wang, Tong Lu, Songhe Zhang, Shaoze Song, Bingyuan Wang, Jiao Li, Huijuan Zhao, and Feng Gao. **A calibrated iterative reconstruction for quantitative photoacoustic tomography using multi-angle light-sheet illuminations.** In *Photons Plus Ultrasound: Imaging and Sensing 2018*, volume 10494, page 104943C. International Society for Optics and Photonics, 2018b. doi: 10.1117/12.2290215.

- [Wang, 2009] Yingxu Wang. **On abstract intelligence: Toward a unifying theory of natural, artificial, machinable, and computational intelligence.** *International Journal of Software Science and Computational Intelligence (IJSSCI)*, 1(1): 1–17, 2009. doi: 10.4018/978-1-4666-0261-8.ch002.
- [Weight et al., 2006] Ryan M Weight, John A Viator, Paul S Dale, Charles W Caldwell, and Allison E Lisle. **Photoacoustic detection of metastatic melanoma cells in the human circulatory system.** *Optics letters*, 31(20):2998–3000, 2006. doi: 10.1364/OL.31.002998.
- [Welvaert and Rosseel, 2013] Marijke Welvaert and Yves Rosseel. **On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data.** *PLoS one*, 8(11):e77089, 2013. doi: 10.1371/journal.pone.0077089.
- [Wilkinson, 2004] Leland Wilkinson. **Classification and regression trees.** *Systat*, 11:35–56, 2004.
- [Wirkert et al., 2016] Sebastian J Wirkert, Hannes Kenngott, Benjamin Mayer, Patrick Mietkowski, Martin Wagner, Peter Sauer, Neil T Clancy, Daniel S Elson, and Lena Maier-Hein. **Robust near real-time estimation of physiological parameters from megapixel multispectral images with inverse Monte Carlo and random forest regression.** *International journal of computer assisted radiology and surgery*, 11(6):909–917, 2016. doi: 10.1007/s11548-016-1376-5.
- [Wold et al., 1987] Svante Wold, Kim Esbensen, and Paul Geladi. **Principal component analysis.** *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987. doi: 10.1016/0169-7439(87)80084-9.
- [Wu et al., 2014] Dan Wu, Lin Huang, Max S Jiang, and Huabei Jiang. **Contrast agents for photoacoustic and thermoacoustic imaging: a review.** *International journal of molecular sciences*, 15(12):23616–23639, 2014. doi: 10.3390/ijms151223616.
- [Xia and Wang, 2013] Jun Xia and Lihong V Wang. **Small-animal whole-body photoacoustic tomography: a review.** *IEEE Transactions on Biomedical Engineering*, 61(5):1380–1389, 2013. doi: 10.1109/TBME.2013.2283507.
- [Xia et al., 2012] Jun Xia, Muhammad R Chatni, Konstantin I Maslov, Zijian Guo, Kun Wang, Mark A Anastasio, and Lihong V Wang. **Whole-body ring-shaped confocal photoacoustic computed tomography of small animals in vivo.** *Journal of biomedical optics*, 17(5):050506, 2012. doi: 10.1117/1.JBO.17.5.050506.
- [Xu and Wang, 2005] Minghua Xu and Lihong V Wang. **Universal back-projection algorithm for photoacoustic computed tomography.** *Physical Review E*, 71(1):016706, 2005. doi: 10.1103/PhysRevE.71.016706.
- [Yang et al., 2019] Changchun Yang, Hengrong Lan, Hongtao Zhong, and Fei Gao. **Quantitative photoacoustic blood oxygenation imaging using deep residual and recurrent neural network.** In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 741–744. IEEE, 2019. doi: 10.1109/ISBI.2019.8759438.
- [Yang and Chung, 2018] Guang Yang and Kevin C. Chung. **Ulnar Artery to Superficial Arch Bypass with a Vein Graft.** In *Operative Techniques: Hand and Wrist Surgery*, pages 732–737. Elsevier Health Sciences, 2018. doi: 10.1016/B978-0-323-40191-3.00081-0.
- [Yang et al., 2009] Xinmai Yang, Erich W Stein, S Ashkenazi, and Lihong V Wang. **Nanoparticles for photoacoustic imaging.** *Wiley interdisciplinary reviews: nanomedicine and nanobiotechnology*, 1(4):360–368, 2009. doi: 10.1002/wnan.42.
- [Yao and Wang, 2014] Junjie Yao and Lihong V Wang. **Photoacoustic brain imaging: from microscopic to macroscopic scales.** *Neurophotonics*, 1(1):011003, 2014. doi: 10.1117/1.NPh.1.1.011003.
- [Yoon et al., 2018] Yeo Hun Yoon, Shujaat Khan, Jaeyoung Huh, and Jong Chul Ye. **Efficient b-mode ultrasound image reconstruction from sub-sampled rf data using deep learning.** *IEEE transactions on medical imaging*, 38(2):325–336, 2018. doi: 10.1109/TMI.2018.2864821.
- [Yu et al., 2018] Leiming Yu, Fanny Nina-Paravecino, David R Kaeli, and Qianqian Fang. **Scalable and massively parallel Monte Carlo photon transport simulations for heterogeneous computing platforms.** *Journal of biomedical optics*, 23(1):010504, 2018. doi: 10.1117/1.JBO.23.1.010504.
- [Yuan and Jiang, 2006] Zhen Yuan and Huabei Jiang. **Quantitative photoacoustic tomography: Recovery of optical absorption coefficient maps of heterogeneous media.** *Applied physics letters*, 88(23):231101, 2006. doi: 10.1117/12.697079.
- [Yuan and Jiang, 2009] Zhen Yuan and Huabei Jiang. **Quantitative photoacoustic tomography.** *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1900):3043–3054, 2009. doi: 10.1098/rsta.2009.0083.
- [Zackrisson et al., 2014] Sophia Zackrisson, SMWY Van De Ven, and SS Gambhir. **Light in and sound out: emerging translational strategies for photoacoustic imaging.** *Cancer research*, 74(4):979–1004, 2014. doi: 10.1158/0008-5472.CAN-13-2387.
- [Zemp, 2010] Roger J Zemp. **Quantitative photoacoustic tomography with multiple optical sources.** *Applied optics*, 49(18):3566–3572, 2010. doi: 10.1364/BIOMED.2010.BWG4.

