Dissertation

submitted to the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Kris Kristoffer Dreher

born in: St. Wendel
Oral examination: 12.11.2025

Data-driven synthesis of photoacoustic images

Referees: Prof. Dr. Jürgen Hesser
Prof. Dr. Lena Maier-Hein

I know what things are.
Sir Andrew Maxwell Dwyer KBE

Abstract

Methodological progress in the age of Artificial Intelligence (AI) is increasingly driven by large-scale, high-quality datasets. Medical imaging is no exception. As the amount of imaging data grows exponentially, the development and validation of data-driven algorithms critically depend on realistic, well-annotated datasets. Photoacoustic imaging (PAI), a hybrid modality combining optical contrast with ultrasound resolution, holds great promise for non-invasive, functional biomedical imaging with the potential for quantitative measurement of blood oxygen saturation. Prospective clinical applications range from cardiovascular imaging to early diagnosis and treatment monitoring in oncology. However, as a comparatively young modality, PAI faces a fundamental challenge: the absence of sufficiently large, annotated datasets for the development and validation of quantitative methods.

This thesis addresses the data scarcity problem in quantitative PAI through a comprehensive pipeline for the generation and validation of synthetic spectral imaging data. First, the thesis reviews the current role of deep learning in PAI, identifying the generation of realistic simulated data and domain adaptation as key techniques to address data scarcity. Consequently, SIMPA is introduced, the open-source simulation and image processing for photonics and acoustics framework enabling large-scale and reproducible PA image simulation. To overcome the remaining domain gap between simulated and experimental data, a conditional invertible neural network-based method is proposed for unsupervised domain transfer. For empirical validation of data-driven PA methods, anatomically realistic tissue-mimicking phantoms are fabricated and characterized, enabling the creation of a comprehensive dataset, including experimental and simulated multispectral images. This resource allows, for the first time, rigorous benchmarking of quantitative methods such as oximetry.

This thesis contributes not only technically through novel methods and validation strategies, but also promotes reproducibility and open science by making all software tools and datasets publicly available. In doing so, it lays the foundation for systematic and transparent development and validation of data-driven methods in PAI and supports the broader translation of PAI into preclinical and clinical applications.

ZUSAMMENFASSUNG

Methodologischer Fortschritt im Zeitalter der Künstlichen Intelligenz (KI) wird zunehmend durch groß angelegte, qualitativ hochwertige Datensätze vorangetrieben. Die medizinische Bildgebung bildet hierbei keine Ausnahme. Mit dem exponentiellen Wachstum bildgebender Daten hängt die Entwicklung und Validierung datengetriebener Algorithmen maßgeblich von realistischen, gut annotierten Datensätzen ab. Die photoakustische (PA) Bildgebung, eine hybride Modalität, die optischen Kontrast mit Ultraschallauflösung kombiniert, verspricht eine nicht-invasive, funktionelle Bildgebung mit dem Potenzial zur quantitativen Messung der Blutsauerstoffsättigung. Zukünftige klinische Anwendungen reichen von der kardiovaskulären Bildgebung über die Frühdiagnostik bis hin zur Therapieüberwachung in der Onkologie. Als vergleichsweise junge Modalität steht die PA Bildgebung jedoch vor einer grundlegenden Herausforderung: dem Fehlen ausreichend großer, annotierter Datensätze für die Entwicklung und Validierung quantitativer Verfahren.

Diese Arbeit widmet sich dem Problem der Datenknappheit in der quantitativen PA Bildgebung durch die Entwicklung einer umfassenden Pipeline zur Generierung und Validierung synthetischer spektraler Bilddaten. Zunächst wird der aktuelle Stand der Anwendung von Deep Learning in der PA Bildgebung systematisch aufgearbeitet, wobei die realistische Simulation von Daten sowie Domänenantransfer als Schlüsseltechniken zur Bewältigung der Datenknappheit identifiziert werden. In der Folge wird SIMPA (Simulation and Image Processing for Photonics and Acoustics) vorgestellt, ein Open-Source Framework zur umfangreichen und reproduzierbaren Simulation von PA Bildern. Zur Überbrückung der verbleibenden Domänenlücke zwischen simulierten und experimentellen Daten wird eine auf konditionalen invertierbaren neuronalen Netzwerken basierende Methode für den unüberwachten Domänen Transfer eingeführt. Für die empirische Validierung datengetriebener PA-Methoden werden anatomisch realistische, gewebeähnliche Phantome gefertigt und charakterisiert, wodurch der Aufbau eines Open-Source Datensatzes ermöglicht wird, bestehend aus experimentellen und simulierten multispektralen Bilddaten. Dieser Datensatz erlaubt erstmals eine systematische Validierung quantitativer Verfahren wie etwa der Oximetrie.

Diese Arbeit leistet nicht nur technische Beiträge durch neue Methoden und Validierungsstrategien, sondern fördert durch die Veröffentlichung aller Softwaretools und Datensätze auch Reproduzierbarkeit und Open Science. Damit legt sie das Fundament für eine transparente Entwicklung und Validierung datengetriebener Methoden in der PA Bildgebung und unterstützt deren breitere Überführung in präklinische und klinische Anwendungen.

Contents

Lists

List of acronyms

List of figures

Part I

Foundations and context

Introduction 1

1.1 Motivation

Photoacoustic imaging (PAI) is an emerging biomedical imaging modality that uniquely combines the physics of light and sound. In biomedical PAI, tissue is illuminated with short pulses of light, and the absorbed optical energy is converted into heat, causing rapid thermoelastic expansion and the emission of ultrasonic pressure waves [1]. PAI bridges a long-standing gap in imaging: it achieves the rich molecular contrast of optical imaging while maintaining the spatial resolution of ultrasound, allowing clinicians and researchers to peer several centimeters into tissue with high resolution [2]. This combination positions PAI as a powerful tool for *in vivo* structural, functional, and molecular imaging beyond the shallow limits of traditional optical microscopy.

The phenomenon underlying PAI is the photoacoustic (PA) effect, first observed by Alexander Graham Bell in 1880. In a seminal experiment, Bell discovered that illuminating materials with rapidly modulated sunlight could generate audible sound waves [3]. This effect, initially a scientific curiosity, essentially laid dormant for decades. It was not until nearly a century later, in 1976, that Rosencwaig and Gersho developed the field of PA spectroscopy, providing a detailed theoretical understanding of PA signal generation in gases and solids [4]. Even then, early PA experiments were limited by available technology. Only the development of pulsed lasers enabled short, intense light pulses at selectable wavelengths, which was "key to making PAI a viable imaging tool" [5]. In practical terms, a modern PAI system typically consists of a pulsed laser for illumination and a set of ultrasound detectors for capturing the generated acoustic waves, often arranged in a tomographic configuration. This straightforward hardware setup, essentially coupling a laser to an ultrasound machine, means PAI can be relatively easily integrated into clinical settings.

One particularly compelling application of PAI is in functional imaging, where it can noninvasively assess physiological parameters such as blood oxygenation [6]. The ability to map blood oxygenation is crucial because oxygenation is a key indicator of tissue function and pathology, making it highly relevant for cancer diagnosis and treatment monitoring, particularly in breast and prostate cancer. Hypoxia, a condition of low oxygen levels in tissue, is a well-known hallmark of many diseases, including cancer. Tumors often have regions of inadequate blood supply and oxygenation; an estimated 50–60%

of advanced solid tumors develop hypoxic zones that can drive malignant progression and therapy resistance [7]. Beyond oncology, PAI holds promise across a broad range of clinical applications [8]. It shows strong potential in vascular imaging, including the assessment of peripheral artery disease, microvascular complications in diabetes, and cerebral hemodynamics. Dermatology and inflammation imaging, such as for arthritis or skin lesions, are emerging areas of interest. Additionally, PAI is being explored for real-time guidance during minimally invasive procedures and surgery, where its label-free nature and depth penetration offer advantages over conventional optical methods. These diverse applications underscore the modality's potential impact on diagnostics, treatment planning, and intraoperative decision-making. In Figure 1.1, a brief overview of the core principal of PAI and potential applications is are illustrated. However, several challenges must still be addressed to enable successful clinical translation of PAI. This thesis contributes toward overcoming these obstacles, with the goal of supporting its mid-term integration into clinical practice.

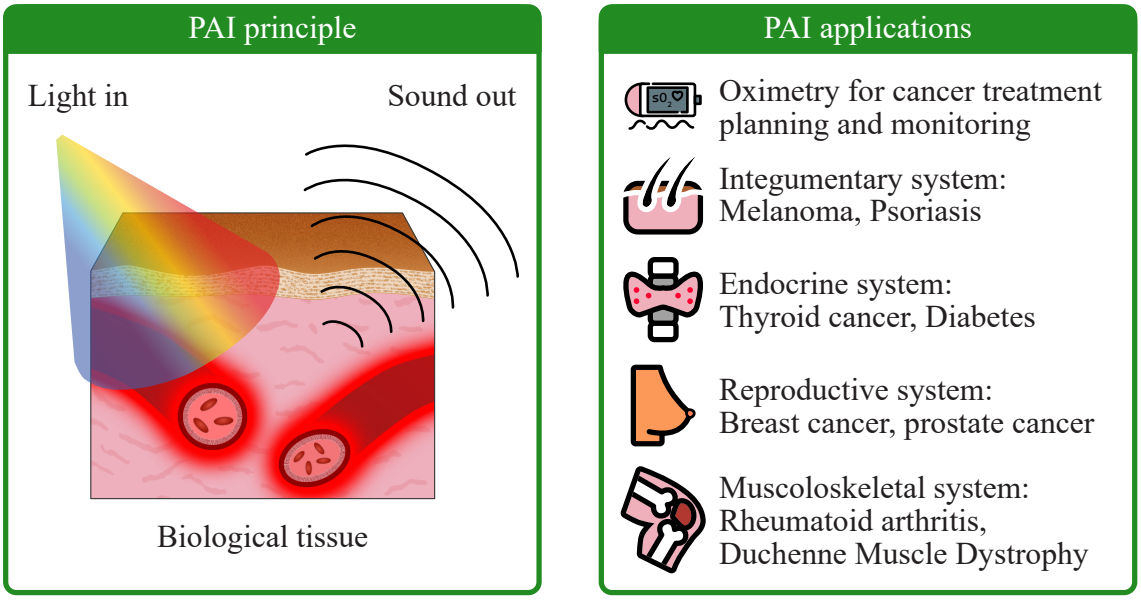


Figure 1.1: There are numerous potential applications for photoacoustic imaging (PAI). Left: Illustration of the multispectral light in-sound out principle of biomedical PAI. Right: Potential applications and previous imaging results of anatomic and physiological systems of the human body. Applications and categorizations from Park et al. (2024) [8].

1.2 Current challenges

PAI has demonstrated substantial potential in small-scale studies for applications such as oximetry, cancer diagnosis, and treatment monitoring [8, 9, 10]. Yet, despite this promise, PAI has not achieved widespread clinical adoption.
In a comprehensive assessment by the International Photoacoustic Standardisation Consortium (IPASC), Assi et al. [11] outline several barriers to the clinical translation of PAI. Among these, key limitations include the lack of reliable quantification of optical and functional tissue properties and the difficulty in enabling data-driven PAI analysis.
Just like in the broader field of machine learning, progress in PAI is increasingly driven by access to large, high-quality datasets. The same review emphasizes: "Data-driven methods bring a wide range of potential benefits for PAI, however, the field is currently fragmented and limited by a scarcity of openly available high-quality data." [11] This is also reflected in the PAI community's efforts to explore data-driven approaches (cf. Figure 1.2).
Nonetheless, PAI remains a relatively young modality. Unlike more established imaging techniques such as computed tomography (CT) or magnetic resonance imaging (MRI), it lacks large-scale clinical datasets. More critically, PAI faces a unique challenge: the absence of reference measurements for many of the properties it seeks to quantify, such as optical absorption and blood oxygen saturation. This issue directly motivates the central research focus of this thesis:

Central research challenge of this thesis

There is a fundamental lack of high-quality, openly available data with reference annotations for the development and validation of data-driven photoacoustic imaging methods.

In this thesis, reference annotations can be understood as pixel-wise labels for optical or functional properties of tissue such as blood oxygen saturation. Given the complexity of this overarching limitation, the thesis addresses it through three interrelated key subchallenges C.1 – C.3, introduced in the following section.

Key subchallenge 1 (C.1)

Need for an efficient and customizable PA image simulation pipeline.

The development of data-driven methods in PAI is increasingly constrained by the limited availability of annotated datasets, particularly in the context of emerging generative and foundation models, which require large volumes of diverse, high-quality data. A key bottleneck is the absence of reliable reference annotations needed to train and evaluate robust algorithms for quantitative PAI.
In this setting, PA image simulation has emerged as an essential tool. Unlike experimental

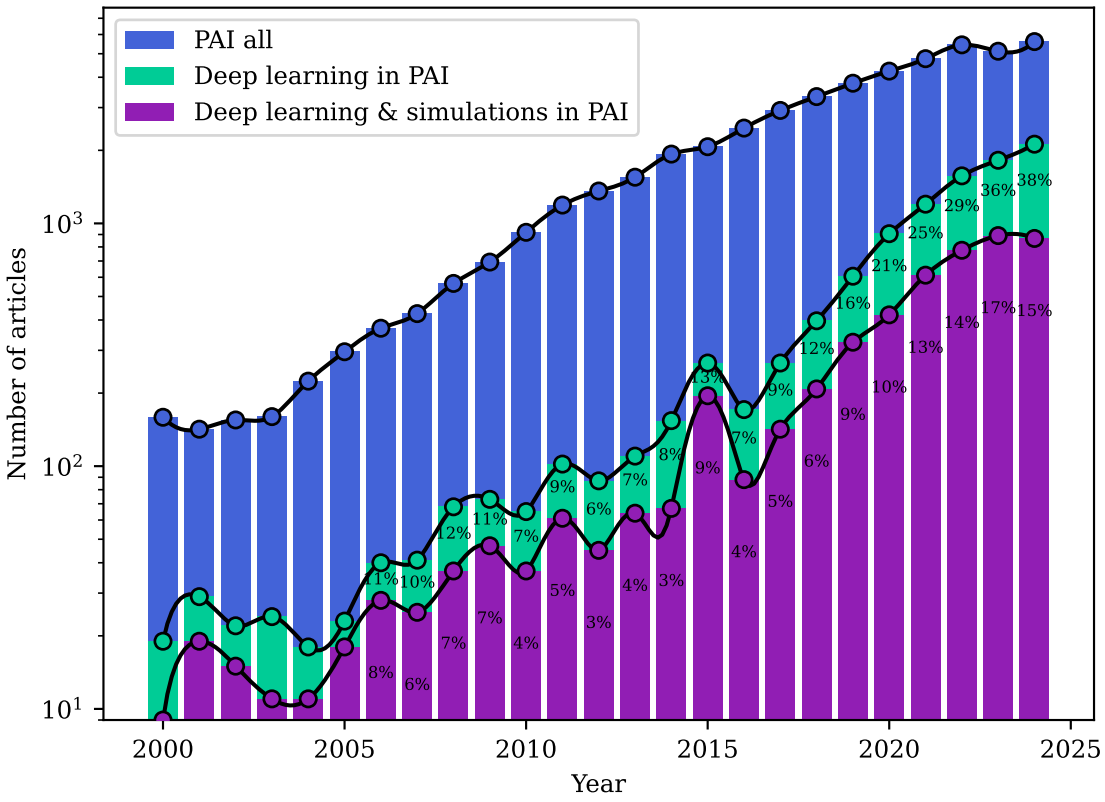


Figure 1.2: The number of publications in the PAI field is exponentially growing. The publications in the entire field are shown in blue, the fraction of learning-related papers within PAI in green and the learning-related papers using some form of simulated data in purple. Results were obtained by querying Google Scholar with the following search terms: PAI all: (”photoacoustic” OR ”optoacoustic”) AND (”imaging” or ”tomography”), Deep Learning in PAI: PAI all AND (”machine learning” OR ”deep learning” OR ”neural network” OR ”neural networks”), and Deep Learning & Simulations in PAI: Deep Learning in PAI AND (”simulation” OR ”simulations”).

data, simulated datasets inherently include ground truth values for all underlying optical, acoustic, and functional properties, making them uniquely valuable for supervised learning and benchmarking. As noted by Park et al. [8], “Improved simulation pipelines, including both optical and acoustic sides of the PAI problem, could add value in future algorithm development.”

To be effective, however, simulation pipelines must satisfy several criteria. They must be efficient enough to enable the generation of large-scale datasets within reasonable computational limits. They must be customizable, allowing adaptation to diverse experimental conditions and application-specific requirements. And critically, they must be realistic, producing synthetic data that enables trained models to generalize effectively to *in vivo* scenarios.

Key subchallenge 2 (C.2)

Remaining domain gap between synthetic and real images.

Despite advances in simulation techniques, there is a broad consensus in the PAI community that a significant domain gap remains between synthetic and experimental data [12] (also referred to as simulated-to-real (Sim2Real) domain gap). Even highly realistic simulations are typically distinguishable from real images by domain experts. More critically, data-driven methods trained solely on simulated data have consistently shown poor generalization to *in vivo* applications [13].
To address this, a variety of domain adaptation techniques have been proposed, most commonly based on cycle-consistent generative adversarial networks (CycleGANs) [14]. While these methods have shown potential, they are prone to known issues such as mode collapse and training instability, limiting their reliability in practice [15]. Moreover, unlike conventional computer vision tasks that rely on three-channel red, green, and blue (RGB) images, PAI involves rich spectral information across multiple wavelengths, necessitating domain transfer methods that explicitly preserve spectral integrity.

Key subchallenge 3 (C.3)

Lack of comprehensive quantitative validation strategy for data-driven PA methods.

Validating the faithfulness of synthetic PA data and quantitative PAI methods, particularly those driven by data and trained on such synthetic data, remains a significant open problem in the field. The development of such methods in the absence of validated reference data presents a fundamental dilemma: when trying to benchmark data generation methods without suitable datasets for benchmarking, progress becomes circular, a classic chicken-and-egg problem. To break this cycle, researchers often turn to tissue-mimicking phantoms that offer controlled, measurable properties for validation purposes [16]. Yet this subfield faces its own challenges: faithfully replicating the anatomical and optical complexity of biological tissues, ensuring structural and spectral stability, and achieving reproducibility across fabrication batches. Importantly, while various phantoms have been proposed for system calibration or hardware testing, there has been no established framework or consensus on how to design phantoms specifically for the development and validation of data-driven photoacoustic methods [8].

Naturally, there are additional crucial challenges to be tackled towards clinical translation, such as uncertainty estimation, detecting and mitigating biases and confounders, generalizability to other devices, cohorts, etc. This thesis, however, addresses the key subchallenges C.1 - C.3 with the contributions outlined in Figure 1.3.

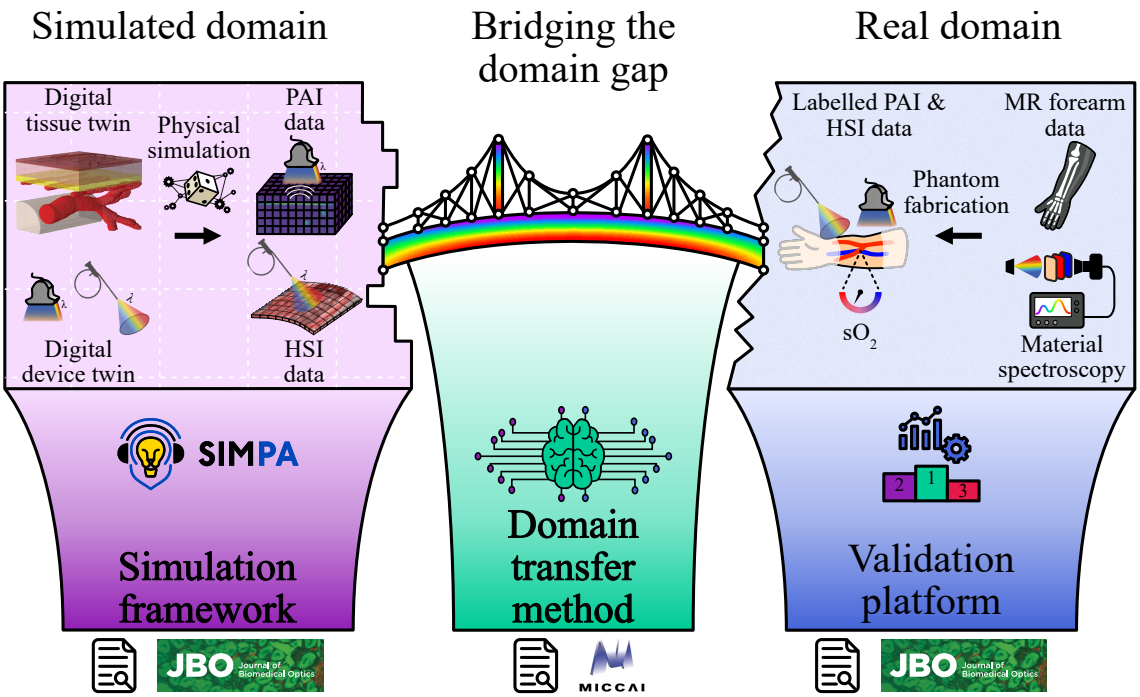


Figure 1.3: Concept and contributions of this thesis. The contributions of this thesis can be summarized by three main pillars, each addressing a different part of the central research challenge of this thesis: the lack of annotated reference data. The first pillar represents a simulation framework that combines digital twins for device and tissue with computational models to enable realistic simulations of photoacoustic imaging (PAI) and hyperspectral imaging (HSI) data. In order to bridge the domain gap between simulated and real data, a multispectral domain transfer method was developed based on conditional invertible neural networks. The third pillar represents a validation platform utilizing experimental data that was acquired after fabricating anthropomorphic phantoms (based on magnetic resonance (MR) data) with known, measured material properties.

This thesis is presented in a cumulative form, meaning that publications carried out during the thesis project make up the main body of this thesis. While the first publication featured in this thesis (cf. 4.1) lays out the general landscape for data-driven methods at the time, the other four publications can be specifically assigned to address one of the key subchallenges C.1-C.3. The second featured publication (cf. 4.2) addresses C.1, the third publication (cf. 4.3) addresses C.2, and both the fourth and fifth publications (cf. 4.4 and 4.5) address C.3.

1.3 Research questions

The research questions (RQs) of this thesis have been specifically designed to match the key subchallenges C.1 - C.3 summarized in the previous section. After identifying the overarching challenge of this thesis, the fundamental lack of reference data in PAI, this thesis tackles this challenge and its subchallenges with a sophisticated pipeline for synthetic data generation, analysis, and quality assessment for validation. The three RQs guiding this thesis address their corresponding numbered key subchallenge:

Research question 1 (RQ1)

How to facilitate multiphysics simulations for the generation of PA images?

Born from an extremely heterogeneous software landscape for the computational modeling of PA images, this RQ involves investigating how various existing tools can be chained together into a unified pipeline that is both user-friendly for the PA community and highly customizable for domain experts. Additionally challenging is how to generate digital twins for both PA devices and tissue, serving as additional input for such a pipeline. These digital twins should ideally be usable across different simulators.

Research question 2 (RQ2)

How can unsupervised domain transfer close the domain gap between simulated and experimental PA images while preserving spectral fidelity?

Generalizability of data-driven methods is an essential requirement for potential applications in patient care. The remaining domain gap between simulated and experimental PA data strongly limits this generalizability. In this thesis, this challenge is reframed as a distribution matching problem. Each image can be viewed as a multidimensional sample from an underlying image distribution. Since simulated and experimental data are drawn from distinct distributions, domain transfer aims to bridge this gap by transforming samples from the simulated distribution into the experimental distribution.
With the ever-increasing momentum of the artificial intelligence (AI) field in general and generative models in particular, domain transfer methods have been shown to perform exceptionally well in these problems [17, 18]. These methods, however, have been mainly applied to natural images acquired with RGB cameras. Since one of the main advantages of PAI lies in its potential for multispectral imaging, such domain transfer methods need to take the spectral dimension into consideration. This RQ examines how they can be adapted to function as a domain transfer method that is designed to be capable of handling spectral data.

Research question 3 (RQ3)

How can anthropomorphic phantoms be designed to provide reliable validation for data-driven methods in PAI?

Tissue-mimicking phantoms that are fabricated from materials whose optical properties have been carefully measured provide the only possibility in PAI where experimental data can be annotated with reference values. These references are crucial for a rigorous validation of any method that estimates one or multiple of their properties. In the designing process of such phantoms, many questions need to be answered, such that they can provide the validation they are fabricated for. Essentially, it is a two-step process where two complementary questions can be asked: First, where is the middle ground between the structural complexity of the phantoms and large-scale reproducibility? Their complexity should be high enough to mimic human tissue but low enough such that reproducible fabrication in a sufficient amount is still possible to satisfy the needs of data-driven methods. Second, once the phantoms are fabricated, how can data be extracted or generated and appropriately annotated such that data-driven methods can be validated using this data?

Background 2

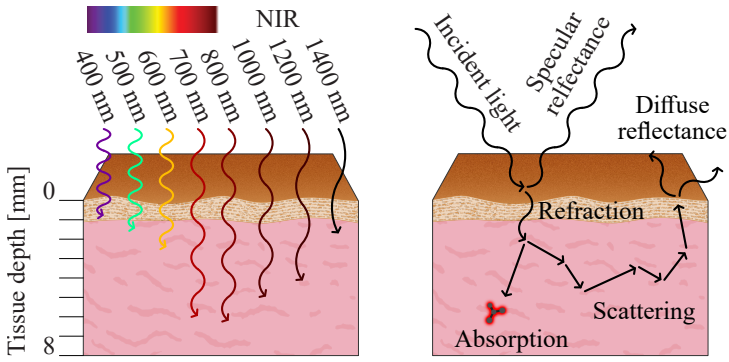
A central challenge in PAI is the lack of annotated reference data, which significantly limits the development and validation of data-driven methods. This thesis addresses that gap through the synthesis of PA images using computational modeling and machine learning. To understand the approaches taken in this work, it is essential to first review the underlying principles of PAI and used machine learning methods. This chapter is therefore structured into three main parts: section 2.1 introduces the fundamental physics of the photoacoustic effect, including interactions of light with biological tissue. Section 2.2 outlines the computational methods used to model PAI systems and simulate corresponding data. Finally, section 2.3 provides an overview of data-driven approaches, with a particular focus on deep learning (DL) techniques relevant to the methods employed in this thesis. This chapter conceptually follows my Master's thesis [19].

2.1 Photoacoustic imaging

This section details the underlying physics needed to understand the image formation process of PAI. It starts by introducing the interactions of light with biological tissue, then goes on to explain how light propagation within tissue is generally described. It finishes by outlining how the PA effect generates acoustic waves and how these waves propagate through tissue. This chapter follows the general concepts of the textbooks *Biophotonics: Concepts to Applications* by Gerd Keiser [20] and *Biomedical Optics: Principles and Imaging* by Wang and Wu [1].

Light propagation

For most biomedical PAI applications, the relevant wavelengths lie within the 600–1300 nm range [21], which corresponds to the end of the visible and near-infrared (NIR) spectrum. This is due to the increased penetration depth of light in this so-called NIR-window as illustrated in Figure 2.1a. Within this range, the dominant interactions between light and tissue include reflection, refraction, absorption, and scattering. These processes are illustrated schematically in Figure 2.1.



(a) Penetration depth **(b)** Interaction phenomena

Figure 2.1: Light-tissue interactions. a) Wavelength-dependent penetration depth of light in biological tissue in the visible and near-infrared (NIR) regions. These are generic values from Clement et al. [22] and Pezzi et al. [23]. The exact penetration depth depends on the light source and tissue properties. b) Light-tissue interaction phenomena in the relevant wavelength range.

Reflection and refraction

When light encounters the boundary between two media, such as air with refractive index n_a and the biological tissue with refractive index n_t, a portion of the incident light is reflected, while the remainder is transmitted into the tissue. This process is known as specular reflection, and the corresponding reflected fraction is referred to as the specular reflectance R_{sp}. The value of R_{sp} depends on the angle of incidence θ_a relative to the surface normal and is described by the Fresnel equation [20]:

$$R_{\mathrm{sp}}(\cos\theta_a) = \frac{1}{2}\left(\frac{n_{\mathrm{a}}\cos\theta_t - n_{\mathrm{t}}\cos\theta_a}{n_{\mathrm{a}}\cos\theta_t + n_{\mathrm{t}}\cos\theta_a}\right)^2 + \frac{1}{2}\left(\frac{n_{\mathrm{a}}\cos\theta_a - n_{\mathrm{t}}\cos\theta_t}{n_{\mathrm{a}}\cos\theta_a + n_{\mathrm{t}}\cos\theta_t}\right)^2. \tag{2.1}$$

Here, θ_t is the angle of the transmitted portion to the surface normal within the tissue. This refracted angle can be determined using Snell's law [24]

$$n_{\mathrm{t}}\sin\,\theta_t = n_{\mathrm{a}} sin\,\theta_a. \tag{2.2}$$

The transmitted fraction, $T = 1 - R_{\mathrm{sp}}$, is known as the transmittance. However, the Fresnel equations only apply when the angle of incidence satisfies $0 \leq \theta < \theta_{\mathrm{c}}$ where θ_{c} is the critical angle. For angles larger than θ_{c}, total internal reflection occurs:

$$R_{\mathrm{sp}}(\cos\theta) = 1 \quad \text{for} \quad \theta_{\mathrm{c}} \leq \theta \leq \frac{\pi}{2}. \tag{2.3}$$

Once inside the tissue, light can undergo multiple scattering and absorption events. The portion of light that re-emerges after experiencing at least one scattering event is termed diffuse reflectance R_{diff}. R_{diff} is also often used for reflections on uneven surfaces without actually entering the tissue. For this thesis, however, R_{diff} is used for their combined effect as most measurement devices can not distinguish these two.

Absorption

Absorption occurs when photons transfer their energy to molecules, promoting them to an excited state. These molecules may subsequently return to their ground state via radiative relaxation, such as fluorescence, or non-radiative relaxation mechanisms, such as the PA effect [25]. The extent of absorption is quantified by the absorption coefficient μ_a, which defines the probability of photon absorption per unit path length. According to the Lambert–Beer law, the attenuation of light intensity I over a path length x in an absorbing medium is given by:

$$I(x) = I_0 e^{-\mu_\mathrm{a} x}, \tag{2.4}$$

where I_0 represents the initial intensity. The reciprocal of the absorption coefficient, μ_a^{-1}, corresponds to the mean free absorption path. Because biological tissue typically contains a mixture of different absorbers, referred to as chromophores, the total absorption coefficient at a spatial position $\vec{r}$ is computed as the sum of contributions from each chromophore at the given wavelength λ:

$$\mu_\mathrm{a}(\vec{r}, \lambda) = \sum_{\mathrm{k}=1}^{\mathrm{K}} c_\mathrm{k}(\vec{r}) \alpha_\mathrm{k}(\lambda). \tag{2.5}$$

Here, c_k denotes the concentration of the k-th chromophore and α_k its absorption spectrum. Figure 2.5 illustrates typical absorption spectra for major chromophores found in biological tissue.

Scattering

Scattering arises from interactions between light and microscopic structures within biological tissue that exhibit a refractive index mismatch relative to the surrounding medium. An exact theoretical treatment of light scattering by spherical particles is provided by Mie theory, which reduces to Rayleigh scattering when the wavelength significantly exceeds the particle size [1]. However, for the purposes of this work, it is sufficient to describe scattering in analogy to absorption, using a bulk scattering coefficient μ_s. Given the highly scattering nature of biological tissue, individual scattering events do not need to be modeled explicitly.

The scattering coefficient μ_s represents the probability of photon scattering per unit path length, with its reciprocal defining the mean free scattering path. In a purely scattering medium of infinitesimal thickness, the attenuation of light intensity can again be described by the Lambert–Beer law:

$$I(x) = I_0 e^{-\mu_\mathrm{s} x}. \tag{2.6}$$

The angular distribution of scattering is characterized by the anisotropy factor g, defined as

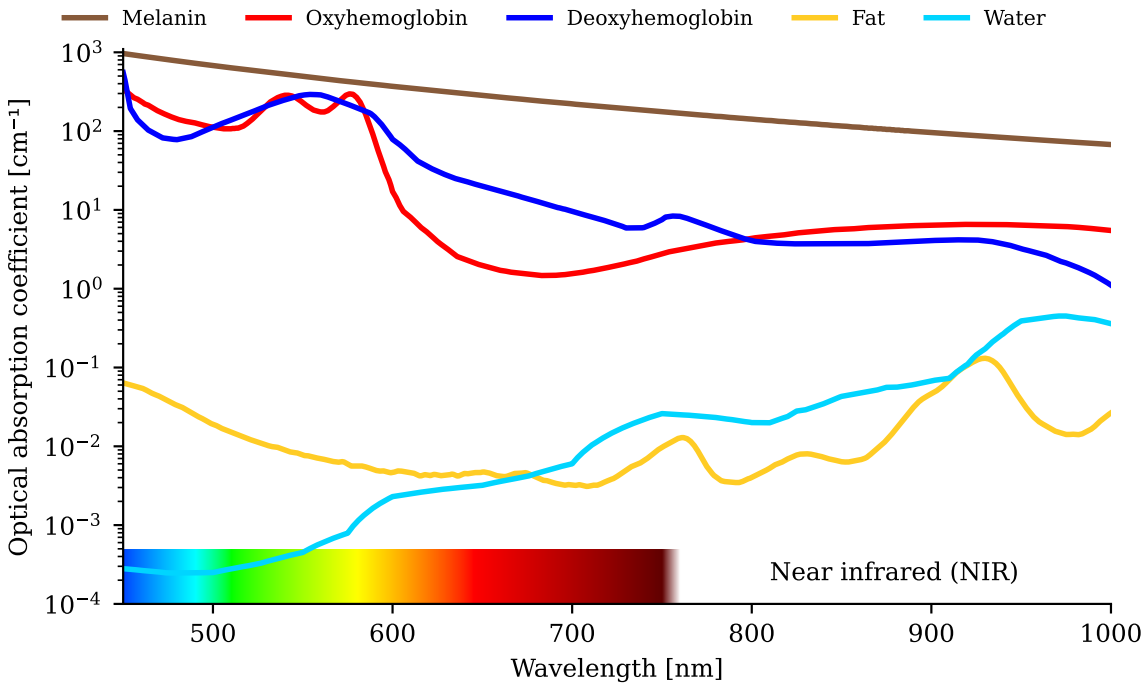


Figure 2.2: Tissue chromophores and their respective absorption spectra. The spectra of deoxyhemoglobin and oxyhemoglobin were calculated with a concentration of 150 g/L. Data from omlc.org.

the average cosine of the scattering angle θ:

$$g = \langle \cos\theta \rangle \begin{cases} -1 \leq g < 0 & \text{for mostly backward-directed scattering.} \\ g = 0 & \text{for perfectly isotropic scattering.} \\ 0 < g \leq 1 & \text{for mostly forward-directed scattering.} \end{cases} \tag{2.7}$$

In most human tissues, scattering is predominantly forward-directed, with typical anisotropy values of $g \gtrsim 0.9$ [26]. To account for this behavior, the reduced scattering coefficient μ_s' is introduced as:

$$\mu_s' = (1 - g)\mu_s. \tag{2.8}$$

This simplification allows multiple highly forward-directed scattering events described by μ_s to be approximated as fewer isotropic events using μ_s'.
In media where both absorption and scattering occur, their combined effect is described by the extinction coefficient μ_t, defined as:

$$\mu_t = \mu_a + \mu_s, \tag{2.9}$$

The reciprocal of μ_t corresponds to the mean free path, indicating the average distance a photon travels before being either scattered or absorbed. For biological tissue, this is often in the range of 0.1 - 1 μm [1].

Radiative transfer theory

The radiative transfer theory utilizes all the knowledge covered in the previous section in order to describe light propagation in biological tissue. In this section, the scalar time-independent radiative transfer equation (RTE) will be derived from energy conservation. It is, however, also possible to derive it from Maxwell's equations [27]. Also, as will be described in more detail in section 2.1, the time scale of light propagation within tissue is much shorter than the conversion of absorbed optical energy to heat, causing the PA effect. Therefore, only the total absorbed energy density is of interest for the PA effect, thus, the scalar time-independent RTE suffices for this thesis.
The RTE is formulated in terms of the radiance $L(\vec{r}, \hat{s}, t)$ $\left(\frac{W}{m^2 sr}\right)$. It quantifies the energy flux per unit normal area, per unit solid angle, and per unit time t. It is a directional quantity that depends on the spatial position $\vec{r}$ and the propagation direction $\hat{s}$, a unit vector orthogonal to the local surface normal. The infinitesimal change in energy dP within a differential volume element dV, subtended by a differential solid angle $d\Omega$ over a time interval dt, can thus be expressed as:

$$dP = \frac{\partial L(\vec{r}, \hat{s}, t)/c}{\partial t} dV d\Omega = -dP_{\text{div}} - dP_{\text{ext}} + dP_{\text{scat}} + dP_{\text{src}}. \tag{2.10}$$

Four contributions make up the change in energy, namely dP_{div}, dP_{ext}, dP_{scat}, and dP_{src}. They will be explained in the following:

Divergence For non-collimated photon beams, the divergent energy dP_{div} radiating outward from the considered volume element is given by:

$$dP_{\text{div}} = \hat{s} \cdot \nabla L(\vec{r}, \hat{s}, t) d\Omega \, dV. \tag{2.11}$$

Extinction The energy loss dP_{ext} within the considered volume element due to extinction (absorption and scattering processes) is given by

$$dP_{\text{ext}} = (\mu_t ds) L(\vec{r}, \hat{s}, t) dA \, d\Omega. \tag{2.12}$$

Scattering Photons scattered into the considered volume element from external directions $\hat{s}'$ within the solid angle $d\Omega'$ contribute to an energy gain, which can be expressed as:

$$dP_{\text{scat}} = (\mu_s dV) \left[\int_{4\pi} L(\vec{r}, \hat{s}', t) \cdot P(\hat{s}', \hat{s}) d\Omega' \right] d\Omega. \tag{2.13}$$

Here, $P(\hat{s}', \hat{s})$ denotes the scattering phase function. This phase function is usually modeled as the Henyey–Greenstein phase function [28].

Photon source The photon source $S(\vec{r}, \hat{s}, t)$ $\left(\frac{W}{m^3 sr}\right)$ must also be included as an additional term contributing to the energy gain:

$$dP_{\text{src}} = S(\vec{r}, \hat{s}, t) dV \, d\Omega. \tag{2.14}$$

Using the four contributions 2.11, 2.12, 2.13 and 2.14 and inserting them into 2.10 yields the RTE:

$$\frac{\partial L(\vec{r}, \hat{s}, t)/c}{\partial t} = -\hat{s} \cdot \nabla L(\vec{r}, \hat{s}, t) - \mu_t L(\vec{r}, \hat{s}, t) + \mu_{\text{s}} \int_{4\pi} L(\vec{r}, \hat{s}', t) \cdot P(\hat{s}', \hat{s}) d\Omega' + S(\vec{r}, \hat{s}, t). \tag{2.15}$$

Solutions to the RTE exist for simplified cases analytically or are calculated numerically. One popular approximation for the RTE is the diffusion approximation.

Diffusion theory

The diffusion equation (DE) relies on the assumption that the medium is highly scattering ($\mu_{\text{a}} \ll \mu_{\text{s}}$) and that, after a sufficient number of scattering events, the radiance becomes nearly isotropic. Under these conditions, the RTE can be simplified to the diffusion equation:

$$\frac{\partial \Phi(\vec{r}, t)}{c \partial t} + \mu_{\text{a}} \Phi(\vec{r}, t) + -\nabla \cdot [D \nabla \Phi(\vec{r}, t)] = S(\vec{r}, t). \tag{2.16}$$

Here, D denotes the diffusion coefficient:

$$D = \frac{1}{3(\mu_{\text{a}} + \mu_{\text{s}}')} = \frac{1}{3\mu_t'} \tag{2.17}$$

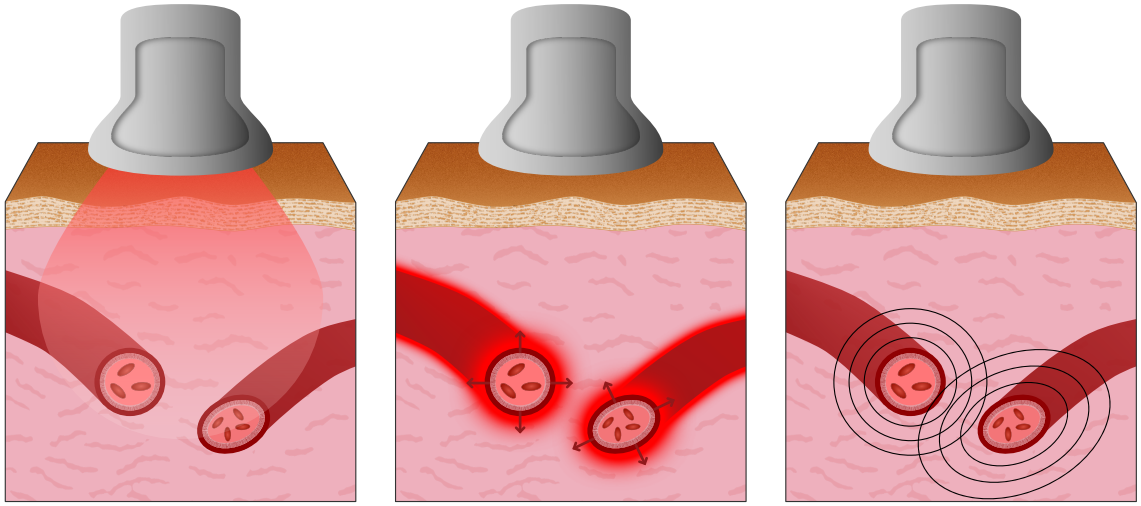
and $S(\vec{r}, t)$ represents the isotropic source term.
The DE is expressed in terms of the fluence rate $\Phi(\vec{r}, t)$ $\left(\frac{W}{m^2}\right)$ which is defined as:

$$\Phi(\vec{r}, t) = \int_{4\pi} L(\vec{r}, \hat{s}, t) d\Omega. \tag{2.18}$$

The conditions for the DE are typically met in scattering-dominated media at distances exceeding several scattering lengths from the light source. However, in many PAI applications, the imaging depth may be comparable to or even smaller than the average scattering length. In such cases, the DE may no longer be an appropriate simplification.

Photoacoustic effect

The PA effect essentially describes the phenomenon of a material converting optical light energy into acoustic waves. This principle is illustrated in Figure 2.3.



(a) Pulsed light absorption **(b)** Thermoelastic expansion **(c)** Acoustic wave generation

Figure 2.3: Schematic representation of the main steps in the photoacoustic effect: (a) light propagation and absorption, (b) thermoelastic expansion, (c) acoustic wave generation.

For this thesis, it is sufficient to describe the PA effect within biological tissue, assuming that the light energy comes from a pulsed laser. Therefore, in this section, the processes that occur within a volume element dV within biological tissue when it absorbs light energy will be described.
The analysis begins with the first law of thermodynamics for a simple compressible medium in differential form [29]:

$$dU = TdS - pdV, \tag{2.19}$$

which expresses conservation of energy in differential form. Here, U is the internal energy, T the temperature, S the entropy, and p the pressure. The first term on the right-hand side represents the transfer of heat into the system, while the second term accounts for mechanical work performed by the system through volume change. For the photoacoustic effect, the laser pulse duration τ_p has to be chosen to be much shorter than the characteristic thermal and stress relaxation times:

$$\tau_{\text{th}} = \frac{d^2}{\alpha_{\text{th}}}, \quad \tau_{\text{s}} = \frac{d}{v_{\text{s}}}, \tag{2.20}$$

where d is the characteristic size of the heated region, α is the thermal diffusivity, and v_s is the speed of sound in the medium. Thermal confinement ($\tau_p \ll t_{th}$) ensures that negligible heat diffuses out of the heated region during the laser pulse. Stress confinement ($\tau_p \ll t_s$) ensures that negligible pressure-driven expansion occurs during this same interval. As an example, for a target resolution of 0.1 mm, $\alpha_{\text{th}} = 1.3 \cdot 10^{-7}$ m and $v_{\text{s}} = 1500$ m/s, the expected values are approximately $\tau_{\text{s}} \approx 77$ ms and $\tau_{\text{th}} \approx 67$ ns. Consequently, typical laser pulse durations are on the order of a few nanoseconds.

Under these confinement conditions, the optical energy deposited in the target during the laser pulse is converted to internal energy without macroscopic volume change and eq. 2.19 simplifies to:

$$dU = TdS, \quad dV \approx 0. \tag{2.21}$$

In absorbing media, part of the optical energy may be lost via radiative relaxation processes such as fluorescence. The photoacoustic effect corresponds to non-radiative relaxation, in which the absorbed photon energy is converted into vibrational energy of the lattice and subsequently thermalized into heat. The fraction of absorbed energy converted to heat is accounted for through a heating efficiency factor η_{th}.
In thermodynamic equilibrium, the volume is a state function of temperature and pressure, $V = V(T, p)$. Its total differential is

$$dV = \left(\frac{\partial V}{\partial T}\right)_p dT + \left(\frac{\partial V}{\partial p}\right)_T dp. \tag{2.22}$$

The two partial derivatives in eq. 2.22 can be related to the material properties of the isothermal compressibility κ_T and the thermal coefficient of volume expansion α [29] by:

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T, \quad \beta = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p. \tag{2.23}$$

Therefore, using $dV \approx 0$ and these relations, eq. 2.22 expressed for dp becomes:

$$dp = \frac{\beta}{\kappa_T} dT. \tag{2.24}$$

During the light pulse, the local temperature rise dT can be expressed as

$$dT = \frac{\eta_{\text{th}} \mu_{\text{a}} \Phi}{\rho C_{\text{V}}} dt. \tag{2.25}$$

Here, η_{th} describes the aforementioned efficiency with which internal energy is converted to heat, called non-radiative relaxation, instead of radiative relaxation in e.g. fluorophores. ρ denotes the mass density, C_{V} the specific heat capacity at constant volume, and Φ the optical fluence rate defined in eq. 2.18. Using the dimensionless Grüneisen parameter describing the PA efficiency

$$\Gamma = \frac{\beta}{\kappa_T \rho C_V}, \tag{2.26}$$

which at body temperature is $\Gamma(\text{T}_0 = 37^\circ\,\text{C}) \approx 0.20$ [1], eq. 2.24 becomes:

$$dp = \Gamma \eta_{\text{th}} \mu_{\text{a}} \Phi dt. \tag{2.27}$$

Finally, after integrating over the time of the light pulse, the so-called initial pressure distribution $p_0(\vec{x})$ at $t = 0$ for a given position $\vec{x}$ after the laser pulse can be calculated with:

$$p_0(\vec{x}) = p(\vec{x}, t = 0) = \Gamma \eta_{\text{th}} \mu_{\text{a}} \phi(\vec{x}). \tag{2.28}$$

Here, the optical fluence $\phi(\vec{x})$ $\left(\frac{J}{m^2}\right)$ is defined as

$$\phi(\vec{x}) = \int_{\tau_p} \Phi(\vec{x}, t) dt. \tag{2.29}$$

As a general reference, the typical rises in pressure and temperature are on the order of $\sim 10\,\text{mbar}$ and $\sim 10\,\text{mK}$, respectively, which are not harmful to biological tissue [30]. This initial pressure rise causes an acoustic wave to propagate through the tissue.

Acoustic wave propagation

Acoustic waves are longitudinal in nature, propagating through alternating variations in pressure p and density ρ at the speed of sound v_s [31]. As a first-order approximation, their relationship can be expressed as:

$$p = v_s^2 \rho. \tag{2.30}$$

For an isotropic medium without acoustic attenuation, the conservation of momentum yields:

$$\nabla p + \rho \frac{\partial \vec{v}}{\partial t} = 0 \tag{2.31}$$

and the conservation of mass leads to:

$$\rho \nabla \cdot \vec{v} + \frac{\partial \rho}{\partial t} = 0, \tag{2.32}$$

where $\vec{v}$ represents the acoustic particle velocity. By taking the divergence of eq. 2.31, the partial time derivative of eq. 2.32, and applying the relationship given in eq. 2.30, these expressions can be combined to form the well-known acoustic wave equation [31]:

$$\nabla^2 p(\vec{x}, t) - \frac{1}{v_s^2} \frac{\partial^2 p(\vec{x}, t)}{\partial t^2} = 0. \tag{2.33}$$

The initial conditions given in eq. 2.28, along with $\partial_t p(\vec{x}, t = 0) = 0$, must be applied to account for the source of the acoustic wave, namely, the initial pressure rise induced by optical heating. Incorporating these conditions yields the PA wave equation:

$$\left(\nabla^2 - \frac{1}{v_s^2} \frac{\partial^2}{\partial t^2}\right) p(\vec{x}, t) = -\frac{\Gamma}{v_s^2} \frac{\partial H(\vec{x}, t)}{\partial t}. \tag{2.34}$$

Here, $H(\vec{x}, t)$ $\left(\frac{W}{m^3}\right)$ is called the heating function and it is defined as

$$H(\vec{x}, t) = \eta_{th} \mu_a \Phi. \tag{2.35}$$

An acoustic wave generated by the initial pressure distribution propagates according to the PA wave equation 2.34. Such a wave can be detected with a regular ultrasound transducer on the surface of the tissue, based on which a PA image can be reconstructed by estimating the initial pressure distribution p_0.

Inverse problems in photoacoustic imaging

Having discussed the physics underlying the forward processes in PAI, this section turns to the steps required to extract the desired quantitative information. As outlined earlier, PAI enables functional imaging and related applications; achieving this, however, requires solving a sequence of inverse problems. Specifically, two main challenges arise: (1) the acoustic inverse problem, and (2) the optical inverse problem. Both have been extensively addressed and reviewed in previous literature [32, 33, 34, 35, 36, 37]. In the following, a concise overview relevant to the context of this thesis is provided.

Acoustic inverse problem

The objective of the acoustic inverse problem is to reconstruct the initial pressure distribution p_0 in the spatial domain as accurately as possible from the recorded measurement data $S(t)$. Recordings typically begin immediately after the excitation light pulse, capturing the arrival of acoustic pressure waves at the detector.
In a tomographic setup, the recorded data take the form $S_n(t)$ for N detector elements, commonly referred to as time-series data in the shape of an $N \times T$ matrix, where T denotes the number of recorded time steps. This establishes a relationship between the arrival time of an acoustic wave at a detector and its point of origin in p_0:

$$S(t) \propto p_0 = \Gamma \eta_{\text{th}} \mu_{\text{a}} \phi. \tag{2.36}$$

Several reconstruction approaches exist, including backprojection-based methods [38], time-reversal techniques [39], model-based inversions [40], and data-driven methods [41]. However, all are subject to notable challenges:

1. **Detection limitations and information loss:** Most tomographic systems have a limited view of the target region. For example, in handheld PA acquisition of a vessel inside an arm, only the fraction of the acoustic wave propagating toward the transducer is recorded. Furthermore, due to tissue heterogeneity, a portion of the wavefront traveling toward the detector may be attenuated or reflected before reaching it. Even for the portion that arrives, transducers possess a finite bandwidth, so frequency components outside this range are not captured.

2. **Device calibration:** Accurate reconstruction requires precise knowledge of detector element positions and characteristics, including bandwidth, acceptance angle, and detection efficiency. Ideally, p_0 should be reconstructed quantitatively in pressure units, which requires a well-characterized transfer function from the raw detector output (digital counts) to physical pressure. In practice, these parameters are often known only approximately or measured with uncertainty.

3. **Model assumptions in reconstruction algorithms:** Accurate modeling of the relationship between time-series data and p_0 requires assumptions about tissue

properties, most importantly the speed of sound, but also density and acoustic attenuation, as these influence reflection and refraction. Ideally, these parameters should be known a priori for each location in the tissue, which is rarely the case in practice.

Optical inverse problem

The optical inverse problem in PAI aims to recover the spatial distribution of $\mu_a(\vec{x})$ or related functional parameters such as blood oxygen saturation sO_2 from the reconstructed initial pressure distribution $p_0(\vec{x})$. Here, the acoustic inverse problem is assumed to be solved, i.e., p_0 is known, thus using the relationship in eq. 2.28, μ_a can be calculated with

$$\mu_a = \frac{p_0}{\phi\,\Gamma}. \tag{2.37}$$

Approaches to solving the optical inverse problem include model-based spectral inversion using light transport models (e.g., Monte Carlo simulations, diffusion approximation, finite element solutions to the RTE) [42], as well as data-driven or hybrid techniques [43]. Regardless of the method, several major challenges remain:

1. **Ambiguity and non-linearity:** Multiple combinations of μ_a and μ_s can yield the same p_0, making the inversion inherently ambiguous. Furthermore, the fluence at a given location $\vec{x}$ depends not only on the optical properties μ_a and μ_s at that location, but also on their distributions throughout the entire tissue volume. For example, a highly absorbing layer located above $\vec{x}$ will reduce the amount of light reaching it.

2. **Spectral coloring and noise:** The fluence also depends on the illumination wavelength, and wavelength-dependent absorption or scattering can further complicate the inversion. Continuing the previous example, a highly absorbing layer might attenuate certain wavelengths more strongly than others. This spectral dependence is compounded by noise sources in the photoacoustic measurement chain, including detector noise and fluctuations in laser pulse energy, both of which can introduce additional uncertainty.

3. **Model assumptions:** Inversion typically requires prior assumptions about light propagation, source characteristics, and optical or thermodynamic parameters such as μ_s and Γ. In practice, these quantities are often assumed constant, despite known dependencies on spatial position, temperature, and wavelength. For instance, Γ varies not only with $\vec{x}$ but also with the local tissue temperature and wavelength.

Taking these dependencies into account, the relationship from eq. 2.28 can be reformulated

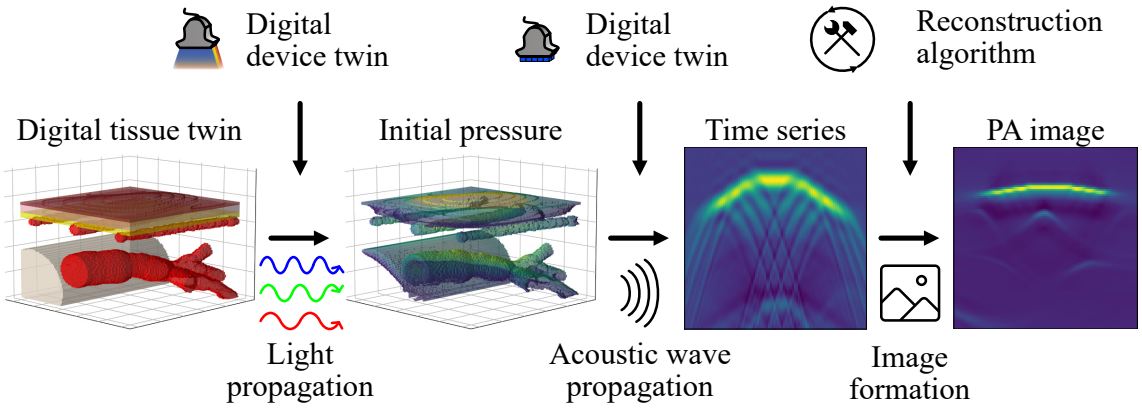


Figure 2.4: Core pipeline elements for the computational modeling of photoacoustic imaging (PAI). First, a digital tissue twin, the tissue geometry, optical, and acoustic parameters are defined, serving as basis for the rest of the pipeline. Using a digital device twin, modeling of light propagation yields the initial pressure distribution, which in turn, is the input for acoustic wave propagation, resulting in recorded time series data. Finally, using an image reconstruction algorithm, a photoacoustic (PA) image is formed.

as:

$$\begin{aligned} p_0(\vec{x},\lambda) = \mu_a(\vec{x},\lambda) \\ &\cdot \phi(\vec{x},\lambda,\mu_a(\vec{x},\lambda),\mu_s(\vec{x},\lambda)) \\ &\cdot \Gamma(\vec{x},\mu_a(\vec{x},\lambda),T(\vec{x})) \\ &\cdot \eta_{mul}(\vec{x},\lambda) \\ &+ \eta_{add}(\vec{x},\lambda), \end{aligned} \tag{2.38}$$

where $\eta_{mul}(\vec{x},\lambda)$ represents multiplicative noise, such as light source variations, and $\eta_{add}(\vec{x},\lambda)$ represents additive noise, such as noise offsets in the detector. Eq. 2.38 highlights the intricate interdependencies between optical, thermodynamic, and illumination parameters, illustrating the overall complexity of the optical inverse problem in PAI.

2.2 Computational modeling of photoacoustic imaging

Accurate computational modeling of PA image formation is essential for simulating experimental conditions, optimizing system design, or training and validating data-driven algorithms. Relevant to this work are the numerical description of light propagation in tissue, the generation of initial pressure distributions, and the simulation of subsequent acoustic wave propagation and detection. The entire computational pipeline is illustrated in Figure 2.4.
As can be seen in Figure 2.4, generating both a digital tissue twin and a digital device twin is also crucial to the computational modeling pipeline of PAI. As this, however, is a major

contribution of this thesis, this is omitted from this section and explained in more detail in section 4.2. Furthermore, the exact thermodynamic modeling of heating as a result of light absorption and non-radiative relaxation is usually neglected by assuming 100 % efficiency and a constant Grüneisen parameter Γ. In the following, the modeling of photon transport is described first, followed by the modeling of acoustic propagation, noise characteristics, and the transformation of recorded time series data into PA images.

Light propagation modeling

Light propagation in tissue is governed by the RTE introduced in section 2.1. As discussed there, the RTE is computationally demanding to solve numerically. A common simplification is the DE, which approximates the RTE under conditions of highly scattering, near-homogeneous media and can be computed with significantly lower cost [44]. The fluence obtained from such models can be converted to the initial pressure p_0 via the laser energy, the Grüneisen parameter Γ, and 2.28. However, the DE becomes inaccurate in the quasi-ballistic regime, particularly near the photon source. For this reason, Monte Carlo (MC) methods are considered the gold standard for modeling light transport [1].
Several open-source MC implementations exist with varying levels of anatomical and optical complexity [45]. In this thesis, photon transport is simulated with Monte Carlo eXtreme (MCX) [46], a graphics processing unit (GPU)-accelerated code capable of modeling volumes with arbitrary spatial distributions of absorption and scattering coefficients. With suitable modifications, arbitrary illumination geometries can also be implemented.
In MCX, photons are represented as energy packets with an initial weight $w = 1$, launched from the defined source position and direction according to the illumination geometry. Photon propagation is modeled in discrete steps:

1. The step size s_i between scattering events is sampled as $s_\mathrm{i} = -\frac{\ln \xi_\mathrm{i}}{\mu_{\mathrm{s_i}}}$, where $\mu_{\mathrm{s_i}}$ is the scattering coefficient of the current voxel i and $\xi_\mathrm{i} \in (0, 1]$ is a uniformly distributed random number.

2. The photon packet is advanced by $s_{\mathrm{i+1}}$ along its direction vector to the next voxel.

3. The residual step size is updated for the change in scattering coefficient from $\mu_{\mathrm{s_{i+1}}}$ to $\mu_{\mathrm{s_{i+2}}}$: $s_{\mathrm{i+2}} = \frac{(s_\mathrm{i} - s_{\mathrm{i+1}})\,\mu_{\mathrm{s_{i+1}}}}{\mu_{\mathrm{s_{i+2}}}}$.

4. Absorbed energy $\Delta w = w\,(1 - e^{-\mu_{\mathrm{a_i}} s_\mathrm{i}})$ is deposited in the current voxel (W_{vox}) and the packet weight is reduced accordingly.

5. If w falls below a threshold w_{th}, a Russian roulette scheme is applied: the photon survives with probability $1/m$ (weight set to $m\,w_{\mathrm{th}}$) or is terminated otherwise, preventing endless propagation with negligible energy.

6. Repeat from step 2 until $s = 0$.

7. A new scattering direction is sampled. The azimuthal angle ϕ is drawn as $\phi = 2\pi\xi$ with $\xi \in [0, 1)$. The polar deflection angle θ is drawn from the Henyey–Greenstein phase function:

$$\cos(\theta) = \begin{cases} \frac{1}{2g}\left[1 + g^2 - \left(\frac{1-g^2}{1-g+2g\xi}\right)\right] & \text{if } g \neq 0, \\ 2\xi - 1 & \text{if } g = 0, \end{cases} \tag{2.39}$$

   with $\xi \in [0, 1]$ and anisotropy factor $g \in [-1, 1]$.

8. Continue until the photon is terminated or exits the simulation volume V.

The process is repeated until the desired number of photons is simulated. Energy leaving the volume is accumulated in W_{exit}. After all packets are traced, the deposited energy is normalized by the total simulated energy:

$$W'_{\text{vox}} = \frac{W_{\text{vox}}}{\sum\limits_{\text{vox}\in V} W_{\text{vox}} + W_{\text{exit}}}. \tag{2.40}$$

The number of photons required depends strongly on voxel spacing. Halving the voxel size increases the number of voxels by a factor of eight (in three-dimensional (3D)), leading to increased computational load and higher photon counts to suppress photon noise, which is a random fluctuation in voxel fluence due to insufficient photon sampling. The required photon count for each spacing is determined empirically.
For accurate modeling of any PA system, an accurate model of the light source, a so-called digital device twin, needs to be implemented as photon launching distribution.

Acoustic wave propagation modeling

As derived in 2.1, the PA wave equation 2.34 governs acoustic propagation resulting from the PA effect. For absorption-free media, this equation is exact. Acoustic absorption can be incorporated using a power-law frequency dependence:

$$\alpha = \alpha_0 f^y, \tag{2.41}$$

where α (dB/cm) is the absorption coefficient, α_0 (dB/MHz^y cm) its prefactor, and y the power-law exponent.
To account for heterogeneity and attenuation, the governing equations (2.30)–(2.32) are modified as:

$$p = v_s^2 \left(\rho + \vec{d} \cdot \nabla\rho_0 - L\rho\right), \tag{2.42}$$

$$0 = \nabla p + \rho \frac{\partial \vec{v}}{\partial t}, \tag{2.43}$$

$$0 = \rho \nabla \cdot \vec{v} + \vec{v} \cdot \nabla\rho_0 + \frac{\partial \rho}{\partial t}, \tag{2.44}$$

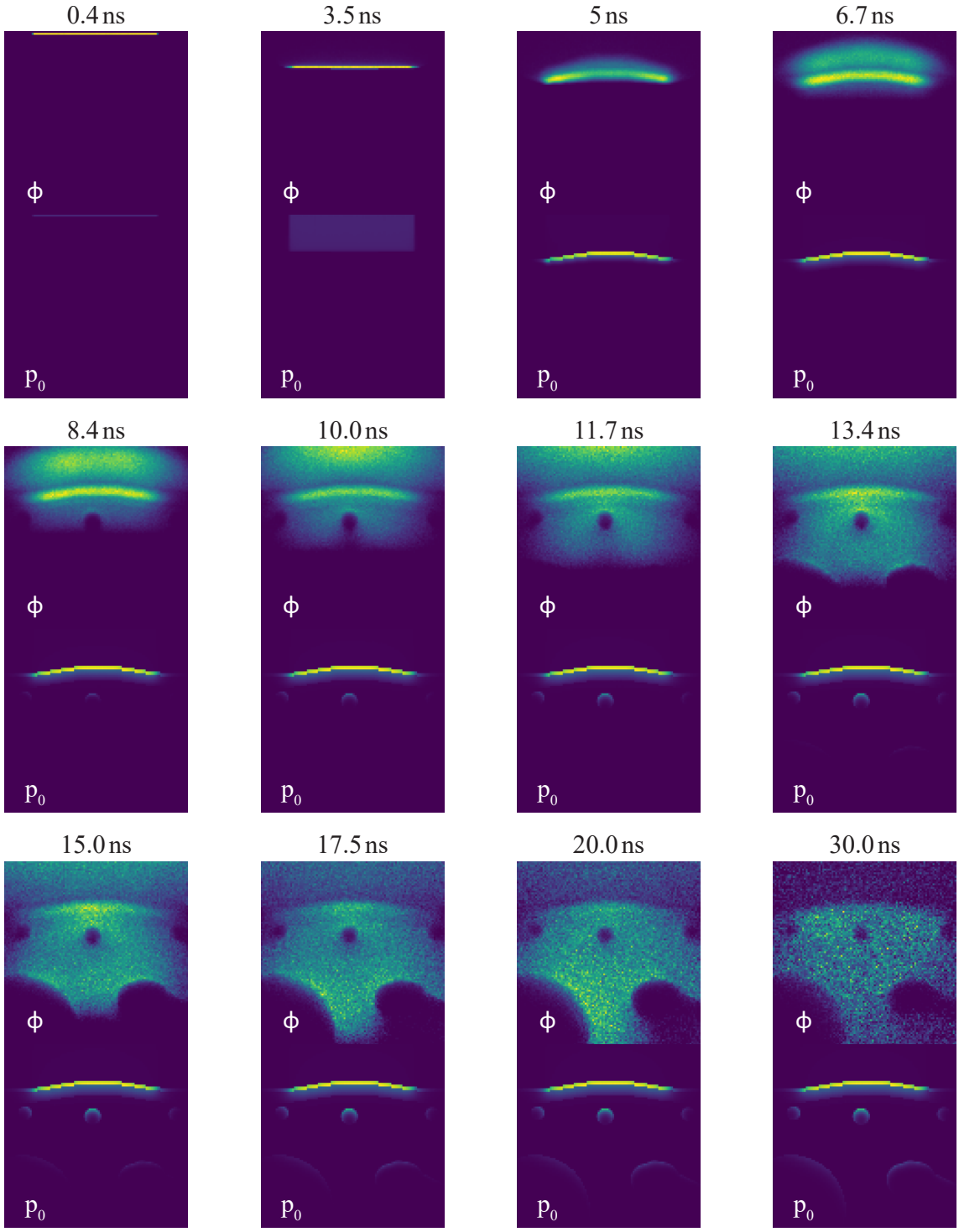


Figure 2.5: Time-resolved photon propagation in an example tissue model with photon fluence ϕ in the top parts and corresponding initial pressure p_0 in the bottom parts of each row. The time starts after the light pulse incident on the tissue from the top.

with particle displacement $\vec{d}$ and absorption–dispersion operator L:

$$L = \tau \frac{\partial}{\partial t}(-\nabla^2)^{\frac{y+1}{2}-1} + \eta(-\nabla^2)^{\frac{y+1}{2}-1}, \tag{2.45}$$

where $\tau = -2\alpha_0 v_s^{y-1}$ and $\eta = 2\alpha_0 v_s^y \tan\left(\frac{\pi}{2}y\right)$ are proportionality constants valid for $0 < y < 3$, $y \neq 1$.
The MATLAB k-Wave toolbox [47] is widely used for solving (2.30)–(2.32) using a k-space pseudospectral method. It supports one-dimensional (1D), two-dimensional (2D), and 3D geometries (planar, cylindrical, spherical) and incorporates detector geometry and characteristics. Spatial discretization dx strongly influences computational cost: analogously to MCX, halving dx increases the number of grid points by 2^D in D dimensions, raising the cost correspondingly. The Nyquist criterion [48, 49] limits the maximum resolvable spatial frequency:

$$f_{\max} = \frac{\min_V(v_s)}{2\,dx}, \tag{2.46}$$

with $\min_V(v_s)$ the minimum sound speed in V.
Similar to the light propagation modeling, detector properties, including center frequency and bandwidth, have to be modeled via a digital device twin to simulate time-series pressure recordings.

Noise modeling

PA images are inherently affected by noise from laser pulse fluctuations, detector sensitivity variations, and probe or target motion [50]. Here, noise is modeled as additive Gaussian noise proportional to the peak amplitude of the measured time series [51]. To estimate system-specific noise, in this thesis, 100 example images were acquired with an experimental system over wavelengths 700–850 nm in 10 nm steps. The standard deviation std of each time series $p_i(\lambda)$ was normalized by its peak value and averaged:

$$\mathrm{psnr}(\lambda)_{\mathrm{exp}} = \frac{1}{100}\sum_{i=1}^{100} \log\left(\frac{\max(p_i(\lambda))}{\mathrm{std}(p_i(\lambda))}\right). \tag{2.47}$$

Here, $\mathrm{std}(p_i(\lambda))$ is estimated by the *estimate_sigma* function provided by the Python scikit-image package. The function is based on the work by Donoho and Johnstone (1994) [52]. This defines the wavelength-dependent noise fraction. Simulated noisy data are generated as:

$$p_{\mathrm{noisy}}(\lambda) = p(\lambda) + \mathcal{N}_{p(\lambda)}\,(0,\ \mathrm{std}_{\mathrm{noise}})\,, \tag{2.48}$$

where $\mathcal{N}_{p(\lambda)}(\mu, \sigma)$ is a Gaussian random array matching the shape of $p(\lambda)$ and $\mathrm{std}_{\mathrm{noise}}$ is optimized such that $\mathrm{psnr}(\lambda)_{\mathrm{sim}}$ of 100 randomly selected simulated images match $\mathrm{psnr}(\lambda)_{\mathrm{exp}}$ of 100 experimentally acquired images.

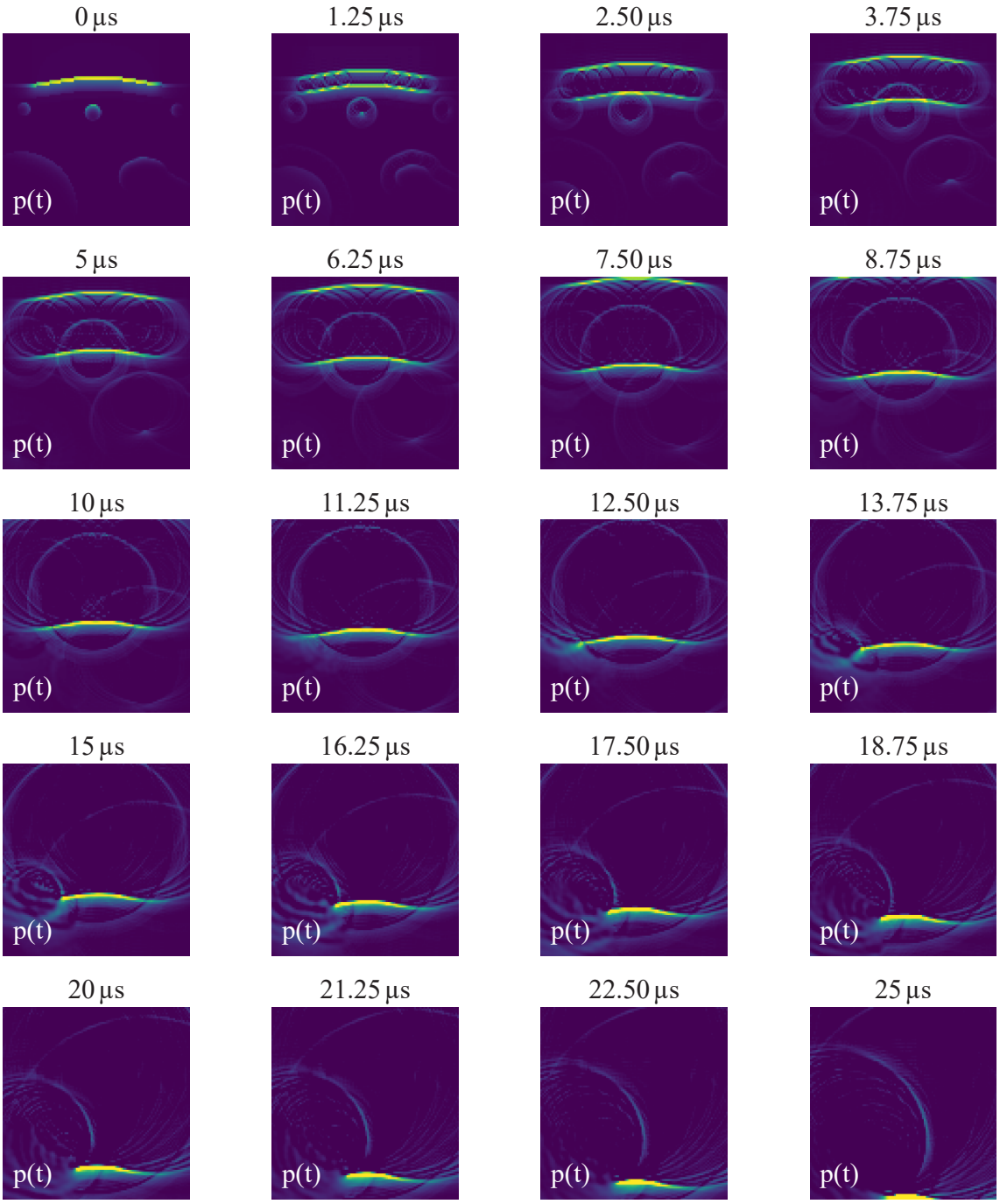


Figure 2.6: Time-resolved acoustic pressure wave p(t) propagation in an example tissue model. The time starts after the light pulse has ended. Notably, the pressure waves from the initial pressure distribution ($p_0 = p$(t=0)) have almost fully exited the field of view at t=13.75 μs but are reflected in the bottom left. This reflected wave travels toward the top right for the rest of the time steps.

Image formation

The final step in the simulation pipeline reconstructs the initial pressure distribution $p_0(\vec{x})$ from the recorded time series. The universal backprojection algorithm [38] is expressed as:

$$p_0(\vec{x}) = \int_{\Omega_0} \frac{d\Omega_0}{\Omega_0} \left[2p(\vec{x}_0, v_s t) - 2v_s t \frac{\partial p(\vec{x}_0, v_s t)}{\partial (v_s t)} \right] \Bigg|_{t=\frac{|\vec{x}-\vec{x}_0|}{v_s}}, \tag{2.49}$$

with

$$d\Omega_0 = \frac{dS_0}{|\vec{x} - \vec{x_0}|^2} \frac{\hat{n}_0^S \cdot (\vec{x} - \vec{x_0})}{|\vec{x} - \vec{x_0}|}. \tag{2.50}$$

This formulation assumes continuous detection, exact speed of sound, and ideal sensor characteristics—conditions rarely met in practice. A simpler, robust alternative is delay-and-sum (DAS) beamforming [53], also common in ultrasound imaging.

For an array of N elements at positions $\vec{r}_j(x, y)$, $j = 1, \ldots, N$, recording a signal from a source at $\vec{a}(x, y)$, the spherical propagation delay is:

$$\tau(x, y, j) = \frac{\|\vec{r}_j(x, y) - \vec{a}(x, y)\|_2}{v_s}. \tag{2.51}$$

The reconstructed pixel value is then:

$$S_{\mathrm{DAS}}(x, y) = \sum_{j=1}^{N} S\big(j, \tau(x, y, j)\big) \cdot A(j), \tag{2.52}$$

where $S(j, \tau)$ is the time series data of sensor element j at time τ and $A(j) \in [0, 1]$ is an apodization weight to reduce side lobes and beamforming artifacts [54].

2.3 Deep learning

DL is a subfield of machine learning in which artificial neural networks with multiple layers learn to represent data at multiple levels of abstraction. In DL, each layer of the network learns a transformation of its input into a higher-level, more abstract representation. This approach falls under representation learning, which means the system discovers the features or representations needed for tasks directly from raw data. A major appeal of deep neural networks is that they can learn these representations automatically without manual feature engineering, often achieving state-of-the-art results in domains like image recognition, speech, and language processing [55]. In this section, a concise introduction to the core DL concepts relevant to this thesis is provided, drawing primarily on the review by LeCun et al. (2015) [56] and the textbook by Goodfellow et al. (2016) [57]. First, learning paradigms are outlined with emphasis on supervised and unsupervised learning. Then, the principles

of representation learning that underpin modern image models are summarized. Next, the network classes used in this work are reviewed. Finally, the generative architectures employed in this thesis, generative adversarial networks (GANs) and invertible neural networks (INNs), are introduced. Throughout, the focus remains on the concepts required for the methods and experiments presented in later chapters.

Learning paradigms

DL models can be trained under different learning schemes depending on the availability of labeled data. In supervised learning, the network learns from example input-output pairs (X, Y), where it is given input data X along with the corresponding output labels or targets Y. The objective is to adjust the model's parameters θ so that its predictions $\hat{Y}$ closely match the known labels. In other words, the model is supposed to represent a function f_θ that maps X to Y as accurately as possible. In this thesis, supervised and unsupervised learning techniques are employed, the core principles of which are compared in Figure 2.7.

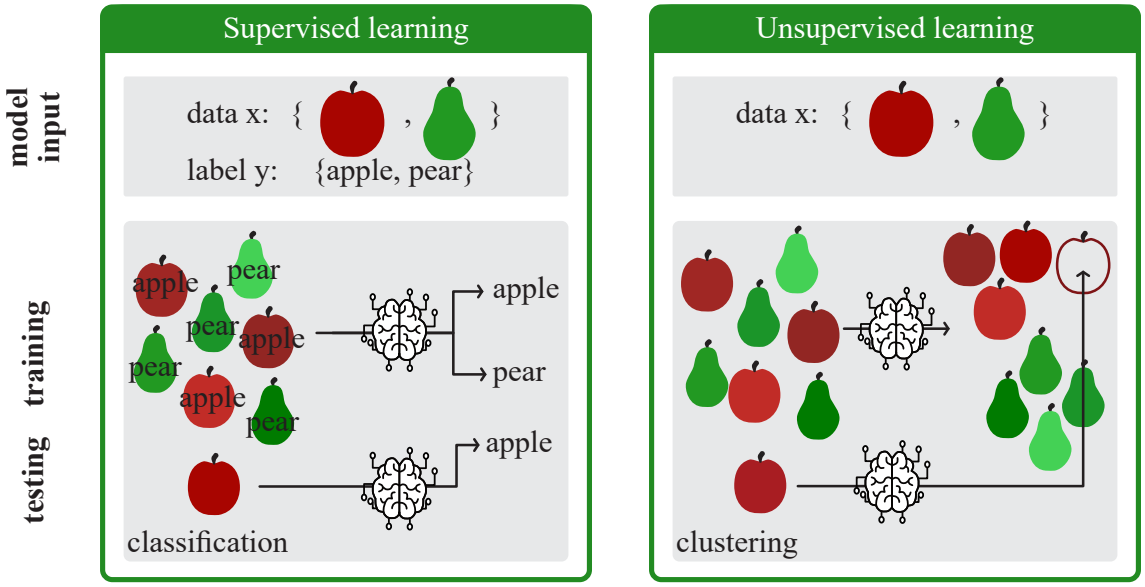


Figure 2.7: Schematic illustration of supervised and unsupervised learning. In supervised learning, models are trained on data x paired with corresponding labels y, as is typical in classification tasks. In the example shown, a classifier learns to associate the labels "apple" or "pear" with input data. During inference, an unseen apple is correctly assigned the label "apple." In contrast, unsupervised learning operates without labeled data. Here, the model infers patterns within the data structure. In the illustrated clustering scenario, the model identifies characteristic features of apples and pears, forming two distinct clusters. A new apple sample is then placed closer to the appropriate "apple cluster". Figure modified from Schellenberg 2024 [58].

Classic supervised tasks for deep networks include image classification, where a network might take an image as input $x \in X$ and predict a label $y \in Y$ (e.g., apple or pear). Training

proceeds by minimizing a loss function that measures error between predictions and true labels across many examples. For classification, this could be, for example, the binary cross entropy [59] or for regression, a simple and often used loss function is the mean squared error (MSE), which is defined by

$$L_{\text{MSE}} = \frac{1}{N}\sum_{\text{n}=1}^{N}(y_{\text{n}} - \hat{y}_{\text{n}})^2 \quad \text{with} \quad \hat{y}_{\text{n}} = f_\theta(x_{\text{n}}), \tag{2.53}$$

where x_n and are the n-th entries of N-dimensional vectors.

In unsupervised learning, by contrast, the model is not provided explicit labels. Instead, it explores the structure of the input data itself to learn useful representations or to generate new data. The goal might be to discover latent patterns, cluster similar inputs, or model the data distribution [60]. For example, an autoencoder network [61] compresses input data into a lower-dimensional code and then reconstructs the input, learning a representation in the process.

Neural network types

Neural networks are inspired by the neurons in the human brain, which receive an input signal from multiple other neurons. The received signals are then weighted according to previously learned experiences and combined. The output signal of the neuron to one or more other neurons is the combined signal after passing an activation potential.

Multilayer perceptron

The multilayer perceptron (MLP) is the quintessential DL architecture, consisting of layers of interconnected neurons. An MLP is a feedforward neural network, meaning information flows from the input layer through any number of hidden layers to the output layer in one direction. Each neuron in an MLP performs a two-step computation: first, a weighted sum of its inputs, then an application of a non-linear activation function. In mathematical terms, for a single neuron, if x_n are the inputs, w_n the weights, and b a bias that form the model parameters $\theta = \{w_n, b\}$, the neuron computes its output z as follows:

$$z = \sigma\left(\sum_n w_n x_n + b\right), \tag{2.54}$$

where σ is an activation function like sigmoid [62] or rectified linear unit (ReLU) [63]. This basic building block, a linear combination followed by a non-linear activation, allows even a simple two-layer network to approximate arbitrary continuous functions [64]. Using multiple hidden layers (i.e., increasing the network depth) can be more parameter-efficient, allowing the network to reuse and refine features at successive layers.

Training an MLP involves finding the weights w_m and biases b that minimize a chosen loss function L (e.g. MSE for regression or cross-entropy for classification) over the training data. This is typically done with stochastic gradient descent (SGD) or its variants such as Adam [65], using the backpropagation algorithm to compute gradients. Backpropagation efficiently computes the derivative of the loss L with respect to each weight in the network by recursive application of the chain rule through the layers. In essence, the algorithm calculates errors from the output layer backward to the input, layer by layer, hence "back-propagating" the error. Each weight $w_m^{(i+1)}$ of a model with M weights is then updated from its previous value w_m^i in the direction that most reduces the error (gradient descent step) given by an n-dimensional loss function L_n:

$$w_\mathrm{m}^{(i+1)} = w_\mathrm{m}^{(i)} - \gamma_i \sum_{\mathrm{n}=1}^{N} \frac{\partial L_n}{\partial w_\mathrm{m}^{(i)}}, \tag{2.55}$$

where γ is the learning rate. Backpropagation thus allows the network to learn from end to end, adjusting all layers' parameters to better achieve the output target. Through iterative training, the deep network gradually discovers internal representations (in its hidden layers) that are useful for mapping inputs to the desired outputs.

Convolutional neural networks

When working with image data, convolutional neural networks (CNNs) are the dominant architecture. CNNs are specialized neural networks that leverage the spatial structure of images through the convolution operation. In a convolutional layer, rather than connecting every input pixel to every neuron (as in an MLP), a set of local filters is learned (also called kernels) that are convolved with the image. A filter is typically a small matrix (for example, 3×3 or 5×5) that is slid across the image. At each spatial location, an element-wise multiplication is performed between the filter weights and the input pixels under the filter, and the results are summed up to produce one output pixel. This process of moving the filter and computing weighted sums is the convolution operation in the discrete case. Formally, the convolution of an image I and a 2D kernel K can be discretely defined as

$$S(m,n) = (K * I)(m,n) = \sum_{i,j} I(m-i, n-j)K(i,j). \tag{2.56}$$

Several key properties make CNNs powerful for image tasks. Local connectivity means each filter operates on a small patch of the image, capturing local spatial patterns (like edges, textures) which are meaningful in images. The same filter (set of weights) is used over all image locations, so a single learned feature (e.g., "horizontal edge") is detected everywhere in the image. This so-called "weight sharing" greatly reduces the number of parameters compared to an MLP. As the network learns, filters in early layers typically become detectors for simple patterns (edges, orientations), while filters in deeper layers assemble those into higher-level features (corners, object parts), reflecting an increasingly

abstract representation of the image. After one or more convolutional layers, CNNs often include pooling layers that down-sample the feature maps (e.g., taking a maximum or average over a small region) to reduce spatial resolution and achieve some invariance to small translations.

Generative architectures

Generative architectures or models are classes of neural networks that model the distribution of data p_{data} in a way that allows the generation of new samples following that distribution $\hat{x} \sim p_{\text{data}}$. This section focuses on the generative models most relevant to this thesis, GANs and INNs.

Generative adversarial networks

GANs are a class of DL models introduced by Goodfellow et al. (2014) [66] for unsupervised learning of complex data distributions. A GAN consists of two neural networks, a Generator (G) and a Discriminator (D), which are trained in tandem, effectively playing a two-player game. The generator network takes random noise as input and produces a synthetic data sample (for example, a fake image). The discriminator network takes an input (real or generated) and outputs a probability (or score) indicating whether the input is real (from the true data distribution) or fake (produced by the generator). During training, the generator is trying to fool the discriminator by producing ever more realistic data, while the discriminator is trying to become more accurate at telling generated fakes apart from real data. There are multiple ways of phrasing this training scheme as an optimization problem. In contrast to the original minimax optimization problem proposed by Goodfellow et al., in this thesis, the least squares training scheme proposed by Mao et al. (2017) [67] is used with the losses $\underset{G}{L}$ and $\underset{D}{L}$ for G and D, respectively:

$$\underset{G}{L} = \mathop{\mathbb{E}}_{x \sim p_z(z)} \left[(D(G(z)) - 1)^2 \right] \tag{2.57}$$

$$\underset{D}{L} = \mathop{\mathbb{E}}_{x \sim p_{\text{data}}(x)} \left[(D(x) - 1)^2 \right] + \mathop{\mathbb{E}}_{x \sim p_z(z)} \left[(D(G(z)))^2 \right]. \tag{2.58}$$

Here, $p_{\text{data}}(x)$ is the distribution of real data x and $p_z(z)$ is the noise prior (e.g., uniform or Gaussian) used as input to G. This training scheme has been shown to be more stable than the original minimax objective. Notably, GANs have yielded impressive results in image generation, creating photo-realistic images of faces, objects, and scenes after training on datasets of examples. However, GANs training can be challenging; issues like the aforementioned unstable convergence, mode collapse (where G produces limited varieties of outputs), and hallucinations are common [15], requiring careful architecture design and training heuristics.

Invertible neural networks

Another approach to generative modeling is to design networks that are invertible, meaning the transformation from input to output can be reversed. Normalizing flows are a family of deep generative models based on INNs [68]. The core idea is to start with a simple known probability distribution (e.g., a standard Gaussian) and apply a sequence of invertible transformations to produce a more complex target distribution. Each transformation is typically a relatively simple neural network layer that is invertible and has a tractable Jacobian determinant. By chaining multiple such flows, a flexible yet exactly invertible function f can be created that maps a latent variable z (drawn from the simple base distribution $p(z)$) to a sample $x = f(z)$ in the space of the target distribution.

$$\log p(x) = \log p(z) + \log \left| \det \frac{\partial f}{\partial x} \right|. \tag{2.59}$$

This formula shows how an INN can exactly compute the likelihood of each data example x by accounting for how volumes change through the transformation f.
In practice, normalizing flow architectures are designed so that both $f^{-1}(x)$ (the inverse mapping) and the Jacobian determinant $\det(\partial f^{-1}/\partial x)$ are efficient to compute. Common architectures use coupling layers or autoregressive transformations that have triangular Jacobian matrices, whose determinants are just the product of diagonal entries. Examples of popular flow-based models include RealNVP [69], NICE [70], and Glow [71], which employ such invertible coupling layers to model image data. An example for a coupling block introduced by Dinh et al. (2016) [69] is illustrated in Figure 2.8.
Coupling blocks operate by splitting the input into two components, $\boldsymbol{u} \rightarrow (\boldsymbol{u}_1, \boldsymbol{u}_2)$, which are transformed via learnable functions s_i and t_i in an alternating fashion (see Figure 2.8). In the forward pass, the transformed outputs are recombined to yield $(\boldsymbol{v}_1, \boldsymbol{v}_2) \rightarrow \boldsymbol{v}$. The inverse pass is implemented by reversing the transformations using the same functions s_i and t_i, applied to the corresponding parts of the output $(\boldsymbol{v}_1, \boldsymbol{v}_2)$. Notably, the transformation functions themselves do not need to be invertible; they can be arbitrary neural networks trained via backpropagation. The element-wise division inherent to these transformations may lead to numerical instability. This is typically addressed by applying an exponential function to s_i and t_i. In this procedure, $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ can be computed from $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ in the forward pass:

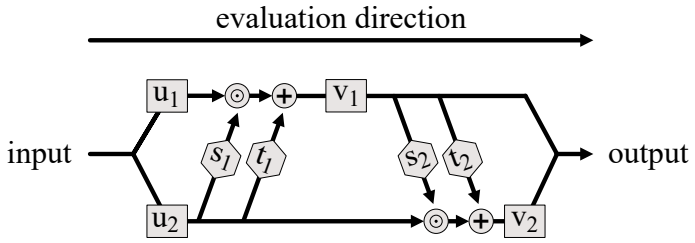
$$\boldsymbol{v}_1 = \boldsymbol{u}_1 \odot \exp(s_2(\boldsymbol{u}_2) + t_2(\boldsymbol{u}_2)) \tag{2.60}$$

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 \odot \exp(s_1(\boldsymbol{v}_1) + t_1(\boldsymbol{v}_1)), \tag{2.61}$$

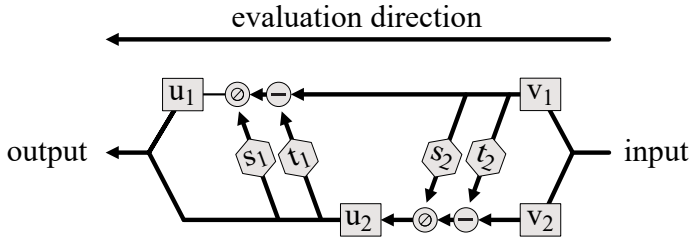
and $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ from $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ in the backward pass:

$$\boldsymbol{u}_2 = (\boldsymbol{v}_2 - t_1(\boldsymbol{v}_1)) \odot \exp(-s_1(\boldsymbol{v}_1)) \tag{2.62}$$

$$\boldsymbol{u}_1 = (\boldsymbol{v}_1 - t_2(\boldsymbol{u}_2)) \odot \exp(-s_2(\boldsymbol{u}_2)). \tag{2.63}$$



(a) Invertible coupling block in forward direction.



(b) Invertible coupling block in backward direction.

Figure 2.8: The fundamental building blocks of Invertible Neural Networks (INNs) are coupling blocks. (a) These blocks divide the input into two components, $\boldsymbol{u} \rightarrow (\boldsymbol{u}_1, \boldsymbol{u}_2)$, which are then alternately transformed using the functions s_i and t_i. (b) The inverse mapping is achieved by reapplying the same functions in a rearranged computational order. Element-wise multiplication and division are denoted by $\odot$ and $\oslash$, respectively. Figure inspired by Ardizonne (2023) [72].

A complete INN is formed by stacking multiple affine coupling layers. To prevent one part of the split (e.g., u_1) from consistently receiving the same input channels across layers, it is standard practice to insert random channel permutations between coupling layers.

Part II

Scientific contributions

Concept overview

3

Quantitative PAI, particularly oximetry, holds great promise for various clinical applications. However, its clinical translation remains limited. Traditional model-based approaches struggle to capture the complexity of the inverse problem, and recent work has highlighted the urgent need for data-driven discovery [11]. A fundamental barrier remains: the field lacks high-quality, annotated datasets with reference values for the optical or functional properties that PAI aims to quantify.

This thesis tackles that barrier through a pipeline of five key contributions as illustrated in Figure 3.1. This chapter provides a conceptual overview of the thesis, outlining how each of the five publications contributes to addressing the overarching challenge of data scarcity in PAI. The purpose is to contextualize and summarize the individual works before presenting each publication.

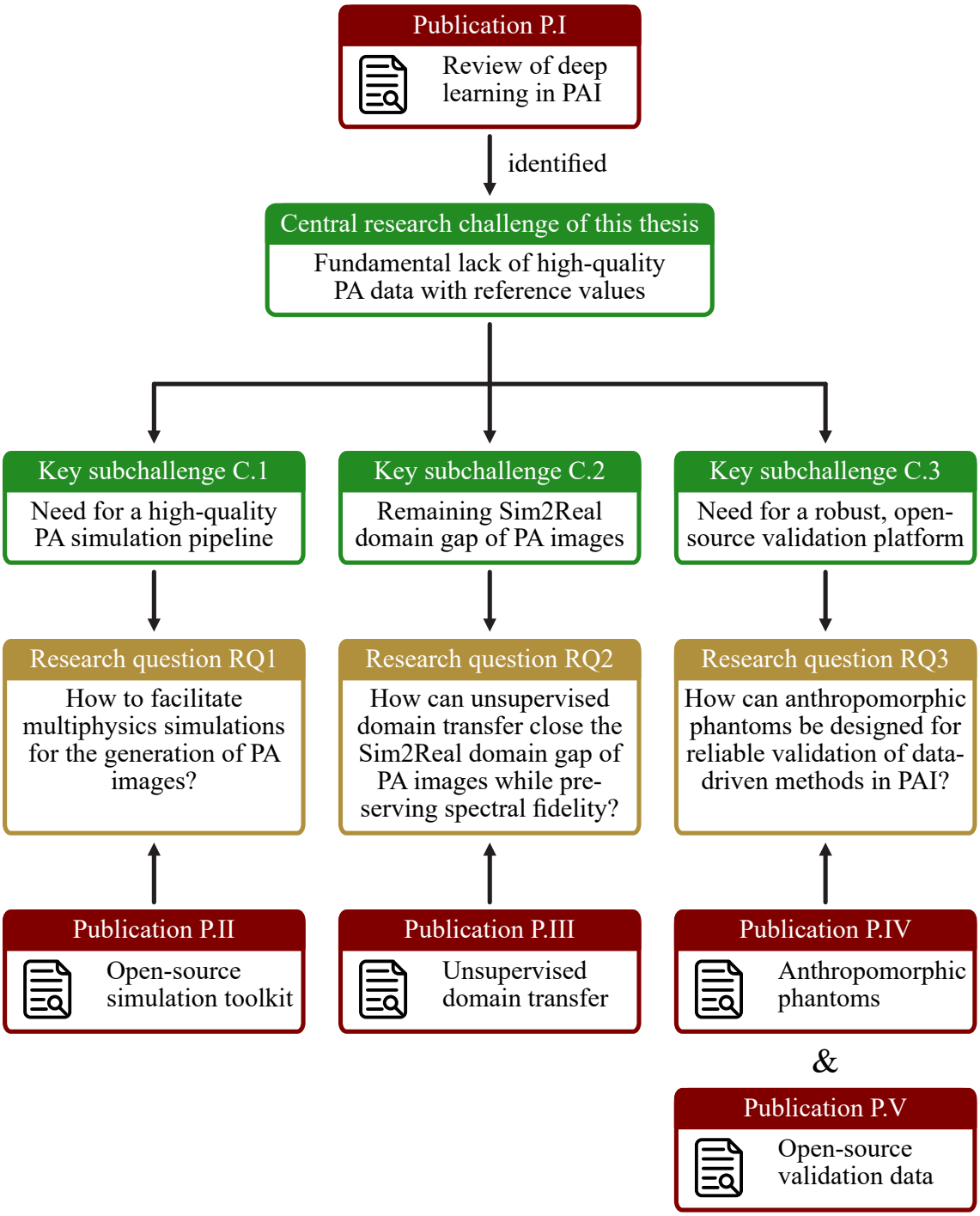


Figure 3.1: This thesis addresses data shortage in photoacoustic imaging (PAI) with five key paper contributions (P.I - P.V). P.I identifies the core challenge tackled by this thesis: The lack of high-quality annotated data. P.II, P.III, and the combination of P.IV and P.V each tackle a specific key subchallenge and address a corresponding research question (RQ) posed in this thesis.

P.I: Review of deep learning in photoacoustic imaging

Establishing the core hypothesis: lack of reference data hinders progress in PAI.

The main body of this thesis begins by systematically investigating the role of deep learning in PAI through a review of the field [12]. This review forms the conceptual foundation for the thesis and identifies critical gaps:

- "One of the core bottlenecks of the application of deep learning algorithms to PAI is the lack of reliable, high-quality experimental training data. For this reason, about 65% of deep learning papers in PAI rely on simulated data for supervised algorithm training."
- "... data-driven approaches to the optical inverse problem have primarily relied on high-fidelity simulations and have not yet been successfully applied in vivo."
- "This is most likely caused by the large domain gap between synthetic and experimental PA images."
- "While the initial results are promising and encouraging, prospective clinical validation studies of such techniques, an integral part of method validation, have not been conducted."

These findings underscore the central hypothesis of the thesis: methodological progress in data-driven PAI is fundamentally limited by the absence of well-annotated, realistic data.

P.II: Open-source simulation toolkit

Digital twin-based open-source simulation framework enables synthetic data generation based on state-of-the-art methods (SIMPA).

To address the first major challenge C.1, generating large-scale, annotated data, the Simulation and Image Processing for Photonics and Acoustics (SIMPA) toolkit [73] was introduced, an open-source framework for PA image simulation. Prior to this work, simulation workflows were fragmented across disconnected toolkits for light transport, acoustic propagation, and image reconstruction. Integrating them into a coherent pipeline was laborious and error-prone.
SIMPA provides a unified, modular simulation pipeline with adapters for all core modeling steps from digital device and tissue twin generation to optical modeling, acoustic propagation, and reconstruction. It is accessible to non-experts while remaining highly customizable for advanced users. As a result, SIMPA enables the efficient generation of fully labeled simulated datasets, laying the groundwork for data-driven algorithm development.

P.III: Unsupervised domain transfer

Conditional invertible neural network (cINN)-based domain transfer closes the simulation-to-real gap while preserving spectral fidelity.

Although simulations provide ground truth data, they exhibit a domain gap to real measurements, limiting the generalizability of trained models. The objective of this work is to close the domain gap between simulation and experimental data.
To overcome limitations of CycleGAN-based domain transfer, such as mode collapse and spectral distortion, an approach based on cINNs combined with GANs [74] was introduced. Unlike adversarial training, cINNs use maximum likelihood estimation to ensure stable training and cycle-consistency by design. The method is conditioned on the image domain and structural features to preserve both realism and spectral fidelity. Applied to two spectral imaging modalities, the approach substantially improves downstream classification performance compared to model-based simulations.

P.IV: Anthropomorphic phantoms

Anthropomorphic phantoms enable reference data for experimental validation.

The third major challenge concerns the absence of a standardized validation strategy for data-driven PAI methods. *In vivo* optical or functional properties can not be measured directly, making rigorous validation difficult.
To address this, anatomically realistic, tissue-mimicking forearm phantoms [75] with five sO_2 levels are fabricated. These phantoms are designed specifically for algorithm validation, enabling repeatable experiments with known measured reference. Starting from clinical magnetic resonance (MR) data, a workflow is developed for generating stable phantoms with blood-mimicking vessel inclusions and realistic optical and acoustic properties. This represents a shift from generic phantom design toward application-specific, validation-oriented fabrication.

P.V: Open-source validation data

Open-source dataset provides common framework for data-driven method validation in PAI.

Building upon the fabrication of the phantoms, Simulated photoacoustic and experimental data for comparing tissue realism and quantitative estimation (SPECTRE) is presented, the first open-access, paired and unpaired multispectral dataset including functional references comprising:

- Experimental PA images from ten anthropomorphic phantoms
- Matched simulations based on known optical properties from the same phantoms
- A probabilistic simulation dataset modeling physiological variability

Two exemplary use cases are demonstrated, domain transfer and oximetry, showing how SPECTRE facilitates validation and performance analysis of data-driven methods.

4 Publications

This thesis is presented in a cumulative form. In accordance with the rules and regulations of the Department of Physics and Astronomy of Heidelberg University, as such, the five publications (P.I - P.V) featured in this work constitute the main body of this thesis. In this section, these five publications are presented with detailed contribution statements. I am the principal or co-principal author of four of the publications. The content and format of the publications have not been altered. For a full body of work published during the project of this thesis, please refer to the *List of publications*.

4.1 P.I: Review of deep learning in photoacoustic imaging

Deep learning for biomedical photoacoustic imaging: A review

Authors: Janek Gröhl, Melanie Schellenberg, **Kris K. Dreher**, Lena Maier-Hein
Status: Published
Journal: Photoacoustics
DOI: doi.org/10.1016/j.pacs.2021.100241

Contributions: JG and LMH conceived and shared administration of the project. **All authors** conceptualized the project and contributed to methodology. JG, MS and **KKD** conducted the investigation for a third of all papers each, i.e. reading through papers, summarizing and categorizing target application, methodology, results, and conclusions.
All authors contributed to reviewing and editing the manuscript. LMH acquired the funding and ensured the project's scientific rigor and relevance.