- [Zeng et al., 2004] Yaguang Zeng, Da Xing, Yi Wang, Bangzhen Yin, and Qun Chen. **Photoacoustic and ultrasonic coimage with a linear transducer array.** *Optics letters*, 29(15):1760–1762, 2004. doi: 10.1364/OL.29.001760.
- [Zhang et al., 2010a] Chi Zhang, Konstantin Maslov, and Lihong V Wang. **Subwavelength-resolution label-free photoacoustic microscopy of optical absorption in vivo.** *Optics letters*, 35(19):3195–3197, 2010a. doi: 10.1364/OL.35.003195.
- [Zhang et al., 2008] Edward Zhang, Jan Laufer, and Paul Beard. **Backward-mode multiwavelength photoacoustic scanner using a planar Fabry-Perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues.** *Applied optics*, 47(4):561–577, 2008. doi: 10.1364/AO.47.000561.
- [Zhang et al., 2009] EZ Zhang, JG Laufer, RB Pedley, and PC Beard. **In vivo high-resolution 3D photoacoustic imaging of superficial vascular anatomy.** *Physics in Medicine & Biology*, 54(4):1035, 2009. doi: 10.1088/0031-9155/54/4/014.
- [Zhang and Dong, 2019] Haimiao Zhang and Bin Dong. **A Review on Deep Learning in Medical Image Reconstruction.** *arXiv preprint arXiv:1906.10643*, 2019. doi: 10.1007/s40305-019-00287-4.
- [Zhang et al., 2018a] Jiayao Zhang, Bin Chen, Meng Zhou, Hengrong Lan, and Fei Gao. **Photoacoustic image classification and segmentation of breast cancer: A feasibility study.** *IEEE Access*, 7:5457–5466, 2018a. doi: 10.1109/ACCESS.2018.2888910.
- [Zhang et al., 2018b] Shanshan Zhang, Jian Yang, and Bernt Schiele. **Occluded pedestrian detection through guided attention in CNNs.** In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6995–7003, 2018b. doi: 10.1109/CVPR.2018.00731.
- [Zhang et al., 2010b] Yu Zhang, Xin Cai, Sung-Wook Choi, Chulhong Kim, Lihong V Wang, and Younan Xia. **Chronic label-free volumetric photoacoustic microscopy of melanoma cells in three-dimensional porous scaffolds.** *Biomaterials*, 31(33):8651–8658, 2010b. doi: 10.1016/j.biomaterials.2010.07.089.
- [Zhao et al., 2017] Lingyi Zhao, Meng Yang, Yuxin Jiang, and Changhui Li. **Optical fluence compensation for handheld photoacoustic probe: An in vivo human study case.** *Journal of Innovative Optical Health Sciences*, 10(04):1740002, 2017. doi: 10.1142/S1793545817400028.
- [Zhu and Zabararas, 2018] Yinhao Zhu and Nicholas Zabararas. **Bayesian deep convolutional encoder–decoder networks for surrogate modeling and uncertainty quantification.** *Journal of Computational Physics*, 366:415–447, 2018. doi: 10.1016/j.jcp.2018.04.018.
- [Zhu et al., 2018] Yunhao Zhu, Guan Xu, Jie Yuan, Janggun Jo, Girish Gandikota, Hakan Demirci, Toshitaka Agano, Naoto Sato, Yusuke Shigeta, and Xueding Wang. **Light emitting diodes based photoacoustic imaging and potential clinical applications.** *Scientific reports*, 8(1):9885, 2018. doi: 10.1038/s41598-018-28131-4.
- [Zou et al., 2019] James Zou, Mikael Huss, Abubakar Abid, Pejman Mohammadi, Ali Torkamani, and Amalio Telenti. **A primer on deep learning in genomics.** *Nature genetics*, 51(1):12–18, 2019. doi: 10.1038/s41588-018-0295-5.