Photoacoustics 22 (2021) 100241

Contents lists available at ScienceDirect

Photoacoustics

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

Review article

Deep learning for biomedical photoacoustic imaging: A review

Janek Gröhl [a,b], Melanie Schellenberg [a], Kris Dreher [a,c], Lena Maier-Hein [a,b,d,*]

[a] German Cancer Research Center, Computer Assisted Medical Interventions, Heidelberg, Germany
[b] Heidelberg University, Medical Faculty, Heidelberg, Germany
[c] Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany
[d] Heidelberg University, Faculty of Mathematics and Computer Science, Heidelberg, Germany

ARTICLE INFO

Keywords:
Photoacoustic imaging
Photoacoustic tomography
Optoacoustic imaging
Deep learning
Signal quantification
Image reconstruction

ABSTRACT

Photoacoustic imaging (PAI) is a promising emerging imaging modality that enables spatially resolved imaging of optical tissue properties up to several centimeters deep in tissue, creating the potential for numerous exciting clinical applications. However, extraction of relevant tissue parameters from the raw data requires the solving of inverse image reconstruction problems, which have proven extremely difficult to solve. The application of deep learning methods has recently exploded in popularity, leading to impressive successes in the context of medical imaging and also finding first use in the field of PAI. Deep learning methods possess unique advantages that can facilitate the clinical translation of PAI, such as extremely fast computation times and the fact that they can be adapted to any given problem. In this review, we examine the current state of the art regarding deep learning in PAI and identify potential directions of research that will help to reach the goal of clinical applicability.

1. Introduction

Photoacoustic imaging (PAI) is a comparatively young and rapidly emerging imaging modality that promises real-time, noninvasive, and radiation-free measurement of optical tissue properties [1]. In contrast to other optical imaging modalities, PAI induces the emergence of acoustic signals to enable structural imaging of chromophores – molecular structures that absorb light – up to several centimeters deep into the tissue. Typical reported penetration depths range from 1 cm up to 6 cm [2–4], heavily depending on the imaged tissue. This high depth penetration is possible because the acoustic scattering of the arising sound waves is orders of magnitude smaller than the optical scattering of the incident light in biological tissue. The underlying physical principle for signal generation is the *PA effect* [5]. It is induced by extremely short light pulses that cause an initial pressure rise p_0 inside the tissue. The initial pressure $p_0 = \Gamma \cdot \mu_a \cdot \phi$ is proportional to the optical absorption coefficient μ_a, the local light fluence ϕ, and the temperature-dependent Grüneisen parameter Γ. The deposited energy is released in form of sound waves that can be measured as time-series pressure data $p(t) = A(p_0, \theta)$ with appropriate acoustic detectors, such as ultrasound transducers. Here, acoustic forward operator A operates on the initial pressure distribution p_0 taking into consideration the acoustic properties θ (such as the speed of sound, density, or the acoustic attenuation) of the medium.

Due to its rapid development, PAI has seen various clinical application attempts over the last few years. Among these, cancer research is a field where PAI shows serious potential [3,6–15]. In this use case, hemoglobin is the enabling endogenous chromophore, due to amplified and sustained angiogenesis [16] being one of the hallmarks of cancer and due to the cancer cells' increased metabolism, which potentially induces a decrease in local blood oxygenation [17]. Furthermore, because inflammatory processes also change the hemodynamic behavior of tissue, PAI is also used for imaging of inflamed joints [18–20] or staging of patients with Crohn's disease [21–23]. To further increase the potential of PAI, it is also applied in combination with other imaging modalities, especially ultrasound imaging [15,24–28]. PAI is further used for brain imaging [29–33] or surgical and interventional imaging applications, such as needle tracking [34,35].

The signal contrast of PAI is caused by distinct wavelength-dependent absorption characteristics of the chromophores [36]. But to exploit information of μ_a for answering clinical questions, open research questions remain that can be categorized into four main areas. In the following, we explain these four major categories and summarize their principal ideas.

Acoustic inverse problem. The most pressing problem concerns the reconstruction of an image S from recorded time-series data by

* Corresponding authors at: German Cancer Research Center, Computer Assisted Medical Interventions, Heidelberg, Germany.
E-mail address: l.maier-hein@dkfz-heidelberg.de (L. Maier-Hein).

https://doi.org/10.1016/j.pacs.2021.100241
Received 5 November 2020; Received in revised form 18 January 2021; Accepted 20 January 2021
Available online 2 February 2021

estimating the initial pressure distribution p_0 from $p(t)$. This problem is referred to as the acoustic inverse problem. To this end, an inverse function A^{-1} for the acoustic operator A needs to be computed in order to reconstruct a signal image $S = A^{-1}(p(t)) \approx p_0 = \mu_a \cdot \phi \cdot \Gamma$ that is an approximation of p_0. Typical examples of algorithms to solve this problem are the universal back-projection [37], delay-and-sum [38], time reversal [39], or iterative reconstruction schemes [40]. While the acoustic inverse problem can be well-posed in certain scenarios (for example by using specific detection geometries) and thus can have a unique solution, several factors lead to considerable difficulties in solving it. These include wrong model assumptions [41], limited-view [42] and limited-bandwidth detectors [43], or device modeling [44] and calibration errors [45].

Image post-processing. PAI, in theory, has exceptionally high contrast and spatial resolution [46]. Because the acoustic inverse problem is ill-posed in certain settings and because of the presence of noise, many reconstructed PA images suffer from distinct artifacts. This can cause the actual image quality of a PA image to fall short of its theoretical potential. To tackle these problems, image post-processing algorithms are being developed to mitigate the effects of artifacts and noise and thus improve overall image quality.

Optical inverse problem. Assuming that a sufficiently accurate reconstruction of p_0 from $p(t)$ has been achieved, the second principle problem that arises is the estimation of the underlying optical properties (most importantly the absorption coefficient μ_a). It is an inverse problem and is referred to as the optical inverse problem. Furthermore, the problem has proven to be exceptionally involved, which can be derived by the fact that methods to solve the problem have not been successfully applied to *in vivo* data yet. It belongs to the category of ill-posed inverse problems, as it does not necessarily possess a unique solution. Furthermore, several other factors make it hard to tackle, including wrong model assumptions [41], non-uniqueness and non-linearity of the problem [47], spectral coloring [48], and the presence of noise and artifacts [49]. Quantification of the absorption coefficient has, for example, been attempted with iterative reconstruction approaches [50], via fluence estimation [51], or by using machine learning-based approaches [52].

Semantic image annotation. Based on the diagnostic power of optical absorption it is possible to generate semantic image annotations of multispectral PA images and a multitude of methods for it are being developed to specifically tackle questions of clinical relevance. To this end, algorithms are being developed that are able to classify and segment multispectral PA images into different tissue types and that can estimate clinically relevant parameters that are indicative of a patient's health status (such as blood oxygenation). Current approaches to create such semantic image annotations suffer from various shortcomings, such as long computation times or the lack of reproducibility in terms of accuracy and precision when being applied to different scenarios.

Simultaneously to the rapid developments in the field of PAI, deep learning algorithms have become the *de facto* state of the art in many areas of research [53] including medical image analysis. A substantial variety of medical applications include classical deep learning tasks such as disease detection [54], image segmentation [55], and classification [56]. Recently, deep learning has also found entrance into the field of PAI, as it promises unique advantages to solve the four listed problems, thus promoting clinical applicability of the developed methods. One further prominent advantage of deep learning is the extremely fast inference time, which enables real-time processing of measurement data.

This review paper summarizes the development of deep learning in PAI from the emergence of the first PA applications in 2017 until today and evaluates progress in the field based on the defined task categories. In Section 2, we outline the methods for our structured literature research. General findings of the literature review including the topical foci, data acquisition techniques, used simulation frameworks as well as network architectures are presented in Section 3. The reviewed literature is summarized according to the four principal categories in Sections 4–7. Finally, the findings are discussed and summarized in Section 8.

2. Methods of literature research

Above, we described the PAI-specific challenges that deep learning can be applied to and thus divided the topic into four major categories: *I. Acoustic inverse problem*, *II. Image post-processing*, *III. Optical inverse problem*, and *IV. Semantic image annotation*. We conducted a systematic literature review for the period between January 2017 and September 2020 and assigned each identified paper to the most suitable categories. For the search, we used several scientific search engines: *Google Scholar*, *IEEE Xplore*, *Pubmed*, *Microsoft Academic Search Engine*, and the *arXiv* search function with the search string **("Deep learning" OR "Neural Network") AND ("photoacoustic" OR "optoacoustic")**. The search results were then refined in a multi-step process (see Fig. 1).

First, potential candidates were identified based on an initial search using their title, as well as the overview presented by the search engine. The search results were complemented by adding additional papers found by means other than the named search engines. For this purpose, we specifically examined proceedings of relevant conferences, websites of key authors we identified, and websites of PA device vendors. Finally, non-relevant papers were excluded by removing Journal/Proceeding paper duplicates and by abstract scanning to determine whether the found papers match the scope of this review. Using the abstract, we excluded papers that did not apply deep learning, and those that did not

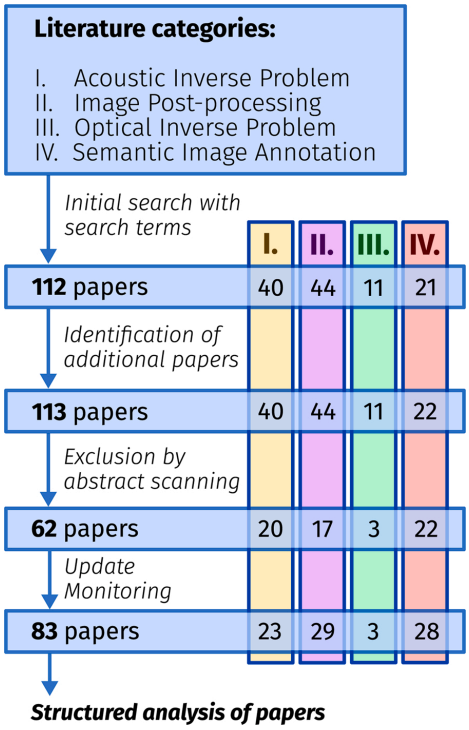


Fig. 1. Overview of the literature review algorithm. First, potentially fitting papers are identified based on an initial search. The search results are complemented by adding additional papers found by other means than the search engines, and finally, non-relevant papers are excluded by removing duplicates and by abstract scanning.

match the scope of *biomedical* PAI. The remaining articles were systematically examined by the authors using a questionnaire to standardize the information that was to be extracted. While writing the paper, we continuously monitored the mentioned resources for new entries until the end of December 2020.

In total, applying the search algorithm as detailed above, we identified 83 relevant papers (excluding duplicates and related, but out-of-scope work) that have been published since 2017.

3. General findings

The application of deep learning techniques to the field of PAI has constantly been accelerating over the last three years and has simultaneously generated a noticeable impact on the field.

Topical foci. After categorization of the papers into the predetermined four areas, the papers were arranged into thematic subcategories (see Fig. 2). Papers related to the Acoustic Inverse Problem (Section 4) generally focus on the reconstruction of PA images from raw time-series pressure data but also related topics, such as dealing with limited-view or limited data settings, as well as the estimation of the speed of sound of tissue. The image post-processing (Section 5) category entails papers that deal with image processing algorithms that are being applied after image reconstruction. The aim of such techniques usually is to improve the image quality, for example by noise reduction or artifact removal. The three papers assigned to the optical inverse problem (Section 6) deal with the estimation of absolute chromophore concentrations from PA measurements. Finally, papers dealing with semantic image annotation (Section 7) tackle use cases, such as the segmentation and classification of tissue types or the estimation of functional tissue parameters, such as blood oxygenation.

Data. Data is key to successfully apply machine learning techniques to any given problem. We analyzed the usage of data in the reviewed papers and summarized the findings in Fig. 3.

Training. The number of training data ranged from 32 to 286,300 samples with a median number of training samples of 2000. As evident from these findings, one of the core bottlenecks of the application of deep learning algorithms to PAI is the lack of reliable experimental training data. This can in particular be caused by a lack of ground truth information on the underlying optical tissue properties or the underlying initial pressure distribution when acquiring experimental measurements. To address this issue, researchers make heavy use of simulated data and as a matter of fact, approximately 65% of papers relied exclusively on these for training the neural network. Table 1 shows the distribution of papers that use experimental data. The table shows that the lack of experimental training data is particularly emphasized for the optical and acoustic inverse problem. In contrast to the other tasks, where manual image annotations can be used as a ground truth reference, the underlying initial pressure distribution or optical tissue properties are generally not known in experimental settings. We have identified three main strategies for generating synthetic training data in this review: random, model-based, and reference-based data generation:

1. *Random data generation.* The first and simplest strategy generates data by creating completely random distributions of the optical and acoustic properties that are necessary for the simulation framework [57]. Here, usually, a Gaussian distribution of the parameters in question is assumed and no dedicated structures are added to the data.
2. *Model-based data generation.* Training data is created by defining geometrical structures that are assigned optical and acoustic properties according to a hand-crafted model [58]. Such a model might include literature references e.g. for the size, shape, and properties of typical absorbers in tissue. For the generation of training data, many different instances of the model are created that all yield different distributions of chromophores.
3. *Reference-based data generation.* For the reference-based approach, reference images of different imaging modalities are taken as the basis for data generation [59]. They are processed in a way that allows for their direct usage to either create distinct segmentation patterns of, for example, vessels or as the initial pressure distribution for subsequent acoustic forward modeling.

Naturally, researchers also utilized combinations of these approaches, including training on a large data set of simulated data and utilizing a smaller experimental data set to adjust the neural network to the experimental data distribution in a process called *transfer learning* [59,60].

Testing. In the field of medical imaging, only few prospective studies warrant reliable insights into the fidelity of deep learning methods [54]. One of the major problems is that algorithms are not directly usable by clinicians due to technical or bureaucratic limitations [61]. Given the fact that most approaches use simulated data to train their algorithms, there is a high probability that many of the presented algorithms – while yielding superb results on the publication data – could fail in a clinical scenario. This can be attributed to the fact that training data can suffer from several shortcomings compared to the data distribution in reality, such as a significant difference in the data (domain gap) [62], an insufficient number of samples (sparsity) [63], or a selection bias [64]. Fig. 3 shows that in PAI 50% of papers tested their deep learning approaches on multiple data sets that are significantly different from the training data distribution. Nearly all of these papers test their approaches on experimental data, and about 35% of the examined papers test on *in vivo* data.

Neural network architectures. Convolutional neural networks

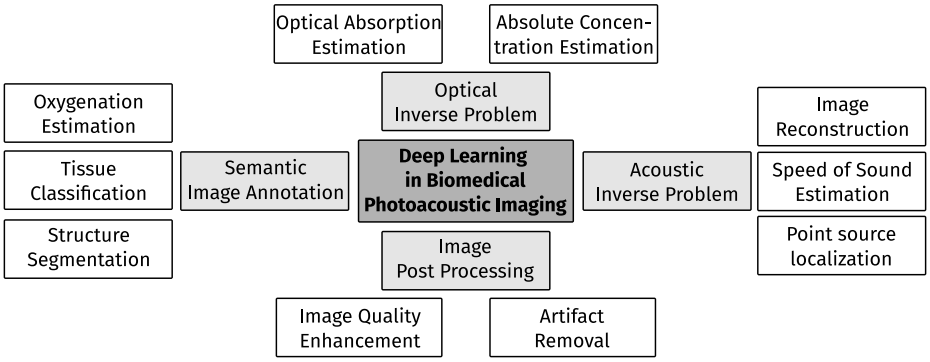


Fig. 2. Overview over the topical foci of current research towards applying deep learning algorithms to problems in biomedical PAI.

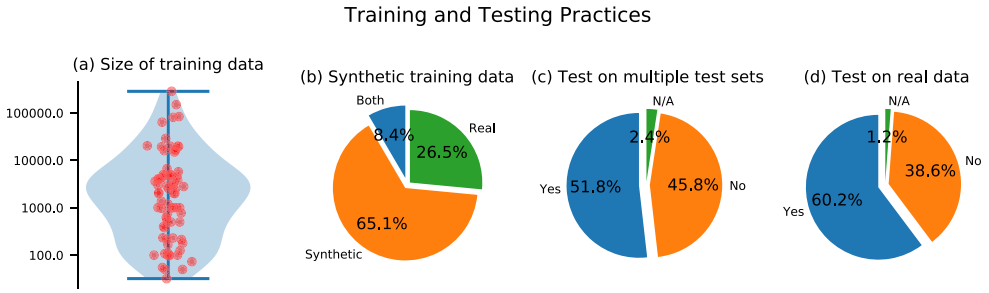


Fig. 3. Analysis of the data used in the reviewed papers. (a) The distribution of the number of samples in the training data set, (b) the percentage of papers working with synthetic or experimental training data, (c) the percentage of papers that tested their approaches on multiple data sets including test data from a data distribution different than the training data and (d) how many papers tested their approach on real data.

Table 1
Overview of the findings for training and test data used in the reviewed papers. The table shows the absolute and relative number of papers that use experimental data for testing or for training.

Problem	Exp. test data	Exp. train data
Acoustic inverse problem	10 (43%)	1 (4%)
Image post-processing	23 (79%)	14 (48%)
Optical inverse problem	1 (33%)	1 (33%)
Semantic image annotation	14 (54%)	11 (42%)

(CNNs) [65] are currently the state-of-the-art method in deep learning-based PAI. Here, especially the U-Net [66] architecture is highly popular and has been used or compared to in most of the reviewed papers.

Table 2 shows the frequency of the main model architecture employed in each paper. It should be noted that often modified versions of the base architectures have been used to specifically fit the target application or to encompass novel ideas.

Simulation frameworks. Given the necessity to create synthetic data sets for algorithm training, it is crucial to realistically simulate the physical processes behind PAI. To this end, we have identified several eminent open-source or freely available frameworks that are being utilized in the field and briefly present five of them here:

1) The *k-Wave* [39] toolbox is a third-party MATLAB toolbox for the simulation and reconstruction of PA wave fields. It is designed to facilitate realistic modeling of PAI including the modeling of detection devices. As of today, it is one of the most frequently used frameworks in the field and is based on a k-space pseudo-spectral time-domain solution to the PA equations.

Table 2
The frequency of the main deep learning architecture that was used in each paper. CNN = Convolutional neural network; FCNN = fully-connected neural network; ResNet = residual neural network; DenseNet = dense neural network; RNN = recurrent neural network; INN = invertible neural network.

Architecture	Frequency
U-Net	43 (52%)
CNN	18 (22%)
FCNN	9 (11%)
ResNet	5 (6%)
DenseNet	3 (4%)
RNN	1 (1%)
INN	1 (1%)
Other	3 (4%)

2) The *mcxyz* [67] simulation tool uses a Monte Carlo model of light transport to simulate the propagation of photons in heterogeneous tissue. With this method, the absorption and scattering properties of tissue are used to find probable paths of photons through the medium. The tool uses a 3-dimensional (3D) Cartesian grid of voxels and assigns a tissue type to each voxel, allowing to simulate arbitrary volumes.
3) The *Monte Carlo eXtreme* [68] (MCX) tool is a graphics processing unit (GPU)-accelerated photon transport simulator. It is also based on the Monte Carlo model of light transport and supports the simulation of arbitrarily complex 3D volumes using a voxel domain, but is also capable of simulating photon transport for 3D mesh models (in the MMC version). Its main advantage is the support of GPU acceleration using a single or multiple GPUs.
4) The *NIRFAST* [69] modeling and reconstruction package was developed to model near-infrared light propagation through tissue. The framework is capable of single-wavelength and multi-wavelength optical or functional imaging from simulated and measured data. It recently integrated the NIRFAST optical computation engine into a customized version of 3DSlicer.
5) The *Toast++* [70] software suite consists of a set of libraries to simulate light propagation in highly scattering media with heterogeneous internal parameter distribution. Among others, it contains numerical solvers based on the finite-element method, the discontinuous Galerkin discretization scheme, as well as the boundary element method.

4. Acoustic inverse problem

The acoustic inverse problem refers to the task of reconstructing an image of the initial pressure distribution from measured time-series pressure data. Reconstructing a PA image from time-series data constitutes the main body of work, either by *enhancing existing model-based approaches* (39% of papers) or by performing *direct image reconstruction* (39% of papers). Furthermore, auxiliary tasks are being examined as well, such as the *localization of wavefronts* from point sources (13% of papers) and the *estimation of the speed of sound* of the medium (9% of papers). Information on these parameters is important to achieve optimal image reconstruction, thereby enhancing the image quality and improving the image's usefulness in clinical scenarios. The evaluation metrics that were used to assess the quality of the reconstruction results and their relative frequencies are shown in Table 3.

In total, we identified 23 papers that tackle the acoustic inverse problem, all of which use simulated PA data for training. Surprisingly, approximately 43% of papers presented results on experimental data either using phantoms or *in vivo* (animal or human) measurements. In the following, we summarize the literature partitioned into the already mentioned sub-topics: deep learning-enhanced model-based image

Table 3
Listing of the metrics that were reported in the literature to evaluate the quality of the image reconstruction. PSNR = Peak signal-to-noise ratio; SSIM = structural similarity; MSE = mean squared error; SNR = signal-to-noise ratio; MAE = mean absolute error; RMSE = root mean squared error; FWHM = full width at half maximum; CC = correlation coefficient.

Metric	Frequency
PSNR	12 (52%)
SSIM	9 (39%)
Relative MSE	5 (22%)
SNR	4 (17%)
MAE	4 (17%)
Relative MAE	3 (13%)
MSE	2 (9%)
RMSE	1 (4%)
FWHM	1 (4%)
CC	1 (4%)

reconstruction, direct image reconstruction, point source localization, and speed of sound estimation.

4.1. Deep learning-enhanced model-based image reconstruction

The central idea is to leverage the flexibility of deep learning to enhance already existing model-based reconstruction algorithms [71,72], by introducing learnable components. To this end, Schwab et al. [73] proposed an extension of the weighted universal back-projection algorithm. The core idea is to add additional weights to the original algorithm, with the task of the learning algorithm then being to find optimal weights for the reconstruction formula. By learning the weights, the authors were able to reduce the error introduced from limited view and sparse sampling configurations by a factor of two. Furthermore, Antholzer et al. [74,75] and Li et al. [76] leveraged neural networks to learn additional regularization terms for an iterative reconstruction scheme. Hauptmann et al. [59] demonstrated the capability for CNNs to perform iterative reconstruction by training a separate network for each iteration step and integrating it into the reconstruction scheme. The authors showed that since their algorithm was trained on synthetic data, several data augmentation steps or the application of transfer learning techniques were necessary to achieve satisfactory results. Finally, Yang et al. [77] demonstrated the possibility to share the network weights across iterations by using recurrent inference machines for image reconstruction.

Key insights: An interesting insight shared in one of the papers by Antholzer et al. [74] was that model-based approaches seem to work better for "exact data", while deep learning-enhanced methods outperform purely model-based approaches on noisy data. This makes the application of deep learning techniques very promising for the typically noisier and artifact-fraught experimental data [49,78]. On the other hand, currently employed deep learning models do not seem to generalize well from simulated to experimental data as evident from the fact that only 43% of papers tested their method on experimental data (cf. Table 1). Ideally, the algorithms would have to be trained on experimental data.

4.2. Direct image reconstruction

The principal idea of direct image reconstruction with deep learning is to reconstruct a photoacoustic image directly from the time series data. In addition to the time series data, some approaches also use hand-crafted features or reference reconstructions obtained from a conventional reconstruction algorithm for regularization. An overview of these principle ideas is summarized in Fig. 4.

The first approaches to direct image reconstruction with CNNs were proposed in 2018 by Waibel et al. [42] and Anas et al. [79]. Modified U-Net architectures were used by Waibel et al. [42] and Lan et al. [80] to estimate the initial pressure distribution directly from time-series pressure data, whereas Anas et al. [79] used a CNN architecture with dense blocks. Furthermore, Lan et al. [81–83] proposed a method based on a generative adversarial network [84] approach that – in addition to time-series data – also uses a reconstructed PA image as additional information to regularize the neural network. Guan et al. [85] compared implementations of all these techniques to assess their merit in brain imaging within a neurological setting. They compared an algorithm that directly estimates the reconstructed image from time-series data, a post-processing approach, as well as a custom approach with hand-rafted feature vectors for the model. Their results show that adding additional information improves the quality of the reconstructed image, that iterative reconstruction generally worked best for their data, and that deep learning-based reconstruction was faster by 3 orders of magnitude. Kim et al. [86] propose a method that uses a look-up-table-based image transformation to enrich the time series data

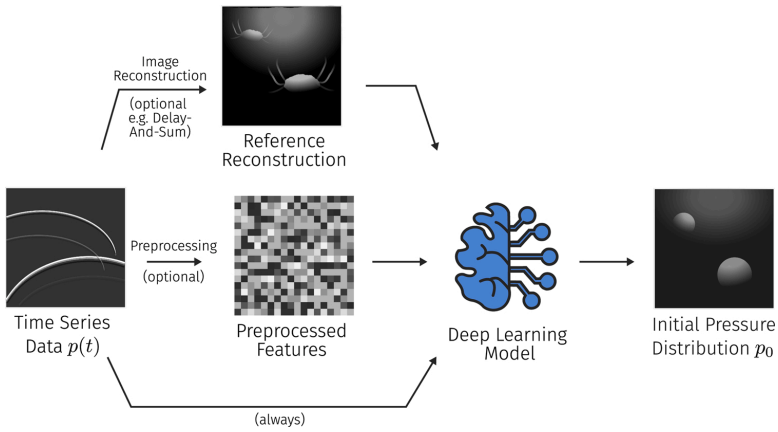


Fig. 4. Visualization of the principal approaches to deep learning-based PA image reconstruction. The time-series data is either given directly to a neural network, or after preprocessing steps, such as reference reconstructions or the calculation of hand-crafted feature maps. The goal of the reconstruction is to estimate the underlying initial pressure distribution.

before image reconstruction with a U-Net. With this method they are able to achieve convincing reconstruction results both on phantom and in vivo data sets.

Key insights: In contrast to deep learning-enhanced model-based reconstruction, direct deep learning reconstruction schemes are comparatively easy to train and most of the papers utilize the U-Net as their base architecture. In several works it was demonstrated that the *infusion of knowledge* by regularizing the network with reference reconstructions or additional data from hand-crafted preprocessing steps led to very promising results [81,82], generalizing them in a way that led to first successes on *in vivo* data. Considering that deep learning-based image reconstruction outperforms iterative reconstruction techniques in terms of speed by orders of magnitude (the median inference time reported in the reviewed papers was 33 ms), it is safe to say that these methods can be a promising avenue of further research. It has to be noted that, especially in terms of robustness and uncertainty-awareness, the field has much room for improvement. For example, Sahlström et al. [44] have modeled uncertainties of the assumed positions of the detection elements for model-based image reconstruction, but no comparable methods were applied in deep learning-based PAI as of yet.

4.3. Point source localization

The localization of the spatial position of point sources from time-series PA measurements was identified as a popular sub-task concerning PA image reconstruction. An algorithm for this could for example be used for the automatic detection and localization of point absorbers, such as needle tips, in a PA image. The general idea is to take time-series data to either regress numerical values for the pixel coordinates of the sources of the wavefronts or to output a two-dimensional map of the possible source locations (see Fig. 5).

To this end, Reiter et al. [87] presented an approach that uses a CNN to transform the time-series data into an image that identifies the 2-dimensional (2D) point localization of the wavefront origin. They further use this approach to distinguish between signals and artifacts in time-series data. Johnstonbaugh et al. [88] also use a CNN in an encoder-decoder configuration to reconstruct the PA signal into an image containing a single point source. A similar architecture proposed by the same group [89] is also used to process the time-series data and output cartesian coordinates of the point source location.

Key insights: Similar to the deep learning-based direct reconstruction methods, methods for point source localization are exceptionally easy to train and can even be trained on *in vitro* experimental data. This ease of accessibility made this task the first application of deep learning in PAI [87]. However, the integrability of these methods into clinical practice and their future impact beyond certain niche applications is questionable because *in vivo* scenarios do typically not exclusively consist of point sources but comprise a very complex and heterogeneous distribution of chromophores.

4.4. Speed of sound estimation

A correct estimate of the speed of sound within the medium is an important constituent to successful image reconstruction. We identified two papers that explicitly incorporated the estimation of the speed of sound into their reconstruction. Shan et al. [90] used a CNN to reconstruct the initial pressure distribution as well as the speed of sound simultaneously from the time-series data and Jeon et al. [91] trained a U-Net to account for the speed of sound aberrations that they artificially introduced to their time-series data.

Key insights: Automatically integrating estimates of the speed of sound into the image reconstruction algorithm can substantially enhance image quality and hence is an interesting direction of research. Nevertheless, the formulation of a corresponding optimization problem is inherently difficult, as it is not straightforward to assess the influence of a speed of sound mismatch on a reconstruction algorithm. Furthermore, the validation of these methods is difficult, as there typically is no *in vivo* ground truth available.

5. Image post-processing

Being a comparatively young imaging modality, PAI still suffers from distinct artifacts [49]. These can have multiple origins and are primarily caused by hardware limitations such as light absorption in the transducer membrane or fluctuations in the pulse laser energy [78]. Other issues can also lead to decreased image quality, such as under-sampling or limited-view artifacts, as well as other influences such as motion artifacts or artifacts specific to the reconstruction algorithm (see Fig. 6). Research in the field of using post-processing algorithms can broadly be divided into two main areas: the elimination of artifacts (Section 5.1) which mostly encompass systematic error sources and the enhancement of image quality (Section 5.2) which is lost mainly through stochastic error sources.

The evaluation metrics that were used to assess the quality of the post-processing results and their relative frequencies are shown in Table 4.

5.1. Artifact removal

One principal approach to speed up image reconstruction is to use sparse data that only contains a fraction of the available time-series data. While this potentially leads to a significant increase in reconstruction speed, it comes with a cost in form of the deterioration of the image quality and the introduction of characteristic under-sampling artifacts. Several groups [92–99] have shown that a large portion of these artifacts

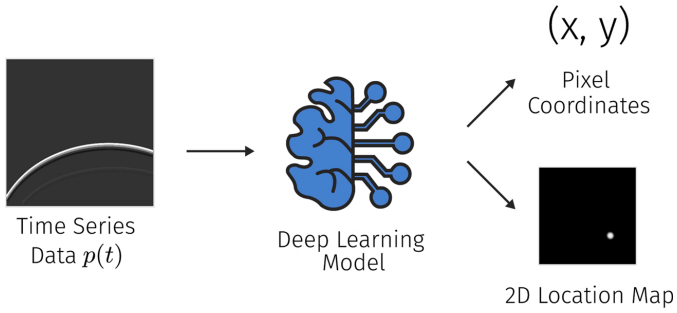


Fig. 5. Approaches for point source localization use time-series data as input to estimate either the pixel coordinates of the point of origin of the pressure wave or a heat map containing the probability of the source being in a certain location of the image.

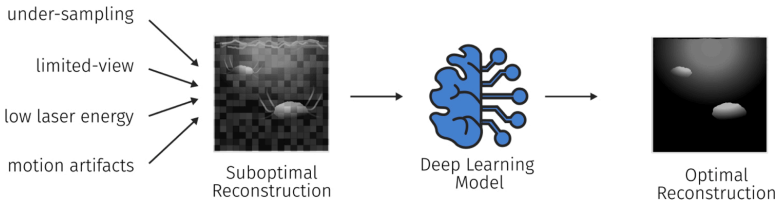


Fig. 6. Post-processing techniques are tasked to improve the image quality of a reconstructed PA image. The image quality can be reduced by many factors including under-sampling, limited-view artifacts, low laser energy, or the presence of motion during the measurement.

Table 4
Listing of the metrics that were reported in the literature to evaluate the quality of the post processed images. SSIM = Structural similarity; PSNR = peak signal-to-noise ratio; MSE = mean squared error; SNR = signal-to-noise ratio; CNR = contrast-to-noise ratio; MAE = mean absolute error; RMSE = root mean squared error; PCC = Pearson correlation coefficient; EPI = edge preserving index; NCC = normalized correlation coefficient.

Metric	Frequency
SSIM	21 (72%)
PSNR	19 (66)
MSE	6 (21%)
SNR	6 (21%)
CNR	5 (17%)
MAE	3 (10%)
Relative MSE	3 (10%)
RMSE	3 (10%)
PCC	2 (7%)
EPI	1 (3%)
NCC	1 (3%)

can be recovered using deep learning techniques. A core strength of such approaches is that experimental PA data can be utilized for training, by artificially undersampling the available channels and training the algorithm to predict the reconstructions from (1) full data, (2) sparse data, or (3) limited-view data [100–106].

Reflection artifact can be introduced by the presence of acoustic reflectors in the medium (for example air). Allman et al. [107] showed that deep learning can be used to distinguish between artifacts and true signals and Shan et al. [108] demonstrated that the technology is also capable of removing such artifacts from the images. Furthermore, Chen et al. [109] introduced a deep learning-based motion correction approach for PA microscopy images that learns to eliminate motion-induced artifacts in an image.

Key insights: For limited-view or limited-data settings, experimental training data can easily be created by artificially constraining the available data, for example, by under-sampling the number of available time series data. On the other hand, for the task of artifact removal, it can be comparatively difficult to train models on *in vivo* experimental settings for different sources of artifacts. This is, because artifacts can have various origins and are also dependent on the specific processing steps. Nevertheless, impressive results of the capability of learning algorithms to remove specific artifacts were demonstrated.

5.2. Image quality enhancement

The quality and resolution of PA images are also limited by several other factors including the limited bandwidth of PA detectors, the influence of optical and acoustic scattering, the presence of noise due to the detection hardware, and fluctuations in the laser pulse-energy.

To remedy this, Gutta et al. [110] and Awasthi et al. [111] proposed methods that aim to recover the full bandwidth of a measured signal. This is achieved by obtaining pairs of full bandwidth and limited bandwidth data using simulations that are used to train a neural network. Since experimental systems are always band-limited, the authors of these works rely on the presence of simulated data. On the other hand, more classical deep learning-based super-resolution algorithms were proposed by Zhao et al. [112,113] to enhance the resolution of PA devices in the context of PA microscopy. For training of super-resolution approaches, the authors are theoretically not restricted by the domain of application and as such can also use data from sources unrelated to PAI. In a similar fashion, Rajendran et al. [114] propose the use of a fully-dense U-Net architecture to enhance the tangential resolution of measurements acquired with circular scan imaging systems. To this end, they use simulated data for training of the algorithm and test their approach on experimental data *in vivo*.

Several approaches have been proposed to enhance the image quality by improving the signal-to-noise ratio of image frames acquired with low energy illumination elements, such as LED-based systems. This has generally been done using CNNs to improve a single reconstructed image, for example by Vu et al. [115], Singh et al. [116], Anas et al. [117], Sharma et al. [118], Tang et al. [119], and Hariri et al. [120] or by using a neural network to fuse several different reconstructions into a higher-quality version, as proposed by Awasthi et al. [121].

Key insights: For the enhancement of image quality, common deep learning tasks from the field of computer vision [122] can be translated to PA images relatively easily, as the algorithms are usually astonishingly straightforward to train and validate. We believe that applying well-established methods from fields adjacent to PAI can be of excellent benefit to the entire field.

6. Optical inverse problem

The optical inverse problem is concerned with estimating the optical tissue properties from the initial pressure distribution. The first method proposed to solve this inverse problem was an iterative reconstruction scheme to estimate the optical absorption coefficient [50]. Over time, the iterative inversion schemes have become more involved [123] and Buchmann et al. [45] achieved first successes towards experimental validation. Recently, data-driven approaches for the optical inverse problem have emerged, including classical machine learning [52] as well as deep learning approaches. A tabulated summary of the identified papers can be found in Table 5.

For the identified deep learning-based approaches, the key objective is to estimate the optical absorption coefficients and subsequently the absolute concentrations of chromophores from the initial pressure distribution (see Fig. 7).

Cai et al. [57] trained a U-Net with residual blocks to estimate the absolute concentration of indocyanine green (ICG) alongside the relative ratio of Hb and HbO_2. To this end, they simulated random smoothed maps of optical tissue properties for training and tested their approach on several simulated data sets, including one created from a digital mouse model [126]. Gröhl et al. [124] trained a total of four U-Net models to estimate fluence and optical absorption from the initial pressure distribution as well as directly from time-series pressure data.

Table 5
Tabulated overview of the identified literature regarding the optical inverse problem. The table shows the publication, the base network architecture, the range of absorption and scattering parameters used for the training data, and the type of data that the approach was validated with.

Publication	Base architecture	Target μ_a [cm^{-1}]	Background μ_a [cm^{-1}]	Background μ_s' [cm^{-1}]	Validation data
Cai et al. [57]	U-Net with residual blocks	N/A	0.2–0.4	5–10	*In silico*
Gröhl et al. [124]	U-Net	2–10	Const. 0.1	Const. 1.5	*In silico*
Chen et al. [125]	U-Net	0.1–2	0.1–0.4	Const. 10	*In vitro*

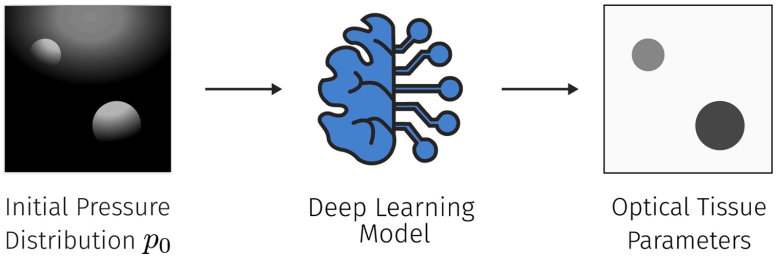


Fig. 7. To solve the optical inverse problem, a neural network is tasked to estimate the underlying optical tissue parameters, primarily the optical absorption coefficient, from the initial pressure distribution p_0.

They also presented a method to estimate the expected error of the inversion, yielding an indicator for the model uncertainty. Their approach was trained and tested on simulated data, which contained tubular structures in a homogeneous background. Finally, Chen et al. [125] trained a U-Net to directly estimate optical absorption from simulated images of initial pressure distribution. They trained and tested their model on synthetic data comprising geometric shapes in a homogeneous background, as well as another model on experimental data based on circular phantom measurements.

Key insights: Model-based methods to tackle the optical inverse problem suffer from the fact that many explicit assumptions have to be made that typically do not hold in complex scenarios [52]. With data-driven approaches, many of these assumptions are only implicitly made within the data distribution, leaving room for a substantial improvement. Obtaining ground truth information on the underlying optical tissue properties *in vivo* can be considered impossible and is exceptionally involved and error-prone even *in vitro* [124]. As such, there has been no application yet to *in vivo* data, leaving the optical inverse problem as one of the most challenging problems in the field of PAI, which is reflected by the comparatively low amount of published research on this topic.

7. Semantic image annotation

While the topical areas until now have mostly considered PA data at a single wavelength, the power of PAI for clinical use cases lies in its

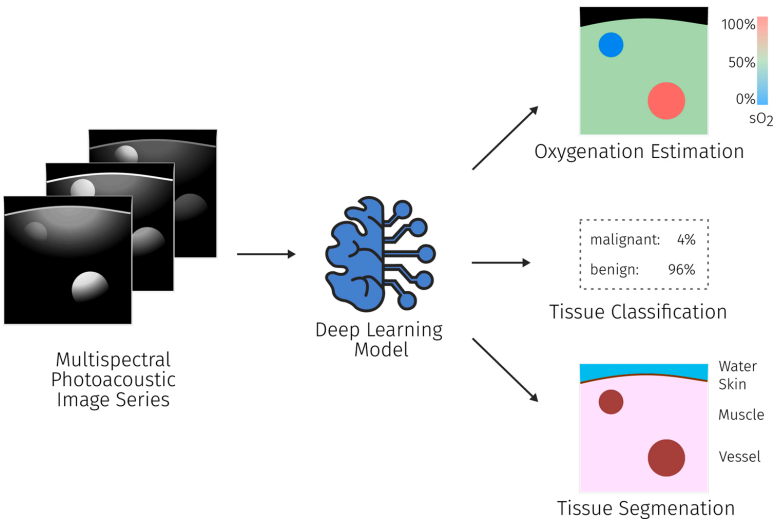


Fig. 8. For semantic tissue annotation (typically multispectral) PA measurements are used as the input and the algorithm is tasked to estimate the desired parameters, such as tissue oxygenation or segmentation maps of different tissue types. The black color in the oxygenation estimation denotes areas where oxygenation values cannot be computed.

ability to discern various tissue properties through analysis of the changes in signal intensity over multiple wavelengths (see Fig. 8). This allows for the estimation of functional tissue properties, especially blood oxygenation (Section 7.1), but also for the classification and segmentation (Section 7.2) of tissues and tissue types.

7.1. Functional property estimation

The estimation of local blood oxygenation sO_2 is one of the most promising applications of PAI. In principle, information on the signal intensities of at least two wavelengths are needed to unmix the relative signal contributions of oxyhemoglobin HbO_2 and deoxyhemoglobin Hb: $sO_2 = HbO_2/(HbO_2 + Hb)$. For PAI, wavelengths around the isosbestic point (≈800 nm) are commonly chosen for this task. Because linear unmixing should not directly be applied to the measured signal intensities due to the non-linear influence of the fluence, a lot of work has been conducted to investigate the applicability of neural networks to tackle this problem. Due to the unavailability of ground truth oxygenation information, the networks are currently being trained exclusively on simulated data. The problem was approached using fully-connected neural networks [58] as well as CNNs [127].

The use of feed-forward fully-connected neural networks was demonstrated by Gröhl et al. [58] to be capable to yield accurate oxygenation estimations *in silico* from single-pixel p_0 spectra. In addition, it was demonstrated that the application of the method to experimental *in vitro* and *in vivo* data yielded plausible results as well. This was also confirmed in an *in silico* study by Nölke et al. [128] using invertible neural networks, showing that using multiple illumination sources can potentially resolve ambiguity in the sO_2 estimates. Olefir et al. [129] demonstrated that introducing prior knowledge to the oxygenation estimation can improve performance. Specifically, the authors introduced two sources of information for regularization. On the one hand, they introduced fluence eigenspectra which they obtained from simulated data and on the other hand, they also estimated their results based on spectra from a larger patch of tissue thus introducing spatial regularization. They demonstrated the applicability of the method to *in vitro* and *in vivo* data in several experiments.

To make full use of the spatial context information, CNNs were employed to estimate blood oxygenation using the spectra of entire 2D images rather than single-pixel spectra. This was demonstrated by Yang et al. [130,131], Luke et al. [132], and Hoffer-Hawlik et al. [133]. Furthermore, Bench et al. [127] showed the feasibility to estimate oxygenation from multispectral 3D images. It has to be noted that there exists a trade-off regarding the spatial extent and number of voxels of the input images and the number of training images that can feasibly be simulated for algorithm training. The approaches demonstrate the feasibility of using CNNs for estimating oxygenation with high accuracy (for reported values in the publication see Table 6), however, a successful application of these methods *in vitro* or *in vivo* has not yet been shown, which is most probably caused by the large domain gap between simulated and experimental PA images.

The estimation of other functional tissue properties has also been investigated, such as the detection and concentration estimation of glucose by Ren et al. [134] and Liu et al. [135], the size of fat adipocytes by Ma et al. [136], as well as the unmixing of arbitrary chromophores in an unsupervised manner by Durairaj et al. [137].

Key insights: The estimation of functional tissue properties is closely related to the optical inverse problem, as functional properties can be derived from the optical properties of the tissue. But the direct estimation of the desired properties without quantification of the optical properties in an intermediate step is very popular. One reason for this is that there exist reference methods that can measure the functional properties and can be used to validate the results [129]. This potentially also enables training an algorithm on experimental data, when using reference measurements as the ground truth. Taking the estimation of tissue oxygenation as an example showcases the potential rewards of comprehensively solving this family of problems, as it would enable a lot of promising applications, such as oxygen-level specific dosimetry in radiotherapy [138] or cancer type classification based on local variations in blood oxygenation in the tumor's microenvironment [139].

Table 6
Overview of some of the reported errors on sO_2 estimation. Standard deviations or interquartile ranges (IQR) are shown if reported. It has to be noted that the used metrics as well as the conventions which pixels the error metrics are calculated on vary drastically between papers. As such the numbers are not directly comparable. For more detailed results, please refer to the linked papers. MedRE = Median relative error; MedAE = median absolute error; MRE = mean relative error; MAE = mean absolute error; RSME = root mean squared error; *depending on dataset.

Publication	Reported sO_2 estimation error
Bench et al. [127]	4.4% ± 4.5% MAE
Cai et al. [57]	0.8% ± 0.2% MRE
Gröhl et al. [58]	6.1% MedRE, IQR: [2.4%, 18.7%]
Hoffer-Hawlik et al. [133]	RSME consistently below 6%
Luke et al. [132]	5.1% MedAE at 25dB SNR
Olefir et al. [129]	0.9%, IQR [0.3%, 1.9%] to 2.5%, IQR [0.5%, 3.5%] MedAE*
Yang et al. [131]	1.4% ± 0.2% MRE
Yang and Gao [130]	4.8% ± 0.5% MAE
Nölke et al. [128]	2%, IQR [0%, 4%] to 9%, IQR [3%, 19%] MedAE*

7.2. Tissue classification and segmentation

Multispectral image information can also be used to differentiate between different tissue types or to identify and detect tissue pathologies. In such cases, strategies for dataset curation differ depending on the use case but using experimental datasets is possible with manual data annotation. In the work of Moustakidis et al. [140] *in vivo* images of a raster-scan optoacoustic mesoscopy (RSOM) system were utilized to automatically differentiate between different skin structures, while Lafci et al. [141] used neural networks to segment the entire imaged object. Furthermore, Nitkunanantharajah et al. [142] used deep learning to automatically differentiate photoacoustic nailfold images from patients with systemic sclerosis and from a healthy control group, Wu et al. [143] imaged *ex vivo* tissue samples to monitor lesion formation during high-intensity focused ultrasound (HIFU) therapy, and Jnawali et al. [144–146] also analyzed *ex vivo* tissue to differentiate cancer tissue from normal tissue in pathology samples.

On the other hand, we identified several papers that used simulated data to train their networks. Typically, simulations of acoustic waves are conducted using pre-processed images of a different modality, such as CT images, and treating the intensity distribution as the initial pressure distribution. This was done by Zhou et al. [147] to investigate the feasibility to differentiate healthy bone tissue from pathologies such as hyperosteogeny and osteoporosis. Further work by Zhang et al. [148] also examined the feasibility of DL-based breast cancer classification, Lin et al. [149] investigated the feasibility of endometrial cancer detection, and several groups including Zhang et al. [150], Luke et al. [132], Chlis et al. [151], and Boink et al. [152] examined the segmentation of blood vessels. Finally, Allman et al. [153] conducted feasibility experiments that demonstrated the capability of neural networks to automatically segment needle tips in PA images. Yuan et al. [154] proposed an approach for vessel segmentation based on manually annotated images of mouse ear measurements, achieving best results with a hybrid U-Net/Fully CNN architecture and Song et al. [155] used a classification algorithm for the detection of circulating melanoma tumor cells in a flow cytometry setup.

Key insights: Semantic image annotation enables intuitive and fast interpretation of PA images. Given the number of potential applications of PAI, we believe that semantic image annotation is a promising future research direction. Because the modality is comparatively young, high-

quality reference data for algorithm training that are annotated by healthcare professionals are very rare. Furthermore, the cross-modality and inter-institutional performance of PAI devices has to our knowledge not been examined as of yet. This makes validation of the proposed algorithms *in vivo* difficult, as reflected by some of the presented work. As discussed throughout this review, the image quality of photoacoustic (PA) images relies heavily on the solutions to the acoustic and optical inverse problems. This potentially introduces difficulties for manual data annotation and thus makes it more difficult to integrate developed methods into clinical practice.

8. Discussion

The clinical translation of deep learning methods in PAI is still in its infancy. Even though many classical image processing tasks, as well as PA-specific tasks, have already been tackled using deep learning techniques, vital limitations remain. For instance, many researchers resort to simulated data due to the lack of high-quality annotated experimental data. Accordingly, none of the proposed techniques were validated in a large-scale clinical PAI study. In this section we will discuss the challenges for clinical translation of deep learning methods in PAI and will conclude by summarizing the key findings of this review.

The challenges of clinical translation of spectroscopic optical imaging techniques have previously been extensively examined by Wilson et al. [156]. While their work focused primarily on the general challenges and hurdles in translating biomedical imaging modalities into clinical practice, in this review, we focused on the extent to which the application of deep learning in particular could facilitate or complicate the clinical integration of PAI. To this end, we have summarized important features that a learning algorithm should fulfill, based on the findings in other literature [156–158]:

Generalizability. In general, training data *must* be representative of the data encountered in clinical practice to avoid the presence of biases [159] in the trained model. The acquisition of high-quality experimental data sets in PAI is extremely problematic due to, for example, the high intra-patient signal variability caused by changes in local light fluence, the small amount of clinically approved PAI devices, and the lack of reliable methods to create ground truth annotations.

The lack of experimental data can be attributed to the comparative youth of PAI, but also to the fact that semantic information for images is only available via elaborate reference measurement setups or manual data annotations, which are usually created by healthcare professionals. But even in commonplace imaging modalities like computed tomography (CT) or magnetic resonance imaging (MRI), high-quality annotations are extremely costly, time-consuming and thus sparse compared to the total number of images that are taken. Finally, annotations of optical and acoustic properties are intrinsically not available *in vivo*, as there currently exist no gold standard methods to non-invasively measure, for example, optical absorption, optical scattering, or the speed of sound.

To account for the lack of available experimental data, approximately 65% of models were trained on simulated photoacoustic data. This practice has led to many insights into the general applicability and feasibility of deep learning-based methods. But methods trained purely on simulated data have shown poor performance when being applied to experimental data. Systematic differences between experimental PA images and those generated by computational forward models are very apparent. These differences are commonly referred to as the *domain gap* and can cause methods to fail on *in vivo* data despite thorough validation *in silico*, since deep learning methods cannot easily generalize to data from different distributions. Closing this gap can make the *in silico* validation of deep learning methods more meaningful. We see several approaches to tackle this problem:

1. Methods to create more realistic simulations. This comprises the implementation of digital twins of the respective PA devices or the simulation of anatomically realistic geometric variations of tissue.
2. Domain adaptation methods that are currently developed in the field of computer vision [122] could help translate images from the synthetic to the real PA image domain.
3. Methods to refine the training process, such as extensive data augmentation specific for PAI, the weighting of training data [160], content disentanglement [161] or domain-specific architecture changes [81], as well as the tight integration of prior knowledge into the entire algorithm development pipeline [59].

A promising opportunity could lie in the field of *life-long learning* [162,163]. This field strives to develop methods that have the ability to continuously learn over time by including new information while retaining prior knowledge [164]. In this context, research is conducted towards developing algorithms that efficiently adapt to new learning tasks (*meta-learning*) [165] and can be trained on various different but related tasks (*multi-task learning*) [166]. The goal is to create models that can continue to learn from data that becomes available after deployment. We strongly believe that the integration of such methods into the field of PAI can have exceptional merit, as this young imaging modality can be expected to undergo frequent changes and innovations in the future.