# 7

## Own contributions

The purpose of this chapter is to give an overview of my contributions and to differentiate these from whole team efforts. Every section already discloses the team members that have been working on the presented results.

### 7.1 Own share in data acquisition and data analysis

Throughout my Ph.D. work, I was part of a multi-disciplinary team of scientists in the department of Computer Assisted Medical Interventions (CAMI) headed by Prof. Dr. Lena Maier-Hein. She was also the primary supervisor for my thesis, and she supervised all experiments. Throughout the entire time of my thesis, I was closely collaborating with Lena's photoacoustics team that was comprised of Thomas Kirchner and Melanie Schellenberg and various bachelor and master students (Angelika Laha, Dominik Waibel, Niklas Holzwarth, and Kris Dreher). I was also collaborating with scientists and Ph.D. students from the multispectral imaging team, especially Tim Adler, Leonardo Ayala, and Anant Vemuri.

Most of the results of this thesis were reproduced from the original work in the published papers on new data sets and sometimes with slight changes or improvements regarding the methodology.

#### Own share in data acquisition

I simulated all *in silico* data sets. The design parameters (for example, the acoustic and optical parameter ranges) for these data sets were usually proposed by me and then discussed in the team. I was part of the design and planning of all experiments for *in vivo* data acquisitions. Still, the porcine brain measurements were lead by Thomas Kirchner, and they were typically acquired by our collaborating surgeons and medical students from the university clinic (E. Santos, M. Herrera, and A. Hernández-Aguilera). The forearm study was designed by the entire photoacoustics team and also served the purpose of testing the newly arrived MSOT acuity echo

system. It was conducted with N=15 healthy human volunteers on a single day, and the leading operators of the imaging system were Thomas Kirchner and Niklas Holzwarth. The phantom data was provided by Lina Hacker and Sarah Bohndiek from Cambridge University, and I had no part in the data acquisition.

### Own share in data analysis

I developed and implemented the methods for all work presented in this thesis with regular input from my colleagues. Here it should be noted that the CE-qPAI study was conducted in close collaboration, where both Thomas Kirchner and myself were serving the role of the leading investigator. A complete list of people that participated in each project is disclosed at the start of each result section.

As it is common in larger research groups, many ideas on methods, experiments, or implementation details were exchanged in the entire CAMI team and the photoacoustics group in particular. However, for all the results presented in this Ph.D. thesis, the implementation of the methods and the data analysis was always performed by myself.

## 7.2 Own publications

This section will list all papers that I was a part of and contributed to during my Ph.D. work. It is subdivided into First authorships and co-authorships, and each of these categories is further subdivided into publications in scientific journals and other contributions, such as arXiv preprints, abstracts, or conference proceedings.

### First authorships - peer reviewed journal publications

**Thomas Kirchner, Janek Gröhl,** and Lena Maier-Hein. Context encoding enables machine learning-based quantitative photoacoustics. *Journal of biomedical optics*, 23(5):056008, 2018.

**Janek Gröhl,** Thomas Kirchner, Tim Adler, and Lena Maier-Hein. Confidence estimation for machine learning-based quantitative photoacoustics. *Journal of Imaging*, 4(12):147, 2018.

### First authorships - other

**Janek Gröhl,** Thomas Kirchner, Tim Adler, and Lena Maier-Hein. Estimation of blood oxygenation with learned spectral decoloring for quantitative photoacoustic imaging (lsd-qpai). *arXiv preprint arXiv:1902.05839*, 2019.

**Janek Gröhl,** Thomas Kirchner, and Lena Maier-Hein. Confidence estimation for quantitative photoacoustic imaging. In *Photons Plus Ultrasound: Imaging and Sensing 2018, volume 10494, page 104941C*. International Society for Optics and Photonics, 2018.

**Janek Gröhl.** International Photoacoustic Standardisation Consortium (IPASC): recommendations for standardized data exchange in photoacoustic imaging (conference presentation). In *Photons Plus Ultrasound: Imaging and Sensing 2019, volume 10878, page 108781S*. International Society for Optics and Photonics, 2019.



**Janek Gröhl**, Thomas Kirchner, and Lena Maier-Hein. Quantitative Photoakustische Tomografie durch lokale Kontextkodierung. In *Bildverarbeitung für die Medizin 2017* Springer, 2017.

### Co-Authorships - peer reviewed journal publications

Martin Klemm, Thomas Kirchner, **Janek Gröhl**, Dominique Cheray, Marco Nolden, Alexander Seitel, Harald Hoppe, Lena Maier-Hein, and Alfred M Franz. Mitk-openigtlink for combining open-source toolkits in real-time computer-assisted interventions. *International journal of computer assisted radiology and surgery*, 12(3):351–361, 2017.

Thomas Kirchner, Franz Sattler, **Janek Gröhl**, and Lena Maier-Hein. Signed real-time delay multiply and sum beamforming for multispectral photoacoustic imaging. *Journal of Imaging*, 4(10):121, 2018.

Tim J Adler, Lynton Ardizzone, Anant Vemuri, Leonardo Ayala, **Janek Gröhl**, Thomas Kirchner, Sebastian Wirkert, Jakob Kruse, Carsten Rother, Ullrich Köthe, et al. Uncertainty-aware performance assessment of optical imaging modalities with invertible neural networks. *International journal of computer assisted radiology and surgery*, pages 1–11, 2019.