Uncertainty estimation. We strongly believe that methods should be uncertainty-aware, since gaining insight into the confidence of deep learning models can serve to avoid blindly assuming estimates to be accurate [167,168]. The primary goal of uncertainty estimation methods is to provide the confidence interval for measurements, for example, by calculating the average and standard deviation over a multitude of estimation samples [169]. On the other hand, such metrics might not be sufficient and a current direction of research is to recover the full posterior probability distribution for the estimate given the input, which for instance enables the automatic detection of multi-modal posteriors [170]. Uncertainty estimation and Bayesian modeling of the inverse problems is an active field of research in PAI [44,171,172]. While a first simulation study [124] has demonstrated the benefits of exploiting uncertainty measures when using deep learning methods, the potential of this branch of research remains largely untapped.

Out-of-distribution detection. A major risk of employing deep learning-based methods in the context of medical imaging can be seen in their potentially undefined behavior when facing out-of-distribution (OOD) samples. In these situations, deep learning-based uncertainty metrics do not have to be indicative of the quality of the estimate [173] and methods that identify OOD situations should be employed to avoid trusting wrong estimations. In the field of multispectral imaging, OOD metrics were used to quantify domain gaps between data that a deep learning algorithm was trained on and newly acquired experimental data [174,175]. We expect the investigation of well-calibrated methods for uncertainty estimation and the automatic detection of OOD scenarios to be integral towards the clinical translation of deep learning-based PAI methodologies.

Explainability. The goal of the field of explainable deep learning is to provide a trace of the inference of developed algorithms [176]. The estimates of a deep learning algorithm should be comprehensible to domain experts in order to verify, improve, and learn from the system [177]. In combination with techniques for uncertainty estimation and OOD detection, we believe that the explainability of algorithms will play an important role in the future of deep learning-based algorithms for PAI, especially in the biomedical context.

Validation. Thorough validation of methods is an integral part of clinical translation and as such plays a crucial role in the regulatory processes of medical device certification [156]. To this end, algorithms can be validated on several different levels, including in-distribution and out-of-distribution test data, as well as clinical validation in large-scale prospective studies [178]. However, there is a systematic lack of prospective studies in the field of medical imaging with deep learning [54], and to our knowledge, there have been no such prospective deep

learning studies in the field of PAI yet. Some of the most impressive clinical trials in the field to date include the detection of Duchenne muscular dystrophy [179] and the assessment of disease activity in Crohn's disease [21]. At least half of the reviewed papers have validated their methods on experimental data, but only approx. 35% of papers have validated their methods on *in vivo* data and even less on human measurements. This further demonstrates the vast differences in complexity within data obtained from phantoms versus living organisms. We expect that before deep learning methods for PAI can reliably be used in a clinical context, much more pre-clinical work is needed to mature the proposed methodologies.

Another crucial aspect that we noticed during this review is the difficulty to compare many of the reported results. This is partly due to the fact that no standardized metrics or common data sets have so far been established in the field. Furthermore, the developed algorithms are tested only on in-house data sets that are usually not openly accessible. We have high hopes that these problems can be mitigated to a certain extent by the ongoing standardization efforts of the PA community, as promoted by the International Photoacoustic Standardisation Consortium (IPASC) [180]. Amongst other issues, this consortium is working on standardized methods to assess image quality and characterize PAI device performance, on the organization of one of the first multi-centric studies in which PA phantoms are imaged all across the globe, as well as a standardized data format that facilitates the vendor-independent exchange of PA data.

Computational efficiency. Depending on the clinical use case, time can be of the essence (with stroke diagnosis being a prominent example [181]) and the speed of the algorithm can be considered an important factor. PAI is capable of real-time imaging [182–184] and the inference of estimates with deep learning can be exceptionally fast due to the massive parallelization capabilities of modern GPUs. The combination of these two factors can enable the real-time application of complex algorithms to PAI. In the reviewed literature, it was demonstrated that entire high-resolution 2D and 3D images can be evaluated in a matter of milliseconds [185]. In comparison to model-based methods, deep learning-based methods take a long time to train and fully optimize before they are ready to use. We believe that the drastic run-time performance increase could enable many time-critical applications of PAI that might otherwise remain unfeasible.

Clinical workflow integration. Deep learning methods have already found success in several medical applications, especially in the field of radiology [178,186,187]. Nevertheless, we believe that the integrability of deep learning methods in PAI heavily depends on the target clinical use case. The deep learning algorithm needs to have a clear impact on clinical practice, for example in terms of benefits for patients, personnel, or the hospital. Furthermore, the methods need to be easy to use for healthcare professionals, ideally being intuitive and introducing no significant time-burdens. PAI is very promising for a multitude of clinical applications [188], which are mostly based on the differences in contrast based on local blood perfusion and blood oxygen saturation. To unleash to the full potential of PAI, the inverse problems need to be solved to gain quantitative information on the underlying optical tissue properties. Deep learning can potentially enable an accurate, reliable, uncertainty-aware, and explainable estimation of the biomarkers of interest from the acquired PA measurements and thus provide unique opportunities towards the clinical translation of PAI. Nevertheless, thorough validation of the developed methods constitutes an essential first step in this direction.

8.1. Conclusion

This review has shown that deep learning methods possess unique advantages when applied to the field of PAI and have the potential to facilitate its clinical translation in the long term. We analyzed the current state of the art of deep learning applications as pertaining to several open challenges in photoacoustic imaging: the acoustic and optical inverse problem, image post-processing, and semantic image annotation.

Summary of findings:

- Deep learning methods in PAI are currently still in their infancy. While the initial results are promising and encouraging, prospective clinical validation studies of such techniques, an integral part of method validation, have not been conducted.
- One of the core bottlenecks of the application of deep learning algorithms to PAI is the lack of reliable, high-quality experimental training data. For this reason, about 65% of deep learning papers in PAI rely on simulated data for supervised algorithm training.
- A commonly used workaround to create suitable experimental training data for image post-processing is to artificially introduce artifacts, for example, by deliberately using less information for image reconstruction.
- Because the underlying optical tissue properties are inherently difficult to measure *in vivo*, data-driven approaches towards the optical inverse problem have primarily relied on the presence of high-fidelity simulated data and have not yet successfully been applied *in vivo*.
- While direct image reconstruction with deep learning shows exceptional promise due to the drastic speed increases compared to model-reconstruction schemes, deep learning methods that utilize additional information such as reconstructions from reference methods or hand-crafted feature vectors have proven much more generalizable.
- Nearly 50% of papers test the presented methods on simulated data only and do not use multiple test sets that are significantly different from the training data distribution.
- A successful application of oxygenation estimation methods using entire 2D or 3D images has not yet been shown *in vitro* or *in vivo*. This is most probably caused by the large domain gap between synthetic and experimental PA images.
- Deep learning in PAI has considerable room for improvement, for instance in terms of, generalizability, uncertainty estimation, out-of-distribution detection, or explainability.

CRediT authorship contribution statement

Janek Gröhl: Conceptualization, Investigation, Methodology, Project administration, Writing - original draft, Writing - review & editing. **Melanie Schellenberg:** Conceptualization, Investigation, Writing - review & editing. **Kris Dreher:** Conceptualization, Investigation, Writing - review & editing. **Lena Maier-Hein:** Conceptualization, Funding acquisition, Methodology, Project administration, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programmethrough the ERC starting grant COMBIOSCOPY under grant agreement No. ERC-2015-StG-37960. The authors would like to thank M. D. Tizabi and A. Seitel for proofreading the manuscript.

References

[1] L.V. Wang, L. Gao, Photoacoustic microscopy and computed tomography: from bench to bedside, Annu. Rev. Biomed. Eng. 16 (2014) 155–185.
[2] C. Kim, T.N. Erpelding, L. Jankovic, et al., Deeply penetrating in vivo photoacoustic imaging using a clinical ultrasound array system, Biomed. Opt. Express 1 (1) (2010) 278–284.

4 Publications

[3] S. Mallidi, G.P. Luke, S. Emelianov, Photoacoustic imaging in cancer detection, diagnosis, and treatment guidance, Trends Biotechnol. 29 (5) (2011) 213–221.
[4] P. Beard, Biomedical photoacoustic imaging, Interface Focus 1 (4) (2011) 602–631.
[5] A. Rosencwaig, A. Gersho, Theory of the photoacoustic effect with solids, J. Appl. Phys. 47 (1) (1976) 64–69.
[6] H. Li, S. Zhuang, D.-a. Li, et al., Benign and malignant classification of mammogram images based on deep learning, Biomed. Signal Process. Control 51 (2019) 347–354.
[7] J. Zhang, B. Chen, M. Zhou, et al., Photoacoustic image classification and segmentation of breast cancer: a feasibility study, IEEE Access 7 (2018) 5457–5466.
[8] I. Quiros-Gonzalez, M.R. Tomaszewski, S.J. Aitken, et al., Optoacoustics delineates murine breast cancer models displaying angiogenesis and vascular mimicry, Br. J. Cancer 118 (8) (2018) 1098.
[9] J.-T. Oh, M.-L. Li, H.F. Zhang, et al., Three-dimensional imaging of skin melanoma in vivo by dual-wavelength photoacoustic microscopy, J. Biomed. Opt. 11 (3) (2006) 034032.
[10] R.M. Weight, J.A. Viator, P.S. Dale, et al., Photoacoustic detection of metastatic melanoma cells in the human circulatory system, Opt. Lett. 31 (20) (2006) 2998–3000.
[11] C. Zhang, K. Maslov, L.V. Wang, Subwavelength-resolution label-free photoacoustic microscopy of optical absorption in vivo, Opt. Lett. 35 (19) (2010) 3195–3197.
[12] Y. Zhang, X. Cai, S.-W. Choi, et al., Chronic label-free volumetric photoacoustic microscopy of melanoma cells in three-dimensional porous scaffolds, Biomaterials 31 (33) (2010) 8651–8658.
[13] K.H. Song, E.W. Stein, J.A. Margenthaler, et al., Noninvasive photoacoustic identification of sentinel lymph nodes containing methylene blue in vivo in a rat model, J. Biomed. Opt. 13 (5) (2008) 054033.
[14] T.N. Erpelding, C. Kim, M. Pramanik, et al., Sentinel lymph nodes in the rat: noninvasive photoacoustic and us imaging with a clinical us system, Radiology 256 (1) (2010) 102–110.
[15] A. Garcia-Uribe, T.N. Erpelding, A. Krumholz, et al., Dual-modality photoacoustic and ultrasound imaging system for noninvasive sentinel lymph node detection in patients with breast cancer, Sci. Rep. 5 (2015) 15748.
[16] D. Hanahan, R.A. Weinberg, Hallmarks of cancer: the next generation, Cell 144 (5) (2011) 646–674.
[17] L.V. Wang, H.-i. Wu, Biomedical Optics: Principles and Imaging, John Wiley & Sons, 2012.
[18] X. Wang, D.L. Chamberland, D.A. Jamadar, Noninvasive photoacoustic tomography of human peripheral joints toward diagnosis of inflammatory arthritis, Opt. Lett. 32 (20) (2007) 3002–3004.
[19] J.R. Rajian, G. Girish, X. Wang, Photoacoustic tomography to identify inflammatory arthritis, J. Biomed. Opt. 17 (9) (2012) 096013.
[20] J. Jo, G. Xu, Y. Zhu, et al., Detecting joint inflammation by an led-based photoacoustic imaging system: a feasibility study, J. Biomed. Opt. 23 (11) (2018) 110501.
[21] F. Knieling, C. Neufert, A. Hartmann, et al., Multispectral optoacoustic tomography for assessment of Crohn's disease activity, N. Engl. J. Med. 376 (13) (2017) 1292.
[22] M.J. Waldner, F. Knieling, C. Egger, et al., Multispectral optoacoustic tomography in Crohn's disease: noninvasive imaging of disease activity, Gastroenterology 151 (2) (2016) 238–240.
[23] H. Lei, L.A. Johnson, K.A. Eaton, et al., Characterizing intestinal strictures of Crohn's disease in vivo by endoscopic photoacoustic imaging, Biomed. Opt. Express 10 (5) (2019) 2542–2555.
[24] J.J. Niederhauser, M. Jaeger, R. Lemor, et al., Combined ultrasound and optoacoustic system for real-time high-contrast vascular imaging in vivo, IEEE Trans. Med. Imaging 24 (4) (2005) 436–440.
[25] A. Aguirre, Y. Ardeshirpour, M.M. Sanders, et al., Potential role of coregistered photoacoustic and ultrasound imaging in ovarian cancer detection and characterization, Transl. Oncol. 4 (1) (2011) 29.
[26] A. Needles, A. Heinmiller, J. Sun, et al., Development and initial application of a fully integrated photoacoustic micro-ultrasound system, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 60 (5) (2013) 888–897.
[27] P. Elbau, L. Mindrinos, O. Scherzer, Quantitative reconstructions in multi-modal photoacoustic and optical coherence tomography imaging, Inverse Probl. 34 (1) (2017) 014006.
[28] S. Mandal, M. Mueller, D. Komljenovic, Multimodal priors reduce acoustic and optical inaccuracies in photoacoustic imaging, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 108781M.
[29] X. Wang, Y. Pang, G. Ku, et al., Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain, Nat. Biotechnol. 21 (7) (2003) 803.
[30] G. Ku, X. Wang, X. Xie, et al., Imaging of tumor angiogenesis in rat brains in vivo by photoacoustic tomography, Appl. Opt. 44 (5) (2005) 770–775.
[31] S. Hu, K.I. Maslov, V. Tsytsarev, et al., Functional transcranial brain imaging by optical-resolution photoacoustic microscopy, J. Biomed. Opt. 14 (4) (2009) 040503.
[32] J. Yao, L.V. Wang, Photoacoustic brain imaging: from microscopic to macroscopic scales, Neurophotonics 1 (1) (2014) 011003.
[33] L. Mohammadi, R. Manwar, H. Behnam, et al., Skull's aberration modeling: towards photoacoustic human brain imaging, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 108785W.
[34] C. Kim, T.N. Erpelding, K.I. Maslov, et al., Handheld array-based photoacoustic probe for guiding needle biopsy of sentinel lymph nodes, J. Biomed. Opt. 15 (4) (2010) 046010.
[35] J.L. Su, A.B. Karpiouk, B. Wang, et al., Photoacoustic imaging of clinical metal needles in tissue, J. Biomed. Opt. 15 (2) (2010) 021309.
[36] P.K. Upputuri, M. Pramanik, Recent advances toward preclinical and clinical translation of photoacoustic tomography: a review, J. Biomed. Opt. 22 (4) (2016) 041006.
[37] M. Xu, L.V. Wang, Universal back-projection algorithm for photoacoustic computed tomography, Phys. Rev. E 71 (1) (2005) 016706.
[38] M. Mozaffarzadeh, A. Mahloojifar, M. Orooji, et al., Double-stage delay multiply and sum beamforming algorithm: application to linear-array photoacoustic imaging, IEEE Trans. Biomed. Eng. 65 (1) (2017) 31–42.
[39] B.E. Treeby, B.T. Cox, k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields, J. Biomed. Opt. 15 (2) (2010) 021314.
[40] C. Huang, K. Wang, L. Nie, et al., Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media, IEEE Trans. Med. Imaging 32 (6) (2013) 1097–1110.
[41] B. Cox, J. Laufer, P. Beard, The challenges for quantitative photoacoustic imaging, Photons Plus Ultrasound: Imaging and Sensing 2009, vol. 7177 (2009) 717713.
[42] D. Waibel, J. Gröhl, F. Isensee, et al., Reconstruction of initial pressure from limited view photoacoustic images using deep learning, Photons Plus Ultrasound: Imaging and Sensing 2018, vol. 10494 (2018) 104942S.
[43] J. Buchmann, J. Guggenheim, E. Zhang, et al., Characterization and modeling of Fabry-Perot ultrasound sensors with hard dielectric mirrors for photoacoustic imaging, Appl. Opt. 56 (17) (2017) 5039–5046.
[44] T. Sahlström, A. Pulkkinen, J. Tick, et al., Modeling of errors due to uncertainties in ultrasound sensor locations in photoacoustic tomography, IEEE Trans. Med. Imaging 39 (6) (2020) 2140–2150.
[45] J. Buchmann, B. Kaplan, S. Powell, et al., Quantitative pa tomography of high resolution 3-D images: experimental validation in a tissue phantom, Photoacoustics 17 (2020) 100157.
[46] M. Xu, L.V. Wang, Photoacoustic imaging in biomedicine, Rev. Sci. Instrum. 77 (4) (2006) 041101.
[47] P. Shao, B. Cox, R.J. Zemp, Estimating optical absorption, scattering, and Grueneisen distributions with multiple-illumination photoacoustic tomography, Appl. Opt. 50 (19) (2011) 3145–3154.
[48] S. Tzoumas, A. Nunes, I. Olefir, et al., Eigenspectra optoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues, Nat. Commun. 7 (2016) 12121.
[49] A. Kazakeviciute, C.J.H. Ho, M. Olivo, Multispectral photoacoustic imaging artifact removal and denoising using time series model-based spectral noise estimation, IEEE Trans. Med. Imaging 35 (9) (2016) 2151–2163.
[50] B.T. Cox, S.R. Arridge, K.P. Köstli, et al., Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method, Appl. Opt. 45 (8) (2006) 1866–1875.
[51] F.M. Brochu, J. Brunker, J. Joseph, et al., Towards quantitative evaluation of tissue absorption coefficients using light fluence correction in optoacoustic tomography, IEEE Trans. Med. Imaging 36 (1) (2016) 322–331.
[52] T. Kirchner, J. Gröhl, L. Maier-Hein, Context encoding enables machine learning-based quantitative photoacoustics, J. Biomed. Opt. 23 (5) (2018) 056008.
[53] D. Shen, G. Wu, H.-I. Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (2017) 221–248.
[54] X. Liu, L. Faes, A.U. Kale, et al., A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis, Lancet Digit. Health 1 (6) (2019) e271–e297.
[55] N. Tajbakhsh, L. Jeyaseelan, Q. Li, et al., Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation, Med. Image Anal. 63 (2020) 101693.
[56] G.S. Tandel, M. Biswas, O.G. Kakde, et al., A review on a deep learning perspective in brain cancer classification, Cancers 11 (1) (2019) 111.
[57] C. Cai, K. Deng, C. Ma, et al., End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging, Opt. Lett. 43 (12) (2018) 2752–2755.
[58] J. Gröhl, T. Kirchner, T. Adler, et al., Estimation of Blood Oxygenation with Learned Spectral Decoloring for Quantitative Photoacoustic Imaging (LSD-qPAI), arXiv preprint arXiv:1902.05839, 2019.
[59] A. Hauptmann, F. Lucka, M. Betcke, et al., Model-based learning for accelerated, limited-view 3-D photoacoustic tomography, IEEE Trans. Med. Imaging 37 (6) (2018) 1382–1393.
[60] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (10) (2009) 1345–1359.
[61] T. Panch, H. Mattie, L.A. Celi, The "inconvenient truth" about AI in healthcare, NPJ Digit. Med. 2 (1) (2019) 1–3.
[62] T. Ross, D. Zimmerer, A. Vemuri, et al., Exploiting the potential of unlabeled endoscopic video data with self-supervised learning, Int. J. Comput. Assist. Radiol. Surg. 13 (6) (2018) 925–933.
[63] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, et al., 3D U-Net: learning dense volumetric segmentation from sparse annotation, International Conference on Medical Image Computing and Computer-Assisted Intervention (2016) 424–432.
[64] M. Kato, T. Teshima, J. Honda, Learning from positive and unlabeled data with a selection bias, International Conference on Learning Representations (2018).

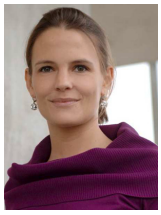
[65] J. Gu, Z. Wang, J. Kuen, et al., Recent advances in convolutional neural networks, Pattern Recognit. 77 (2018) 354–377.
[66] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional networks for biomedical image segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention (2015) 234–241.
[67] S.L. Jacques, Coupling 3D Monte Carlo light transport in optically heterogeneous tissues to photoacoustic signal generation, Photoacoustics 2 (4) (2014) 137–142.
[68] Q. Fang, D.A. Boas, Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units, Opt. Express 17 (22) (2009) 20178–20190.
[69] H. Dehghani, M.E. Eames, P.K. Yalavarthy, et al., Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction, Commun. Numer. Methods Eng. 25 (6) (2009) 711–732.
[70] M. Schweiger, S.R. Arridge, The Toast++ software suite for forward and inverse modeling in optical tomography, J. Biomed. Opt. 19 (4) (2014) 040801.
[71] J. Schwab, S. Antholzer, R. Nuster, et al., Deep learning of truncated singular values for limited view photoacoustic tomography, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 1087836.
[72] J. Schwab, S. Antholzer, R. Nuster, et al., Real-time Photoacoustic Projection Imaging Using Deep Learning, arXiv preprint arXiv:1801.06693, 2018.
[73] J. Schwab, S. Antholzer, M. Haltmeier, Learned backprojection for sparse and limited view photoacoustic tomography, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 1087837.
[74] S. Antholzer, J. Schwab, J. Bauer-Marschallinger, et al., NETT regularization for compressed sensing photoacoustic tomography, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 108783B.
[75] S. Antholzer, M. Haltmeier, Discretization of learned NETT regularization for solving inverse problems, arXiv preprint arXiv:2011.03627, 2020.
[76] H. Li, J. Schwab, S. Antholzer, et al., NETT: solving inverse problems with deep neural networks, Inverse Probl. 36 (6) (2020).
[77] C. Yang, H. Lan, F. Gao, Accelerated photoacoustic tomography reconstruction via recurrent inference machines, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2019) 6371–6374.
[78] R. Manwar, M. Hosseinzadeh, A. Hariri, et al., Photoacoustic signal enhancement: towards utilization of low energy laser diodes in real-time photoacoustic imaging, Sensors 18 (10) (2018) 3498.
[79] E.M.A. Anas, H.K. Zhang, C. Audigier, et al., Robust photoacoustic beamforming using dense convolutional neural networks, Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation (2018) 3–11.
[80] H. Lan, C. Yang, D. Jiang, et al., Reconstruct the photoacoustic image based on deep learning with multi-frequency ring-shape transducer array, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2019) 7115–7118.
[81] H. Lan, K. Zhou, C. Yang, et al., Ki-GAN: knowledge infusion generative adversarial network for photoacoustic image reconstruction in vivo, International Conference on Medical Image Computing and Computer-Assisted Intervention (2019) 273–281.
[82] H. Lan, K. Zhou, C. Yang, et al., Hybrid neural network for photoacoustic imaging reconstruction, 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2019) 6367–6370.
[83] H. Lan, D. Jiang, C. Yang, et al., Y-net: hybrid deep learning image reconstruction for photoacoustic tomography in vivo, Photoacoustics 20 (2020) 100197.
[84] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., Generative adversarial nets, Advances in Neural Information Processing Systems (2014) 2672–2680.
[85] S. Guan, A.A. Khan, S. Sikdar, et al., Limited View and Sparse Photoacoustic Tomography for Neuroimaging with Deep Learning, arXiv preprint arXiv: 1911.04357, 2019.
[86] M.W. Kim, G.-S. Jeng, I. Pelivanov, et al., Deep-learning image reconstruction for real-time photoacoustic system, IEEE Trans. Med. Imaging (2020).
[87] A. Reiter, M.A.L. Bell, A machine learning approach to identifying point source locations in photoacoustic data, Photons Plus Ultrasound: Imaging and Sensing 2017, vol. 10064 (2017) 100643J.
[88] K. Johnstonbaugh, S. Agrawal, D. Abhishek, et al., Novel deep learning architecture for optical fluence dependent photoacoustic target localization, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 108781L.
[89] K. Johnstonbaugh, S. Agrawal, D.A. Durairaj, et al., A deep learning approach to photoacoustic wavefront localization in deep-tissue medium, IEEE Transactions on Ultrasonics Ferroelectrics, and Frequency Control (2020).
[90] H. Shan, G. Wang, Y. Yang, Simultaneous reconstruction of the initial pressure and sound speed in photoacoustic tomography using a deep-learning approach, Novel Optical Systems, Methods, and Applications XXII, vol. 11105 (2019) 1110504.
[91] S. Jeon, C. Kim, Deep learning-based speed of sound aberration correction in photoacoustic images, Photons Plus Ultrasound: Imaging and Sensing 2020, vol. 11240 (2020) 112400J.
[92] N. Davoudi, X.L. Deán-Ben, D. Razansky, Deep learning optoacoustic tomography with sparse data, Nat. Mach. Intell. 1 (10) (2019) 453–460.
[93] S. Antholzer, M. Haltmeier, J. Schwab, Deep learning for photoacoustic tomography from sparse data, Inverse Probl. Sci. Eng. 27 (7) (2019) 987–1005.
[94] S. Guan, A. Khan, S. Sikdar, et al., Fully dense UNet for 2D sparse photoacoustic tomography artifact removal, IEEE J. Biomed. Health Inform. 24 (2019).
[95] P. Farnia, M. Mohammadi, E. Najafzadeh, et al., High-quality photoacoustic image reconstruction based on deep convolutional neural network: towards intra-operative photoacoustic imaging, Biomed. Phys. Eng. Express 6 (4) (2020).
[96] T. Tong, W. Huang, K. Wang, et al., Domain transform network for photoacoustic tomography from limited-view and sparsely sampled data, Photoacoustics 19 (2020) 100190.
[97] T. Vu, A. DiSpirito III, D. Li, et al., Deep Image Prior for Sparse-Sampling Photoacoustic Microscopy, arXiv preprint arXiv:2010.12041, 2020.
[98] J. Zhou, D. He, X. Shang, et al., Photoacoustic Microscopy with Sparse Data Enabled by Convolutional Neural Networks for Fast Imaging, arXiv preprint arXiv:2006.04368, 2020.
[99] A. DiSpirito III, D. Li, T. Vu, et al., Reconstructing Undersampled Photoacoustic Microscopy Images Using Deep Learning, arXiv preprint arXiv:2006.00251, 2020.
[100] S. Antholzer, M. Haltmeier, R. Nuster, et al., Photoacoustic image reconstruction via deep learning, Photons Plus Ultrasound: Imaging and Sensing 2018, vol. 10494 (2018) 104944U.
[101] S. Antholzer, J. Schwab, M. Haltmeier, Deep Learning Versus ℓ^1-Minimization for Compressed Sensing Photoacoustic Tomography, 2018 IEEE International Ultrasonics Symposium (IUS) (2018) 206–212.
[102] H. Deng, X. Wang, C. Cai, et al., Machine-learning enhanced photoacoustic computed tomography in a limited view configuration, Advanced Optical Imaging Technologies II, vol. 11186 (2019) 111860J.
[103] H. Zhang, L. Hongyu, N. Nyayapathi, et al., A new deep learning network for mitigating limited-view and under-sampling artifacts in ring-shaped photoacoustic tomography, Comput. Med. Imaging Graph. 84 (2020) 101720.
[104] G. Godefroy, B. Arnal, E. Bossy, Solving the Visibility Problem in Photoacoustic Imaging with a Deep Learning Approach Providing Prediction Uncertainties, arXiv preprint arXiv:2006.13096, 2020.
[105] H. Lan, C. Yang, F. Gao, et al., Better Than Ground-Truth? Photoacoustic Image Reconstruction Beyond Supervised to Compensate Limit-View and Remove Artifacts, arXiv preprint arXiv:2012.02472, 2020.
[106] T. Lu, T. Chen, F. Gao, et al., LV-GAN: a deep learning approach for limited-view optoacoustic imaging based on hybrid datasets, J. Biophotonics (2020) e202000325.
[107] D. Allman, A. Reiter, M.A.L. Bell, Photoacoustic source detection and reflection artifact removal enabled by deep learning, IEEE Trans. Med. Imaging 37 (6) (2018) 1464–1477.
[108] H. Shan, G. Wang, Y. Yang, Accelerated correction of reflection artifacts by deep neural networks in photo-acoustic tomography, Appl. Sci. 9 (13) (2019) 2615.
[109] X. Chen, W. Qi, L. Xi, Deep-learning-based motion-correction algorithm in optical resolution photoacoustic microscopy, Vis. Comput. Ind. Biomed. Art 2 (1) (2019) 12.
[110] S. Gutta, V.S. Kadimesetty, S.K. Kalva, et al., Deep neural network-based bandwidth enhancement of photoacoustic data, J. Biomed. Opt. 22 (11) (2017) 116001.
[111] N. Awasthi, G. Jain, S.K. Kalva, et al., Deep neural network based sinogram super-resolution and bandwidth enhancement for limited-data photoacoustic tomography, IEEE Trans. Ultrason. Ferroelectr. Freq. Control (2020).
[112] H. Zhao, Z. Ke, N. Chen, et al., A new deep learning method for image deblurring in optical microscopic systems, J. Biophotonics 13 (3) (2020) e201960147.
[113] H. Zhao, Z. Ke, F. Yang, et al., Deep learning enables superior photoacoustic imaging at ultralow laser dosages, Adv. Sci. (2020) 2003097.
[114] P. Rajendran, M. Pramanik, Deep learning approach to improve tangential resolution in photoacoustic tomography, Biomed. Opt. Express 11 (12) (2020) 7311–7323.
[115] T. Vu, M. Li, H. Humayun, et al., A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer, Exp. Biol. Med. 245 (2020), 1535370220914285.
[116] M.K.A. Singh, K. Sivasubramanian, N. Sato, et al., Deep learning-enhanced LED-based photoacoustic imaging, Photons Plus Ultrasound: Imaging and Sensing 2020, vol. 11240 (2020) 1124038.
[117] E.M.A. Anas, H.K. Zhang, J. Kang, et al., Enabling fast and high quality LED photoacoustic imaging: a recurrent neural networks based approach, Biomed. Opt. Express 9 (8) (2018) 3852–3866.
[118] A. Sharma, M. Pramanik, Convolutional neural network for resolution enhancement and noise reduction in acoustic resolution photoacoustic microscopy, Biomed. Opt. Express 11 (12) (2020) 6826–6839.
[119] K. Tang, B. Li, J. Zhang, et al., Denoising method for photoacoustic microscopy using deep learning, SPIE Future Sensing Technologies, vol. 11525 (2020) 115252P.
[120] A. Hariri, K. Alipour, Y. Mantri, et al., Deep Learning Improves Contrast in Low-Fluence Photoacoustic Imaging, arXiv preprint arXiv:2004.08782, 2020.
[121] N. Awasthi, K.R. Prabhakar, S.K. Kalva, et al., PA-Fuse: deep supervised approach for the fusion of photoacoustic images with distinct reconstruction characteristics, Biomed. Opt. Express 10 (5) (2019) 2227–2243.
[122] A. Voulodimos, N. Doulamis, A. Doulamis, et al., Deep learning for computer vision: a brief review, Comput. Intell. Neurosci. 2018 (2018).
[123] B. Cox, J. Laufer, P. Beard, Quantitative photoacoustic image reconstruction using fluence dependent chromophores, Biomed. Opt. Express 1 (1) (2010) 201–208.
[124] J. Gröhl, T. Kirchner, T. Adler, et al., Confidence estimation for machine learning-based quantitative photoacoustics, J. Imaging 4 (12) (2018) 147.
[125] T. Chen, T. Lu, S. Song, et al., A deep learning method based on U-Net for quantitative photoacoustic imaging, Photons Plus Ultrasound: Imaging and Sensing 2020, vol. 11240 (2020) 112403V.
[126] B. Dogdas, D. Stout, A.F. Chatziioannou, et al., Digimouse: a 3D whole body mouse atlas from CT and cryosection data, Phys. Med. Biol. 52 (3) (2007) 577.
[127] C. Bench, A. Hauptmann, B.T. Cox, Toward accurate quantitative photoacoustic imaging: learning vascular blood oxygen saturation in three dimensions, J. Biomed. Opt. 25 (8) (2020) 085003.

[128] J.-H. Nölke, T. Adler, J. Gröhl, et al., Invertible Neural Networks for Uncertainty Quantification in Photoacoustic Imaging, arXiv preprint arXiv:2011.05110, 2020.
[129] I. Olefir, S. Tzoumas, C. Restivo, et al., Deep learning based spectral unmixing for optoacoustic imaging of tissue oxygen saturation, IEEE Trans. Med. Imaging 39 (2020).
[130] C. Yang, F. Gao, Eda-net: dense aggregation of deep and shallow information achieves quantitative photoacoustic blood oxygenation imaging deep in human breast, International Conference on Medical Image Computing and Computer-Assisted Intervention (2019) 246–254.
[131] C. Yang, H. Lan, H. Zhong, et al., Quantitative photoacoustic blood oxygenation imaging using deep residual and recurrent neural network, 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (2019) 741–744.
[132] G.P. Luke, K. Hoffer-Hawlik, A.C. Van Namen, et al., O-Net: A Convolutional Neural Network for Quantitative Photoacoustic Image Segmentation and Oximetry, arXiv preprint arXiv:1911.01935, 2019.
[133] K. Hoffer-Hawlik, G.P. Luke, absO2luteU-Net: Tissue Oxygenation Calculation Using Photoacoustic Imaging and Convolutional Neural Networks, 2019.
[134] Z. Ren, G. Liu, Y. Ding, Effects of multiple factors on the photoacoustic detection of glucose based on artificial neural network, Optics in Health Care and Biomedical Optics VIII, vol. 10820 (2018) 108201E.
[135] T. Liu, Z. Ren, C. Zhang, et al., Combined effect of multiple factors on the photoacoustic detection of glucose solution based on neural networks, AOPC 2020: Optical Spectroscopy and Imaging; and Biomedical Optics, vol. 11566 (2020) 115660W.
[136] X. Ma, M. Cao, Q. Shen, et al., Adipocyte size evaluation based on photoacoustic spectral analysis combined with deep learning method, Appl. Sci. 8 (11) (2018) 2178.
[137] D.A. Durairaj, S. Agrawal, K. Johnstonbaugh, et al., Unsupervised deep learning approach for photoacoustic spectral unmixing, Photons Plus Ultrasound: Imaging and Sensing 2020, vol. 11240 (2020) 112403H.
[138] M. Matsuo, S. Matsumoto, J.B. Mitchell, et al., Magnetic resonance imaging of the tumor microenvironment in radiotherapy: perfusion, hypoxia, and metabolism, Semin. Radiat. Oncol. 24 (3) (2014) 210–217.
[139] M.R. Horsman, Measurement of tumor oxygenation, Int. J. Radiat. Oncol. Biol. Phys. 42 (4) (1998) 701–704.
[140] S. Moustakidis, M. Omar, J. Aguirre, et al., Fully automated identification of skin morphology in raster-scan optoacoustic mesoscopy using artificial intelligence, Med. Phys. 46 (9) (2019) 4046–4056.
[141] B. Lafci, E. Merćep, S. Morscher, et al., Efficient segmentation of multi-modal optoacoustic and ultrasound images using convolutional neural networks, Photons Plus Ultrasound: Imaging and Sensing 2020, vol. 11240 (2020) 112402N.
[142] S. Nitkunanantharajah, K. Haedicke, T.B. Moore, et al., Three-dimensional optoacoustic imaging of nailfold capillaries in systemic sclerosis and its potential for disease differentiation using deep learning, Sci. Rep. 10 (1) (2020) 1–9.
[143] X. Wu, J. Sanders, M. Dundar, et al., Multi-wavelength photoacoustic imaging for monitoring lesion formation during high-intensity focused ultrasound therapy, Photons Plus Ultrasound: Imaging and Sensing 2017, vol. 10064 (2017) 100644A.
[144] K. Jnawali, B. Chinni, V. Dogra, et al., Deep 3D convolutional neural network for automatic cancer tissue detection using multispectral photoacoustic imaging, Medical Imaging 2019: Ultrasonic Imaging and Tomography, vol. 10955 (2019) 109551D.
[145] K. Jnawali, B. Chinni, V. Dogra, et al., Transfer learning for automatic cancer tissue detection using multispectral photoacoustic imaging, Medical Imaging 2019: Computer-Aided Diagnosis, vol. 10950 (2019) 109503W.
[146] K. Jnawali, B. Chinni, V. Dogra, et al., Automatic cancer tissue detection using multispectral photoacoustic imaging, Int. J. Comput. Assist. Radiol. Surg. 15 (2) (2020) 309–320.
[147] X. Zhou, T. Feng, H. Zhan, et al., Analysis of photoacoustic signals of hyperosteogeny and osteoporosis, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 1087856.
[148] J. Zhang, B. Chen, M. Zhou, et al., Photoacoustic image classification and segmentation of breast cancer: a feasibility study, IEEE Access 7 (2018) 5457–5466.
[149] Y. Lin, H. Song, R. Zheng, et al., Computer-aided classification system for early endometrial cancer of co-registered photoacoustic and ultrasonic signals, Optics in Health Care and Biomedical Optics IX, vol. 11190 (2019) 111901R.
[150] J. Zhang, K. Deng, B. Chen, et al., Pathology study for blood vessel of ocular fundus images by photoacoustic tomography, 2018 IEEE International Ultrasonics Symposium (IUS) (2018) 1–4.
[151] N.K. Chlis, A. Karlas, N.-A. Fasoula, et al., A Sparse Deep Learning Approach for Automatic Segmentation of Human Vasculature in Multispectral Optoacoustic Tomography, bioRxiv 833251, 2019.
[152] Y.E. Boink, S. Manohar, C. Brune, A partially-learned algorithm for joint photo-acoustic reconstruction and segmentation, IEEE Trans. Med. Imaging 39 (1) (2019) 129–139.
[153] D. Allman, F. Assis, J. Chrispin, et al., A deep learning-based approach to identify in vivo catheter tips during photoacoustic-guided cardiac interventions, Photons Plus Ultrasound: Imaging and Sensing 2019, vol. 10878 (2019) 108785E.
[154] A.Y. Yuan, Y. Gao, L. Peng, et al., Hybrid deep learning network for vascular segmentation in photoacoustic imaging, Biomed. Opt. Express 11 (11) (2020) 6445–6457.
[155] Z. Song, Y. Fu, J. Qu, et al., Application of convolutional neural network in signal classification for in vivo photoacoustic flow cytometry, Optics in Health Care and Biomedical Optics X, vol. 11553 (2020) 115532W.
[156] B.C. Wilson, M. Jermyn, F. Leblond, Challenges and opportunities in clinical translation of biomedical optical spectroscopy and imaging, J. Biomed. Opt. 23 (3) (2018) 030901.
[157] R. Miotto, F. Wang, S. Wang, et al., Deep learning for healthcare: review, opportunities and challenges, Brief. Bioinform. 19 (6) (2018) 1236–1246.
[158] C.J. Kelly, A. Karthikesalingam, M. Suleyman, et al., Key challenges for delivering clinical impact with artificial intelligence, BMC Med. 17 (1) (2019) 195.
[159] C.J. Pannucci, E.G. Wilkins, Identifying and avoiding bias in research, Plast. Reconstruct. Surg. 126 (2) (2010) 619.
[160] S.J. Wirkert, A.S. Vemuri, H.G. Kenngott, et al., Physiological parameter estimation from multispectral images unleashed, International Conference on Medical Image Computing and Computer-Assisted Intervention (2017) 134–141.
[161] M. Ilse, J.M. Tomczak, C. Louizos, et al., Diva: domain invariant variational autoencoders, Medical Imaging with Deep Learning (2020) 322–348.
[162] P. Ruvolo, E. Eaton, Ella: an efficient lifelong learning algorithm, International Conference on Machine Learning (2013) 507–515.
[163] Z. Chen, B. Liu, Lifelong machine learning, Synth. Lect. Artif. Intell. Mach. Learn. 12 (3) (2018) 1–207.
[164] G.I. Parisi, R. Kemker, J.L. Part, et al., Continual lifelong learning with neural networks: a review, Neural Netw. 113 (2019) 54–71.
[165] C. Finn, P. Abbeel, S. Levine, Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, arXiv preprint arXiv:1703.03400, 2017.
[166] R. Caruana, Multitask learning, Mach. Learn. 28 (1) (1997) 41–75.
[167] A. Kendall, Y. Gal, What uncertainties do we need in bayesian deep learning for computer vision? in: I. Guyon, U.V. Luxburg, S. Bengio (Eds.), Advances in Neural Information Processing Systems, vol. 30 Curran Associates, Inc., 2017, pp. 5574–5584.
[168] K. Osawa, S. Swaroop, M.E.E. Khan, et al., Practical deep learning with bayesian principles, Advances in Neural Information Processing Systems (2019) 4287–4299.
[169] Y. Gal, Z. Ghahramani, Dropout as a bayesian approximation: representing model uncertainty in deep learning, International Conference on Machine Learning (2016) 1050–1059.
[170] L. Ardizzone, J. Kruse, S. Wirkert, et al., Analyzing Inverse Problems with Invertible Neural Networks, arXiv preprint arXiv:1808.04730, 2018.
[171] J. Tick, A. Pulkkinen, T. Tarvainen, Image reconstruction with uncertainty quantification in photoacoustic tomography, J. Acoust. Soc. Am. 139 (4) (2016) 1951–1961.
[172] J. Tick, A. Pulkkinen, T. Tarvainen, Modelling of errors due to speed of sound variations in photoacoustic tomography using a bayesian framework, Biomed. Phys. Eng. Express 6 (1) (2019) 015003.
[173] V. Kuleshov, N. Fenner, S. Ermon, Accurate Uncertainties for Deep Learning Using Calibrated Regression, arXiv preprint arXiv:1807.00263, 2018.
[174] T.J. Adler, L. Ardizzone, L. Ayala, et al., Uncertainty handling in intra-operative multispectral imaging with invertible neural networks, International Conference on Medical Imaging with Deep Learning-Extended Abstract Track (2019).
[175] T.J. Adler, L. Ayala, L. Ardizzone, et al., Out of distribution detection for intra-operative functional imaging, Uncertainty for Safe Utilization of Machine Learning in Medical Imaging and Clinical Image-Based Procedures (2019) 75–82.
[176] A. Holzinger, G. Langs, H. Denk, et al., Causability and explainability of artificial intelligence in medicine, Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. 9 (4) (2019) e1312.
[177] W. Samek, T. Wiegand, K.-R. Müller, Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models, arXiv preprint arXiv:1708.08296, 2017.
[178] D.F. Steiner, R. MacDonald, Y. Liu, et al., Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer, Am. J. Surg. Pathol. 42 (12) (2018) 1636.
[179] A.P. Regensburger, L.M. Fonteyne, J. Jüngert, et al., Detection of collagens by multispectral optoacoustic tomography as an imaging biomarker for duchenne muscular dystrophy, Nat. Med. 25 (12) (2019) 1905–1915.
[180] S. Bohndiek, Addressing photoacoustics standards, Nat. Photonics 13 (5) (2019) 298.
[181] M. Fisher, K. Takano, The penumbra, therapeutic time window and acute ischaemic stroke, Bailliere's Clin. Neurol. 4 (2) (1995) 279–295.
[182] J. Gamelin, A. Maurudis, A. Aguirre, et al., A real-time photoacoustic tomography system for small animals, Opt. Express 17 (13) (2009) 10489–10498.
[183] J. Kim, S. Park, Y. Jung, et al., Programmable real-time clinical photoacoustic and ultrasound imaging system, Sci. Rep. 6 (2016) 35137.
[184] T. Kirchner, F. Sattler, J. Gröhl, et al., Signed real-time delay multiply and sum beamforming for multispectral photoacoustic imaging, J. Imaging 4 (10) (2018) 121.
[185] R. Shang, K. Hoffer-Hawlik, G.P. Luke, A Two-Step-Training Deep Learning Framework for Real-Time Computational Imaging without Physics Priors, arXiv preprint arXiv:2001.03493, 2020.
[186] Z. Akkus, J. Cai, A. Boonrod, et al., A survey of deep-learning applications in ultrasound: artificial intelligence-powered ultrasound for improving clinical workflow, J. Am. Coll. Radiol. 16 (9) (2019) 1318–1328.
[187] S. Tonekaboni, S. Joshi, M.D. McCradden, et al., What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use, arXiv preprint arXiv:1905.05134, 2019.
[188] A.B.E. Attia, G. Balasundaram, M. Moothanchery, et al., A review of clinical photoacoustic imaging: current and future trends, Photoacoustics 16 (2019) 100144.

Janek Gröhl received his MSc degree in medical informatics from the University of Heidelberg and Heilbronn University of Applied Sciences in 2016. He submitted his PhD thesis to the medical faculty of the University of Heidelberg in January 2020. In 2020, he worked as a postdoctoral researcher at the Division of Computer Assisted Medical Interventions (CAMI), German Cancer Research Center (DKFZ) and is now working as a research associate at the Cancer Research UK Cambridge Institute and does research in software engineering and computational biophotonics focusing on machine learning-based signal quantification in photoacoustic imaging.

Melanie Schellenberg received her MSc degree in Physics from the University of Heidelberg in 2019. She is currently pursuing an interdisciplinary PhD in computer science at the Division of Computer Assisted Medical Interventions (CAMI), German Cancer Research Center (DKFZ) and aiming for quantitative photoacoustic imaging with a learning-to-simulate approach.

Kris Dreher received his MSc degree in Physics from the University of Heidelberg in 2020. He is currently pursuing a PhD at the Division of Computer Assisted Medical Interventions (CAMI), German Cancer Research Center (DKFZ) and does research in deep learning-based domain adaptation methods to tackle the inverse problems of photoacoustic imaging.



Lena Maier-Hein received her doctoral degree from the Karlsruhe Institute of Technology in 2009 and conducted her postdoctoral research in the Division of Medical and Biological Informatics, German Cancer Research Center (DKFZ), and at the Hamlyn Centre for Robotics Surgery, Imperial College London. She is now head of the Division of Computer Assisted Medical Interventions (CAMI) at the DKFZ with a research focus on surgical data science and computational biophotonics.

4.2 P.II: Open-source simulation toolkit

SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics

Authors: Janek Gröhl, **Kris K. Dreher**, Melanie Schellenberg, Tom Rix, Niklas Holzwarth, Patricia Vieten, Leonardo Ayala, Sarah E. Bohndiek, Alexander Seitel, Lena Maier-Hein

Status: Published
Journal: Journal of Biomedical Optics
DOI: doi.org/10.1117/1.JBO.27.8.083010

Contributions: JG and **KKD** are shared principal authors of this work. JG, **KKD** and LMH conceived and conceptualized the project. JG, SEB, AS and LMH shared administration of the project. JG and **KKD** defined methodology and the experiments to be conducted, and wrote the first version of the manuscript.
In the following, the contributions to the codebase are described: JG and **KKD** co-developed the codebase of the package and **KKD** wrote the main body of the code for the experiments. MS contributed to the volume and tissue creation. TR contributed to the reconstruction algorithms. NH and PV contributed to linear unmixing and quantification processing components, respectively. LA contributed to visualization. The exact coding contributions can be seen at github.com/IMSY-DKFZ/simpa/graphs/contributors and github.com/IMSY-DKFZ/simpa_paper_experiments/graphs/contributors.
All authors contributed to reviewing and editing the manuscript. LMH acquired the funding and ensured the project's scientific rigor and relevance.

SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics

Janek Gröhl,[a,†,‡] **Kris K. Dreher**,[a,b,*,‡] **Melanie Schellenberg**,[a,c,d] **Tom Rix**,[a,c] **Niklas Holzwarth**,[a] **Patricia Vieten**,[a,b] **Leonardo Ayala**,[a,e] **Sarah E. Bohndiek**,[f,g] **Alexander Seitel**,[a,§] **and Lena Maier-Hein**[a,c,e,§]

[a]German Cancer Research Center (DKFZ), Division of Intelligent Medical Systems, Heidelberg, Germany
[b]Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany
[c]Heidelberg University, Faculty of Mathematics and Computer Science, Heidelberg, Germany
[d]HIDSS4Health - Helmholtz Information and Data Science School for Health, Heidelberg, Germany
[e]Heidelberg University, Medical Faculty, Heidelberg, Germany
[f]University of Cambridge, Cancer Research UK Cambridge Institute, Robinson Way, Cambridge, United Kingdom
[g]University of Cambridge, Department of Physics, Cambridge, United Kingdom

Abstract

Significance: Optical and acoustic imaging techniques enable noninvasive visualisation of structural and functional properties of tissue. The quantification of measurements, however, remains challenging due to the inverse problems that must be solved. Emerging data-driven approaches are promising, but they rely heavily on the presence of high-quality simulations across a range of wavelengths due to the lack of ground truth knowledge of tissue acoustical and optical properties in realistic settings.

Aim: To facilitate this process, we present the open-source simulation and image processing for photonics and acoustics (SIMPA) Python toolkit. SIMPA is being developed according to modern software design standards.

Approach: SIMPA enables the use of computational forward models, data processing algorithms, and digital device twins to simulate realistic images within a single pipeline. SIMPA's module implementations can be seamlessly exchanged as SIMPA abstracts from the concrete implementation of each forward model and builds the simulation pipeline in a modular fashion. Furthermore, SIMPA provides comprehensive libraries of biological structures, such as vessels, as well as optical and acoustic properties and other functionalities for the generation of realistic tissue models.

Results: To showcase the capabilities of SIMPA, we show examples in the context of photoacoustic imaging: the diversity of creatable tissue models, the customisability of a simulation pipeline, and the degree of realism of the simulations.

Conclusions: SIMPA is an open-source toolkit that can be used to simulate optical and acoustic imaging modalities. The code is available at: https://github.com/IMSY-DKFZ/simpa, and all of the examples and experiments in this paper can be reproduced using the code available at: https://github.com/IMSY-DKFZ/simpa_paper_experiments.

Keywords: simulation; open-source; photoacoustics; optical imaging; acoustic imaging.

Paper 210395SSR received Dec. 20, 2021; accepted for publication Feb. 28, 2022; published online Apr. 4, 2022.

*Address all correspondence to Kris K. Dreher, k.dreher@dkfz-heidelberg.de

†Present address: Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way, Cambridge, CB2 0RE, United Kingdom, and Department of Physics, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, United Kingdom.

‡These authors contributed equally.

§Shared last authorship.

1 Introduction

Optical and acoustic imaging techniques enable real-time and noninvasive visualisation of structural and functional tissue properties without exposing the patient to harmful ionizing radiation. Nevertheless, the applicability of purely optical and acoustic imaging techniques is limited, for example, by the low penetration depth of near-infrared spectroscopy[1] or by the difficulties of measuring functional tissue properties with ultrasound imaging.[2] Furthermore, quantitative measurements are challenging as the state-of-the-art model-based approaches to solve the underlying inverse problems rely on assumptions that might not hold when applied to *in vivo* measurements.

Data-driven approaches can be chosen to address these inverse problems. To this end, high-quality well-annotated data are needed, for example, to train deep learning algorithms[3–5] or to optimize device design.[6,7] In living subjects, the acquisition of such data is extremely difficult because the underlying optical and acoustic tissue properties are generally not well known.[8] As such, for algorithm training, many researchers instead use simulated data, which are comparatively easy to obtain, have known underlying optical and acoustic properties, and can be used for both algorithm training and validation.[9–13] Nevertheless, the application of algorithms trained exclusively on synthetic training data to experimental measurements is challenging due to systematic differences between synthetic and experimental data.[14]

Photoacoustic imaging (PAI) combines the advantages of optical and acoustic imaging by exploiting the photoacoustic (PA) effect, resulting in optical contrast with scalable high spatial resolution down to microns as a function of imaging depth, which can be up to several centimeters.[15] PAI enables the recovery of functional tissue properties, such as blood oxygen saturation.[16] To quantitatively recover such parameters, two inverse problems have to be solved: the acoustic inverse problem, which constitutes the accurate and quantitative reconstruction of the initial pressure distribution, and the optical inverse problem, which constitutes the quantitative recovery of the optical absorption coefficient.[8] To generate realistic PA simulations for the purpose of training a data-driven method, all physical and computational aspects of signal formation need to be considered;[17] these include synthetic volume generation, photon propagation, acoustic wave propagation, and image reconstruction.

In recent years, a heterogeneous software landscape has emerged with various open-source or free-to-use tools to cover each of these physical and computational aspects. For example, for volume generation, there exist open access resources, such as the Digimouse[18] annotated digital mouse phantom, digital breast phantoms (available at: https://github.com/DIDSR/VICTRE, last visited March 22, 2022),[19,20] and the multimodal imaging-based detailed anatomical model of the human head and neck atlas MIDA.[21] But usually, researchers use pseudorandom distributions of light-absorbing molecules (chromophores) to create tissue-mimicking *in silico* phantoms.[13,22] For optical modeling of photon transport in tissue, numerous approaches have been established; these focus in general either on (1) Monte Carlo methods including, for example, mcxyz,[23] MCX,[24] or ValoMC,[25] which uses a Monte Carlo approach to light transport to simulate the propagation of photons in heterogeneous tissue, or (2) analytical methods to solve the radiative transfer equation, including diffusion approximation or finite element solvers as implemented in, for example, NIRFAST[26] or Toast++.[27] For acoustic modeling, there exists the popular k-Wave[28] toolbox, which is a third-party MATLAB toolbox for the simulation and reconstruction of PA wave fields and is one of the most frequently used frameworks in the field. For image reconstruction, there are many different approaches, including backprojection algorithms,[29–31] model-based algorithms,[32,33] and fast Fourier transform-based reconstruction algorithms.[34,35]

To navigate these tools and integrate them into a complete pipeline, the user must transform the output of each toolkit into an appropriate form for input to the next[36,37] or model the entire process in a joint computational framework.[38,39] Each step in assembling these pipelines can be time-consuming or error-prone, especially including correct consideration of the physical quantities and their units. They are typically limited to the toolkits that are currently integrated in their respective framework and thus lack broad applicability to other simulators. Furthermore, a seamless exchange from, e.g., a finite element method optical forward simulator to a Monte Carlo simulator is not straightforward in existing frameworks.

To tackle these challenges, we developed the open-source simulation and image processing for photonics and acoustics (SIMPA) Python toolkit, which features a modular design that allows for easy exchange and combination of simulation pipeline elements. In its first version, the toolkit facilitates the simulation and processing of PA data and provides a straightforward way to adapt a simulation to meet the specific needs of a given researcher or project. It can easily be extended to support simulations corresponding to other optical and acoustic imaging modalities. The core idea of the framework is to standardize the information flow between different computational models by providing a central software architecture that abstracts from the individual requirements of external libraries. SIMPA achieves this by defining abstract implementations of the simulation steps based on adapters that can be implemented, such that specific toolkits can easily be integrated into the SIMPA ecosystem. SIMPA is tested using both Windows (specifically Windows 10) and Linux (specifically Ubuntu 20.04) operating systems. Third-party toolkits are executed on the GPU by default if this is supported by the respective toolkit and a compatible GPU is installed. Furthermore, SIMPA offers the possibility of exporting simulated time-series data compliant to the data format proposed by the International Photoacoustic Standardisation Consortium (IPASC).[40]

In this paper, we first outline the purpose and the software details of SIMPA in Sec. 2. Here, we give an overview of the software development process, the software architecture, the modeling of digital device twins, and computational tissue generation. Afterward, there is an extensive simulation and image processing examples section (Sec. 3) in which we show the possibilities that SIMPA offers. We demonstrate the modularity of SIMPA by showcasing the results of example simulations including an overview of how parameter choices can affect the results and the degree of realism of the simulations that is achievable with SIMPA.

2 SIMPA Toolkit

SIMPA aims to facilitate realistic image simulation for optical and acoustic imaging modalities by providing adapters to crucial modeling steps, such as volume generation, optical modeling, acoustic modeling, and image reconstruction (Fig. 1). SIMPA provides a communication layer between various modules that implement optical and acoustic forward and inverse models.

Non-experts can use the toolkit to create sensible simulations from default parameters in an end-to-end fashion. Domain experts are provided with the functionality to set up a highly customisable pipeline according to their specific use cases and tool requirements.

The following high-level requirements are key to meeting the above purpose:

1. Modularity: The different modules of the simulation pipeline should be implemented such that each of them can be paired with arbitrary implementations of preceding or succeeding modules. Specific module implementations can seamlessly be exchanged without

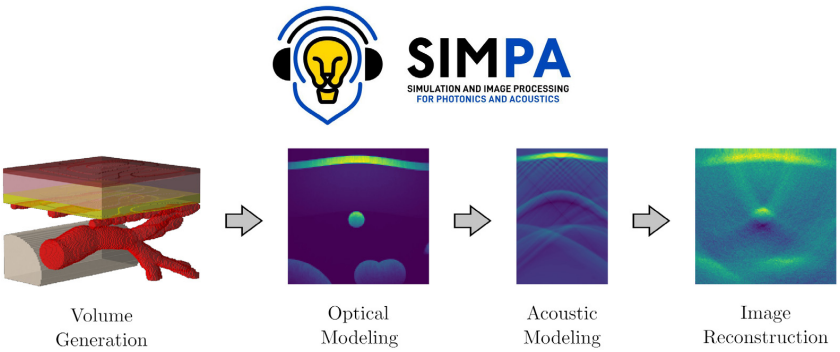


Fig. 1 The simulation and image processing for photonics and acoustics (SIMPA) toolkit.

breaking the simulation pipeline. The user has the freedom to arrange the elements of the simulation pipeline in exactly the way that they choose.
2. Extensibility: The user should have the freedom to add any custom elements to the pipeline and to implement custom module adapters. There should exist documentation that shows how custom adapters for each module and completely new modules can be implemented.
3. Physical correctness: Each module implementation should have a single purpose, produce plausible results, and not alter other parts of the pipeline. Physical quantities (i.e., units) should be correctly handled by the information flow between separate modules.
4. Independence: Arbitrarily many sequentially executed SIMPA simulations should not influence the results of subsequent simulations.
5. Usability: The entry for new users must be as easy as possible such that sensible PA images can be simulated without prior knowledge. A simulation with default parameters can be started using only a few lines of code.

The following sections of this paper introduce the software development life cycle in Sec. 2.1 and SIMPA's software architecture in Sec. 2.2, as well as another prominent contribution of SIMPA: a volume creation adapter that enables the user to create diverse spatial distributions of tissue properties as detailed in Sec. 2.4.

2.1 *Software Development Life Cycle*

SIMPA is developed using the Python programming language (Python Software Foundation),[41] version 3.8 because it is currently one of the most commonly used programming languages. We use git[42] as the version control system, and the code is maintained on GitHub (available at: https://github.com/IMSY-DKFZ/simpa, last visited March 22, 2022). Stable versions of the develop branch are integrated into the main branch and then form a release with an increase in the version number according to the Semantic Versioning Specification (SemVer) scheme.[43]

SIMPA code is written using a quality-controlled development process. Every feature request or bug fix is assigned an issue on the SIMPA GitHub page (available at: https://github.com/IMSY-DKFZ/simpa/issues, last visited March 22, 2022). Issues can be opened and commented on by any SIMPA user, and the code is written in separate branches that are only integrated into the develop branch after a successful code review by a member of the SIMPA core developer team. To ensure good code quality, the code reviews are designed to check whether the code follows the SIMPA developer guide:

1. The code is executable and yields the expected result in a typical use case.
2. The code is accompanied by an automatic or manual test.
3. The code is written using the Python Enhancement Proposal (PEP) 8 style guide for Python code.[44]
4. The code documents its intended use case, input parameters, and expected output.

More specifically, each new feature and bug fix must add a unit test that tests the functionality of the feature. If automatic unit testing is not possible (e.g., because required third-party binaries are not available in the integrated testing environment), then a manual integration test is defined in which the feature is being used within a SIMPA simulation run. The output of the manual test is then reviewed by a SIMPA developer as a sanity check. Using such a mixture of automatic and manual tests, we aim to provide tests for every intended use case of SIMPA to ensure that the toolkit is stable and working as intended.

2.2 *Software Architecture*

SIMPA provides a unified abstract data structure that combines existing simulation tools to represent the full signal generation process of a given optical and/or acoustic imaging

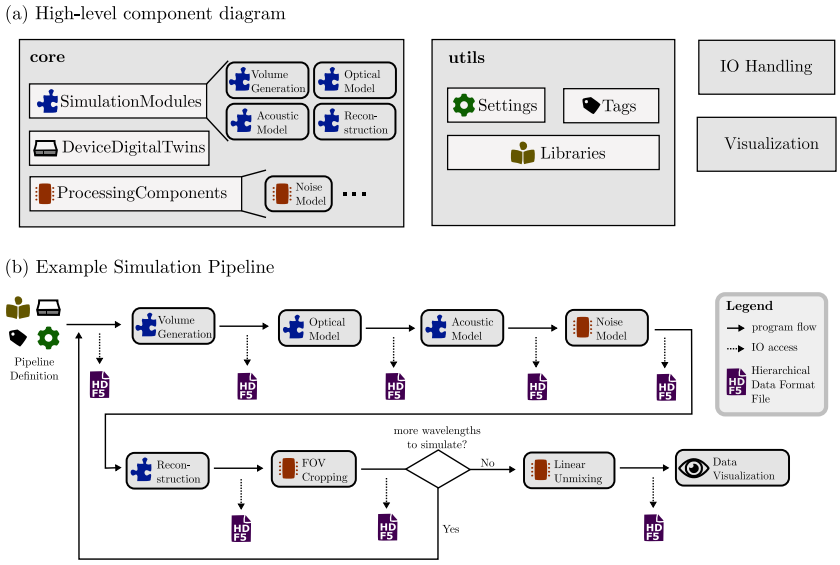


Fig. 2 Software components of SIMPA. (a) The main software components of SIMPA's software architecture. The toolkit consists of two main components, `core` and `utils`, as well as several smaller components (e.g., `io_handling`, `visualisation`), which are each composed of several subcomponents. The `core` contains all `SimulationModules`, `DeviceDigitaltwins`, and `ProcessingComponents`. The `utils` component contains the `Settings` dictionary, a standardized list of `Tags`, various `Libraries`, and other utility and helper classes to facilitate using the toolkit. (b) An example simulation pipeline. The pipeline is defined via a `Settings` dictionary using a standardized list of `Tags`. During the pipeline execution, each pipeline element (which can be either a `SimulationModule` or a `ProcessingComponent`) is called sequentially. After each step, the new results are amended to a hierarchical data format 5 (HDF5) file. The pipeline is repeated for each wavelength; afterwards, all multispectral `ProcessingComponents` are executed, and the results can be visualised. In this example, the included pipeline elements are volume generation, optical modeling, acoustic modeling, noise modeling, image reconstruction, field of view (FOV) cropping, linear unmixing, and result visualisation.

modality. Specifically, SIMPA handles the data/information flow from and to each simulation tool and provides an infrastructure to use these tools in an integrated pipeline. Figure 2 shows the main components of SIMPA and visualises their interactions in an example simulation pipeline.

SIMPA contains two primary Python modules: `core` and `utils`. Furthermore, the toolkit features several smaller Python modules: `io_handling`, `log`, `visualisation`, `algorithms`, `examples`, and `tests`. The SIMPA `core` defines a centralized structure to provide simulation tool-specific adapters to the abstract modules for each step in the simulation process. The `utils` package provides a collection of libraries and convenience methods to help a researcher set up a customized simulation pipeline.

2.2.1 *Core*

The `core` is organized into three Python submodules. The `SimulationModules` submodule provides interfaces for all simulation modules (e.g., the ones that are required for complete PA forward modeling). To meet the modularity criterion (Sec. 2), it contains abstract module definitions for the major parts: `VolumeCreationModule`, `OpticalForwardModule`, `AcousticForwardModule`, and `ReconstructionForwardModule`. Furthermore, the `core` contains a `ProcessingComponents`

submodule that contains a base for components that supplement the main simulation modules, such as a component for noise modeling, which currently supports a number of noise models: salt and pepper noise, Gaussian noise, Poisson noise, uniform noise, and Gamma noise. Finally, the `DeviceDigitalTwins` submodule also contains base classes that enable the definition of digital twins of PA imaging devices such as slit or pencil illuminations combined with circular or linear detector geometries. Details on the digital device representation in SIMPA can be found in Sec. 2.3.

The main entry point for the user is the `simulate` method that is contained in the `core`. This method is responsible for the execution of all desired simulation modules and processing components (referred to as `pipeline_elements`).

To meet the extensibility criterion (Sec. 2), a developer has the freedom to add custom new simulation module adapters, processing components, or digital device twins. Each pipeline element in the simulation pipeline has to be fully self-contained and thus handle its produced result correctly within the information flow of SIMPA. To ensure this, each of the Python submodules provides an abstract class that encapsulates parts of the functionality. For example, a user can define a custom simulation module using the abstract `SimulationModule` class as a blueprint. To implement a Python adapter, it has to inherit from this class and overwrite the `implementation` method. Internally, the representation of the computational grid is defined by isotropic voxels. This does not necessarily exclude external tools that work on differently defined grids such as anisotropic voxels or mesh-based methods if the according adapter translates one into another. The edge size of the voxels is generally defined by the user attribute `SPACING_MM`, but this would not prevent an adapter from resampling the voxel sizes. The grid uses the default unit for length within SIMPA, which is mm. The default unit for time in SIMPA is ms.

2.2.2 *Utils*

The `utils` Python module contains utility classes such as the `Tags` and `Settings` classes. The `Settings` class is a dictionary that contains key-value pairs defining the simulation parameters. To assert standardized naming conventions of the dictionary keys, these keys are globally accessible via the `Tags` class. Furthermore, there is the `Libraries` package that provides both `LiteratureValues` as well as collections of classes that represent, e.g., geometrical shapes and biochemical molecules. The `LiteratureValues` are used to instantiate these classes for the purpose of generating synthetic tissue models. The `Libraries` package provides the following collections:

- `LiteratureValues`: Reference values for optical, acoustic, and morphological tissue properties including the respective online source.
- `SpectrumLibrary`: Classes based on a `Spectrum`. A `Spectrum` represents wavelength-dependent tissue properties such as optical absorption or scattering defined for a specific set of wavelengths (depending on the reference literature).
- `MoleculeLibrary`: Classes based on a `Molecule`. The `Molecule` class is used to represent the optical and acoustic properties of biochemical molecules such as melanin or hemoglobin.
- `TissueLibrary`: Predefined `MolecularComposition` classes. A `MolecularComposition` is a linear mixture of different `Molecules`. The elements of the `TissueLibrary` are defined such that the optical and acoustic properties of the mixed `Molecules` agree with the literature references (e.g., skin or blood).
- `StructureLibrary`: Classes based on a `Structure`. Each `Structure` defines the geometry of a certain volumetric shape (such as cuboids, tubes, or vessel trees) in a voxelized grid.

The interplay of these libraries is described in greater detail in Sec. 2.4. All libraries are designed such that they are easily customisable by the user, for example, by allowing for the addition of spectra, molecules, or tissue types.

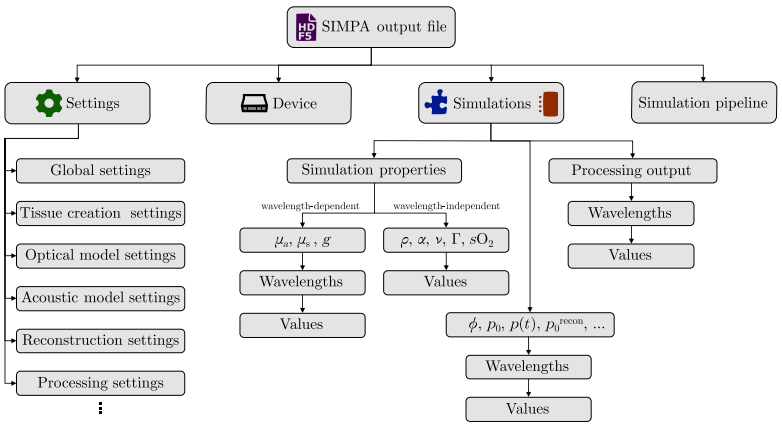


Fig. 3 The SIMPA file data structure is hierarchical. The output file of SIMPA uses the Hierarchical Data Format 5 (HDF5). The top-level fields are (1) `Settings` in which the input parameters for the global simulation pipeline as well as for all pipeline elements are stored in. (2) The `Device` describes the digital device twin with which the simulations are performed. (3) The `Simulations` field stores all of the simulation property maps that serve as input for the pipeline elements, such as the optical absorption (μ_a), scattering (μ_a), and anisotropy (g). These properties are wavelength-dependent and therefore are saved for each wavelength respectively. The density (ρ), acoustic attenuation (α), speed of sound (ν), Grüneisen parameter (Γ) or blood oxygen saturation (sO_2) are wavelength-independent and therefore only stored once. The `Simulations` field also stores the outputs for each wavelength of each processing component and simulation module such as optical fluence (ϕ), initial pressure (p_0), time series pressure data ($p(t)$), or the reconstructed image (p_0^{recon}). (4) The simulation pipeline is a list that stores the specific module adapters that have been combined and their order to form the simulation pipeline.

2.2.3 *IO and data format*

SIMPA uses the *Hierarchical Data Format 5* (HDF5)[45] as it comprises a hierarchical data structure, has interfaces in many commonly used programming languages, and features the possibility of adding metadata. All inputs, settings, and outputs of a SIMPA simulation are stored in a central HDF5 file, and at the end of the simulation, the file contents can be repacked to be saved in a compressed manner. The SIMPA `io_handling` Python module abstracts from the communication with the `h5py` package[46] and contains functionality to save and write data to the hard drive (Fig. 3).

SIMPA also offers the feature to export simulated time-series data into the data format proposed by IPASC.[47] This data format is based on HDF5 as well and defines a standardized list of metadata parameters to include.[48]

2.3 *Digital Device Twins*

SIMPA enables the definition of digital twins of optical and acoustic devices by providing abstract base classes for the implementation of detectors and illuminators (cf. Fig. 4). To this end, SIMPA contains the `DetectionGeometryBase` and `IlluminationGeometryBase` classes, both of which inherit from the `DigitalDeviceTwinBase` class. The `DigitalDeviceTwinBase` class defines the device position and the field-of-view (FOV). The `DetectionGeometryBase` and `IlluminationGeometryBase` classes are responsible for defining the necessary parameters and abstract methods for the implementation of custom devices. To define a detection geometry or an illumination geometry, a class that inherits from the fitting base class and implements the necessary abstract methods needs to be written. A PA device is defined by having both a detection and an illumination geometry.

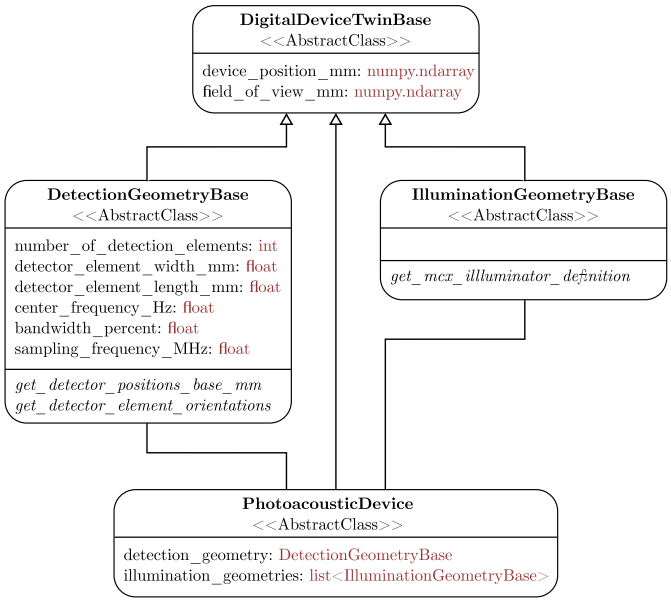


Fig. 4 Unified modeling language (UML) class diagram of the digital device representation in SIMPA. Each box represents a class with the class name in bold. The first set of elements are the fields defined by these classes with their types shown in red, and the italic fields refer to abstract methods. A PA device comprises a detection geometry and an illumination geometry. All classes inherit from the `DigitalDeviceTwinBase` class, which defines common attributes: the device position and the FOV.

SIMPA predefines some commonly used detection and illumination geometries, as well as some PA devices. Currently, SIMPA provides classes for curved arrays, linear arrays, and planar array detection geometries, as well as disk, Gaussian beam, pencil beam, pencil array, and slit illumination geometries. Using these classes, the user can freely combine detection and illumination geometries as well as their relative positions to accurately represent real devices. SIMPA also provides digital twins of some PA devices: the multispectral optoacoustic tomography (MSOT) Acuity Echo, the MSOT InVision 256-TF, or the raster-scan optoacoustic mesoscopy (RSOM) Xplorer P50 from iThera Medical (iThera Medical GmbH, Munich, Germany). Because SIMPA currently only supports MCX as the optical forward model, and MCX only has a limited amount of supported illumination geometries, the MSOT Acuity Echo and the MSOT InVision 256-TF illumination geometries are represented by individual classes, and a version of MCX that supports these geometries is provided as a fork at: https://github.com/IMSY-DKFZ/mcx, (last visited March 22, 2022).

2.4 *Diverse Tissue Modeling*

A core prerequisite for the simulation of realistic PA images is the modeling of diverse tissue geometries by generating plausible distributions of optical and acoustic parameters in a virtual volume. In this context, diversity comprises not only a wide variety of geometrical shapes that might occur in biological tissue but also the accurate modeling of optical and acoustic properties of different tissue types such as skin, blood, or fat. These tissue types are usually mixtures of molecules each with distinct properties, which can be difficult to represent computationally. To meet this need, SIMPA provides a `VolumeCreation` module that enables the convenient generation of custom tissue models. The backbone of the module is the way that the optical and acoustic tissue properties are represented using flexible `MolecularCompositions` (see Fig. 5). Using a hierarchical listing of predefined structures, the user can then create custom

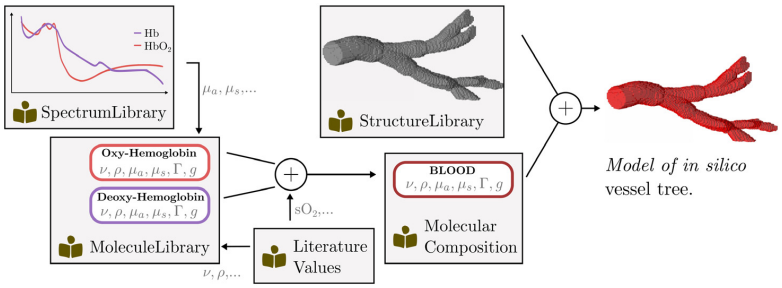


Fig. 5 Overview of the steps involved for modeling an *in silico* vessel tree with SIMPA. The diagram shows the resources that SIMPA provides for users to create custom tissue models. Wavelength-dependent properties such as the optical absorption (μ_a), scattering (μ_s), or scattering anisotropy (g) are provided in the `SpectrumLibrary`, whereas wavelength-independent properties such as the speed of sound (ν), the tissue density (ρ), or the Grüneisen parameter (Γ) are provided by the `LiteratureValues`. A `MolecularComposition` corresponds to a linear mixture of `Molecules` that can be used in combination with a geometrical molecular distribution from the `StructureLibrary` to create an *in silico* model.

spatial distributions of these molecular compositions. The ability to create realistic tissue models depends on several factors, which include (1) the availability of high-quality reference measurements for the optical and acoustic properties, (2) information on the molecular composition of various tissue types, and (3) an accurate representation of their spatial distributions within the region of interest.

2.4.1 *Optical and acoustic molecule properties*

A full list of all tissue properties considered in SIMPA is given in Tables 1–2. Within the SIMPA codebase, these molecular properties are integrated as inherent parts of a `Molecule`. While most properties can be approximated as singular values, the optical absorption, scattering, and anisotropy are wavelength-dependent and therefore represented by a `Spectrum` that linearly interpolates between the nearest known wavelengths to approximate the full spectrum during simulation.

For wavelength-dependent information on the optical properties of the chromophores most commonly found in human tissue, Jacques published an invaluable review article[49] and made information available via the OMLC website.[50] We decided to follow the system of units introduced in the cited literature in SIMPA. For the tissue properties relevant for acoustic forward modeling, we used the IT'IS database for thermal and electromagnetic parameters of biological tissues[51] as it provides information on the mean value and distribution of these properties for many different tissue types. Other literature sources that are being used by SIMPA for representing molecular properties are Kedenburg et al.[52] for heavy water, Zhang et al.[53] for water, or Antunes et al.[54] for the optical properties of bone.

Table 1 Overview of all optical properties that are represented in a SIMPA molecule with their respective units.

Optical properties	Unit
Absorption coefficient	cm^{-1}
Scattering coefficient	cm^{-1}
Scattering anisotropy	Unitless
Grüneisen parameter	Unitless

Table 2 Overview of all acoustic properties that are represented in a SIMPA molecule with their respective units.

Acoustic properties	Unit
Speed of sound	$m\,s^{-1}$
Density	$kg\,m^{-3}$
Acoustic attenuation	$dB\,cm^{-1}\,MHz^{-1}$

2.4.2 *Molecular tissue compositions*

To represent the properties of a tissue type, SIMPA uses a `MolecularComposition`, which is a linear combination of `Molecules`. In the `TissueLibrary`, SIMPA provides several predefined tissue types such as blood, skin, muscle, and bone that are compiled from literature sources. However, the framework can also be used to easily define custom user-specific molecular compositions.

Information on the molecular composition of tissue types is sparse and scattered throughout the literature. SIMPA models the properties of different skin layers and muscle tissue using the review article of Bashkatov et al.,[55] melanin content in the epidermis using Alaluf et al.,[56] the water volume fractions of different tissue types in the human body using Timmins and Wall[57] and Forbes et al.,[58] and the distribution of arterial and venous blood oxygenations using Molnar and Nemeth[59] and Merrick and Hayes.[60]

2.4.3 *Spatial tissue distribution*

Taking the creation of an *in silico* forearm as an example, specialized clinical papers can be used to obtain information on aspects such as the distribution of sizes of the radial and ulnar artery[61] and their accompanying veins,[62] the thickness of skin layers such as the dermis and epidermis,[63] the separation of the radius and ulna bones,[64] and the depth of subcutaneous vessels.[65]

SIMPA offers the ability to create voxelized volumes of molecular compositions and provides two main ways to create their spatial distributions.

Model-based volume generator. The purpose of this `Adapter` is to enable a rule-based creation of voxelized simulation volumes. The generator is given a list of structures that are each represented by a voxelized definition of their shape, a molecular composition, and a priority. In the case of two structures occupying the same voxel, the molecular composition of the structure with the higher priority is chosen for that voxel. Based on their shape and priority, all structures are then merged into a single distribution of optical and acoustic parameters.

SIMPA provides a `StructureLibrary` that contains many basic 3D shapes (`Structures`), such as layers, spheres, elliptical tubes, cuboids, parallelepipeds, or vessel trees. These `Structures` can be mixed to create arbitrary simulation volumes. For the generation of vessel trees, we have integrated a random walk-based algorithm into SIMPA, in contrast to other work that uses Lindenmayer systems to build a grammar with the inclusion of stochastic rules.[66]

Segmentation-based volume generator. The purpose of this `Adapter` is to take voxelized segmentation masks as input and map them to specific tissue types, which allows for the easy inclusion of spatial tissue property distributions from other sources. The user themself is responsible for loading a segmentation mask from a file into memory and transforming it into a numpy array as an input for the SIMPA pipeline.

3 SIMPA Use Cases

The functionality spectrum covered by the SIMPA toolkit is best demonstrated by exemplary use cases. The use cases in this section build upon each other with increasing complexity. Section 3.1

introduces the initiation of a basic simulation pipeline. Section 3.2 shows the convenience of changing smaller hyperparameters of an existing pipeline and the impact on the outcome of the pipeline, and Sec. 3.3 analogously illustrates this for the change of whole simulation modules as well as digital device twins. Section 3.4 showcases the diversity of possible tissue geometries, and Sec. 3.5 compares simulation outcomes of SIMPA with a real PA image. Finally, Sec. 3.6 combines the previous sections to exemplify the generation of a diverse dataset of PA images. The optical and acoustic modeling toolkits used for all experiments in this section were MCX[24] and k-Wave,[28] using SIMPA-provided adapters. MCX uses the Monte Carlo method that repeatedly draws random variables from an underlying model distribution to reach high accuracy.[67] MCX approximates a light transport model using this method. K-Wave is based on the k-space pseudospectral method for modeling nonlinear ultrasound propagation in heterogeneous media.[68] All experiments were conducted using a workstation with an AMD(R) Ryzen 3900x 12-core central processing unit, 64 GB of RAM, and NVIDIA RTX 3090 GPU running Ubuntu 20.04., and they can be reproduced using the code available at https://github.com/IMSY-DKFZ/simpa_paper_experiments. The run times for each executable experiment are mentioned; however, a detailed analysis of SIMPA's run times, computational requirements, and postprocessing examples[69,70] can be found in the Supplementary Material.

3.1 *Running a Simulation Out-of-the-Box*

Simulations are run using the `simulate` function, which is located in the `core` Python module. The function `simulate` takes three input arguments: (1) a `list` with a definition of the simulation pipeline, (2) a `Settings` dictionary, which contains all parameters for the simulation, and (3) a `Device`, which represents a digital twin of a PAI device. The following listing shows how these three input parameters are defined and given to the `simulate` function. For each of the used simulation pipeline elements, a settings dictionary that contains the parameters needs to be defined. An overview of the user-side pseudocode to set up a simulation with SIMPA is given by:

```
import simpa as sp

# Create general settings
settings = sp.Settings(general_settings)

# Create specific settings for each pipeline element
# in the simulation pipeline
settings.set_volume_creation_settings(volume_
  creation_settings)
settings.set_optical_settings(optical_settings)
settings.set_acoustic_settings(acoustic_settings)
settings.set_reconstruction_settings(reconstruction_settings)

# Set the simulation pipeline
simulation_pipeline = [sp.VolumeCreatorModule(settings),
  sp.OpticalForwardModule(settings),
  sp.AcousticForwardModule(settings),
  sp.ReconstructionModule(settings)]

# Choose a PA device with device position in the volume
device = sp.CustomDevice()

# Simulate the pipeline
sp.simulate(simulation_pipeline, settings, device)
```

3.2 *Customising Simulation Parameters*

SIMPA enables easy customization of simulation parameters according to the criterion usability. A wide range of simulation outputs can be achieved by simply changing one parameter, such as

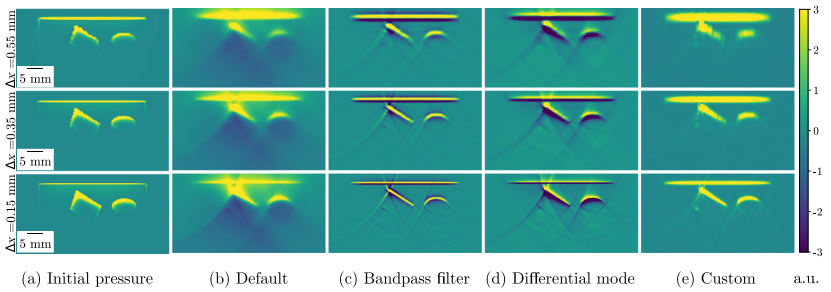


Fig. 6 Simulation results with different hyperparameter configurations using a digital device twin of the MSOT Acuity Echo (iThera Medical GmbH, Munich, Germany). The results are shown for three spacings (Δx) in three rows (0.15, 0.35, and 0.55 mm), and from left to right, the columns show the following: (a) the ground truth initial pressure distribution; (b) the default pipeline with delay-and-sum reconstruction of the time-series pressure data (pressure mode); (c) delay-and-sum reconstruction with a bandpass filter (Tukey window with an alpha value of 0.5 and 1 kHz as high-pass and 8 MHz as low-pass frequencies) applied to the time-series data; (d) delay-and-sum reconstruction with the first derivative of the time-series data (differential mode); and (e) delay-and-sum reconstruction with a bandpass filter with the same configuration as in (c), the first derivative of the time-series data and envelope detection.

Table 3 Mean run times of the optical and acoustic forward modules and image reconstruction for simulation pipelines with different parameter combinations in seconds (s). The mean time was calculated from the run times of the pipelines: default, bandpass filter, differential mode, and custom. The times are reported for three different spacings: 0.15, 0.35, and 0.55 mm.

Spacing (mm)	Optical modeling time (s)	Acoustic modeling time (s)	Image reconstruction time (s)
0.55	1.19	7.96	2.22
0.35	2.80	8.28	2.21
0.15	27.77	11.62	2.22

the spacing or image reconstruction bandpass filter (Fig. 6), with the latter achieved, for example, by setting `Tags.RECONSTRUCTION_PERFORM_BANDPASS_FILTERING` in the reconstruction module settings to `True` instead of `False` as it is by default. To showcase this, a simulation pipeline was executed with three different spacings (0.15, 0.35, and 0.55 mm) and reconstructed with the default settings (delay-and-sum), with an applied bandpass filter, with a "differential mode" (delay-and-sum of the first derivative of the time signal), and finally, with a customized set of hyperparameters. For the bandpass filter, a Tukey window[71] with an alpha value of 0.5 and 1 kHz as high-pass and 8 MHz as low-pass frequencies was applied. The set of hyperparameters was chosen such that the result is most similar to the underlying initial pressure. For different phantom designs, illumination geometries, or detection geometries, a different choice of parameters might be preferable.

The overall run time for these simulations was about 480 s. The run times of the optical and acoustic forward modules as well as the image reconstruction for the specified hardware are reported in Table 3. Only the mean for the different parameter combinations of these times are reported; however, an extensive listing of the run times of each module in each pipeline can be found in Tables S1-S4 in the Supplementary Material.

3.3 *Rapid Prototyping with Multiple Pipelines*

SIMPA facilitates simulation of phantom imaging, which is highly relevant for experimental planning and rapid prototyping. To demonstrate this, a pipeline was executed with two

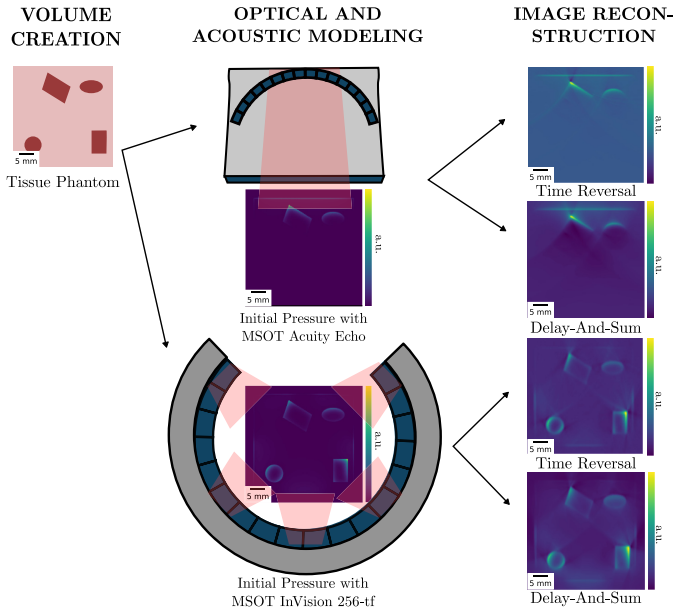


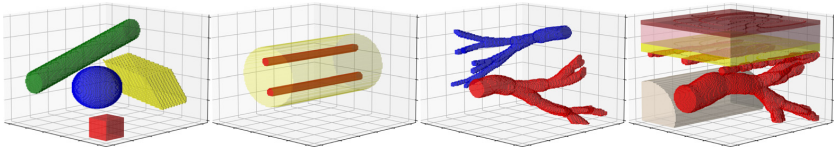
Fig. 7 Demonstration of the versatility of the toolkit. From the same tissue phantom, two initial pressure distributions and time-series data are simulated using completely different PA digital device twin [in this case, the MSOT Acuity Echo and the MSOT InVision 256-TF (iThera Medical GmbH, Munich, Germany)]. The simulated time-series data are then reconstructed using different reconstruction algorithms (time reversal and delay-and-sum), resulting in four distinct simulation results.

commercial PAI systems (MSOT Acuity Echo and the MSOT InVision 256-TF devices, iThera Medical GmbH). For each device, the optical and acoustic forward simulations were executed only once. With the produced time-series data as a result, for each device, two different reconstruction adapters were added to the pipeline to reconstruct the final PA images. Currently, the following reconstruction algorithms are supported: delay-and-sum,[29] delay-multiply-and-sum,[31] signed delay-multiply-and-sum,[72] and time reversal.[73] The results that are shown in Fig. 7 are reconstructed with delay-and-sum or time reversal. The overall run time for these simulations was about 320 s with a spacing of 0.15 mm.

Not only does the user have the ability to easily exchange devices and module adapters in the simulation pipeline but the pipeline can also be designed in such a way that the simulation is executed efficiently. The optical and acoustic forward models had to be simulated only once for each device, and the two image reconstruction algorithms were applied afterward, demonstrating the modularity of SIMPA.

3.4 *Generating Diverse Tissue Geometries*

A wide range of *in silico* tissue models can be generated using SIMPA. For this, we specifically showcase tissue structure distributions aligned to different use cases from the literature. Figure 8(a) shows an arrangement of different geometrical shapes such as cuboids and spheres as used by Cox et al.[69] In Fig. 8(b), a cylindrical phantom with two absorbing inclusions, comparable to the one presented by Hacker et al.,[74] is generated. A volume containing complex vessel trees can easily be generated similar to the human lung vessel dataset acquired from computed tomography used by Bench et al.[14] as shown in Fig. 8(c). Lastly, realistic tissue models such as a human forearm used by Gröhl et al.[13] are possible by combining the previously mentioned structures, which are shown in Fig. 8(d). The overall run time for these simulations was about 10 s.



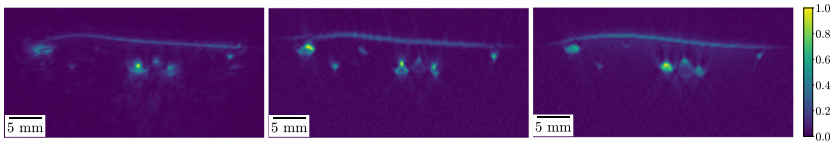
(a) Geometrical shapes (b) Phantom (c) Vessel tree (d) Forearm model

Fig. 8 Examples of chromophore distributions that can be created using the SIMPA volume generation module. (a) Arbitrarily placed and oriented geometrical structures, i.e., a tube (green), a sphere (blue), a parallelepiped (yellow), and a cuboid (red); (b) a cylindrical phantom (yellow) with two tubular inclusions (red); (c) a vessel tree with high blood oxygen saturation (red) and a vessel tree with lower blood oxygen saturation (blue); and (d) a forearm model including the epidermis (brown), dermis (pink), fat (yellow), vessels (red), and a bone (gray).

3.5 *Simulating Realistic Photoacoustic Images*

Being as realistic as possible is key to many applications in which simulated data are needed. Nevertheless, it has been reported multiple times that a domain gap exists between simulated and real PA images.[3–5,14] By its modular nature, SIMPA can be used to simulate PA images with a high degree of realism. To visually demonstrate the capabilities of the current version of SIMPA in this regard, an image of a human forearm was recorded from a volunteer using the MSOT Acuity Echo. The measurement was conducted within a healthy volunteer study that was approved by the ethics committee of the medical faculty of Heidelberg University under reference number S-451/2020, and the study is registered with the German Clinical Trials Register under reference number DRKS00023205. Based on this real image, the model-based, as well as the segmentation-based volume creators were used to synthetically recreate this image with SIMPA to compare the results with the original image (Fig. 9).

For the segmentation-based volume creator, the original image was manually annotated, and the different classes were assigned tissue properties by trial and error, so the image as a whole looks as close to the original image as possible. Using the model-based volume creator, the volume of the original image was recreated using the basic geometrical structures as described in Sec. 2.4. It should be mentioned that the model-based recreation of high-quality images, such as the one depicted in Fig. 9, is relatively time-consuming as it requires substantial manual interaction. To address this resource bottleneck and thus pave the way for the generation of large (training) data sets as required by modern machine learning algorithms, SIMPA also offers the option of generating the simulation volumes from predefined sets of rules (see Sec. 3.6). The results show that both of these methods can lead to images that closely resemble the real PA image. The overall run time for these simulations was about 80 s.



(a) Real image (b) Segmentation-based image (c) Model-based image

Fig. 9 Comparison of simulations using SIMPA with a real PA image of a human forearm. From left to right, the panels show: (a) the normalized reconstructed PA image of a real human forearm acquired with the MSOT Acuity Echo; (b) a simulated image using SIMPA's segmentation-based volume creator with a reference segmentation map of (a); and (c) a simulated image using SIMPA's model-based volume creator. For both volume creators, a digital device twin of the MSOT Acuity Echo was used. For easier comparison, all images were normalized from 0 to 1 in arbitrary units.

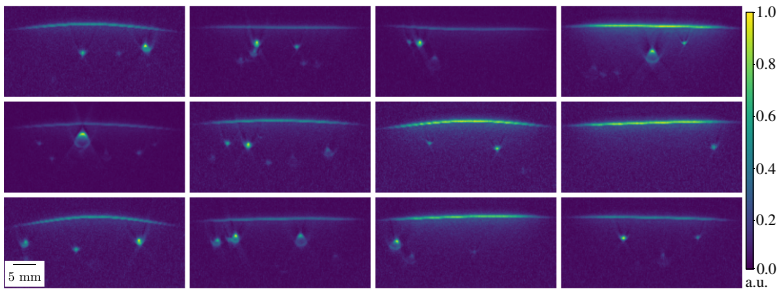


Fig. 10 Example of a diverse dataset of simulated PA images. With randomized settings of amount, location, size, shape, and blood oxygen saturation of vessels as well as the curvature of the skin, 12 diverse PA images were generated and normalized between 0 and 1 in arbitrary units (a.u.). The spacing of all images was 0.15 mm. For all simulations, a digital device twin of the MSOT Acuity Echo was used.

3.6 *Generating a Diverse Dataset of Photoacoustic Images*

For the training and the generalization ability of a complex deep learning model, a large and diverse dataset is crucial. In PAI, however, a vast amount of real PA images with ground truth annotations for their underlying properties such as optical absorption *in vivo* is not feasible. To remedy this, SIMPA can be used to generate an arbitrarily large dataset of simulated PA images with a degree of realism that can be seen in the previous section. In Fig. 10, 12 diverse PA images were simulated using randomized tissue mimicking settings of SIMPA's model-based volume creator.

These randomized settings allow for controlled distributions of, for example, amount of vessels, vessel locations, skin curvature, and blood oxygen saturations. The overall run time for the generation of these 12 images was about 570 s with a spacing of 0.15 mm. An investigation of adverse programming effects in SIMPA when generating such a dataset can be found in the Supplementary Material.

4 Conclusion and Discussion

In this work, we present SIMPA, an open-source software library that allows for the simulation and image processing of optical and acoustic imaging modalities taking into account user-specific requirements common in the community. Core to the toolkit is its modular design, which allows for a flexible definition of simulation and processing pipelines. To this end, SIMPA defines abstract interfaces for the necessary forward modeling steps that allow for the integration of arbitrary third-party simulation tools in addition to modules already implemented in SIMPA. It already includes interfaces to toolkits that are commonly used in the field, such as MCX[24] and k-Wave,[28] is open-source, and is actively maintained and improved. Furthermore, a strong emphasis has been placed on tissue modeling as the basis for each simulation. SIMPA provides methods and functionalities to generate numerical tissue models that incorporate optical and acoustic tissue properties by means of a dynamic definition of molecular compositions. Using PAI as an example, we show the simulation results for several typical SIMPA use cases. By generating a diverse dataset of PA images, we demonstrate that SIMPA can create simulations with a high degree of flexibility suitable for, e.g., training of deep learning algorithms.

The images simulated with SIMPA look realistic (Fig. 9); however, because of the vast number of modeling assumptions both within SIMPA and within the used forward models, throughout all forward modeling steps, there remains a domain gap between simulated and experimental measurements. Steps toward increasing the realism of simulated images have already been taken by including various noise models and diverse tissue geometries such as the deformability of structures. This enables horizontal layers to more closely resemble the deformation of skin and vessels can thus also be squeezed analogously to applying pressure with an imaging device.

Despite these efforts, computational modeling inaccuracies such as device-specific artifacts or a heterogeneous background with, e.g., varying blood volume fraction and oxygen saturation are not yet included.

SIMPA's modular design also facilitates the exchangeability of simulation algorithms without affecting the integrity of the simulation pipeline. Because of the modular design, arbitrary pipeline elements can be added to the simulation. SIMPA provides example scripts to achieve this and comprises an extensive test suite that incorporates unit tests for the code, as well as manual test scripts that can be used to test the integration of forward models. Analysis of over 100 subsequent runs shows that sequential simulations do not affect each other; the detailed results can be found in the Supplementary Material.

Decreasing the potential for user error and lowering the barrier to entry for PA simulation is one of the core ideas behind SIMPA; hence, we show here the simulation and customization of specific use cases. SIMPA itself also contains many example scripts and documentation. The SIMPA developers try to ensure high code quality through its software development life-cycle, which includes the presence of tests, as well as internal code reviews before changes are integrated. Using SIMPA lowers the barrier of entry into the field of PA image simulation by taking over many of the researchers' responsibilities in navigating the respective simulation tools. At the same time, this increased ease of use comes at the cost of a reduced amount of flexibility, as users are limited to the SIMPA interface and do not directly control the third-party tools. Despite the high level of abstraction, there is still room for user errors that can potentially be hard to identify. For support, researchers can open issues in the SIMPA GitHub repository and can also join the SIMPA Slack channel upon request.

Two major contributions of this work are the model-based volume creator that enables the user to create diverse spatial distributions of tissue properties and the segmentation-based volume creator that loads segmentation masks. The model-based approach includes features such as the simulation of partial volume effects and the rendering of the model in different spacings. Furthermore, it is straightforward to create diverse tissue geometries using random variables during the creation process (see Sec. 3.6). SIMPA provides many utility functions that make the model-based volume creator easy to use. Rendering the scene description into a voxelized grid, however, can become computationally expensive for small spacings, and the user is limited by the SIMPA-defined structure primitives (unless they want to implement their own `Structure` classes). The segmentation-based approach addresses this issue by featuring great flexibility in the shapes that it can simulate. Moreover, the creation of the voxelized grid is generally much faster. On the negative, the spacing of the simulation is limited to the spacing of the segmentation, which can lead to hard edges and staircase artifacts.

In addition to the signal simulation steps detailed in this paper, SIMPA also provides post-processing modules for image processing. SIMPA currently provides two algorithms: (1) an iterative qPAI algorithm, implemented based on the publication of Cox et al. from 2006[69] (cf. Fig. S3 in the Supplementary Material 4.1), and (2) a linear spectral unmixing algorithm based on singular value decomposition (cf. Fig. S4 in the Supplementary Material 4.2).

Future work will include supporting more forward models, such as numerical approximations of the radiative transfer equation for photon transport in biological tissue;[75] supporting other optical imaging modalities such as multi-/hyperspectral diffuse reflectance imaging; the addition of more reconstruction algorithms; the capabilities for ultrasound simulation; and the provision of more digital commercial PA devices from a variety of vendors including distinct artifacts that are introduced by different devices. The IPASC is working on a standardiszed data format for PAI (Ref. 48) and has a digital device definition embedded in its format. They are currently planning to integrate support for their definition of the devices into MCX (available at: https://github.com/IPASC/PACFISH/issues/15, last visited March 22, 2022). Once this is achieved, we will support arbitrary illumination geometries within SIMPA. Furthermore, the variety of structures that can be used will be increased by including heterogeneous backgrounds that more closely represent the irregularities within tissue as well as larger, more complex, and connected structures that can represent organs or tumors. A great current challenge is the steep increase of needed computational resources, especially RAM and hard drive space, when decreasing the spacing of the computational grid. To this end, optimization strategies will be investigated to minimize the achievable spacing for a given hardware configuration. We only

tested SIMPA with NVIDIA GPUs for GPU acceleration, but we plan to support a wider variety of computing platforms in the future. We are currently also working toward an interactive visualisation tool for the data and the addition of a graphical user interface for SIMPA, which could further flatten the learning curve. Other interesting avenues of future work could be the consideration of heterogeneous molecular distributions within the structures or the integration of state-of-the-art deep learning-based processing components or module adapters.

Disclosures

The authors have no conflicts of interest to declare that are relevant to the content of this article.

Acknowledgments

The authors would like to thank Minu D. Tizabi for proofreading the manuscript. This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme through the ERC Starting Grant COMBIOSCOPY under Grant Agreement No. ERC-2015-StG-37960 and through the ERC Consolidator Grant NEURAL SPICING under Grant Agreement No. 101002198.

Code, Data, and Materials Availability

The experiments conducted in this paper do not require any external data. The latest release of the SIMPA code can be downloaded from GitHub (https://github.com/IMSY-DKFZ/simpa, last visited 22 March 2022). The code used to generate the results and figures is available in a GitHub repository (https://github.com/IMSY-DKFZ/simpa_paper_experiments).

References

1. R. H. Wilson et al., "Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization," *J. Biomed. Opt.* **20**(3), 030901 (2015).
2. E. Macé et al., "Functional ultrasound imaging of the brain," *Nat. Methods* **8**(8), 662–664 (2011).
3. C. Yang et al., "Review of deep learning for photoacoustic imaging," *Photoacoustics* **21**, 100215 (2021).
4. A. Hauptmann and B. T. Cox, "Deep learning in photoacoustic tomography: current approaches and future directions," *J. Biomed. Opt.* **25**(11), 112903 (2020).
5. J. Gröhl et al., "Deep learning for biomedical photoacoustic imaging: a review," *Photoacoustics* **22**, 100241 (2021).
6. T. Sowers, H. Yoon, and S. Emelianov, "Investigation of light delivery geometries for photoacoustic applications using Monte Carlo simulations with multiple wavelengths, tissue types, and species characteristics," *J. Biomed. Opt.* **25**(1), 016005 (2020).
7. L. A. Ayala et al., "Band selection for oxygenation estimation with multispectral/hyperspectral imaging," arXiv:1905.11297v2 (2021).
8. B. Cox, J. Laufer, and P. Beard, "The challenges for quantitative photoacoustic imaging," *Proc. SPIE* **7177**, 717713 (2009).
9. S. Tzoumas et al., "Eigenspectra optoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," *Nat. Commun.* **7**, 12121 (2016).
10. T. Kirchner, J. Gröhl, and L. Maier-Hein, "Context encoding enables machine learning-based quantitative photoacoustics," *J. Biomed. Opt.* **23**(5), 056008 (2018).
11. C. Cai et al., "End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging," *Opt. Lett.* **43**(12), 2752–2755 (2018).
12. H. Lan et al., "Y-net: hybrid deep learning image reconstruction for photoacoustic tomography *in vivo*," *Photoacoustics* **20**, 100197 (2020).
13. J. Gröhl et al., "Learned spectral decoloring enables photoacoustic oximetry," *Sci. Rep.* **11**, 6565 (2021).