Thomas Kirchner, **Janek Gröhl**, Mildred A Herrera, Tim Adler, Adrián Hernández-Aguilera, Edgar Santos, and Lena Maier-Hein. Photoacoustics can image spreading depolarization deep in gyrencephalic brain. *Scientific reports*, 9(1):8661, 2019.

### Co-Authorships - other

Lena Maier-Hein, Tobias Ross, **Janek Gröhl**, Ben Glocker, Sebastian Bodenstedt, Christian Stock, Eric Heim, Michael Götz, S Wirkert, Hannes Kenngott, et al. Crowd-algorithm collaboration for large-scale endoscopic image annotation with confidence. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 616–623. Springer, 2016.

Dominik Waibel, **Janek Gröhl**, Fabian Isensee, Thomas Kirchner, Klaus Maier-Hein, and Lena Maier-Hein. Reconstruction of initial pressure from limited view photoacoustic images using deep learning. In *Photons Plus Ultrasound: Imaging and Sensing 2018*, volume 10494, page 104942S. International Society for Optics and Photonics, 2018.

Thomas Kirchner, **Janek Gröhl**, Franz Sattler, Moritz S Bischoff, Angelika Laha, Marco Nolden, and Lena Maier-Hein. An open-source software platform for translational photoacoustic research and its application to motion-corrected blood oxygenation estimation. *arXiv preprint arXiv:1901.09781*, 2019.

Thomas Kirchner, **Janek Gröhl**, Niklas Holzwarth, Mildred A Herrera, Adrián Hernández-Aguilera, Edgar Santos, and Lena Maier-Hein. Photoacoustic monitoring of blood oxygenation during neurosurgical interventions. In *Photons Plus Ultrasound: Imaging and Sensing 2019*, volume 10878, page 108780C. International Society for Optics and Photonics, 2019.

Tim J Adler, Lynton Ardizzone, Leonardo Ayala, **Janek Gröhl**, Anant Vemuri, Sebastian J Wirkert, Beat P Müller-Stich, Carsten Rother, Ullrich Köthe, and Lena Maier-Hein. Uncertainty handling in intra-operative multispectral imaging with invertible neural networks. *MIDL Proceedings of Machine Learning Research*, 2019.

Sarah Bohndiek, Joanna Brunker, **Janek Gröhl**, Lina Hacker, James Joseph, William C Vogt, Paolo Armanetti, Hisham Assi, Jeffrey C Bamber, Paul C Beard, et al. International Photoacoustic Standardisation Consortium (IPASC): overview (conference presentation). In *Photons Plus Ultrasound: Imaging and Sensing 2019, volume 10878, page 108781N*. International Society for Optics and Photonics, 2019.

Leonardo A Ayala, Sebastian J Wirkert, **Janek Gröhl**, Mildred A Herrera, Adrian Hernandez-Aguilera, Anant Vemuri, Edgar Santos, and Lena Maier-Hein. Live Monitoring of Haemodynamic Changes with Multispectral Image Analysis. In *OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging, pages 38-46*. Springer, Cham, 2019.

# Janek Matthias Gröhl

## CURRICULUM VITAE

Date of birth 08.10.1991  
Place of birth Hamburg  
Citizenship German  
Email jgroehl@outlook.de



### Education

Since 11/2016 **Ph.D. Student – German Cancer Research Center**  
Ph.D. thesis title: Data-driven quantitative photoacoustic imaging.

09/2014 - 10/2016 **M.Sc. Medical Informatics – Heidelberg University & Heilbronn University (Grade: 1,1)**  
Master thesis title: Machine learning-based quantitative photoacoustic tomography.

09/2011 - 08/2014 **B.Sc. Medical Informatics – Heidelberg University & Heilbronn University (Grade: 1,3)**  
Bachelor thesis title: An analysis of the potential of smart glasses in dermatology using the example of the malignant melanoma.