14. C. Bench, A. Hauptmann, and B. T. Cox, "Toward accurate quantitative photoacoustic imaging: learning vascular blood oxygen saturation in three dimensions," *J. Biomed. Opt.* **25**(8), 085003 (2020).
15. P. Beard, "Biomedical photoacoustic imaging," *Interface Focus* **1**(4), 602–631 (2011).
16. Y. Zhou, J. Yao, and L. V. Wang, "Tutorial on photoacoustic tomography," *J. Biomed. Opt.* **21**(6), 061007 (2016).
17. J. Gröhl, "Data-driven quantitative photoacoustic imaging," PhD Thesis, Heidelberg University (2021).
18. B. Dogdas et al., "Digimouse: a 3d whole body mouse atlas from CT and cryosection data," *Phys. Med. Biol.* **52**(3), 577 (2007).
19. S. Park et al., "Realistic three-dimensional optoacoustic tomography imaging trials using the VICTRE breast phantom of FDA (conference presentation)," *Proc. SPIE* **11240**, 112401H (2020).
20. Y. Lou et al., "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *J. Biomed. Opt.* **22**(4), 041015 (2017).
21. M. I. Iacono et al., "MIDA: a multimodal imaging-based detailed anatomical model of the human head and neck," *PLoS One* **10**(4), e0124126 (2015).
22. S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse Prob. Sci. Eng.* **27**(7), 987–1005 (2019).
23. S. L. Jacques, "Coupling 3D Monte Carlo light transport in optically heterogeneous tissues to photoacoustic signal generation," *Photoacoustics* **2**(4), 137–142 (2014).
24. Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Express* **17**(22), 20178–20190 (2009).
25. A. A. Leino, A. Pulkkinen, and T. Tarvainen, "ValoMC: a Monte Carlo software and Matlab toolbox for simulating light transport in biological tissue," *OSA Continuum* **2**(3), 957–972 (2019).
26. H. Dehghani et al., "Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction," *Commun. Numer. Methods Eng.* **25**(6), 711–732 (2009).
27. M. Schweiger and S. R. Arridge, "The toast++ software suite for forward and inverse modeling in optical tomography," *J. Biomed. Opt.* **19**(4), 040801 (2014).
28. B. E. Treeby and B. T. Cox, "k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**(2), 021314 (2010).
29. M. Xu and L. V. Wang, "Universal back-projection algorithm for photoacoustic computed tomography," *Phys. Rev. E* **71**(1), 016706 (2005).
30. S. Park et al., "Adaptive beamforming for photoacoustic imaging," *Opt. Lett.* **33**(12), 1291–1293 (2008).
31. G. Matrone et al., "The delay multiply and sum beamforming algorithm in ultrasound b-mode medical imaging," *IEEE Trans. Med. Imaging* **34**(4), 940–949 (2015).
32. H. Grün et al., "Photoacoustic tomography using a fiber based Fabry-Perot interferometer as an integrating line detector and image reconstruction by model-based time reversal method," *Proc. SPIE* **6631**, 663107 (2007).
33. A. Hauptmann et al., "Model-based learning for accelerated, limited-view 3-D photoacoustic tomography," *IEEE Trans. Med. Imaging* **37**(6), 1382–1393 (2018).
34. Y. Xu, D. Feng, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography. I. Planar geometry," *IEEE Trans. Med. Imaging* **21**(7), 823–828 (2002).
35. M. Jaeger et al., "Fourier reconstruction in optoacoustic imaging using truncated regularized inverse k-space interpolation," *Inverse Prob.* **23**(6), S51 (2007).
36. N. Akhlaghi et al., "Multidomain computational modeling of photoacoustic imaging: verification, validation, and image quality prediction," *J. Biomed. Opt.* **24**(12), 121910 (2019).
37. S. Agrawal et al., "Modeling combined ultrasound and photoacoustic imaging: simulations aiding device development and artificial intelligence," *Photoacoustics* **24**, 100304 (2021).
38. C. Sowmiya and A. K. Thittai, "Simulation of photoacoustic tomography (PAT) system in comsol and comparison of two popular reconstruction techniques," *Proc. SPIE* **10137**, 101371O (2017).

39. C. Fadden and S.-R. Kothapalli, "A single simulation platform for hybrid photoacoustic and RF-acoustic computed tomography," *Appl. Sci.* **8**(9), 1568 (2018).
40. S. E. Bohndiek et al., "IPASC: a community-driven consensus-based initiative towartandardizationion in photoacoustic imaging," in *IEEE Int. Ultrason. Symp.*, IEEE, pp. 1–4 (2020).
41. Python Software Foundation, http://www.python.org, (accessed 22 March 2022).
42. S. Chacon and B. Straub, *Pro Git*, Apress, Berlin, Germany (2014).
43. T. Preston-Werner, Semantic Versioning 2.0.0, https://semver.org/ (accessed 22 March 2022).
44. Python Enhancement Proposals, https://www.python.org/dev/peps/pep-0008/ (accessed 22 March 2022).
45. M. Folk et al., "An overview of the HDF5 technology suite and its applications," in *Proc. EDBT/ICDT Workshop Array Databases*, pp. 36–47 (2011).
46. A. Collette, *Python and HDF5*, O'Reilly, Sebastopol, California (2013).
47. S. Bohndiek, "Addressing photoacoustics standards," *Nat. Photonics* **13**(5), 298–298 (2019).
48. International Photoacoustic Standardisation Consortium, https://www.ipasc.science/documents/20210916_IPASC_Format_V2.pdf (accessed 22 March 2022).
49. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**(11), R37 (2013).
50. S. Prahl and S. Jacques, https://omlc.org/ (accessed 22 March 2022).
51. P. Hasgall et al., "IT'IS database for thermal and electromagnetic parameters of biological tissues," Version 4.0, May 15, 2018. doi: 10.13099/VIP21000-04-0, www.itis.ethz.ch/database.
52. S. Kedenburg et al., "Linear refractive index and absorption measurements of nonlinear optical liquids in the visible and near-infrared spectral region," *Opt. Mater. Express* **2**, 1588–1611 (2012).
53. X. Zhang, L. Hu, and M.-X. He, "Scattering by pure seawater: effect of salinity," *Opt. Express* **17**, 5698–5710 (2009).
54. A. Antunes et al., "Optical properties on bone analysis: an approach to biomaterials," *Proceedings* **27**(1), 36 (2019).
55. A. N. Bashkatov, E. A. Genina, and V. V. Tuchin, "Optical properties of skin, subcutaneous, and muscle tissues: a review," *J. Innovative Opt. Health Sci.* **4**(01), 9–38 (2011).
56. S. Alaluf et al., "Ethnic variation in melanin content and composition in photoexposed and photoprotected human skin," *Pigment Cell Res.* **15**(2), 112–118 (2002).
57. P. Timmins and J. Wall, "Bone water," *Calcified Tissue Res.* **23**(1), 1–5 (1977).
58. R. Forbes et al., "The composition of the adult human body as determined by chemical analysis," *J. Biol. Chem.* **203**(1), 359–366 (1953).
59. Z. Molnar and M. Nemeth, "Monitoring of tissue oxygenation: an everyday clinical challenge," *Front. Med.* **4**, 247 (2018).
60. E. B. Merrick and T. J. Hayes, "Continuous, non-invasive measurements of arterial blood oxygen levels," *Hewlett-Packard J.* **28**(2), 2–9 (1976).
61. G. Yang and K. C. Chung, "Ulnar artery to superficial arch bypass with a vein graft," in *Operative Techniques: Hand and Wrist Surgery*, K. C. Chung, Ed., pp. 732–737, Elsevier Health Sciences, Amsterdam, Netherlands (2018).
62. M. G. Hubmer et al., "The posterior interosseous artery in the distal part of the forearm. Is the term 'recurrent branch of the anterior interosseous artery'justified?" *Br. J. Plast. Surg.* **57**(7), 638–644 (2004).
63. P. Oltulu et al., "Measurement of epidermis, dermis, and total skin thicknesses from six different body regions with a new ethical histometric technique," *Turkish J. Plast. Surg.* **26**(2), 56 (2018).
64. J. B. Christensen et al., "A study of the interosseous distance between the radius and ulna during rotation of the forearm," *Anat. Rec.* **160**(2), 261–271 (1968).
65. C. Goh et al., "Subcutaneous veins depth measurement using diffuse reflectance images," *Opt. Express* **25**(21), 25741–25759 (2017).
66. M. A. Galarreta-Valverde et al., "Three-dimensional synthetic blood vessel generation using stochastic l-systems," *Proc. SPIE* **8669**, 86691I (2013).

67. R. Y. Rubinstein and D. P. Kroese, *Simulation and the Monte Carlo Method*, John Wiley & Sons, Hoboken, New Jersey (2016).
68. B. E. Treeby et al., "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using AK-space pseudospectral method," *J. Acoust. Soc. Am.* **131**(6), 4324–4336 (2012).
69. B. T. Cox et al., "Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method," *Appl. Opt.* **45**, 1866–1875 (2006).
70. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.* **19**(1), 44–57 (2002).
71. J. W. Tukey, "An introduction to the calculation of numerical spectrum analysis," Spectra Analysis of Time Series, pp. 25–46 (1967).
72. T. Kirchner et al., "Signed real-time delay multiply and sum beamforming for multispectral photoacoustic imaging," *J. Imaging* **4**(10), 121 (2018).
73. B. E. Treeby, E. Z. Zhang, and B. T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," *Inverse Prob.* **26**(11), 115003 (2010).
74. L. Hacker et al., "A copolymer-in-oil tissue-mimicking material with tuneable acoustic and optical characteristics for photoacoustic imaging phantoms," *IEEE Trans. Med. Imaging* **40**(12), 3593–3603 (2021).
75. T. Tarvainen et al., "Utilising the radiative transfer equation in quantitative photoacoustic tomography," *Proc. SPIE* **10064**, 100643E (2017).

Janek Gröhl received his PhD from the University of Heidelberg in April 2021. In 2020, he worked as a postdoctoral researcher at the German Cancer Research Center (DKFZ) and he was working as a research associate at the Cancer Research UK Cambridge Institute in 2021. He was awarded the Walter Benjamin Fellowship by the German Research Foundation (DFG) in 2022. He conducts research on data-driven methods for image processing and signal quantification in photoacoustic imaging.

Kris K. Dreher received his MSc degree in physics from the University of Heidelberg in 2020. He is currently pursuing a PhD at the division of Intelligent Medical Systems (IMSY), DKFZ, and does research in deep learning-based domain adaptation methods to tackle the inverse problems of photoacoustic imaging.

Melanie Schellenberg received her MSc degree in physics from the University of Heidelberg in 2019. She is currently pursuing an interdisciplinary PhD in computer science at the division of IMSY, DKFZ, and aiming for quantitative photoacoustic imaging with a learning-to-simulate approach.

Tom Rix received his MRes degree in medical physics and biomedical engineering from the University College London in 2020. He submitted his MSc thesis in applied computer sciences at Heidelberg University in January 2022, where he worked on photoacoustic image synthesis with deep learning for highly realistic photoacoustic image simulations. He is going to pursue a PhD at the Division of IMSY, DKFZ, in quantitative photoacoustic imaging.

Niklas Holzwarth received his MSc degree in physics from the University of Heidelberg in 2020. He is currently pursuing an interdisciplinary PhD in computer science at the division of IMSY, DKFZ investigating a sensorless 3D photoacoustic approach, referred to as "tattoo tomography."

Patricia Vieten received her BSc degree in physics from Heidelberg University in 2019. She is currently pursuing her MSc degree in physics at the Division of IMSY, DKFZ, and is working on semantic segmentation of multispectral photoacoustic images using deep learning-based methods.

Leonardo Ayala received his MSc degree in physics from Balseiro Institute in 2016, Argentina. He is currently pursuing a PhD at the division of IMSY, DKFZ, and does research in deep learning-based translational biophotonics.

Sarah Bohndiek received her PhD at University College London in 2008 and then worked in both the UK (at Cambridge) and the USA (at Stanford) as a postdoctoral fellow in molecular imaging. Since 2013, she has been a group leader at the University of Cambridge and was appointed as full professor of Biomedical Physics in 2020. She was recently awarded the CRUK Future Leaders in Cancer Research Prize and SPIE Early Career Achievement Award.

Alexander Seitel is a computer scientist currently working as a group lead and deputy head at DKFZ in Heidelberg and holds a doctorate in medical informatics from the University of Heidelberg. His research focusses on computer-assisted interventions and novel imaging methodologies aiming to improve interventional healthcare. In this area, he conducted various international projects at DKFZ and during his two-year postdoctoral fellowship at the University of British Columbia, Vancouver, Canada.

Lena Maier-Hein is a full professor at Heidelberg University (Germany) and division head at the DKFZ. She is managing director of the National Center for Tumor Diseases (NCT) Heidelberg and of the DKFZ Data Science and Digital Oncology cross-topic program. Her research concentrates on machine learning-based biomedical image analysis with a specific focus on surgical data science, computational biophotonics, and validation of machine learning algorithms.

4.3 P.III: Unsupervised domain transfer

Unsupervised domain transfer with conditional invertible neural networks

Authors: **Kris K. Dreher**, Leonardo Ayala, Melanie Schellenberg, Marco Hübner, Jan-Hinrich Nölke, Tim J. Adler, Silvia Seidlitz, Jan Sellner, Alexander Studier-Fischer, Janek Gröhl, Felix Nickel, Ullrich Köthe, Alexander Seitel, Lena Maier-Hein

Status: Published
Journal: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)
DOI: doi.org/10.1007/978-3-031-43907-0_73

Contributions: **KKD**, JG and LMH conceived and conceptualized the project. **KKD** developed the method and wrote the main body of the codebase. **KKD**, LA and LMH defined the experiments to be conducted. **KKD** and LA conducted the experiments and wrote the first version of the manuscript. **KKD** and LMH shared administration of the project.
KKD, MS, JN, and JG contributed to designing the figures and photoacoustic data acquisition. MH contributed to data processing of hyperspectral data. TJA and UK contributed to discussions. SS, JS, ASF, and FN contributed to hyperspectral data acquisition. The exact coding contributions can be seen at IMSY-DKFZ/UDT-cINN/graphs/contributors.
All authors contributed to reviewing and editing the manuscript. LMH acquired the funding and ensured the project's scientific rigor and relevance.

Unsupervised Domain Transfer with Conditional Invertible Neural Networks

Kris K. Dreher[1,2(✉)], Leonardo Ayala[1,3], Melanie Schellenberg[1,4], Marco Hübner[1,4], Jan-Hinrich Nölke[1,4], Tim J. Adler[1], Silvia Seidlitz[1,4,5,6], Jan Sellner[1,4,5], Alexander Studier-Fischer[7], Janek Gröhl[8,9], Felix Nickel[7], Ullrich Köthe[4], Alexander Seitel[1,6], and Lena Maier-Hein[1,3,4,5,6]

[1] Intelligent Medical Systems, German Cancer Research Center (DKFZ), Heidelberg, Germany
{k.dreher,l.maier-hein}@dkfz-heidelberg.de
[2] Faculty of Physics and Astronomy, Heidelberg University, Heidelberg, Germany
[3] Medical Faculty, Heidelberg University, Heidelberg, Germany
[4] Faculty of Mathematics and Computer Science, Heidelberg University, Heidelberg, Germany
[5] Helmholtz Information and Data Science School for Health, Karlsruhe, Heidelberg, Germany
[6] National Center for Tumor Diseases (NCT) Heidelberg a Partnership Between DKFZ and Heidelberg University Hospital, Heidelberg, Germany
[7] Department of General, Visceral, and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany
[8] Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK
[9] Department of Physics, University of Cambridge, Cambridge, UK

Abstract. Synthetic medical image generation has evolved as a key technique for neural network training and validation. A core challenge, however, remains in the domain gap between simulations and real data. While deep learning-based domain transfer using Cycle Generative Adversarial Networks and similar architectures has led to substantial progress in the field, there are use cases in which state-of-the-art approaches still fail to generate training images that produce convincing results on relevant downstream tasks. Here, we address this issue with a domain transfer approach based on conditional invertible neural networks (cINNs). As a particular advantage, our method inherently guarantees cycle consistency through its invertible architecture, and network training can efficiently be conducted with maximum likelihood training. To showcase our method's generic applicability, we apply it to two spectral imaging modalities at different scales, namely hyperspectral imaging (pixel-level) and photoacoustic tomography (image-level). According to

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43907-0_73.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14220, pp. 770–780, 2023.
https://doi.org/10.1007/978-3-031-43907-0_73

comprehensive experiments, our method enables the generation of realistic spectral data and outperforms the state of the art on two downstream classification tasks (binary and multi-class). cINN-based domain transfer could thus evolve as an important method for realistic synthetic data generation in the field of spectral imaging and beyond. The code is available at https://github.com/IMSY-DKFZ/UDT-cINN.

Keywords: Domain transfer · invertible neural networks · medical imaging · photoacoustic tomography · hyperspectral imaging · deep learning

1 Introduction

The success of supervised learning methods in the medical domain led to countless breakthroughs that might be translated into clinical routine and have the potential to revolutionize healthcare [6,13]. For many applications, however, labeled reference data (ground truth) may not be available for training and validating a neural network in a supervised manner. One such application is spectral imaging which comprises various non-interventional, non-ionizing imaging techniques that can resolve functional tissue properties such as blood oxygenation in real time [1,3,23]. While simulations have the potential to overcome the lack of ground truth, synthetic data is not yet sufficiently realistic [9]. Cycle Generative Adversarial Networks (GAN)-based architectures are widely used for domain transfer [12,24] but may suffer from issues such as unstable training, hallucinations, or mode collapse [15]. Furthermore, they have predominantly been used for conventional RGB imaging and one-channel cross-modality domain adaptation, and may not be suitable for other imaging modalities with more channels. We address these challenges with the following contributions:

Domain Transfer Method: We present an entirely new sim-to-real transfer approach based on conditional invertible neural networks (cINNs) (cf. Fig. 1) specifically designed for data with many spectral channels. This approach inherently addresses weaknesses of the state of the art with respect to the preservation of spectral consistency and, importantly, does not require paired images.

Instantiation to Spectral Imaging: We show that our method can generically be applied to two complementary modalities: photoacoustic tomography (PAT; image-level) and hyperspectral imaging (HSI; pixel-level).

Comprehensive Validation: In comprehensive validation studies based on more than 2,000 PAT images (real: ∼1,000) and more than 6 million spectra for HSI (real: ∼6 million) we investigate and subsequently confirm our two main hypotheses: (H1) Our cINN-based models can close the domain gap between simulated and real spectral data better than current state-of-the-art methods

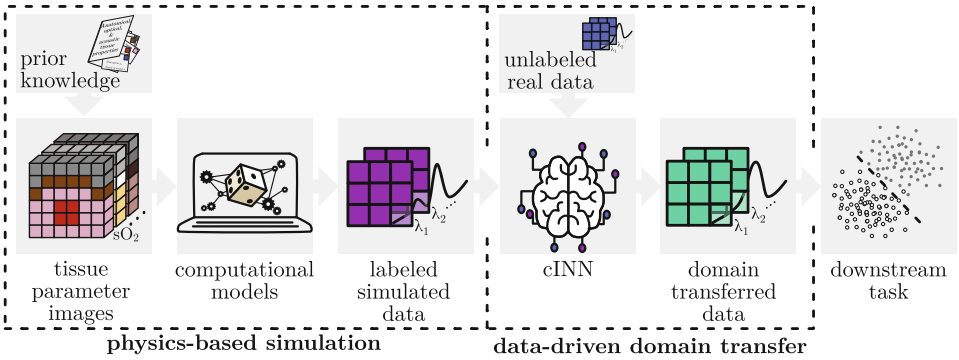


Fig. 1. Pipeline for data-driven spectral image analysis in the absence of labeled reference data. A physics-based simulation framework generates simulated spectral images with corresponding reference labels (e.g. tissue type or oxygenation (sO_2)). Our domain transfer method based on cINNs leverages unlabeled real data to increase their realism. The domain-transferred data can then be used for supervised training of a downstream task (e.g. classification).

regarding spectral plausibility. (H2) Training models on data transferred by our cINN-based approach can improve their performance on the corresponding (clinical) downstream task without them having seen labeled real data.

2 Materials and Methods

2.1 Domain Transfer with Conditional Invertible Neural Networks

Concept Overview. Our domain transfer approach (cf. Fig. 2) is based on the assumption that data samples from both domains carry domain-invariant information (e.g. on optical tissue properties) and domain-variant information (e.g. modality-specific artifacts). The invertible architecture, which inherently guarantees cycle consistency, transfers both simulated and real data into a shared latent space. While the domain-invariant features are captured in the latent space, the domain-variant features can either be filtered (during encoding) or added (during decoding) by utilizing a domain label D. To achieve spectral consistency, we leverage the fact that different tissue types feature characteristic spectral signatures and condition the model on the tissue label Y if available. For unlabeled (real) data, we use randomly generated proxy labels instead. To achieve high visual quality beyond spectral consistency, we include two discriminators Dis_{sim} and Dis_{real} for their respective domains. Finally, as a key theoretical advantage, we avoid mode collapse with maximum likelihood optimization. Implementation details are provided in the following.

cINN Model Design. The core of our architecture is a cINN [2] (cf. Fig. 2), comprising multiple (i) scales of N_i-chained affine conditional coupling (CC)

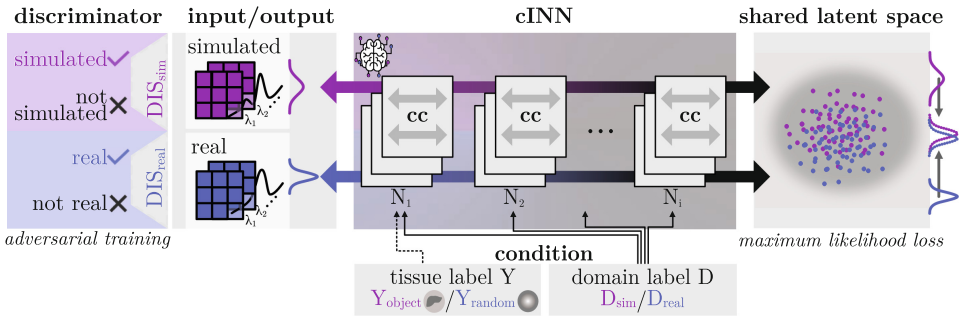


Fig. 2. Proposed architecture based on cINNs. The invertible architecture transfers both simulated and real data into a shared latent space (right). By conditioning on the domain D (bottom), a latent vector can be transferred to either the simulated or the real domain (left) for which the discriminator Dis_{sim} and Dis_{real} calculate the losses for adversarial training.

blocks [7]. These scales are necessary in order to increase the receptive field of the network and are achieved by Haar wavelet downsampling [11]. A CC block consists of subnetworks that can be freely chosen depending on the data dimensionality (e.g. fully connected or convolutional networks) as they are only evaluated in the forward direction. The CC blocks receive a condition consisting of two parts: domain label and tissue label, which are then concatenated to the input along the channel dimension. In the case of PAT, the tissue label is a full semantic and random segmentation map for the simulated and real data, respectively. In the case of HSI, the tissue label is a one-hot encoded vector for organ labels.

Model Training. In the following, the proposed cINN with its parameters θ will be referred to as $f(x, DY, \theta)$ and its inverse as f^{-1} for any input $x \sim p_D$ from domain $D \in \{D_{sim}, D_{real}\}$ with prior density p_D and its corresponding latent space variable z. The condition DY is the combination of domain label D as well as the tissue label $Y \in \{Y_{sim}, Y_{real}\}$. Then the maximum likelihood loss $\mathcal{ML}$ for a training sample x_i is described by

$$\underset{D}{\mathcal{ML}} = \mathbb{E}_i \left[\frac{||f(x_i, DY, \theta)||_2^2}{2} - log|J_i| \right] \text{ with } J_i = det \left(\frac{\partial f}{\partial x}\bigg|_{x_i} \right). \tag{1}$$

For the adversarial training, we employ the least squares training scheme [18] for generator $Gen_D = f_D^{-1} \circ f_{D'}$ and discriminator Dis_D for each domain with $x_{D'}$ as input from the source domain and x_D as input from the target domain:

$$\underset{Gen_D}{\mathcal{L}} = \underset{x_{D'} \sim p_{D'}}{\mathbb{E}} \left[(Dis_D(Gen_D(x_{D'}) - 1))^2 \right] \tag{2}$$

$$\underset{Dis_D}{\mathcal{L}} = \underset{x_D \sim p_D}{\mathbb{E}} \left[(Dis_D(x_D) - 1)^2 \right] + \underset{x_{D'} \sim p_{D'}}{\mathbb{E}} \left[(Dis_D(Gen_D(x_{D'})))^2 \right]. \tag{3}$$

Finally, the full loss for the proposed model comprises the following:

$$\mathcal{L}_{Total_{Gen}} = \mathcal{ML}_{real} + \mathcal{ML}_{sim} + \mathcal{L}_{Gen_{real}} + \mathcal{L}_{Gen_{sim}} \quad \text{and} \quad \mathcal{L}_{Total_{Dis}} = \mathcal{L}_{Dis_{real}} + \mathcal{L}_{Dis_{sim}} . \tag{4}$$

Model Inference. The domain transfer is done in two steps: 1) A simulated image is encoded in the latent space with conditions D_{sim} and Y_{sim} to its latent representation z, 2) z is decoded to the real domain via D_{real} with the simulated tissue label Y_{sim}: $x_{sim \to real} = f^{-1}(\cdot, D_{real}Y_{sim}, \theta) \circ f(\cdot, D_{sim}Y_{sim}, \theta)(x_{sim})$.

2.2 Spectral Imaging Data

Photoacoustic Tomography Data. PAT is a non-ionizing imaging modality that enables the imaging of functional tissue properties such as tissue oxygenation [22]. The **real PAT data** (cf. Fig. 3) used in this work are images of human forearms that were recorded from 30 healthy volunteers using the MSOT Acuity Echo (iThera Medical GmbH, Munich, Germany) (all regulations followed under study ID: S-451/2020, and the study is registered with the German Clinical Trials Register under reference number DRKS00023205). In this study, 16 wavelengths from 700 nm to 850 nm in steps of 10 nm were recorded for each image. The resulting 180 images were semantically segmented into the structures shown in Fig. 3 according to the annotation protocol provided in [20]. Additionally, a full sweep of each forearm was performed to generate more unlabeled images, thus

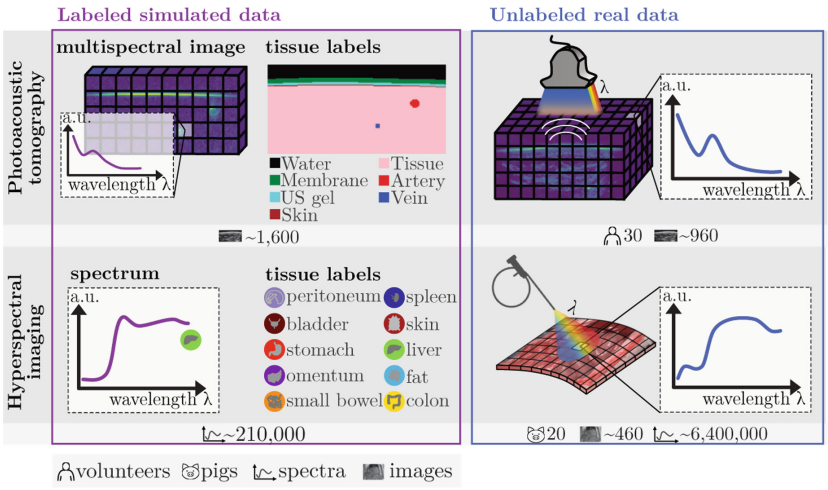


Fig. 3. Training data used for the validation experiments. For PAT, 960 real images from 30 volunteers were acquired. For HSI, more than six million spectra corresponding to 460 images and 20 individuals were used. The tissue labels PAT correspond to 2D semantic segmentations, whereas the tissue labels for HSI represent 10 different organs. For PAT, ∼1600 images were simulated, whereas around 210,000 spectra were simulated for HSI.

amounting to a total of 955 real images. The **simulated PAT data** (cf. Fig. 3) used in this work comprises 1,572 simulated images of human forearms. They were generated with the toolkit for Simulation and Image Processing for Photonics and Acoustics (SIMPA) [8] based on a forearm literature model [21] and with a digital device twin of the MSOT Acuity Echo.

Hyperspectral Imaging Data. HSI is an emerging modality with high potential for surgery [4]. In this work, we performed pixel-wise analysis of HSI images. The **real HSI data** was acquired with the Tivita® Tissue (Diaspective Vision GmbH, Am Salzhaff, Germany) camera, featuring a spectral resolution of approximately 5 nm in the spectral range between 500 nm and 1000 nm nm. In total, 458 images, corresponding to 20 different pigs, were acquired (all regulations followed under study IDs: 35-9185.81/G-161/18 and 35-9185.81/G-262/19) and annotated with ten structures: bladder, colon, fat, liver, omentum, peritoneum, skin, small bowel, spleen, and stomach (cf. Fig. 3). This amounts to 6,410,983 real spectra in total. The **simulated HSI data** was generated with a Monte Carlo method (cf. algorithm provided in the supplementary material). This procedure resulted in 213,541 simulated spectra with annotated organ labels.

3 Experiments and Results

The purpose of the experiments was to investigate hypotheses H1 and H2 (cf. Sect. 1). As comparison methods, a CycleGAN [24] and an unsupervised image-to-image translation (UNIT) network [16] were implemented fully convolutionally for PAT and in an adapted version for the one-dimensional HSI data. To make the comparison fair, the tissue label conditions were concatenated with the input, and we put significant effort into optimizing the UNIT on our data.

Realism of Synthetic Data (H1) : According to qualitative analyses (Fig. 4) our domain transfer approach improves simulated PAT images with respect to key properties, including the realism of skin, background, and sharpness of vessels.

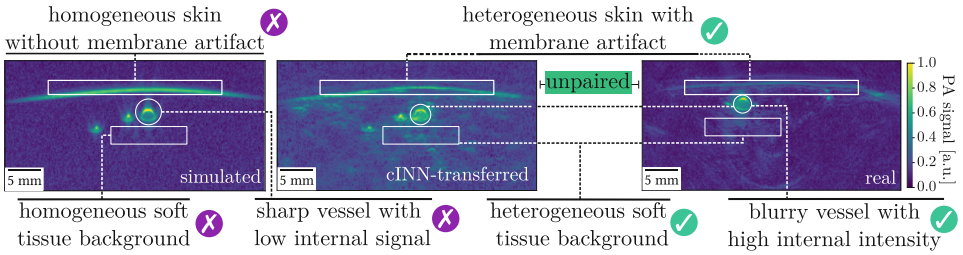


Fig. 4. Qualitative results. In comparison to simulated PAT images (left), images generated by the cINN (middle) resemble real PAT images (right) more closely. All images show a human forearm at 800 nm.

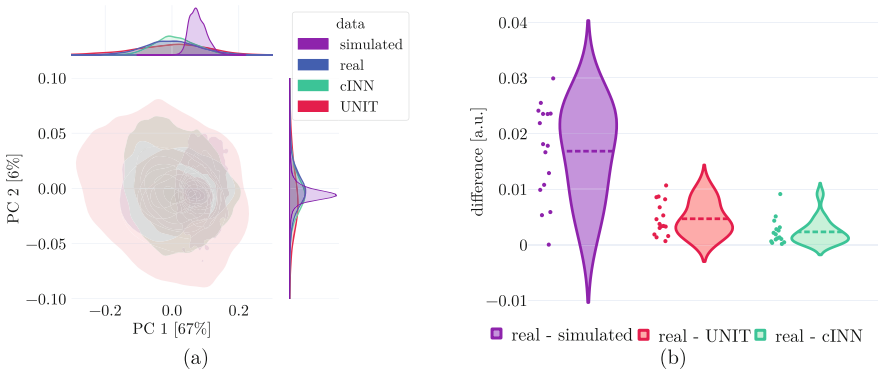


Fig. 5. Our domain transfer approach yields realistic spectra (here: of veins). The PCA plots in a) represent a kernel density estimation of the first and second components of a PCA embedding of the real data, which represent about 67% and 6% of the variance in the real data, respectively. The distributions on top and on the right of the PCA plot correspond to the marginal distributions of each dataset's first two components. b) Violin plots show that the cINN yields spectra that feature a smaller difference to the real data compared to the simulations and the UNIT-generated data. The dashed lines represent the mean difference value, and each dot represents the difference for one wavelength.

A principal component analysis (PCA) performed on all artery and vein spectra of the real and synthetic datasets demonstrates that the distribution of the synthetic data is much closer to the real data after applying our domain transfer approach (cf. Fig. 5a)). The same holds for the absolute difference, as shown in Fig. 5b). Slightly better performance was achieved with the cINN compared to the UNIT. Similarly, our approach improves the realism of HSI spectra, as illustrated in Fig. 6, for spectra of five exemplary organs (colon, stomach, omentum, spleen, and fat). The cINN-transferred spectra generally match the real data very closely. Failure cases where the real data has a high variance (translucent band) are also shown.

Benefit of Domain-Transferred Data for Downstream Tasks (H2): We examined two classification tasks for which reference data generation was feasible: classification of veins/arteries in PAT and organ classification in HSI. For both modalities, we used the completely untouched real test sets, comprising 162 images in the case of PAT and $\sim$ 920,000 spectra in the case of HSI. For both tasks, a calibrated random forest classifier (sklearn [19] with default parameters) was trained on the simulated, the domain-transferred (by UNIT and cINN), and real spectra. As metrics, the balanced accuracy (BA), area under receiver operating characteristic (AUROC) curve, and F1-score were selected based on [17].

As shown in Table 1, our domain transfer approach dramatically increases the classification performance for both downstream tasks. Compared to physics-based simulation, the cINN obtained a relative improvement of 37% (BA), 25% (AUROC), and 22% (F1 Score) for PAT whereas the UNIT only achieved a

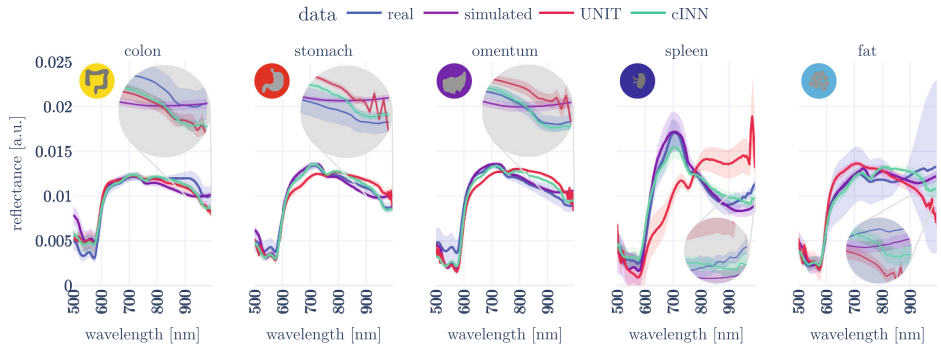


Fig. 6. The cINN-transferred spectra are in closer agreement with the real spectra than the simulations and the UNIT-transferred spectra. Spectra for five exemplary organs are shown from 500 nm to 1000 nm. For each subplot, a zoom-in for the near-infrared region (>900 nm) is shown. The translucent bands represent the standard deviation across spectra for each organ.

Table 1. Classification scores for different training data. The training data refers to real data, physics-based simulated data, data generated by a CycleGAN, by a UNIT without and with tissue labels (UNIT_Y), and by a cINN without (cINN_D) and with (proposed cINN_DY) tissue labels as condition. Additionally, $\text{cINN}_{\text{DY without GAN}}$ refers to a cINN_DY without the adversarial training. The best-performing methods, except if trained on real data, are printed in **bold**.

Classifier training data	PAT			HSI		
	BA	AUROC	F1-Score	BA	AUROC	F1-Score
Real	0.75	0.84	0.82	0.40	0.81	0.44
Simulated	0.52	0.64	0.64	0.24	0.75	0.18
CycleGAN	0.39	0.20	0.16	0.11	0.57	0.06
UNIT	0.50	0.44	0.65	0.20	0.72	0.20
UNIT_Y	0.64	**0.81**	0.77	0.24	0.74	0.25
cINN_D	0.66	0.73	0.72	0.25	0.72	0.20
$\text{cINN}_{\text{DY without GAN}}$	0.65	0.78	0.76	0.28	0.75	**0.26**
cINN_DY (proposed)	**0.71**	0.80	**0.78**	**0.29**	**0.76**	0.24

relative improvement in the range of 20%-27% (depending on the metric). For HSI, the cINN achieved a relative improvement of 21% (BA), 1% (AUROC), and 33% (F1 Score) and it scored better in all metrics except for the F1 Score than the UNIT. For all metrics, training on real data still yields better results.

4 Discussion

With this paper, we presented the first domain transfer approach that combines the benefits of cINNs (exact maximum likelihood estimation) with those

of GANs (high image quality). A comprehensive validation involving qualitative and quantitative measures for the remaining domain gap and downstream tasks suggests that the approach is well-suited for sim-to-real transfer in spectral imaging. For both PAT and HSI, the domain gap between simulations and real data could be substantially reduced, and a dramatic increase in downstream task performance was obtained - also when compared to the popular UNIT approach.

The only similar work on domain transfer in PAT has used a cycle GAN-based architecture on a single wavelength with only photon propagation as PAT image simulator instead of full acoustic wave simulation and image reconstruction [14]. This potentially leads to spectral inconsistency in the sense that the spectral information either is lost during translation or remains unchanged from the source domain instead of adapting to the target domain. Outside the spectral/medical imaging community, Liu et al. [16] and Grover et al. [10] tasked variational autoencoders and invertible neural networks for each domain, respectively, to create the shared encoding. They both combined this approach with adversarial training to achieve high-quality image generation. Das et al. [5] built upon this approach by using labels from the source domain to condition the domain transfer task. In contrast to previous work, which used en-/decoders for each domain, we train a single network as shown in Fig. 2. with a two-fold condition consisting of a domain label (D) and a tissue label (Y) from the source domain, which has the advantage of explicitly aiding the spectral domain transfer.

The main limitation of our approach is the high dimensionality of the parameter space of the cINN as dimensionality reduction of data is not possible due to the information and volume-preserving property of INNs. This implies that the method is not suitable for arbitrarily high dimensions. Future work will comprise the rigorous validation of our method with tissue-mimicking phantoms for which reference data are available.

In conclusion, our proposed approach of cINN-based domain transfer enables the generation of realistic spectral data. As it is not limited to spectral data, it could develop into a powerful method for domain transfer in the absence of labeled real data for a wide range of image modalities in the medical domain and beyond.

Acknowledgements. This project was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (NEURAL SPICING, 101002198) and the Surgical Oncology Program of the National Center for Tumor Diseases (NCT) Heidelberg.

References

1. Adler, T.J., et al.: Uncertainty-aware performance assessment of optical imaging modalities with invertible neural networks. Int. J. Comput. Assist. Radiol. Surg. **14**(6), 997–1007 (2019). https://doi.org/10.1007/s11548-019-01939-9
2. Ardizzone, L., Lüth, C., Kruse, J., Rother, C., Köthe, U.: Conditional invertible neural networks for guided image generation (2020)

3. Ayala, L., et al.: Spectral imaging enables contrast agent-free real-time ischemia monitoring in laparoscopic surgery. Sci. Adv. (2023). https://doi.org/10.1126/sciadv.add6778
4. Clancy, N.T., Jones, G., Maier-Hein, L., Elson, D.S., Stoyanov, D.: Surgical spectral imaging. Med. Image Anal. **63**, 101699 (2020)
5. Das, H.P., Tran, R., Singh, J., Lin, Y.W., Spanos, C.J.: Cdcgen: cross-domain conditional generation via normalizing flows and adversarial training. arXiv preprint arXiv:2108.11368 (2021)
6. De Fauw, J., Ledsam, J.R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., et al.: Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat. Med. **24**(9), 1342–1350 (2018)
7. Dinh, L., Sohl-Dickstein, J., Bengio, S.: Density estimation using real nvp. arXiv preprint arXiv:1605.08803 (2016)
8. Gröhl, J., et al.: Simpa: an open-source toolkit for simulation and image processing for photonics and acoustics. J. Biomed. Opt. **27**(8), 083010 (2022)
9. Gröhl, J., Schellenberg, M., Dreher, K., Maier-Hein, L.: Deep learning for biomedical photoacoustic imaging: a review. Photoacoustics **22**, 100241 (2021)
10. Grover, A., Chute, C., Shu, R., Cao, Z., Ermon, S.: Alignflow: cycle consistent learning from multiple domains via normalizing flows. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 4028–4035 (2020)
11. Haar, A.: Zur theorie der orthogonalen funktionensysteme. Mathematische Annalen **71**(1), 38–53 (1911)
12. Hoffman, J., Tzet al.: Cycada: cycle-consistent adversarial domain adaptation. In: International Conference on Machine Learning, pp. 1989–1998 (2018)
13. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net a self-configuring method for deep learning-based biomedical image segmentation. Nat. Methods **18**(2), 203–211 (2021)
14. Li, J., et al.: Deep learning-based quantitative optoacoustic tomography of deep tissues in the absence of labeled experimental data. Optica **9**(1), 32–41 (2022)
15. Li, K., Zhang, Y., Li, K., Fu, Y.: Adversarial feature hallucination networks for few-shot learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13470–13479 (2020)
16. Liu, M.Y., Breuel, T., Kautz, J.: Unsupervised image-to-image translation networks. Adv. Neural Inf. Process. Syst. **30** (2017)
17. Maier-Hein, L., Reinke, A., Godau, P., Tizabi, M.D., Büttner, F., Christodoulou, E., et al.: Metrics reloaded: pitfalls and recommendations for image analysis validation (2022). https://doi.org/10.48550/ARXIV.2206.01653
18. Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., Paul Smolley, S.: Least squares generative adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2794–2802 (2017)
19. Pedregosa, F., et al.: Scikit-learn: machine learning in python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
20. Schellenberg, M., et al.: Semantic segmentation of multispectral photoacoustic images using deep learning. Photoacoustics **26**, 100341 (2022). https://doi.org/10.1016/j.pacs.2022.100341
21. Schellenberg, M., et al.: Photoacoustic image synthesis with generative adversarial networks. Photoacoustics **28**, 100402 (2022)
22. Wang, X., Xie, X., Ku, G., Wang, L.V., Stoica, G.: Noninvasive imaging of hemoglobin concentration and oxygenation in the rat brain using high-resolution photoacoustic tomography. J. Biomed. Opt. **11**(2), 024015 (2006)

23. Wirkert, S.J., et al.: Physiological parameter estimation from multispectral images unleashed. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) MICCAI 2017. LNCS, vol. 10435, pp. 134–141. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66179-7_16
24. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2223–2232 (2017)

4.4 P.IV: Anthropomorphic phantoms

Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging

Authors: **Kris K. Dreher**, Janek Gröhl, Friso Grace, Leonardo Ayala, Jan-Hinrich Nölke, Christoph J. Bender, Melissa J. Watt, Catherine-Louise White, Ran Tao, Wibke Johnen, Minu D. Tizabi, Alexander Seitel, Lena Maier-Hein, Sarah E. Bohndiek

Status: Published
Journal: Journal of Biomedical Optics
DOI: https://doi.org/10.1117/1.JBO.30.7.076006

Contributions: **KKD** conceived and conceptualized the project. **KKD**, JG, LMH, and SEB developed the methodology and defined the experiments to be conducted. **KKD** and JG developed the phantom fabrication protocol. **KKD** conducted the experiments relevant for optical characterization, fabricated the phantoms (either by himself or with support from JG, MJW, CW, and RT) performed the data analysis and wrote the first version of the manuscript. **KKD**, AS, LMH, and SEB shared administration of the project.
FG, LA, JN, and CJB contributed to data acquisition. JG, MJW, CLW, and RT contributed to phantom fabrication. WJ performed 3D model design and 3D-printing of the phantom mold. MDT did the anatomical check of the forearm segmentations. The exact coding contributions can be seen at IMSY-DKFZ/anthropomorphic-phantoms/graphs/contributors. **All authors** contributed to reviewing and editing the manuscript. LMH and SEB acquired the funding and ensured the project's scientific rigor and relevance.

JBO Journal of Biomedical Optics

RESEARCH PAPER

Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging

Kris K. Dreher,[a,b,*] Janek Gröhl,[c,d] Friso Grace,[a,e] Leonardo Ayala,[a,f] Jan-Hinrich Nölke,[a,g] Christoph J. Bender,[a,g] Melissa J. Watt,[c,d] Katie-Lou White,[c,d] Ran Tao,[c,d] Wibke Johnen,[h,i] Minu D. Tizabi,[a] Alexander Seitel,[a] Lena Maier-Hein,[a,f,g,j,*,†] and Sarah E. Bohndiek[c,d,*,†]

[a]German Cancer Research Center (DKFZ), Division of Intelligent Medical Systems (IMSY), Heidelberg, Germany
[b]Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany
[c]University of Cambridge, CRUK Cambridge Institute, Cambridge, United Kingdom
[d]University of Cambridge, Department of Physics, Cambridge, United Kingdom
[e]University of St Andrews, School of Physics and Astronomy, St Andrews, United Kingdom
[f]NCT Heidelberg, a partnership between DKFZ and University Hospital Heidelberg, National Center for Tumor Diseases (NCT), Heidelberg, Germany
[g]Heidelberg University, Faculty of Mathematics and Computer Science, Heidelberg, Germany
[h]DKFZ, Division of Medical Physics in Radiation Oncology, Heidelberg, Germany
[i]Heidelberg Institute for Radiation Oncology (HIRO), National Center for Radiation Research in Oncology (NCRO), Heidelberg, Germany
[j]Heidelberg University, Medical Faculty, Heidelberg, Germany

ABSTRACT. **Significance:** Optical imaging of blood oxygenation (sO_2) can be achieved based on the differential absorption spectra of oxy- and deoxyhemoglobin. A key challenge in realizing clinical validation of the sO_2 biomarkers is the absence of reliable sO_2 reference standards, including test objects.

Aim: To enable quantitative testing of multispectral imaging methods for assessment of sO_2 by introducing anthropomorphic phantoms with appropriate tissue-mimicking optical properties.

Approach: We used the stable copolymer-in-oil base material to create physical anthropomorphic structures and optimized dyes to mimic the optical absorption of blood across a wide spectral range. Using 3D-printed phantom molds generated from a magnetic resonance image of a human forearm, we molded the material into an anthropomorphic shape. Using both reflectance hyperspectral imaging (HSI) and photoacoustic tomography (PAT), we acquired images of the forearm phantoms and evaluated the performance of linear spectral unmixing (LSU).

Results: Based on 10 fabricated forearm phantoms with vessel-like structures featuring five distinct sO_2 levels (between 0 and 100%), we showed that the measured absorption spectra of the material correlated well with HSI and PAT data with a Pearson correlation coefficient consistently above 0.8. Further, the application of LSU enabled a quantification of the mean absolute error in sO_2 assessment with HSI and PAT.

Conclusions: Our anthropomorphic tissue-mimicking phantoms hold potential to provide a robust tool for developing, standardising, and validating optical imaging of sO_2.

*Address all correspondence to Kris K. Dreher, k.dreher@dkfz-heidelberg.de; Lena Maier-Hein, l.maier-hein@dkfz-heidelberg.de; Sarah E. Bohndiek, seb53@cam.ac.uk

†Shared last authorship

Keywords: anthropomorphic phantoms; optical imaging; oximetry; photoacoustic imaging; hyperspectral imaging

Paper 250083GR received Mar. 19, 2025; revised Jun. 17, 2025; accepted Jun. 19, 2025; published Jul. 17, 2025.

1 Introduction

Oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb) are critical endogenous contrast agents that enable noninvasive measurement of blood oxygen saturation (sO_2), also referred to as oximetry, due to their distinct optical absorption spectra. Oximetry methods are invaluable for a range of clinical applications throughout the patient care pathway, from diagnosis to treatment planning and monitoring of treatment response.[1–5] Pulse oximetry is the most widely available of these tools, which employs red and infrared light to noninvasively estimate sO_2 for bulk tissue at a single measurement site, such as a fingertip.[1]

To obtain spatially resolved information, for example, in image-guided surgery, imaging modalities such as hyperspectral imaging (HSI) and photoacoustic tomography (PAT) are used. HSI and PAT typically involve making wavelength-resolved measurements across the visible and near-infrared spectral range.[1] HSI uses light diffusely reflected from the tissue to map HbO_2 and Hb signals near the tissue surface,[6,7] whereas PAT combines pulsed laser illumination and ultrasonic detection to probe deeper tissue layers.[8,9] In both techniques, deriving sO_2 from spectral data commonly relies on linear spectral unmixing (LSU),[10,11] assuming that the optical absorption responsible for the image contrast is a linear combination of the absorption spectra of all contrast agents present at any given point weighted by their concentration.

Accurately quantifying sO_2 in HSI and PAT is challenging because a range of assumptions are made in both the datacube (x, y, λ) reconstruction and spectral analysis pipeline that can lead to corruption of the measured tissue biomarkers. For example, in HSI, signals can be distorted due to additional optical interactions, such as fluorescence, whereas in PAT, depth-dependent signal attenuation arises, known as spectral coloring. In both modalities, patient motion and skin tone bias can introduce further complexity.[10,12] Consequently, developing robust oximetry calibration methods for HSI and PAT remains an active area of research.[13–16]

For the development and rigorous validation of any scientific method, including oximetry, a reliable performance measure or reference is essential. In oximetry, however, the principal challenge is that an *in vivo* ground truth for sO_2 is not available noninvasively with current technology. Therefore, many studies only rely on qualitative visualizations or measurement of relative changes in the same individual or specimen over time, rather than calibrating for absolute sO_2.[17–19] Validation approaches can take several forms. Validation can be developed by comparison to simulated data; however, these often fail to generalize when applied to *in vivo* tissue.[20] Adding a level of complexity, experimental data obtained from bulk test objects ("phantoms") made with a mixture of blood, hemoglobin, or other biological solutions[21,22] can closely mimic tissue spectra and facilitate validation, but typically do not allow well-controlled adjustment of sO_2, lack internal structure, tend to be unstable over time, and are prone to bacterial contamination.[23,24] Addressing some of these limitations, blood flow phantoms allow for chemical manipulation of sO_2 as blood flows through a circuit while being imaged.[25,26] Nevertheless, blood flow circuits rely on ancillary reference measurements (e.g., partial pressure probes or offline oximeters) to provide a gold standard reference. They are typically simple in their structure, e.g., a tube flowing through a slab of base material and thus lack anatomical variability, which is not only beneficial to train deep learning methods but also allows for a detailed analysis of oximetry methods in depth and their performance dependent on different amounts and positions of absorbing structures.

Structurally stable, anthropomorphic phantoms with tissue-mimicking optical properties over a broad wavelength range remain rare.[27] Obtaining oximetry estimations from anthropomorphic phantoms typically requires embedding a flow circuit within a complex material composition, which is prohibitively challenging. To overcome this limitation and tackle the challenge

of oximetry validation in HSI and PAT, we introduce an anthropomorphic phantom design that mimics the morphology of a human forearm and includes absorbing dyes that replicate the absorption coefficient (μ_a) characteristics of HbO_2 and Hb between 700 nm and 850 nm. We first selected and optimized dye proxies for Hb and HbO_2. We then created 3D-printable molds derived from human forearm imaging to achieve morphologically relevant geometry, which are available open-source. Using a copolymer-in-oil base material, we created structurally stable anthropomorphic forearm phantoms with optical characteristics confirmed via simulation studies and signal correlation analyses. Finally, we demonstrated the value of the created phantoms for oximetry validation in HSI and PAT. These phantoms provide a versatile, morphologically realistic test platform for oximetry validation in the near-infrared. We see another major contribution of this work that in addition to the specific structures and absorptions we targeted, this paper presents a general workflow from external data of human anatomy and a specific target tissue, in our case, magnetic resonance imaging (MRI) data and blood vessels, respectively, to structurally stable tissue-mimicking phantoms.

2 Materials and Methods

Combining dye proxies with anthropomorphic phantom designs in this study enables a new approach to oximetry validation (Fig. 1).

2.1 Forearm Phantom Fabrication

2.1.1 *Phantom material*

The base phantom material was prepared according to protocols outlined by Hacker et al.[28] and Gröhl et al.[29] Briefly, for each ~80 mL batch of base material, 76.5 mg of titanium dioxide (TiO_2, Sigma Aldrich 232033-100 g) was sonicated in a water bath together with 50 mL of mineral oil (Sigma Aldrich 330779-1L) until completely dispersed. Next, 12.57 g of polystyrene-block-poly(ethylene-ran-butylene)-block-polystyrene (SEBS, Sigma Aldrich 200557-250G) and 1 g of butylated hydroxytoluene (HT, Sigma Aldrich W218405-1KG-K) were added to the oil. The mixture was heated in a silicone oil bath at 160°C for about 45 min, stirred every 10 min, and allowed to liquefy fully. Finally, the beaker was placed in a vacuum chamber to remove any residual air bubbles. A more detailed manufacturing protocol is provided in the Supplementary Notes S1 in the Supplementary Material.

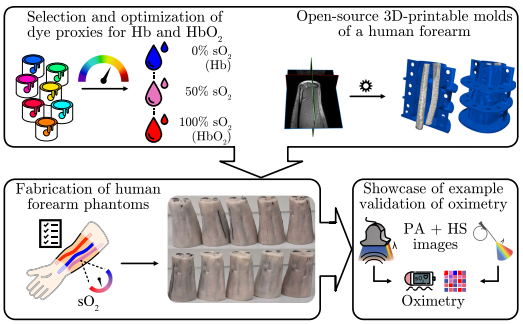


Fig. 1 Workflow and key contributions of this study. (Top) First, 26 dyes were investigated to mimic the absorption spectra of oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb) in the wavelength (λ) range of 700 to 850 nm. After a selection and optimization process, two proxy dyes that could be mixed to five levels of oxygen saturation (sO_2) (0%, 30%, 50%, 70%, and 100%) were found. Second, for realistic tissue morphology, a 3D-printable mold was created based on an open-source magnetic resonance (MR) image of a human forearm. (bottom) Third, ten forearm phantoms were fabricated and finally, photoacoustic (PA) and hyperspectral (HS) images were acquired and we show that example images of these can be used to validate oximetry methods.

2.1.2 *Optical property characterization*

The optical properties of the phantom materials were quantified using a double-integrating sphere (DIS) system (according to the method of Pickering et al.[30] with the system described in Hacker et al.[28]). The DIS system measured total reflectance and transmittance over a wavelength range of 700 to 850 nm, and the resulting data were processed with the inverse adding-doubling (IAD)[31] algorithm to obtain μ_a and scattering coefficient (μ_s).[32] The refractive index was set to $n = 1.4$, and the anisotropy factor to $g = 0.7$, as suggested by Jones and Munro.[33] For each material batch, two optical sample slabs were fabricated. The thickness of each slab was measured five times at three different locations (top, middle, and bottom) using a digital calliper and provided as input to the IAD algorithm. Subsequently, each location was measured from both the front and back with the DIS system, producing 12 individual data points per wavelength.

2.1.3 *Proxy dyes for oxy- and deoxyhemoglobin*

After starting with dyes that were already used in combination with gel-wax phantoms,[34] a total of 26 candidate dyes were evaluated for their optical properties in the abovementioned base material between 700 and 850 nm (Figs. S1–S3 in the Supplementary Material, Table S1 in the Supplementary Material). The μ_a spectra measured with the DIS system served as inputs to a non-negative least square (NNLS) optimization, which aimed to identify a linear combination of dyes whose summed absorption profiles closely match those of HbO_2 and Hb. As target spectra, we used the data available at Ref. 35. Based on these results, the most promising dyes were selected and further refined by iteratively adjusting concentrations and re-measuring μ_a.

2.1.4 *3D forearm model*

A high-resolution MRI dataset of the human forearm was obtained from Kerkhof et al.[36] From these data, muscle and bone structures were segmented, and the segmentation was verified by a physician (MDT). The outer hull and bones were then 3D-printed to serve as mold and embedded features, respectively (Fig. 2). The digital model for the mold was designed and constructed using Autodesk Inventor 2024 and Geomagic Freeform. An Objet 500 Connex (Stratasys, Ltd., Eden Prairie, Minnesota, USA) was used for 3D printing. Specifically, VeroCyan™ was used to

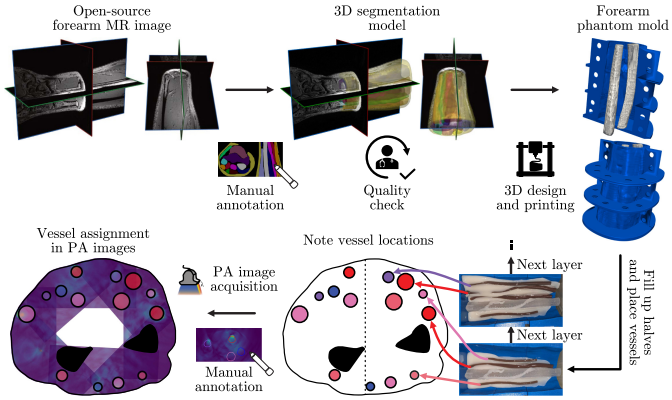


Fig. 2 3D-printable models were constructed from an open-source magnetic resonance (MR) image for a structured phantom fabrication and annotation process. From top left to top right: The MR image was manually segmented for bones, muscles, and fat. A physician performed a quality check ensuring morphological correctness. Based on this segmentation, a 3D-printable model of the outer hull of the forearm has been designed and printed including the two bones, radius and ulna, that will stay in the mold as positives. From bottom right to bottom left: Each half of the mold is filled successively with background material and vessels are placed, which are annotated in a cross-sectional sketch of the mold. Finally, after acquiring photoacoustic (PA) images from multiple angles, each vessel in the PA images can be assigned to its corresponding oxygen saturation during manual annotation.

fabricate the external mold, whereas VeroClear™ was chosen to print the bones, which remained inside each phantom as positive structures. Visual inspections and photoacoustic measurements of VeroClear™ indicate negligible μ_a and μ_s; thus, we did not opt for a challenging and expensive full optical characterization.

2.1.5 *Phantom fabrication process*

Ten forearm phantoms were fabricated, each containing background material and 14 embedded vessels. The vessels were created by drawing phantom material into tubes with diameters of 3, 4, 5, or 6 mm, producing 140 vessels in total (28 per sO_2 level, 7 per diameter). The minimum vessel diameter we could fabricate was 3 mm, which coincides with the upper range of the radial artery, the largest superficial vessel typically visible in the forearm. Therefore, our vessel diameters are larger than anatomically expected. At least one vessel for each sO_2 level was positioned near the surface ("superficial vessel") in each phantom. Dye mixtures within the vessels were determined based on the outcome of the NNLS optimization. Particular attention was paid to achieving an isosbestic point near 800 nm. For mixing intermediate sO_2 levels, the base dyes were accurately weighed, and the resulting ratios were verified by measuring μ_a spectra and applying LSU. This two-step approach ensured reliable verification of the intended mixing ratios.

Nine of the phantoms were manufactured with background materials spanning three sO_2 levels (0%, 50%, and 100%) and their volumes as fractions of the respective material batch (1%, 2.5%, and 4%), yielding 3×3 combinations. An additional out-of-distribution (OOD) phantom was created with three distinct combinations of volume fraction and sO_2 (0.5%, 100%), (5%, 0%), and (3%, 70%) (Table S2 in the Supplementary Material). Optical properties were verified for each background material (Figs. S4–S7 in the Supplementary Material). The fabrication proceeded by splitting the 3D-printed mold into two halves and pouring two 100-mL batches of the phantom mixture in layers. Vessels were placed incrementally within each half after each layer, and their approximate positions were documented. To ensure that the overall coverage of vessel locations matches with human forearms, we compared them by plotting the vessel locations in both humans from an in-house dataset[37] and phantoms in Fig. S8 in the Supplementary Material. Once both halves were filled and all 14 vessels were in place, the two parts of the mold were joined, and any remaining cavity in the center was filled with the residual phantom material.

2.2 Optical Imaging Techniques

2.2.1 *Photoacoustic tomography*

All phantoms were scanned using the MSOT Acuity Echo system (iThera Medical GmbH, Munich, Germany) in a water bath using the wavelengths from 700 to 850 nm in steps of 10 nm. Each of the ten phantoms was imaged at three predefined locations, and for each location, eight angular views were acquired in 45-deg increments around the phantom. This arrangement resulted in 24 images per phantom, for a total of 240 images overall. The phantoms were mounted on a rotational stage within the water bath to facilitate consistent data acquisition across all angles and locations.

The acquired time-series data were corrected for laser energy (Fig. S9 in the Supplementary Material) and filtered using a bandpass with cutoffs at 50 kHz and 20 MHz. The 700-nm wavelength was excluded from subsequent analysis due to laser instability. Reconstructions were performed with a delay-and-sum algorithm implemented in the open-source toolkit for simulation and image processing for photonics and acoustics (SIMPA),[38] specifying a speed of sound of 1497.4 ms^{-1} and a voxel resolution of 0.1 mm. After reconstruction, a Hilbert transform was applied for envelope detection. The speed of sound was chosen to enable coregistration with concurrently acquired ultrasound images. The photoacoustic (PA) images shown in this work are the unnormalized results of this reconstruction algorithm.

2.2.2 *Hyperspectral imaging*

All phantoms were imaged using the Tivita 2.0 camera (Diaspective Vision GmbH, Am Salzhaff, Germany). The wavelengths from 500 to 1000 nm in steps of 5 nm were imaged, but for consistency with PAT, we used the wavelengths from 700 to 850 nm in steps of 10 nm for analysis in

this work. Each phantom was placed on a rotational stage and scanned at eight angular positions in 45-deg increments. Before each capture, the camera was refocused to ensure clear images. This procedure yielded eight hyperspectral (HS) images per phantom, for a total of 80 images across all phantoms. The acquired HS images were automatically corrected for white and dark references using the open-source htc[39] software.

2.3 Quality Assurance

To ensure that our phantoms faithfully replicate the intended optical and acoustic properties, we performed two key evaluations: simulation studies to investigate the effects of air bubbles and speed of sound variations on PA images and signal correlation analyses to confirm the relationship between measured signals and the known absorption spectra at varying sO_2 levels.

2.3.1 *Simulation studies*

Ultrasound segmentations including observed air bubbles were used to simulate phantom images, comparing vessel spectra with and without air inclusions. In addition, a sensitivity analysis on speed of sound variations (± 50 and ± 100 ms^{-1} around the assumed speed of sound 1470 ms^{-1}) was performed. Because the main focus of this work is on the optical properties, refer to Sec. S3 in the Supplementary Material for detailed experiment descriptions and results in Figs. S16–S25. Briefly, all simulations were conducted with SIMPA,[38] employing MCX[40] for photon transport and k-wave[41] for acoustic wave propagation. Each simulation used a digital device twin of the MSOT Acuity Echo, and a digital tissue twin constructed from manual segmentations of five representative phantom PA images. The vessels and background regions in these digital twins were assigned the absorption and scattering coefficients obtained via the DIS system.

2.3.2 *Signal correlation*

PA signals (S) are proportional to the product of the Grüneisen parameter (Γ), μ_a, and the local fluence (Φ): $S \propto \Gamma \mu_a \Phi$.[8] Because the Grüneisen parameter of the mixed dyes could be different such that the PAT signal does not correlate linearly with the μ_a,[42] we investigated whether the measured μ_a for different sO_2 levels correlates linearly with the PA signal. The processing steps included

- Vessel segmentation: Superficial vessels were segmented as regions of interest (ROIs), and the top 5% of brightest pixels in each vessel region were averaged.
- Spectrum fitting: A linear regression (one multiplicative factor plus one offset) was applied to match the expected absorption spectrum from 710 to 850 nm.
- Correlation analysis: The Pearson correlation coefficient (R value) between the measured PA spectrum and the known absorption spectrum was determined for each vessel.

The HS signal (I) is formed from the diffusely reflected fraction of light that is neither absorbed nor scattered out of the detection path. A common method to approximate μ_a from reflectance data uses the Lambert-Beer law, where $\mu_a \propto -\log(I)$.[43] We therefore checked whether the measured μ_a for different sO_2 levels correlated with the HS signal. The processing steps include

- Region selection: ROIs were chosen within superficial vessels, avoiding specular highlights and vessel edges to minimize cross-talk from the surrounding tissue.
- Lambert-Beer approximation: An approximate absorption spectrum was obtained by applying $\hat{\mu}_a = -\log(I)$.
- Spectrum fitting: A linear regression (one multiplicative factor plus one offset) was applied to match the expected absorption spectrum from 700 to 850 nm.
- Correlation analysis: The R value was computed to assess the agreement between the derived absorption spectrum and the known phantom absorption.

2.4 Oximetry Method Validation

To quantify the accuracy of our phantom-based oximetry measurements, the derived spectra (both with and without signal correlation) were used as inputs to an LSU algorithm.[10,11] The

performance of LSU was assessed by calculating the mean absolute error (MAE) in estimated sO_2 levels. In the PAT experiments, LSU performance was also evaluated as a function of depth. In addition, for PAT data, a fluence compensation step was introduced before LSU in simulation studies, where the reconstructed image was divided by the estimated fluence to account for optical attenuation effects in the PA signal. Finally, errors were computed at the most granular level—individual tissue-type instances (e.g., each vessel)—to obtain the mean, confidence interval, and standard deviation. These metrics were then successively aggregated following the hierarchical data structure: across vessels, across tissue types, and ultimately across all phantoms.

3 Results

3.1 Two Dyes Were Found to Mimic the Optical Absorption of Hemoglobin Across the Full Range of Oxygenations

Following the dye testing and optimization process, IR-1061 and Spectrasense-765 were identified as the closest proxies for blood absorption between 700 and 850 nm (Fig. 3, derived μ_s in Fig. S10 in the Supplementary Material) with Spectrasense-765 closely reproducing key features of Hb (local minimum at 730 nm, local maximum at 760 nm). Although IR-1061 exhibited a less pronounced slope than HbO_2, the measured absorption still increased monotonically with wavelength. Notably, the isosbestic point for the two dyes was located at ~800 nm, mirroring that of HbO_2 and Hb. The R values to their respective target spectra can be found in Table S3 in the Supplementary Material.

3.2 Five Oxygen Saturation Levels with Characteristics Similar to Blood Were Derived

By varying the mixing ratios of IR-1061 and Spectrasense-765, five levels of sO_2 were created that approximate blood-like characteristics (Fig. 4; DIS reflectance and transmittance in Fig. S11 in the Supplementary Material). Although the relationship between dye concentrations and resulting μ_a spectra is highly nonlinear, a relatively linear trend was observed when transitioning from a 90:10 to a 100:0 dye ratio (representing 0% to 100% sO_2 levels). Intermediate mixing ratios (e.g., 93:7, 95:5, and 97:3) then produced five distinct sO_2 values: 0%, 30%, 50%, 70%, and 100%. Taking 90:10 and 100:0 as LSU endmembers yielded intermediate values of 30.7%, 52.4%, and 67.4% sO_2.

3.3 Signal Correlation Shows Good Agreement with Measured Absorption Spectra

The HSI setup [Fig. 5(a)] captures a top-down view of each phantom [Fig. 5(b)]. Analysis of the signal in the denoted ROI for the representative case with 50% sO_2 demonstrated an extremely close match with the expected spectra [Fig. 5(c)].

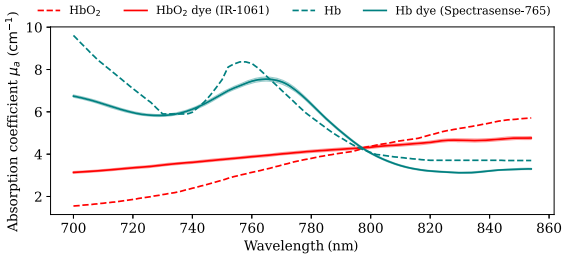


Fig. 3 IR-1061 and Spectrasense-765 can mimic the absorption characteristics of the blood. Dashed lines indicate the absorption coefficient (μ_a) spectra of oxyhemoglobin (HbO_2) and deoxyhemoglobin (Hb). Solid lines indicate the measured absorption spectra of the proxy dyes. Bands around the measured spectra indicate the standard deviation across the 12 measurement points for each optical sample slab. Mean absolute errors between targets and proxy dyes are 0.78 and 0.64 cm^{-1} for HbO_2 and Hb, respectively.

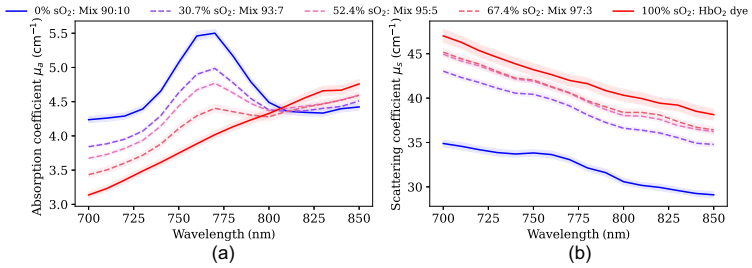


Fig. 4 Five oxygen saturation (sO_2) levels were used for forearm phantom fabrication. Based on IR-1061 and Spectrasense-765, five sO_2 levels (in %) were derived with the respective mixture ratios of 100:0, 97:3, 95:5, 93:7, and 90:10. Panels (a) and (b) represent the absorption coefficient (μ_a) and scattering coefficient (μ_s), respectively. Solid lines are the spectra that are used as end-members (0% sO_2 and 100% sO_2) for linear spectral unmixing (LSU). Dashed lines represent the intermediate levels, and the corresponding percentages in the legend are the LSU results when using the solid lines as endmembers. Bands around the spectra indicate the standard deviation across the 12 measurement points for each optical sample slab.

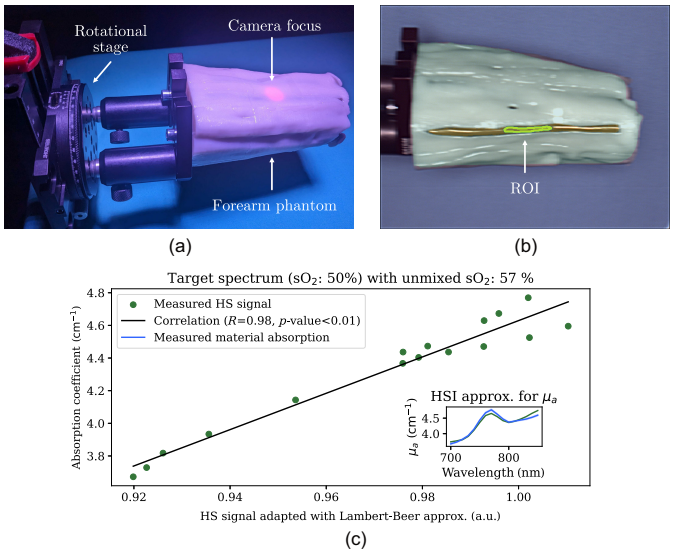


Fig. 5 Measurement setup for hyperspectral imaging (HSI) (a), corresponding image (b), and signal correlation of the 50% oxygen saturation (sO_2) superficial vessel (c). The phantoms were mounted on a rotational stage, and the camera was adjusted for each phantom such that the middle of the phantoms was in focus (red dot). Images from eight angles in steps of 45 deg were acquired per phantom. Panel (b) shows an example of an RGB-reconstructed image using the spectral range of 530 to 725 nm including a region of interest (ROI) indicated by yellow margins in the middle of a superficial vessel. The black solid lines in panel (c) represent the resulting linear regression function with the corresponding Pearson correlation coefficient (R value). The inset plot shows the measured absorption (blue) and estimated absorption (green, using the correlation function) as qualitative confirmation.

The PAT system [Fig. 6(a)] provides cross-sectional images of the phantoms [Fig. 6(b)]. The PAT correlation analysis in Fig. 6(c) also shows high concordance with the expected absorption, although a higher degree of noise is seen in the PAT data compared with HSI. Analogous results were obtained from other imaged forearm phantoms across the range of defined sO_2 levels,

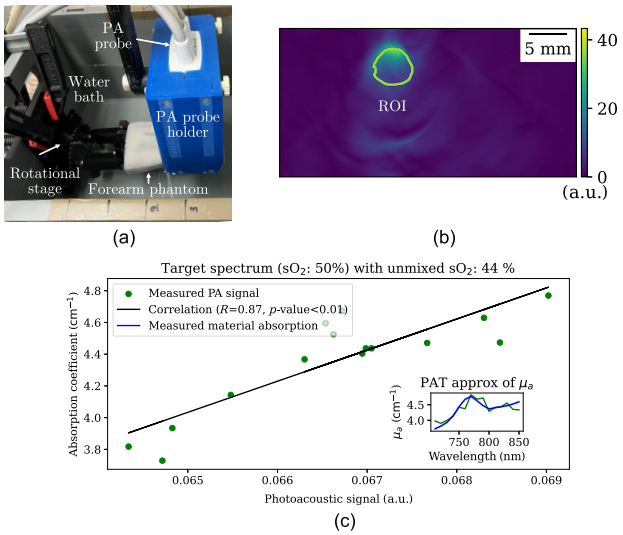


Fig. 6 Measurement setup for photoacoustic tomography (PAT) (a), corresponding image (b), and signal correlation of the 50% oxygen saturation (sO_2) superficial vessel (c). The phantoms were mounted on a rotational stage in a water bath, and images were acquired from eight angles in steps of 45 deg in three distinct locations along the phantoms. For each measurement location, the PAT probe, which was attached to a mechanical arm to minimize motion during image acquisitions, was adjusted such that it was approximately in the middle of the phantoms and 2 mm above the uppermost point of the phantom. Panel (b) shows an example PA image at 800 nm including a region of interest (ROI) indicated by yellow margins in the middle of a superficial vessel. The black solid lines in panel (c) represent the resulting linear regression function with the corresponding Pearson correlation coefficient (R value). The inset plot shows the measured absorption (blue) and estimated absorption (green, using the correlation function) as qualitative confirmation.

Table 1 Pearson correlation coefficients (R value) for linear regression of hyperspectral imaging (HSI) and photoacoustic tomography (PAT) signal and absorption correlation.

	0% sO_2	30% sO_2	50% sO_2	70% sO_2	100% sO_2
HSI	0.85	0.99	0.98	0.90	0.97
PAT	0.85	0.83	0.87	0.93	0.99

including images with more than one embedded vessel in the cross-sectional image (Figs. S12–S15 in the Supplementary Material). The correlation coefficient R value between the imaging signals and the measured absorption spectra in superficial vessels across the five example phantoms all exceeded 0.8, indicating strong agreement (Table 1).

3.4 Dyes Do Not Exhibit Any Significant Long-Term Optical Degradation

To indirectly examine the long-term optical stability of the used dyes, we re-recorded the HS images described above in Sec. 3.3 more than 1 year later and performed the same correlation analysis between the original and the re-recorded HS signals. The R values in Table 2 all exceed 0.95 and thus show very strong agreement. Plots with all measured spectra and their correlations are provided in the Figs. S26–S30 in the Supplementary Material.

Table 2 Pearson correlation coefficients (R values) for linear regression of the original and re-recorded hyperspectral (HS) signal (measured over 1 year later) correlations.

	0% sO_2	30% sO_2	50% sO_2	70% sO_2	100% sO_2
R-value	0.98	0.97	1.00	0.96	0.99

Table 3 Mean absolute error for linear spectral unmixing (LSU) applied on photoacoustic tomography (PAT) images and hyperspectral imaging (HSI) images.

	PAT		HSI	
	Entire phantom	Vessels-only	Entire phantom	Vessels-only
LSU	33.4 ± 19.6	30.4 ± 18.4	39.9 ± 15.0	45.3 ± 19.0
	[32.6, 34.1]	[29.4, 31.3]	[39.6, 40.3]	[44.7, 46.0]
Calibrated LSU	29.9 ± 5.0	27.9 ± 4.9	32.0 ± 0.6	31.1 ± 0.8
	[29.7, 30.1]	[27.6, 28.1]	[32.0, 32.0]	[31.1, 31.1]
Fluence compensation	27.9 ± 17.9	24.2 ± 16.9	Not applicable	
	[27.2, 28.6]	[23.3, 25.1]		

Calibrated LSU indicates that the correlation function was calculated via linear regression in Sec. 3.3. For PAT, fluence compensation means that the reconstructed image was corrected based on simulated fluence. The standard deviation is indicated with "±" and the bounds of the 95% confidence intervals of the mean are in square brackets: [low, high]. All values are in percentage points (p.p.).

3.5 Forearm Phantoms Can Be Used for Validation of Oximetry Methods

The performance for the three tested oximetry methods (LSU, LSU with calibration from signal correlation analysis, and LSU with prior fluence compensation) applied to both PAT and HSI data can be found in Table 3. On average, the calibrated versions of LSU showed lower MAE and reduced standard deviation compared to the uncalibrated approach, suggesting improved performance. Fluence compensation in PAT increased the accuracy of evaluation in vessels, although its benefit for the entire phantom was similar to the uncalibrated method. A key consideration for PAT is evaluation at a function of tissue depths (Fig. 7) through which our findings indicate that the calibrated LSU outperforms the uncalibrated approach at most depths in the entire phantom; the fluence compensation method particularly performs well for the first 5 mm.

4 Discussion

In this study, we presented a set of anthropomorphic forearm phantoms that were designed to replicate the optical absorption of HbO_2 and Hb at five distinct sO_2 levels (0%, 30%, 50%, 70%, and 100%) within a biologically relevant wavelength range (700 to 850 nm). By employing 3D-printed molds and strategically selected dyes, these phantoms achieved a morphologically realistic geometry while reproducing the optical properties HbO_2 and Hb.

Phantom designs of previous work often rely on simplified geometries such as tubular or multilayer structures. Although flow-based setups for controlled sO_2 exist, they generally lack morphological accuracy or long-term stability. Our approach addresses these shortcomings by incorporating forearm morphology and carefully chosen dyes. This strategy yields phantoms that more closely match the morphology and absorption characteristics of human tissue and provides a robust testing platform for oximetry methods across various imaging modalities. It also avoids the need for an *in situ* assessment of ground-truth sO_2, as would be required in a flow circuit. Although there may be instances when testing with blood itself is vital, for example, for optical

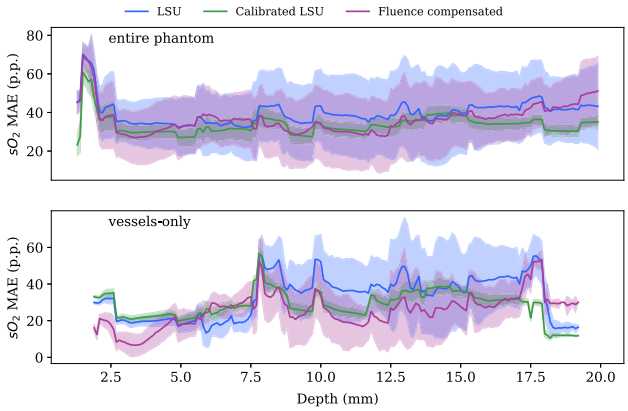


Fig. 7 Error analysis of three oximetry methods for photoacoustic tomography as a function of depth. Mean absolute error (MAE) in percentage points (p.p.) of linear spectral unmixing (LSU) without (blue) and with calibration (green) using a superficial vessel, and with fluence compensation (pink) is plotted against the depth of the evaluation pixel for both the whole phantom (top) and for the vessels only (bottom). The bands around the solid lines indicate the standard deviation.

systems that rely on scattering from moving red blood cells, for many target applications such as intraoperative imaging, optical systems could be more easily and routinely tested using the approach outlined here.

One core contribution of our work is the open-source data and code for both the phantom molds and dye optimization process. Researchers can 3D-print their own molds derived from high-resolution MRI data, ensuring consistent morphology across laboratories. They can also adapt our dye optimization framework to derive new chromophore combinations, extending the utility of our work to other tissue types. Moreover, our validation experiments with HSI and PAT indicate that the fabricated phantoms produce signals in good agreement with measured absorption, thereby offering a reliable, standardized environment not only for validation of different imaging systems between sites but also for evaluating different oximetry algorithms. By providing a standardized but anthropomorphic phantom, the approach described here could accelerate the development and comparison of optical imaging methods, ultimately improving clinical translation.

Our correlation analyses show that both HSI and PAT signals strongly match the known absorption coefficients of each phantom, underscoring the fidelity of the phantoms for experimental validation. In particular, Pearson correlation coefficients consistently exceeded 0.8, and the analysis for the five example forearms demonstrated close alignment with the expected spectral behavior. Some instability was observed when applying oximetry methods at depth, which is likely due to various vessels being embedded at different positions, having substantially larger absorption coefficients than their surroundings. Therefore, when going from background material to vessel material, the error of the oximetry method might experience a sudden drop or increase.

Despite these advantages, several limitations remain. First, the two dyes used exhibit highly nonlinear mixing behavior, complicating the process of achieving intermediate sO_2 levels. Although we successfully generated five distinct levels, further refinement is needed to enhance reproducibility and minimize iteration. Further, even though we showed that the HS signals did not degrade over time, we did not explicitly test the long-term optical and chemical stability of these dyes within the base copolymer-in-oil matrix. Second, the DIS system employed for optical characterization is susceptible to measurement uncertainties, which have been reported in prior studies.[44] Although we attempted to minimize these uncertainties by verifying the mixing ratio both by weighing the dyes and by using LSU on the measured μ_a, minor deviations are still

possible. Because the focus of this paper was on the optical properties of the phantoms, a full acoustic characterization was not performed, and only basic speed of sound measurements, borrowed from previous work,[28] were used. Finally, the phantoms contained small air bubbles despite extensive vacuuming, which may cause acoustic reverberations, particularly for vessels located deeper than 1.5 cm. Therefore, we recommend limiting the validation of oximetry methods with these phantoms to vessels whose centres lie at depths shallower than 1.5 cm. To simplify fabrication and ensure reproducibility, we opted for a homogeneous background material without an additional skin-mimicking layer. Although adding a realistic skin layer is conceptually valuable and would add more realism to the phantoms, it is technically challenging in a phantom of this size without introducing more air entrapments. Finally, for visual comparison, we showed a side-by-side view of a human forearm, an example forearm phantom, and its corresponding simulation in Fig. S31 in the Supplementary Material, including their unmixed images. As there are still some very apparent differences such as the heterogeneity of the bulk background or the size of vessels, future work will tackle these discrepancies.

5 Conclusion

By enabling accurate and repeatable performance assessments, our tissue-mimicking phantoms provide a robust standard for oximetry validation, bridging the gap left by limited *in vivo* reference methods. Looking ahead, the fabrication strategies and dye selection processes described here offer an initial step for developing even more complex phantoms, thereby fostering more reliable and clinically translatable optical imaging techniques.

Disclosures

The authors declare no conflict of interest regarding this work.

Code and Data Availability

The data and code to reproduce the findings of this study are openly available. The data are available under the CC-BY 4.0 license at: https://doi.org/10.5281/zenodo.15102333. The code is available under the MIT license at: https://github.com/IMSY-DKFZ/anthropomorphic-phantoms.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 101002198), Cancer Research UK (Grant Nos. C9545/A29580 and C14478/A27855), and the UKRI Engineering and Physical Sciences Research Council (Grant Nos. EP/L015889/1, EP/X037770/1, and EP/V027069/1). RT acknowledges the financial support of the Trinity Barlow Scholarship, Cambridge Trust International Scholarship, and the Herchel Smith Scholarship. JG acknowledges funding from the Walter Benjamin Stipendium of the Deutsche Forschungsgemeinschaft. For the purpose of open access, the author has applied a Creative Commons Attribution (CC-BY) license to any Author Accepted Manuscript version arising. The authors would like to thank Prof. Dr. Jürgen Hesser for the helpful and insightful discussions. A Large Language Model (ChatGPT) has been used to enhance the language of the manuscript. All model outputs have been thoroughly reviewed and verified for correctness by the authors.

References

1. M. Taylor-Williams et al., "Noninvasive hemoglobin sensing and imaging: optical tools for disease diagnosis," *J. Biomed. Opt.* **27**, 080901 (2022).
2. T. L. Lefebvre et al., "The potential of photoacoustic imaging in radiation oncology," *Front. Oncol.* **12**, 803777 (2022).
3. J. J. M. Riksen, A. V. Nikolaev, and G. V. Soest, "Photoacoustic imaging on its way toward clinical utility: a tutorial review focusing on practical application in medicine," *J. Biomed. Opt.* **28**, 121205 (2023).
4. V. Ntziachristos, "Addressing unmet clinical need with optoacoustic imaging," *Nat. Rev. Bioeng.* **3**, 182–184 (2024).
5. J. Park et al., "Clinical translation of photoacoustic imaging," *Nat. Rev. Bioeng.* **3**, 193–212 (2024).
6. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.* **19**, 010901 (2014).

7. N. T. Clancy et al., "Surgical spectral imaging," *Med. Image Anal.* **63**, 101699 (2020).
8. L. V. Wang and H.-I. Wu, *Biomedical Optics: Principles and Imaging*, John Wiley & Sons (2007).
9. L. Wang, Ed., *Photoacoustic Imaging and Spectroscopy*, CRC Press, Boca Raton (2017).
10. R. Hochuli et al., "Estimating blood oxygenation from photoacoustic images: can a simple linear spectroscopic inversion ever work?" *J. Biomed. Opt.* **24**, 121914 (2019).
11. J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **5**, 354–379 (2012).
12. T. R. Else et al., "Effects of skin tone on photoacoustic imaging and oximetry," *J. Biomed. Opt.* **29**, S11506 (2023).
13. L. Ayala et al., "Band selection for oxygenation estimation with multispectral/hyperspectral imaging," *Biomed. Opt. Express* **13**, 1224–1242 (2022).
14. M. Larsson et al., "Artificial neural networks trained on simulated multispectral data for real-time imaging of skin microcirculatory blood oxygen saturation," *J. Biomed. Opt.* **29**, S33304 (2024).
15. J. Gröhl et al., "Learned spectral decoloring enables photoacoustic oximetry," *Sci. Rep.* **11**, 6565 (2021).
16. J.-H. Nölke et al., "Photoacoustic quantification of tissue oxygenation using conditional invertible neural networks," *IEEE Trans. Med. Imaging* **43**, 3366–3376 (2024).
17. M. Chen et al., "Simultaneous photoacoustic imaging of intravascular and tissue oxygenation," *Opt. Lett.* **44**, 3773–3776 (2019).
18. L. A. Ayala et al., "Live monitoring of haemodynamic changes with multispectral image analysis," *Lect. Notes Comput. Sci.* **11796**, 38–46 (2019).
19. T. Kirchner et al., "Photoacoustics can image spreading depolarization deep in gyrencephalic brain," *Sci. Rep.* **9**, 8661 (2019).
20. J. Gröhl et al., "Deep learning for biomedical photoacoustic imaging: a review," *Photoacoustics* **22**, 100241 (2021).
21. J. Zalev et al., "Opto-acoustic imaging of relative blood oxygen saturation and total hemoglobin for breast cancer diagnosis," *J. Biomed. Opt.* **24**, 121915 (2019).
22. T. Kirchner and M. Frenz, "Multiple illumination learned spectral decoloring for quantitative optoacoustic oximetry imaging," *J. Biomed. Opt.* **26**, 085001 (2021).
23. L. Hacker et al., "Criteria for the design of tissue-mimicking phantoms for the standardization of biophotonic instrumentation," *Nat. Biomed. Eng.* **6**, 541–558 (2022).
24. A. B. Walter and E. D. Jansen, "Development of a platform for broadband, spectra-fitted, tissue optical phantoms," *J. Biomed. Opt.* **28**, 025001 (2023).
25. W. C. Vogt et al., "Photoacoustic oximetry imaging performance evaluation using dynamic blood flow phantoms with tunable oxygen saturation," *Biomed. Opt. Express* **10**, 449–464 (2019).
26. P. Ghassemi et al., "Rapid prototyping of biomimetic vascular phantoms for hyperspectral reflectance imaging," *J. Biomed. Opt.* **20**, 121312 (2015).
27. M. Dantuma et al., "Tunable blood oxygenation in the vascular anatomy of a semi-anthropomorphic photoacoustic breast phantom," *J. Biomed. Opt.* **26**, 036003 (2021).
28. L. Hacker et al., "A copolymer-in-oil tissue-mimicking material with tuneable acoustic and optical characteristics for photoacoustic imaging phantoms," *IEEE Trans. Med. Imaging* **40**, 3593–3603 (2021).
29. J. Gröhl et al., "Moving beyond simulation: data-driven quantitative photoacoustic imaging using tissue-mimicking phantoms," *IEEE Trans. Med. Imaging* **43**, 1214–1224 (2024).
30. J. W. Pickering et al., "Double-integrating-sphere system for measuring the optical properties of tissue," *Appl. Opt.* **32**, 399–410 (1993).
31. S. Prahl, *Everything I Think You should Know about Inverse Adding-Doubling*, Vol. 1344, pp. 1–74, Oregon Medical Laser Center, St. Vincent Hospital (2011).
32. S. L. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**, R37 (2013).
33. C. J. M. Jones and P. R. T. Munro, "Stability of gel wax based optical scattering phantoms," *Biomed. Opt. Express* **9**, 3495–3502 (2018).
34. E. Maneas et al., "Gel wax-based tissue-mimicking phantoms for multispectral photoacoustic imaging," *Biomed. Opt. Express* **9**, 1151–1163 (2018).
35. S. Prahl, W. B. Gratzer, and N. Kollias, "Tabulated molar extinction coefficient for hemoglobin in water," 2018, https://omlc.org/spectra/hemoglobin/summary.html.
36. F. D. Kerkhof, T. van Leeuwen, and E. E. Vereecke, "The digital human forearm and hand," *J. Anatomy* **233**(5), 557–566 (2018).
37. M. Schellenberg et al., "Semantic segmentation of multispectral photoacoustic images using deep learning," *Photoacoustics* **26**, 100341 (2022).
38. J. Gröhl et al., "SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics," *J. Biomed. Opt.* **27**, 083010 (2022).
39. J. Sellner and S. Seidlitz, "Hyperspectral tissue classification," Zenodo (2025).

40. Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Opt. Express* **17**, 20178–20190 (2009).
41. B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**, 021314 (2010).
42. B. T. Cox et al., "Quantitative spectroscopic photoacoustic imaging: a review," *J. Biomed. Opt.* **17**, 061202 (2012).
43. I. Oshina and J. Spigulis, "Beer–Lambert law for optical tissue diagnostics: current state of the art and the main limitations," *J. Biomed. Opt.* **26**, 100901 (2021).
44. R. Tao et al., "Tutorial on methods for estimation of optical absorption and scattering properties of tissue," *J. Biomed. Opt.* **29**, 060801 (2024).

Kris K. Dreher received his MSc degree in physics from the University of Heidelberg in 2020. He is currently pursuing a PhD at the Division of Intelligent Medical Systems (IMSY), German Cancer Research Center (DKFZ). His research focuses on computational modeling of optical imaging techniques and development of deep learning-based methods with application in optical imaging.

Jan-Hinrich Nölke received his MSc degree in physics from Heidelberg University in 2021. He is currently pursuing an interdisciplinary PhD at the division of Intelligent Medical Systems (IMSY), German Cancer Research Center (DKFZ). His research focuses on deep learning-based uncertainty quantification in medical imaging.

Katie-Lou White is a PhD student in Sarah Bohndiek's group at the University of Cambridge, developing multispectral endoscopes for early cancer detection in Barrett's Oesophagus with a focus on hardware innovation and first-in-human clinical trials. This project is funded by the International Alliance for Cancer Early Detection. She completed her bachelor's and master's degrees in natural sciences (physics) at the University of Cambridge, where she specialized in biophysics, optics, and condensed matter.

Ran Tao completed her MRes in connected electronic and photonic systems in 2021 and is currently pursuing her PhD in physics, both from the University of Cambridge. Her research focuses on broadband systems for estimating tissue optical absorption and scattering properties, extending optical measurements from near-infrared-I (650 to 950 nm) to short-wavelength infrared (1000 to 2500 nm).

Minu D. Tizabi is a physician, scientist and writer in the division of intelligent medical systems. In 2007, she graduated from high school at the age of 14, becoming Germany's youngest high school graduate, followed by medical studies at Heidelberg University which she completed as Germany's youngest physician in 2015. She obtained her doctorate of medicine from Heidelberg University in 2017.

Alexander Seitel is a computer scientist currently working as a group lead and deputy head at the German Cancer Research Center (DKFZ) in Heidelberg and holds a doctorate in medical informatics from the University of Heidelberg. His research focusses on computer-assisted interventions and novel imaging methodologies aiming to improve interventional healthcare. In this area, he conducted various international projects at DKFZ and during his two-year postdoctoral fellowship at the University of British Columbia, Vancouver, Canada.

Lena Maier-Hein is a full professor at Heidelberg University (Germany) and division head at the German Cancer Research Center (DKFZ). She is managing director of the National Center for Tumor Diseases (NCT) Heidelberg and of the DKFZ Data Science and Digital Oncology cross-topic program. Her research concentrates on machine learning-based biomedical image analysis with a specific focus on surgical data science, computational biophotonics and validation of machine learning algorithms.

Sarah E. Bohndiek completed her PhD in radiation physics at the University College London in 2008 and then worked in both the United Kingdom (at Cambridge) and the United States (at Stanford) as a postdoctoral fellow in molecular imaging. Since 2013, she has been a group leader at the University of Cambridge, where she is jointly appointed in the Department of

Physics and the Cancer Research UK Cambridge Institute. She was appointed as full professor of biomedical physics in 2020. She was recently awarded the CRUK Future Leaders in Cancer Research Prize and SPIE Early Career Achievement Award in recognition of her innovation in biomedical optics.

Biographies of the other authors are not available.

4.5 P.V: Open-source validation data

SPECTRE: Simulated photoacoustic and experimental data for comparing tissue realism and quantitative estimation

Authors:	**Kris K. Dreher**, Jan-Hinrich Nölke, Christoph J. Bender, Friso Grace, Alexander Seitel, and Lena Maier-Hein
Status:	In preparation
Journal:	N/A
DOI:	N/A
Copyright:	N/A
Contributions:	**KKD** and LMH conceived and conceptualized the project. **KKD** and LMH developed the methodology and defined the experiments to be conducted. **KKD** conducted the experiments for the technical validation of the dataset and wrote the first version of the manuscript. **KKD**, AS, and LMH shared administration of the project. FG, JN, and CJB contributed to data acquisition. JN provided the code for hierarchical bootstrapping. **All authors** contributed to reviewing and editing the manuscript. LMH acquired the funding and ensured the project's scientific rigor and relevance.

SPECTRE: Simulated photoacoustic and experimental data for comparing tissue realism and quantitative estimation

Kris K. Dreher[1,2], Jan-Hinrich Nölke[1,3], Christoph J. Bender[1,4], Friso Grace[1,5], Alexander Seitel[1,6], and Lena Maier-Hein[1,3,4,6,7]

[1] Intelligent Medical Systems, German Cancer Research Center (DKFZ), Heidelberg, Germany k.dreher@dkfz-heidelberg.de, l.maier-hein@dkfz-heidelberg.de
[2] Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany
[3] Faculty of Mathematics and Computer Science, Heidelberg University, Germany
[4] Medical Faculty, Heidelberg University, Heidelberg, Germany
[5] University of St Andrews, School of Physics and Astronomy, St Andrews, United Kingdom
[6] National Center for Tumor Diseases (NCT) Heidelberg, a partnership between DKFZ and Heidelberg University Hospital, Heidelberg, Germany
[7] Helmholtz Information and Data Science School for Health, Karlsruhe/Heidelberg, Germany

Abstract. Photoacoustic imaging (PAI) holds significant clinical potential for its functional imaging capabilities, yet the lack of high-quality, experimental datasets with reference measurements for optical or functional properties such as blood oxygenation has hindered the development and validation of data-driven methods and thus the translation of PAI into clinical routine. To support the validation of such methods, we present SPECTRE (Simulated Photoacoustic and Experimental data for Comparing Tissue Realism and quantitative Estimation), a curated dataset comprising three complementary components. First, a set of multispectral PA images acquired from anatomically realistic tissue-mimicking phantoms provides a controlled experimental reference for benchmarking. Second, corresponding paired simulated images based on the segmented phantom geometries with known optical properties enable pixel-wise comparison and image quality assessment. Third, a probabilistic simulation dataset models anatomical variability within the phantoms, facilitating the training and testing of robust algorithms in analogy to the real, unpaired use case. All simulated images are labeled with optical and functional properties, enabling direct comparison with their experimental counterpart and quantitative evaluation. This resource supports a range of tasks, for two of which (domain transfer and oximetry) we provide example applications. By offering the first open-access resource with both paired and unpaired real and simulated multispectral photoacoustic data alongside known reference values for optical and functional properties, SPECTRE establishes a foundation for benchmarking and accelerating methodological advances in PAI.

Keywords: Photoacoustic imaging · synthetic data · validation · oximetry · data-driven methods · deep learning.

1 Background

Photoacoustic imaging (PAI) is a hybrid imaging modality that combines optical contrast with ultrasound resolution to visualize optical absorption deep within biological tissue, offering the potential for real-time functional imaging [1,2,3]. Despite this promise, PAI has not yet been translated into clinical routine. Traditional model-based approaches for estimating functional parameters such as blood oxygen saturation (sO_2) have shown limited accuracy and robustness in *in vivo* settings [4]. A primary cause is the wavelength- and tissue-dependent attenuation of light, commonly referred to as spectral coloring, which renders the inverse problem nonlinear and ill-posed [5,6].

In response, recent research has focused on the development of more advanced, data-driven methods that aim to learn the inverse mapping from multispectral photoacoustic signals to physiological parameters [7,8,9,10]. However, a major obstacle persists: the absence of reference values for optical and functional tissue properties *in vivo*. This severely limits the ability to train and validate data-driven methods under realistic conditions.

To overcome this limitation, researchers often fabricate tissue-mimicking phantoms with known optical properties [11]. While phantoms offer a controlled environment for benchmarking and validation, their fabrication remains challenging due to the complex, multiscale structure of biological tissues. Moreover, producing phantoms in sufficient quantity and variability to support training of data-driven models is highly labor-intensive. Existing phantom-based datasets, such as flow phantoms or static constructs, typically have predefined, low-variability structures and do not adequately reflect the anatomical diversity needed to train robust algorithms [12]. Another existing dataset has a high number of PA images, the phantoms, however, only include geometrical shapes, thus are not anthropomorphic, and don't enable oximetry as downstream task [13].

In previous work, we presented the fabrication of anatomically realistic tissue-mimicking forearm phantoms and detailed the corresponding multispectral photoacoustic data acquisition process [14]. We demonstrated that the recorded photoacoustic signals reflect optical properties of blood and used the dataset to validate quantitative oximetry methods.

An alternative approach involves the simulation of photoacoustic images using physics-based models that replicate the underlying light and acoustic propagation processes [15]. Simulations provide access to complete ground-truth parameter maps and can, with sufficient computational resources, yield large-scale datasets suitable for training data-driven models. However, a persistent domain gap between simulated and experimental data limits the generalizability of models trained on purely synthetic datasets. This issue is typically addressed through two strategies: (1) improving simulation fidelity by refining the underlying physical models and generating high-quality digital twins of tissue and device con-

figurations [16,17,18]; and (2) applying domain transfer methods to enhance the visual and spectral realism of simulated images. The latter are commonly trained in an unsupervised fashion, given the lack of paired real and simulated images [19,20,21].

In this work, we extend the foundation made in our previous work and address the challenge of missing reference data through a combined approach that leverages both experimentally acquired phantom data and simulations. Specifically, we introduce SPECTRE (Simulated Photoacoustic and Experimental data for Comparing Tissue Realism and Quantitative Estimation), a curated dataset comprising experimental and simulated photoacoustic images derived from ten anthropomorphic forearm phantoms fabricated according to a standardized protocol [14]. SPECTRE includes: (i) experimental multispectral PAI data acquired from the phantoms; (ii) paired simulations based on manually segmented geometries with known optical properties; and (iii) a probabilistic simulation dataset modeling anatomical and physiological variability within the phantom domain for large-scale data-driven method training. An overview of the three datasets and their generation process is provided in Fig. 1.

To demonstrate the applicability of SPECTRE, we provide example validations for two relevant tasks in quantitative PAI: (1) unsupervised domain transfer and (2) blood sO_2 estimation from multispectral photoacoustic images.

2 Methods

This section describes the three datasets comprising SPECTRE. Dataset I provides experimental measurements from tissue-mimicking phantoms. Dataset II contains paired simulations for direct comparison and domain transfer. Dataset III offers probabilistic simulations to model variability for training and validating methods in an unpaired setting.