07/2002 - 06/2011 **A-levels: Lloyd Gymnasium Bremerhaven (Grade: 1,2)**  
Project title: Development of artificial limbs and body parts from 1550 until today (computer game)

### Research Experience

Since 11/2016 **Ph.D. Student at the German Cancer Research Center**

09/2015 - 10/2016 **Research Assistant at the German Cancer Research Center**

09/2013 - 08/2015 **Research Assistant at the Gecko Institute of Heilbronn University**

### Honors and Awards

---

07/2018	Best Poster Award: Annual Ph.D. Retreat of the German Cancer Research Center
04/2017	1 <sup>st</sup> prize conhIT Nachwuchspreis 2017 for the master thesis “Machine learning based quantitative photoacoustic tomography” including a presentation at the conhIT 2017 conference
04/2017	Thomas-Gessmann special award for the master thesis “Machine learning based quantitative photoacoustic tomography”
10/2015	Thomas-Gessmann-Award for the outstanding bachelor thesis „Potentialanalyse von SmartGlasses in der Dermatologie am Beispiel des malignen Melanoms“

### Scholarships

---

11/2016 - 10/2019	PhD Stipend of the Helmholtz Graduate School for Cancer Research at the German Cancer Research Center
09/2015 - 08/2016	National Stipend of Heilbronn University and Dürr Dental AG
09/2012 - 08/2014	National Stipend of Heilbronn University and the Thomas-Gessmann-Stiftung

### Selected publications

---

**T Kirchner, J Gröhl**, and L Maier-Hein. Context encoding enables machine learning-based quantitative photoacoustics. *Journal of biomedical optics*, 23(5):056008, 2018.

**J Gröhl**, T Kirchner, T Adler, and L Maier-Hein. Confidence estimation for machine learning-based quantitative photoacoustics. *Journal of Imaging*, 4(12):147, 2018.

**J Gröhl**, T Kirchner, T Adler, and L Maier-Hein. Estimation of blood oxygenation with learned spectral decoloring for quantitative photoacoustic imaging (LSD-qPAI). *arXiv preprint arXiv:1902.05839*, 2019.

L Maier-Hein, T Ross, **J Gröhl**, B Glocker, S Bodenstedt, C Stock, E Heim, M Götz, S Wirkert, H Kenngott, S Speidel, and K Maier-Hein. Crowd-algorithm collaboration for large-scale endoscopic image annotation with confidence. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 616–623, 2016.

D Waibel, **J Gröhl**, F Isensee, T Kirchner, K Maier-Hein, and L Maier-Hein. Reconstruction of initial pressure from limited view photoacoustic images using deep learning. In *Photons Plus Ultrasound: Imaging and Sensing 2018*, volume 10494, page 104942S, 2018.

#### **Invited Talks**

---

- |         |   |
|---------|---|
| 04/2020 | Invited talk on “Machine learning-based inference of functional tissue properties from multispectral photoacoustic imaging” at the <i>2020 OSA Biophotonics Congress: Biomedical Optics (BioMed 2020)</i> |
| 04/2019 | Seminar talk on “Machine learning-based quantitative photoacoustic imaging” within the <i>Biomedical Photonics Seminar</i> at Bern University   |
| 02/2018 | Talk on “Quantitative photoacoustic imaging” within the <i>Pizza and Talk</i> series organized by the Graduate Office of the German Cancer Research Center  |

#### **Teaching Experience**

---

- |                   |  |
|-------------------|--|
| Since 11/2016     | (Co-)supervision of bachelor’s and master’s theses |
| 06/2012 - 08/2013 | Tutor for Mathematics at Heilbronn University      |

#### **Other activities**

---

- |                       |  |
|-----------------------|--|
| Since 04/2018         | Website design for the <a href="http://www.ipasc.science">www.ipasc.science</a> webpage  |
| Since 04/2018         | Working group lead in the International Photoacoustic Standardisation Consortium (IPASC) |
| 11/2016 - 07/2018     | Organization of a biweekly seminar talk at the DKFZ                                      |
| 02 / 2012 – 08 / 2015 | Voluntary work for the student organization „miandyou“ of the informatics faculty        |
| 01/2005 – 08/2011     | Voluntary work for a protestant youth organization Dionys-Lehe in Bremerhaven            |

Mannheim, 12th of January 2020





# Acknowledgements

The past three years and three months of conducting my Ph.D. research have been an extraordinary experience for me. I have learned a lot about countless different things. My ways through the Ph.D. was paved by the people that advised me, that I was interacting with regularly, that supported me, and that helped during various stages.

First and foremost, I would like to thank my Ph.D. supervisor Prof. Dr. **Lena Maier-Hein** for her guidance during my Ph.D. journey. She was always ready to listen to my questions, concerns, or ideas and never disappointed in giving great advice on any of these. Her dedication and enthusiasm for science was a great source of inspiration and motivation for me.

Many thanks to the members of my *thesis advisory committee*: Dr. **Ralf Floca**, Prof. Dr. **Carsten Rother**, and PD Dr. **Moritz Bischoff**, who were involved at various stages of the projects and offered great methodological and general advice for the project. Thanks to PD Dr. **Klaus Maier-Hein** for participating in these committee meetings. At a crucial point in time, he pointed out that sometimes it is more valuable to pursue one promising aspect instead of branching out into many different and new topics.

I would like to especially thank **Thomas Kirchner**, whom I first interacted with as the supervisor of my master's studies and who later became a great colleague and friend. Without continuous collaboration with him, none of the projects and experiments would have been conducted as they were. *Thank you.*

Many thanks to my office, the **Phantastisch produktiven Positivitätskanonen**. Our efforts for *maximum office personalization* culminated in a workplace that was a hot-spot for creativity and orthogonal thinking. Thank you all for making coming to the office every day a splendid experience.

Many thanks to everyone who was part of the photoacoustics team at some point or another. Thanks to **Dominik Waibel** and **Angelika Laha** for being great master students. Thanks to **Niklas Holzwarth** and **Kris Dreher** for being exceptional bachelor students and later research assistants and master students. Thanks to **Franz Sattler** for his continuous contributions to the group as a research assistant and thanks to **Melanie Schellenberg** and **Tim Adler**, who became great colleagues in no time.

Many thanks also to **all members of the CAMIC group(s)**. Thank you for allowing me to explore

many ideas with you, for your helpful hints, continuous support, the opportunities for countless coffee breaks, the great discussions, and the excellent retreats. "The team spirit is strong with these ones".

Thanks to **Lina Hacker** for your great help, the invaluable contribution of experimental data, and for our weekly and most-efficient conversations.

Great thanks to **Sinan Onogur, Clemens Hentschke, and André Klein**. You have become great friends and the best board game crew I could ever imagine. I enjoyed the numerous discussions and talks we had on anything and everything.

Many thanks to **my family**, who have supported me throughout the whole endeavor. You are the best family one could ever wish for. Thank you for everything.

Last but not least, I owe a great thanks to **Michelle Lipski**. Thank you for enduring my absent-mindedness when being deep in thought every so often and my grumpiness when things did not go the way I planned. I could not imagine someone better by my side.



# Eidesstattliche Versicherung

## Statutory Declaration

1. Bei der eingereichten Dissertation zu dem Thema **Data-driven Quantitative Photoacoustic Imaging** handelt es sich um meine eigenständig erbrachte Leistung.  
*I herewith formally declare that I have written the submitted dissertation **Data-driven Quantitative Photoacoustic Imaging** independently.*
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.  
*I did not use any third party support except for the quoted literature and other sources mentioned in the text. Content from other work, either literally or in content, has been declared as such.*
3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.  
*The thesis has not been submitted to any examination body in this, or similar, form.*
4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.  
*I confirm the correctness of the aforementioned declarations.*
5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt. Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.  
*I am aware of the legal consequences of this declaration. To the best of my knowledge I have told the pure truth and not concealed anything.*

Heidelberg, 12.01.2020

---

Janek Gröhl