2.1 Dataset I: Experimental phantom dataset

To enable systematic validation of quantitative PAI methods, we previously introduced a set of ten anatomically realistic, tissue-mimicking forearm phantoms fabricated using a controlled protocol [14]. Nine of these phantoms, the standard phantoms, contain a combination of three different blood volume fractions and three different sO_2-levels while one additional phantom (phantom 10) serves as an out-of-distribution phantom outside of this 3 x 3 grid of combinations. Each phantom was imaged using the MSOT Acuity Echo system (iThera Medical GmbH, Munich, Germany) under standardized conditions. Specifically, to capture high anatomical and positional variability, each phantom was imaged at three distinct longitudinal positions along the forearm and from eight angular views (spaced at 45° intervals), yielding 24 multispectral images per phantom and a total of 240 images across the dataset. The wavelength range is 710 nm - 850 nm in steps of 10 nm.

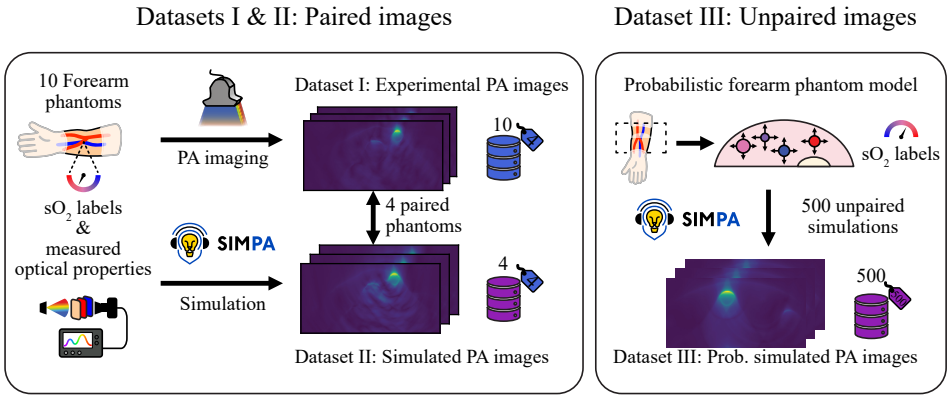


Fig. 1: **SPECTRE contains three datasets, two of them paired and one unpaired.** Dataset I contains experimental photoacoustic (PA) images that were acquired from ten anthropomorphic forearm phantoms with known optical parameters. As indicated by the label symbol, four of these phantoms were fully annotated, resulting in 96 labeled (test split) and 144 unlabeled images (train split). Based on these annotations, simulations could be generated via SIMPA, making up the 96 paired images of Dataset II. A probabilistic model of the forearm phantoms was generated, modeling anatomical and functional variability with, for example, different vessel sizes and positions or oxygen saturations (sO_2), respectively. Based on this probabilistic model, randomized phantom tissue constellations could be drawn and PA images simulated using SIMPA. Dataset III contains unpaired, labeled simulated images of 500 digital forearm phantoms. Blue corresponds to experimental data and manual annotations, and purple to simulated data.

Of the full dataset, a subset of four phantoms, thus 96 images, was manually annotated to establish voxel-level ground truth. This subset was selected to provide representative coverage of phantom properties, where three phantoms had different blood volume fractions and sO_2-levels, and one of these phantoms (phantom 10) was the out-of-distribution sample for robustness evaluation. The purpose of this subset is to serve as a test set throughout this work and thus to enable an initial evaluation of the phantoms' suitability for validating quantitative imaging methods.

Annotations were made manually and the fabrication protocols were used to assign vessels their corresponding sO_2. This resulted in semantic segmentation masks with nine tissue classes: (0) water, (1) bulk material, vessels at (2) 0%, (3) 30%, (4) 50%, (5) 70%, and (6) 100% sO_2, (7) bone, and (8) air bubbles. In addition, sO_2 maps were derived for blood-containing regions (labels 1–6), while sO_2 was set to undefined ("None") for non-perfused classes. These labels, in combination with the fabrication metadata reported in Dreher et al. [14], allow for the complete assignment of optical and acoustic properties across all annotated volumes.

2.2 Dataset II: Paired simulated phantom dataset

The 96 annotated images from Dataset I served as the basis for generating simulated images in Dataset II, thus enabling paired image comparison. Segmentation maps were loaded into the SIMPA framework [22] using its segmentation-based dataloader. Optical and acoustic properties, experimentally measured during phantom fabrication, were assigned to each segmented class. SIMPA was used to manage the simulation pipeline: Photon transport was modeled using MCX [16], and acoustic wave propagation was simulated with k-Wave [17]. Noise was added to the simulated time series data by iteratively optimizing Gaussian noise magnitude until the peak signal-to-noise ratio matched that of the corresponding experimental images for each wavelength. Image reconstruction was performed using a delay-and-sum algorithm. For further details on the simulation process, see [14]. This procedure yielded one multispectral simulated image per annotated experimental image, resulting in 96 paired simulations.

2.3 Dataset III: Unpaired simulated phantom dataset

In *in vivo* applications such as forearm oximetry, reference values for optical and functional tissue properties are not directly measurable. To address this limitation, a probabilistic simulation approach is commonly employed [23,6] based on anatomical priors from literature and statistics. Analogously, derived from the statistics of the experimental phantoms, we designed a probabilistic model of the phantoms to generate digital phantoms with high variability.

For each simulation, one of the three predefined imaging positions was randomly selected, and a circular base geometry for the phantom was generated using the mean radius observed in the corresponding subset of real phantoms. Background optical properties were uniformly sampled from those of the nine standard phantoms. With a probability of 40%, an elliptical structure representing a bone was inserted into the deeper region of the volume. This inclusion was modeled with negligible optical absorption and scattering but with acoustic properties approximating those of human cortical bone.

The number of blood vessels was randomly chosen between one and four, reflecting the distribution observed in the experimental data. Vessel radii were uniformly sampled between 3 and 6 mm, and their optical properties were drawn from the measurements obtained during phantom fabrication. Additionally, to mimic fabrication artifacts observed in the real phantoms, a random number (between 0 and 10) of small air inclusions was added, each with a diameter sampled between 0.1 mm and 0.25 mm.

Using this model, 500 multispectral photoacoustic images were simulated with the SIMPA framework, forming the unpaired, variability-rich Dataset III.

3 Technical validation

This section presents two example use cases demonstrating how SPECTRE can be used to validate relevant tasks in the field of PAI. An overview of the experimental design and validation strategy is provided in Fig. 2. First, the paired

datasets (I and II) are qualitatively and quantitatively compared to assess the visual and spectral realism of simulated images. Dataset II is transferred to the real domain (called Dataset II: Sim2Real) by a domain transfer method that was trained on datasets I (experimental) and III (unpaired simulations). Subsequently, Dataset II: Sim2Real is also compared analogously to Dataset I. Second, two data-driven oximetry methods, one operating on 2D image-level input and the other on 1D pixel-wise spectral information, are evaluated. Both models are trained on Dataset III and Dataset III: Sim2Real, and tested on Dataset I to assess real-world performance. Before presenting the results, the methods used in these validations are briefly introduced.

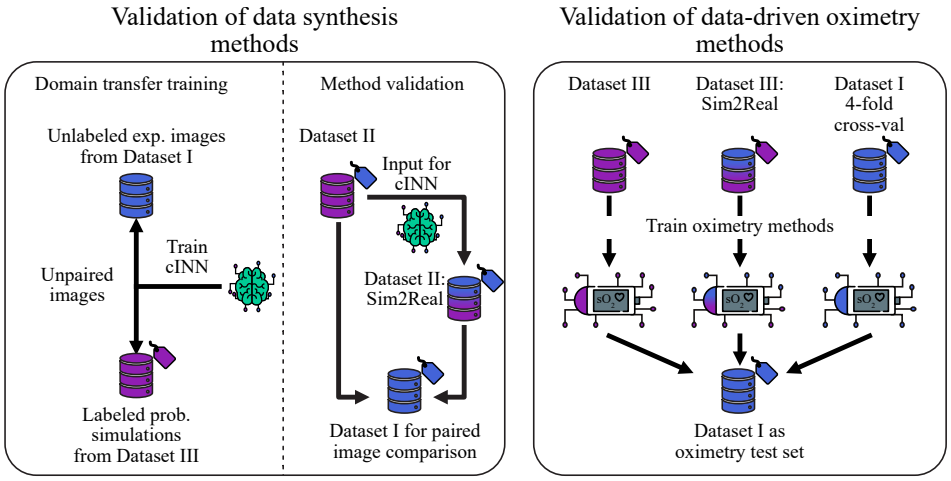


Fig. 2: **Various tasks can be validated using SPECTRE.** Domain transfer or assessment of synthetic data and data-driven oximetry are exemplarily validated. For domain transfer, first, unlabeled real data (Dataset I in purple) and labeled prob. simulated data (Dataset II in blue) can be used to train a domain transfer model, in this case a conditional invertible neural network (cINN)-based model. The model can use the simulated data to generate a more realistic Sim2Real dataset (Dataset II domain transferred with purple-blue gradient). For its validation, the generated dataset and Dataset II are compared to Dataset I in a paired manner. Data-driven oximetry methods can be trained on Dataset III and its domain transferred version as well as a 4-fold cross-validation of Dataset I. These three models can be applied to Dataset I, and their performances can be compared.

3.1 Data-driven methods

For domain transfer, we employed a method based on conditional invertible neural networks (cINNs), as introduced in [20]. Briefly, the cINN is trained to map

both simulated (s) and real (r) images into a shared Gaussian latent space (z_s and z_r), conditioned on the image domain. The segmentation maps are used as additional conditioning for simulated data, whereas noise is used for the real data. During inference, a simulated image is encoded into its latent representation z_s, which is then decoded into the real domain $z_{s \to r}$ using the real domain conditioning. This process enables the generation of realistic, labeled images aligned with the experimental data distribution.

For oximetry, two methods with different spatial modeling capacities were evaluated. The first, learned spectral decoloring (LSD) [6], is a fully connected neural network trained on 1D pixel-wise spectra extracted from vessels at various depths. By learning from this variability, LSD captures the effects of spectral decoloring implicitly. However, due to the inherent ambiguity in the mapping between spectra and sO_2 values caused by variations in depth and surrounding tissue, multiple spectra may correspond to the same sO_2 value, or vice versa.

To explicitly incorporate spatial context and improve robustness, we additionally trained a 2D U-Net [24] on the same oximetry task. This model, adapted from the architecture proposed by Schellenberg et al. [23], processes full multispectral image slices, allowing it to account for local spatial features that influence spectral appearance.

3.2 Realism of synthetic data

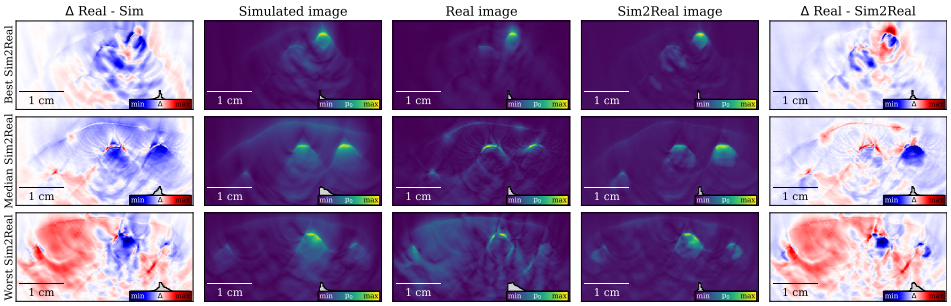


Fig. 3: **Domain transferred images look more similar to real images.** The images in the first, second, and third row correspond to the phantom image with best, median, and worst mean absolute error (MAE) of the domain-transferred images from Dataset II: Sim2Real compared to the real images from Dataset I. In the middle, the real images are shown and on the left and right, the simulated and Sim2Real images, respectively, are presented with their corresponding difference to the real images. All images correspond to 800 nm.

Fig. 3 presents a qualitative comparison of the best, median, and worst cases, ranked by mean absolute error (MAE) (of Dataset II: Sim2Real) between simulated images from Dataset II, their domain-transferred counterparts from Dataset

II: Sim2Real, and the corresponding experimental images from Dataset I. The domain-transferred images show reduced error in background tissue regions and, notably, exhibit lower signal intensity within vessels compared to their simulated counterparts, aligning more closely with experimental observations.

Fig. 4 shows the quantitative MAE comparison between the simulated and domain-transferred images to the experimental reference. While the domain-transferred images achieve a slightly lower mean and comparable median of the MAE distribution, they demonstrate substantially narrower 95% confidence intervals and interquartile ranges, indicating improved consistency across the dataset.

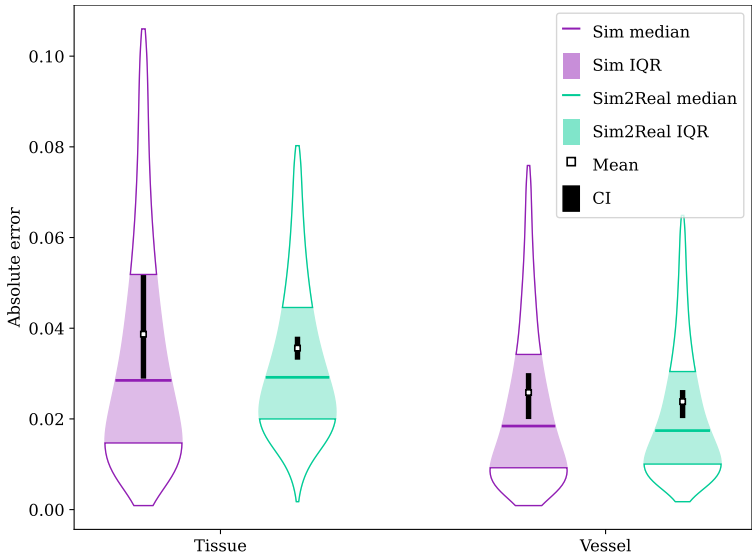


Fig. 4: **Domain transferred images have a smaller difference to real images.** The raw pixel distributions for absolute errors for both in the whole phantom image (tissue) and in the vessels only are shown as violinplots left and right, respectively, for both segmentation-based simulations (purple) and domain-transferred data (green). For both domains, the mean absolute error (squares), including 95% confidence intervals (CIs), as well as the median absolute error (horizontal colored lines) with the interquartile ranges (IQRs), are plotted. All descriptive statistics were calculated using hierarchical bootstrapping.

3.3 Data-driven oximetry

This section presents the results of the data-driven oximetry methods described in section 3.1 are presented. These methods share the lerning-based modeland the difference in the following is introduced by the training datasets. In all cases,

the experimental images of Dataset I are the test set to assess generalization to real data.

Each method was trained on three different datasets: (1) Dataset I, (2) Dataset III, and (3) Dataset III: Sim2Real. Training on Dataset I serves to approximate an upper bound on model performance under in-distribution conditions. As only four labeled forearm phantoms were available, four-fold cross-validation was employed for experiments involving Dataset I to for robustness while maximizing data use.

Oximetry on image-level 2D U-Nets were trained on each of the three datasets to estimate blood sO_2 from multispectral photoacoustic images with 15 wavelengths and a spatial resolution of 128 × 256 pixels. As shown in Table 1, the model trained on the Sim2Real dataset achieved the lowest absolute error (AE) on the experimental test set. However, representative examples in Fig. 5 illustrate that none of the models reliably capture the anatomical structure or accurately estimate sO_2 across the image.

Training dataset	Entire phantom	Vessels-only
Simulated	Mean: 37.1, CI: 25.4, 49.3	38.5, CI: 32.4, 43.9
	Median: 35.5, IQR: 11.8, 51.9	32.6, IQR: 12.5, 57.7
Sim2Real	Mean: 33.7, CI: 21.9, 46.9	31.4, CI: 27.6, 35.7
	Median: 29.9, IQR: 15.8, 50.0	29.6, IQR: 14.0, 45.7
Real	Mean: 39.6, CI: 21.5, 55.1	32.8, CI: 29.5, 37.0
	Median: 49.5, IQR: 12.9, 50.5	29.9, IQR: 19.5, 30.4

Table 1: Absolute errors in percentage points (p.p.) for U-Nets trained on simulated, Sim2Real, and real data for both the entire phantoms and vessels-only. The top values for each row indicate the mean absolute error with corresponding 95% confidence intervals (CIs), and the bottom values indicate the median absolute errors with corresponding interquartile ranges (IQRs). All descriptive statistics have been calculated using hierarchical bootstrapping.

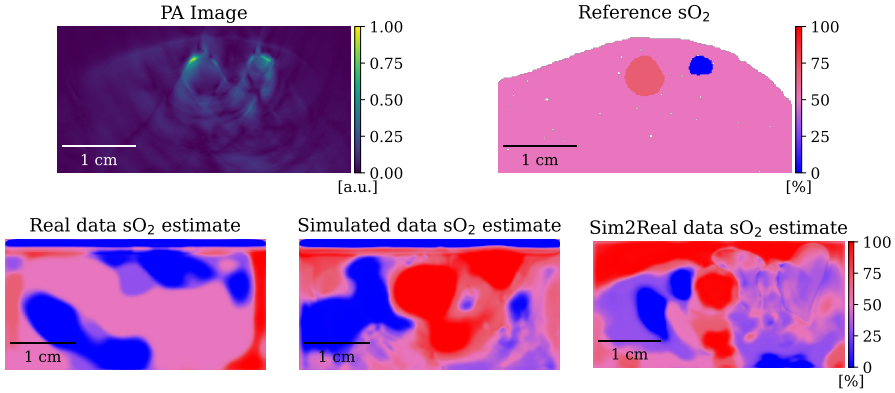


Fig. 5: **Example of blood oxygen saturation (sO_2) estimations by U-Net models.** In the top row, a representative exemplary PA image at 800 nm is shown as input to the trained U-Net with its corresponding sO_2 label. The images in the second row are in order the sO_2 estimations made by the U-Net models trained on real data, simulated data, and Sim2Real data.

Oximetry on pixel-level Pixel-wise LSD models were trained on each of the three datasets to estimate sO_2 from spectra consisting of 15 wavelengths. As shown in Table 2, the model trained on Sim2Real data achieved the second-lowest AE, following the model trained on Dataset I.

Training dataset	Vessels-only
Simulated	Mean: 59.4, CI: 48.4, 73.7
	Median: 67.6, IQR: 29.4, 90.3
Sim2Real	Mean: 32.0, CI: 29.3, 34.8
	Median: 27.0, IQR: 14.8, 46.7
Real	Mean: 29.5, CI: 26.9, 32.5
	Median: 26.1, IQR: 13.1, 41.4

Table 2: Absolute errors for LSD models trained on simulated, Sim2Real, and real data applied to spectra of vessels of real data only. The top values for each row indicate the mean absolute error with corresponding 95% confidence intervals (CIs) and the bottom values of each row indicate the median absolute errors with corresponding interquartile ranges (IQRs). All descriptive statistics have been calculated using hierarchical bootstrapping. All values are in percentage points (p.p.).

A more detailed analysis of the LSD models was performed by first investigating the estimated sO2 values for all test spectra, depending on their label, as shown in Fig. 6a, where the pixel error distributions are plotted. Second, the mean absolute error is plotted along depth for all LSD models in Fig. 6b.

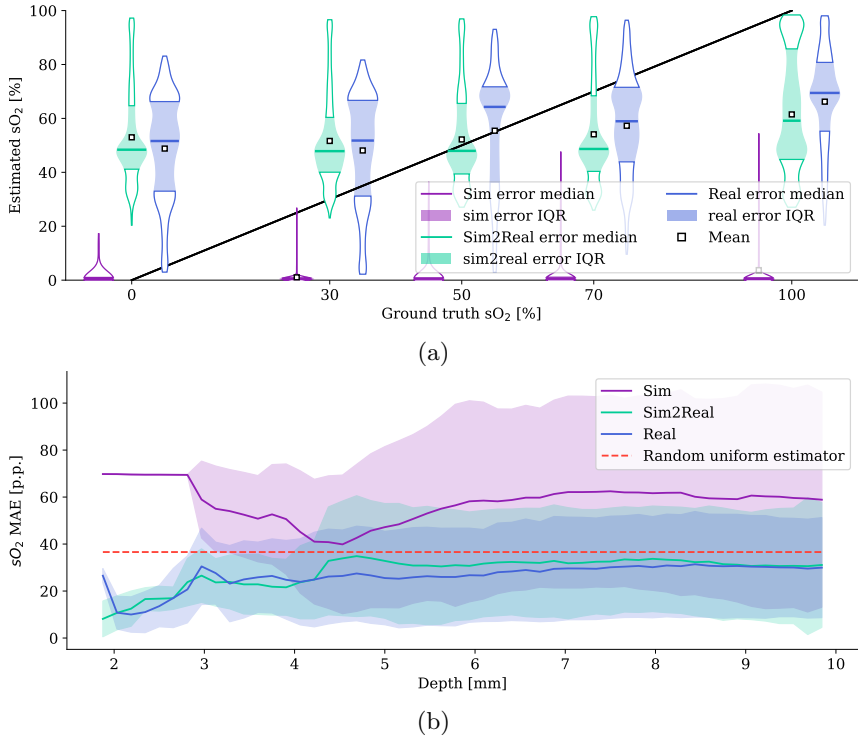


Fig. 6: **Sim2Real training data yields results more similar to real data compared to simulated data.** In subfigure 6a, for each of the five sO_2 levels within the phantoms (0%, 30%, 50%, 70%, 100%), the estimated pixel distributions for all three learned spectral decoloring (LSD) models are plotted. The median and interquartile range (IQR) are indicated in white and black, respectively. Subfigure 6b shows the mean absolute pixel error along depth within the phantoms for all three LSD models including a random uniform estimator with a mean absolute error (MAE) of 36.6 p.p. Bands around the mean indicate the standard deviation.

Figure 6a illustrates that the $\text{LSD}_{\text{Sim2Real}}$ model exhibits behavior more similar to the LSD_{real} model compared to the LSD_{sim} model. The LSD_{sim} model consistently predicts 0% sO_2 across all input spectra, indicating a complete failure to generalize to experimental data. While the $\text{LSD}_{\text{Sim2Real}}$ and LSD_{real} models predominantly predict sO_2 values near 50%, suggesting limited generalization, the presence of additional modes near 30% and 100% in the $\text{LSD}_{\text{Sim2Real}}$ predictions, and across most ground truth classes in the LSD_{real} model, suggests improved but still insufficient performance.

4 Discussion

In this work, we present **SPECTRE**, the first paired dataset of multispectral simulated and experimental photoacoustic images designed for validating quantitative tasks in PAI. Until now, the lack of experimental data with voxel-level annotations has made it challenging to rigorously evaluate data-driven methods. Most existing approaches have relied exclusively on simulated data, limiting their ability to generalize to real measurements. Domain transfer methods, in particular, have often been validated either indirectly via downstream tasks with available labels or solely on synthetic data.

SPECTRE addresses this limitation by leveraging data from anatomically realistic, tissue-mimicking forearm phantoms described in previous work [14], enabling paired and unpaired validation in a setting representative of *in vivo* use cases. In addition to introducing the dataset, we demonstrated its utility through two example tasks: domain transfer and sO_2 estimation. The dataset is also applicable to other tasks, such as segmentation or quantitative optical property estimation.

From the validation experiments, we draw several observations. In the domain transfer task, Sim2Real images consistently showed a moderate reduction in absolute error compared to the original simulations in Dataset II. Qualitatively, key discrepancies between simulated and experimental data, such as lower vessel signal intensity, were partially mitigated by the domain transfer approach. In the oximetry task, domain-transferred training data also yielded lower AE in sO_2 estimation compared to models trained on purely simulated data, suggesting improved generalization.

However, the results do not yet support definitive conclusions. The 2D U-Net model exhibited limited generalization capacity, likely due to the small training set size (72 images per fold) relative to the complexity of the input ($15 \times 128 \times 256$ dimensions). The pixel-wise LSD model trained on simulated data failed to generalize, consistently predicting 0% sO_2. In contrast, the LSD model trained on Sim2Real data produced a broader output distribution, with multiple modes of the prior distribution. Notably, its behavior resembled that of the LSD model trained on experimental data, suggesting that the domain-transferred and real data distributions are more similar.

Still, both LSD models displayed broad sO_2 distributions (cf. Fig. 6a), consistent with prior observations in [6], where spectral ambiguity in complex forearm tissue geometries limited the suitability of the LSD method. This further illustrates the intrinsic difficulty of the inverse problem and the need for methods that explicitly account for spectral decoloring and model uncertainty.

Based on these observations, one main conclusion is that the original experimental design regarding data splits resulted in too few data points in both training and test splits for the data-driven methods to be both trained and validated on. A 70:30 train:test split is fairly common in the ML community, and since in this work, an out-of-distribution sample was introduced as well, a 60:40 split was chosen. The 40% (96 images) were annotated and used as a held-out test set, which is also common practice. Since the baseline method trained on

these experimental data also could not perform substantially better than a random guesser, we can conclude that this training data was not enough for this task. Future work will focus on extending the annotated dataset to include all ten phantoms, thereby addressing current data limitations. A different training scheme for the Sim2Real models will also be explored. So far, only 60% of the dataset was used for domain transfer. An alternative approach could involve applying domain transfer to the full dataset while withholding ground-truth annotations for downstream tasks such as segmentation or oximetry during training. This would allow models to be trained on all domain-transferred data and evaluated across the complete set of annotated experimental images, improving statistical power without introducing data leakage.

In addition, we plan to evaluate more advanced oximetry models incorporating uncertainty estimation, such as those proposed by Nölke et al. [10], to better handle the inherent ambiguities in spectral photoacoustic data.

Acknowledgements. This project was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (NEURAL SPICING, 101002198).

References

1. Ntziachristos, V.: Addressing unmet clinical need with optoacoustic imaging. Nature Reviews Bioengineering pp. 1–3 (Sep 2024). https://doi.org/10.1038/s44222-024-00242-w
2. Park, J., Choi, S., Knieling, F., Clingman, B., Bohndiek, S., Wang, L.V., Kim, C.: Clinical translation of photoacoustic imaging. Nature Reviews Bioengineering pp. 1–20 (Sep 2024). https://doi.org/10.1038/s44222-024-00240-y
3. Lefebvre, T.L., Brown, E., Hacker, L., Else, T., Oraiopoulou, M.E., Tomaszewski, M.R., Jena, R., Bohndiek, S.E.: The Potential of Photoacoustic Imaging in Radiation Oncology. Frontiers in Oncology **12** (Mar 2022). https://doi.org/10.3389/fonc.2022.803777
4. Hochuli, R., An, L., Beard, P.C., Cox, B.T.: Estimating blood oxygenation from photoacoustic images: can a simple linear spectroscopic inversion ever work? Journal of Biomedical Optics **24**(12), 121914 (Dec 2019). https://doi.org/10.1117/1.JBO.24.12.121914
5. Cox, B.T., Laufer, J.G., Beard, P.C., Arridge, S.R.: Quantitative spectroscopic photoacoustic imaging: a review. Journal of Biomedical Optics **17**(6), 061202 (Jun 2012). https://doi.org/10.1117/1.JBO.17.6.061202
6. Gröhl, J., Kirchner, T., Adler, T.J., Hacker, L., Holzwarth, N., Hernández-Aguilera, A., Herrera, M.A., Santos, E., Bohndiek, S.E., Maier-Hein, L.: Learned spectral decoloring enables photoacoustic oximetry. Scientific Reports **11**(1), 6565 (Mar 2021). https://doi.org/10.1038/s41598-021-83405-8
7. Bench, C., Hauptmann, A., Cox, B.T.: Toward accurate quantitative photoacoustic imaging: learning vascular blood oxygen saturation in three dimensions. Journal of Biomedical Optics **25**(8), 085003 (Aug 2020). https://doi.org/10.1117/1.JBO.25.8.085003

8. Kirchner, T., Jaeger, M., Frenz, M.: Machine learning enabled multiple illumination quantitative optoacoustic oximetry imaging in humans. Biomedical Optics Express **13**(5), 2655–2667 (May 2022). https://doi.org/10.1364/BOE.455514
9. Gröhl, J., Yeung, K., Gu, K., Else, T.R., Golinska, M., Bunce, E.V., Hacker, L., Bohndiek, S.E.: Distribution-informed and wavelength-flexible data-driven photoacoustic oximetry. Journal of Biomedical Optics **29**(S3), S33303 (Jun 2024). https://doi.org/10.1117/1.JBO.29.S3.S33303
10. Nölke, J.H., Adler, T.J., Schellenberg, M., Dreher, K.K., Holzwarth, N., Bender, C.J., Tizabi, M.D., Seitel, A., Maier-Hein, L.: Photoacoustic Quantification of Tissue Oxygenation Using Conditional Invertible Neural Networks. IEEE Transactions on Medical Imaging **43**(9), 3366–3376 (Sep 2024). https://doi.org/10.1109/TMI.2024.3403417
11. Vogt, W.C., Zhou, X., Andriani, R., Wear, K.A., Pfefer, T.J., Garra, B.S.: Photoacoustic oximetry imaging performance evaluation using dynamic blood flow phantoms with tunable oxygen saturation. Biomedical Optics Express **10**(2), 449–464 (Feb 2019). https://doi.org/10.1364/BOE.10.000449
12. Dantuma, M., Kruitwagen, S., Julia, J.O., Meerdervoort, R.P.P.v., Manohar, S.: Tunable blood oxygenation in the vascular anatomy of a semi-anthropomorphic photoacoustic breast phantom. Journal of Biomedical Optics **26**(3), 036003 (Mar 2021). https://doi.org/10.1117/1.JBO.26.3.036003
13. Gröhl, J., Else, T.R., Hacker, L., Bunce, E.V., Sweeney, P.W., Bohndiek, S.E.: Moving Beyond Simulation: Data-Driven Quantitative Photoacoustic Imaging Using Tissue-Mimicking Phantoms. IEEE Transactions on Medical Imaging **43**(3), 1214–1224 (Mar 2024). https://doi.org/10.1109/TMI.2023.3331198
14. Dreher, K.K., Gröhl, J., Grace, F., Ayala, L., Nölke, J.H., Bender, C.J., Watt, M.J., White, K.L., Tao, R., Johnen, W., Tizabi, M.D., Seitel, A., Maier-Hein, L., Bohndiek, S.E.: Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging. Journal of Biomedical Optics **30**(7), 076006 (Jul 2025). https://doi.org/10.1117/1.JBO.30.7.076006
15. Gröhl, J., Schellenberg, M., Dreher, K., Maier-Hein, L.: Deep learning for biomedical photoacoustic imaging: A review. Photoacoustics **22**, 100241 (Jun 2021). https://doi.org/10.1016/j.pacs.2021.100241
16. Fang, Q., Boas, D.A.: Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units. Optics Express **17**(22), 20178–20190 (Oct 2009). https://doi.org/10.1364/OE.17.020178
17. Treeby, B.E., Cox, B.T.: k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields. Journal of Biomedical Optics **15**(2), 021314 (Mar 2010). https://doi.org/10.1117/1.3360308
18. Gröhl, J., Kunyansky, L., Poimala, J., Else, T.R., Cecio, F.D., Bohndiek, S.E., Cox, B.T., Hauptmann, A.: Digital twins enable full-reference quality assessment of photoacoustic image reconstructions (May 2025). https://doi.org/10.48550/arXiv.2505.24514
19. Li, J., Li, J., Wang, C., Chen, T., Lu, T., Li, S., Sun, B., Sun, B., Sun, B., Gao, F., Gao, F., Gao, F., Ntziachristos, V., Ntziachristos, V., Ntziachristos, V., Ntziachristos, V.: Deep learning-based quantitative optoacoustic tomography of deep tissues in the absence of labeled experimental data. Optica **9**(1), 32–41 (Jan 2022). https://doi.org/10.1364/OPTICA.438502, publisher: Optica Publishing Group
20. Dreher, K.K., Ayala, L., Schellenberg, M., Hübner, M., Nölke, J.H., Adler, T.J., Seidlitz, S., Sellner, J., Studier-Fischer, A., Gröhl, J., Nickel, F., Köthe, U., Seitel, A., Maier-Hein, L.: Unsupervised Domain Transfer with Conditional Invertible

Neural Networks. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. pp. 770–780. Springer Nature Switzerland, Cham (2023). https://doi.org/10.1007/978-3-031-43907-0_73

21. Wang, Z., Tao, W., Zhang, Z., Zhao, H.: Towards bridging the synthetic-to-real gap in quantitative photoacoustic tomography via unsupervised domain adaptation. Photoacoustics **45**, 100736 (Oct 2025). https://doi.org/10.1016/j.pacs.2025.100736
22. Gröhl, J., Dreher, K.K., Schellenberg, M., Rix, T., Holzwarth, N., Vieten, P., Ayala, L., Bohndiek, S.E., Seitel, A., Maier-Hein, L.: SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics. Journal of Biomedical Optics **27**(8), 083010 (Apr 2022). https://doi.org/10.1117/1.JBO.27.8.083010
23. Schellenberg, M., Gröhl, J., Dreher, K.K., Nölke, J.H., Holzwarth, N., Tizabi, M.D., Seitel, A., Maier-Hein, L.: Photoacoustic image synthesis with generative adversarial networks. Photoacoustics **28**, 100402 (Dec 2022). https://doi.org/10.1016/j.pacs.2022.100402
24. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. pp. 234–241. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

Part III

Reflections and closing

5 Discussion and future perspective

AI is advancing at an extraordinary pace, continuously reshaping scientific research and technological development across disciplines. In the context of PAI, the influence of AI is increasingly evident as shown by P.I, the first featured publication of this thesis. Among the first comprehensive reviews on DL in PAI [13, 76, 77], it remains the most-cited review of DL in the field. The review identified the lack of annotated, high-quality reference data as the principal obstacle to the broader adoption and effectiveness of data-driven methods in the field. Since the publication of P.I, interest in DL for PAI has grown significantly. Numerous subsequent reviews have expanded on the landscape, covering both general applications [78, 79] and domain-specific tasks such as artifact reduction or oximetry [80, 81].

Compared to traditional model-based methods, data-driven approaches offer several compelling advantages. They excel at capturing complex, nonlinear relationships between input and output, making them particularly effective for tasks such as solving partial differential equations in physics [82] or predicting protein folding in molecular biology [83]. In the physical sciences, these models are often employed under the paradigm of physics-informed neural networkss (PINNs), which integrate physical laws into the learning process to varying degrees, from purely data-driven learning to tightly constrained physics-guided inference [84, 85, 86]. Once trained, these models can serve as fast surrogate solvers for complex simulations, achieving significant reductions in computational time [87]. Moreover, the rapid inference enabled by data-driven models is a major advantage for real-time or intraoperative applications in medical imaging [88].

Recent breakthroughs, including diffusion models, transformer-based architectures, and large-scale foundation models, are beginning to steer methodological innovation within the PAI community. These models have already demonstrated promising results in a range of tasks, including image synthesis, oximetry, artifact reduction, and image reconstruction [89, 90, 91, 92, 93, 94].

However, these advantages come with important limitations. Training large-scale models often demands considerable computational resources, contributing to growing environmental concerns due to their substantial carbon footprint [95]. Additionally, many AI models operate as “black boxes”, making it difficult to interpret their internal decision-making processes. Despite the development of explainable AI techniques [96] and progress in

reasoning-based modeling [97], the lack of transparency remains a critical concern, particularly in high-stakes domains such as medicine, where model interpretability is essential for clinical acceptance [98].
Ultimately, there remains a broad consensus that the scarcity of high-quality data continues to represent one of the primary obstacles in the field. This fundamental limitation emerged as the central challenge addressed by this thesis. To address this challenge, three subchallenges and their corresponding research questions have been derived resulting in a comprehensive data-driven synthetic image generation pipeline including its validation. This section discusses the contributions made towards each individual RQ in the broader context of data-driven PAI.

5.1 Facilitating multiphysics simulations for photoacoustic image generation

The challenge of a comprehensive photoacoustic simulation pipeline

Computational modeling in PAI is fundamentally challenging due to the inherently multiphysics nature of the modality. Simulating PA images requires accurate modeling of light propagation through heterogeneous biological tissue, the conversion of absorbed optical energy into initial pressure via the PA effect, and subsequent acoustic wave propagation toward the detection system. Each of these steps involves complex, spatially varying physical properties and nonlinear interactions across multiple scales.
One of the most significant difficulties does not lie in the availability of simulation algorithms themselves. Highly capable simulators for photon transport [46, 99] (e.g., MC methods or finite element methods) and acoustic wave propagation [47, 100] (e.g., k-space pseudospectral methods) already exist. The problem lies in managing the complexity of integrating these tools into a coherent and extensible simulation pipeline. Moreover, the realism and generalizability of the simulations hinge critically on the quality of the input data, including device configurations, anatomical structures, and tissue properties.
To address these challenges, SIMPA was introduced, an open-source simulation toolkit for PAI, as described in *P.V: Open-source simulation toolkit*. SIMPA was developed to facilitate flexible, modular, and high-fidelity synthetic PA image generation. Rather than implementing custom physical solvers, SIMPA integrates the state-of-the-art simulation backends MCX [46] for light transport and k-Wave [47] for acoustic wave propagation, through a modular architecture that ensures compatibility with updates and allows straightforward substitution or extension with other simulation tools.
A core contribution of SIMPA is its standardized framework for defining digital twins of both biological tissue and imaging devices. This abstraction layer enables the reuse of complex anatomical models across experiments and fosters reproducibility. SIMPA supports the composition of application-specific, heterogeneous tissue models by combining literature-derived properties and analytical models (e.g., for vasculature).

Modeling realistic biological tissue and its photoacoustic effect

Significant limitations persist in modeling realistic tissue and device properties [37]. Creating anatomically plausible, multi-scale 3D tissue volumes remains a labor-intensive task, requiring accurate structural information as well as detailed knowledge of spatially and spectrally varying optical and acoustic parameters. Even widely used resources such as Jacques (2013) [101] provide only limited spectral coverage and lack consistency across tissue types. High-resolution *in vivo* characterization of tissue properties remains challenging, and even *ex vivo* measurements may not generalize due to differences in temperature, perfusion, or structural integrity. As a result, soft tissue is often modeled as homogeneous bulk, limiting physiological fidelity. To improve the multiscale modeling accuracy of bio-

logical tissue in PA simulations, two key aspects must be addressed. First, the anatomical realism of tissue geometries must be significantly enhanced. While procedural generation of simple homogeneous regions is common, accurate modeling of complex structures, such as vascular networks, remains a major challenge. Several specialized toolkits and data-driven approaches have been developed to generate realistic vascular trees using parallelized algorithms [102] and image-based realistic tissue [58], which could be integrated into existing pipelines to enable more faithful modeling of features like tumor vascularization or organ-specific microstructures. Second, the quantitative characterization of tissue properties, particularly optical parameters, is critical. While ongoing research continues to develop and refine measurement techniques for optical property estimation [103], each method carries specific limitations in terms of accuracy or spatial resolution. In addition to this, there is currently no open-source library providing a curated collection of tissue-specific optical properties, annotated with detailed metadata about experimental conditions and acquisition protocols. In contrast, such a resource does exist for acoustic properties, offering a useful model for how an open, community-driven database of optical parameters might be generated [104]. Establishing such a standardized, openly available database for optical properties would represent a major step forward in enabling more physiologically accurate and reproducible PA simulations.
Finally, accurate modeling of the PA effect itself, particularly the conversion of the absorbed energy to initial pressure via the Grüneisen parameter, remains underexplored. In most simulation pipelines, including SIMPA, this parameter is assumed constant and wavelength-independent, despite known variability across tissue types and conditions. Incorporating spatially and spectrally resolved models of the Grüneisen parameter represents a critical future step for enhancing realism.

Addressing research question 1 (RQ1)

How to facilitate multiphysics simulations for the generation of PA images?

In summary, this thesis addresses RQ1 by presenting SIMPA, the first open-source simulation toolkit tailored to the specific needs of PA image generation. By decoupling the simulation workflow from specific simulation toolkits and imaging devices, SIMPA significantly lowers the technical barrier for building high-fidelity, reproducible simulation workflows in PAI. SIMPA has proven to be a robust and extensible solution to facilitating multiphysics simulations in PAI. It has been adopted in several international initiatives, including projects under the umbrella of the IPASC [105, 106], and its growing impact on the community is also reflected by over 100 stars on its GitHub repository. While challenges such as accurate modeling of anatomical complexity and tissue-specific properties persist, SIMPA provides a flexible and robust foundation on which increasingly realistic simulations can be constructed.

5.2 Closing the domain gap while preserving spectral fidelity

Domain transfer methods

The limited generalizability of data-driven methods from simulated to experimental data in PAI is largely attributed to the domain gap between these two distributions. In this thesis, this problem is addressed through the use of unsupervised domain transfer, aiming to align the distributions of simulated and experimental data. Domain adaptation techniques have shown notable success in bridging such distributional gaps, particularly in the field of natural image processing [17, 18]. Consistent with this, the results presented here demonstrate that, following domain transfer, simulated PA images more closely resemble experimental data in terms of visual characteristics and spectral behaviour.
However, whether this improved similarity corresponds to physically meaningful corrections or merely reflects superficial adjustments remains uncertain. Given the complexity of PA image formation, governed by nonlinear light propagation, absorption, thermal expansion, and acoustic wave propagation, it is unlikely that the model has captured the underlying physics. Instead, it may have learned distribution-level features and introduced generalized patterns, such as redistributing intensities across channels or adding background clutter, which improve statistical resemblance but may lack physiological or physical plausibility.
A domain transfer model's capability of automatically learning to transform samples from a source distribution to a target distribution without direct supervision is not restricted to the transfer from simulated to real images. Their general applicability extends well beyond this use case. In principle, such models can be trained to align any two data distributions that share underlying structural or statistical commonalities. This flexibility could be particularly relevant for addressing systemic biases in PAI, such as those introduced by variations in skin tone, sex, or population-specific features. For instance, it is well-documented that differences in melanin absorption can result in skin tone bias in PAI [107, 108]. The cINN-based method has already shown promising capabilities in addressing structured domain shifts, as demonstrated in P.XXV, where simulated differences in skin tone and blood volume in soft tissue between toy datasets were successfully mitigated. Similarly, imbalances in training datasets, for example, due to underrepresentation of certain demographic groups, can lead to shortcut learning and biased model behavior in deployment [109]. Furthermore, shifts in prevalence or data distribution at test time are common challenges in real-world clinical applications that could be addressed by such models [110].
The practical utility of such domain transfer models lies in their demonstrated ability to improve downstream performance. In preliminary results from the SPECTRE dataset, oximetry models trained on domain-transferred data (Sim2Real) produced outputs more consistent with those trained on experimental data. Additionally, in P.III, substantial performance gains were achieved in artery-vein and organ classification tasks. In this light, domain transfer may even be viewed as an advanced data augmentation strategy: If it

improves task performance, it is useful. Ultimately, if domain transfer methods can reliably boost the performance or reliability of clinical tasks and thereby contribute to better patient outcomes, they hold clear value. Moving forward, however, integrating stronger physical priors, beyond the image data and segmentation masks, into the domain transfer process may foster more interpretable and trustworthy models that better respect the underlying physics of PA image formation.

Recent generative model developments

The cINN- and GAN-based domain transfer approach introduced in *P.III: Unsupervised domain transfer* represents the first method to combine the exact maximum likelihood training of cINNs with the high image fidelity typically associated with GANs. At the time of publication, GANs were still on the verge of being considered to be the state-of-the-art image generation, while the release of DALL-E 2 in late 2022 marked a turning point in generative modeling towards diffusion models and transformer architectures [111, 112, 113]. While recent ChatGPT models or emerging foundation models are likely still too data-hungry to be applied to a low-data modality such as PAI [114], methods that are specifically developed for low data regimes [115] hold promise. However, a systematic comparison between the cINN-based method and these emerging alternatives is needed to establish their relative advantages in the context of domain transfer for spectral imaging. With the speed of developments, it remains only a matter of time before newer network architectures or training strategies can outperform the proposed approach.

Sepctral fidelity

One of the main advantages of PAI in comparison to other modalities like ultrasound imaging lies in its potential for functional imaging due to its multispectral capabilities. Naturally, all image processing methods in PAI need to be designed in a way that they accommodate this multispectral data. For domain transfer from simulated to experimental PA images, this means that tissue types might have different spectral signatures depending on the domain. For instance, as shown in P.III, a systematic spectral shift between simulated and experimental vessel data was observed. This is likely due to differences in skin tone modeling or bulk tissue properties, both of which contribute to spectral coloring through light-tissue interactions.
Ideally, a domain transfer method would detect these differences and infer meaningful and physically plausible transformations without direct supervision. As discussed before, asserting that a domain transfer model understands this process correctly is challenging. Therefore, specific attention has to be paid to this spectral fidelity, which is not common in the general field of computer vision working with natural images where most innovations originate from.
The cINN-based method demonstrated the ability to transfer spectral information across domains, learning domain-specific differences such as subtle spectral shifts of arteries and

veins between experimental and simulated data. Its successful application to pixel-level spectra in the hyperspectral imaging (HSI) domain also suggests the method's potential generalizability not only beyond PAI to other spectral imaging modalities but also to natural RGB images.

Addressing research question 2 (RQ2)

How can unsupervised domain transfer close the domain gap between simulated and experimental PA images while preserving spectral fidelity?

Until simulations reach a level of realism sufficient for direct generalization, there will remain a need to bridge the gap between domains. Domain transfer techniques provide a practical and effective means of doing so, enabling the use of abundant synthetic data to support model development. Overall, this thesis addresses RQ2 by demonstrating that unsupervised domain transfer models can improve the visual and statistical alignment between synthetic and experimental PA images without requiring paired training data. While these models do not explicitly learn the underlying physics, they effectively capture domain-specific differences and introduce corrections that enhance downstream task performance. The cINN-based method was able to adjust spectral characteristics in a manner that aligns better with experimental observations, suggesting a promising degree of spectral preservation. Although perfect spectral fidelity cannot be guaranteed, the results indicate that such methods can serve as a powerful tool to reduce the domain gap and improve model generalization in low-data settings.

5.3 Design of phantoms for validation platform

Phantom design

Designing anthropomorphic phantoms for the validation of quantitative PAI methods is a highly complex task, requiring careful consideration of both biological realism and practical feasibility [116]. In addition to the challenges already discussed in the context of computational tissue modeling, such as accurately capturing anatomical geometry and assigning tissue-specific optical and acoustic properties, physical phantom fabrication is considerably more demanding and intricate.
Fortunately, prior advances in phantom material development [117] and community-driven initiatives, such as IPASC's efforts to disseminate standardized practices through workshops and seminars, have enabled more consistent and reproducible fabrication protocols. Leveraging these developments, the phantoms introduced in this thesis were created within a use case-specific workflow: anatomical data were acquired through MRI of a human forearm, from which a 3D digital model was extracted and subsequently 3D printed to ensure anthropomorphic accuracy. The optical properties of the embedded tissue-mimicking regions were tailored by adjusting dye concentrations to reproduce the absorption spectra of oxygenated and deoxygenated hemoglobin. These phantoms incorporate five distinct levels of blood oxygen saturation sO_2 in both the background material and vessel-like inclusions. While the presented workflow was specifically developed for oximetry in forearm-based applications, broader adoption across the field will require further refinement due to several limitations. One notable challenge is the presence of small air bubbles, which can introduce acoustic artifacts and significantly degrade signal quality, particularly at greater imaging depths. These artifacts are primarily caused by air inclusions at material interfaces, such as between surrounding soft tissue and vessels or bones. Addressing this issue will require improvements in degassing procedures, such as more effective vacuuming techniques, as well as redesigning the phantom molds to minimize air entrapment during casting. Another potential area for improvement is the integration of contrast agents for MRI or CT into the phantom material. In the current workflow, PA images were manually annotated by matching visible vessel structures to the fabrication protocol. This process was both time-consuming and prone to human error. Embedding MRI- or CT-visible contrast agents would allow high-resolution structural imaging of the internal phantom architecture, thereby enabling more accurate and semi-automated annotation of corresponding PA data.
Overall, further advances in phantom fabrication are needed to enable the development of reproducible, scalable, and stable anthropomorphic phantoms. Data-driven methods require large quantities of high-quality data for both training and validation, and the corresponding acquired images must closely resemble the intended clinical use case, in this thesis, oximetry in human forearms. The phantoms used here were fabricated using 3D-printed molds, similar to the approach described by Grasso et al. (2023) [118], into which the phantom material was cast. A promising future direction lies in the direct 3D-printing of phantom materials, as demonstrated by Ragunathan et al. (2025) [119], which could enable

the fabrication of highly complex and anatomically detailed 3D structures within larger, heterogeneous phantoms.

Validating data-driven photoacoustic methods

In the broader computer vision community, segmentation and classification have been studied since at least the 1960s. However, the pace of innovation in machine learning has recently outstripped the development of equally robust validation frameworks. Only recently, in 2024, a large consortium proposed a general framework for selecting performance measures for classification and segmentation tasks [120]. These tasks, which involve assigning discrete labels to entire images, regions, or individual pixels, are relatively well-structured in terms of evaluation. In contrast, regression-based tasks in PAI, such as image reconstruction or oximetry, require continuous value prediction at the pixel level and are more difficult to validate due to the absence of reliable *in vivo* reference.
Efforts like IPASC's standardised image reconstruction project have attempted to standardize evaluation for reconstruction algorithms. However, findings suggest that optimal metrics vary considerably by use case [121]. Some applications may require high contrast, while others prioritize structural accuracy.
A recent approach by Gröhl et al. [122] and Breger et al. [123] introduced a dataset in which domain experts rated reconstructed images based on visual interpretability. These were acquired from optically characterized phantoms, offering a promising quantitative validation method that integrates both expert opinion and controlled reference conditions.
Quantitative tasks such as oximetry present even greater challenges. Accurate ground truth requires reference values for each pixel, which current *in vivo* imaging cannot provide. Tunable flow phantoms can serve as validation tools, but their limited structural diversity restricts generalizability. To address this, this thesis introduces SPECTRE, the first comprehensive validation framework for quantitative PAI, based on anatomically realistic, tissue-mimicking phantoms with reference annotations.
Although initial oximetry results from SPECTRE were inconclusive, they provide a valuable foundation. It remains to be determined whether performance limitations arise from model shortcomings or residual fabrication artifacts. Upon completion of annotations for all ten phantoms, the evaluation presented in *P.XII: Open-source validation data* will be repeated on a larger and more diverse dataset. The final release of SPECTRE is intended to serve as a standardized benchmark for the community to rigorously compare state-of-the-art quantitative PAI methods.
Ultimately, further research is required to bridge the gap between phantom-based validation and *in vivo* applicability. Faithfully modeling the structural, optical, and acoustic complexity of biological tissue remains a significant challenge. However, until accurate reference values can be obtained *in vivo*, well-characterized phantoms will remain indispensable for the rigorous validation of quantitative PAI methods.

Addressing research question 3 (RQ3)

How can anthropomorphic phantoms be designed to provide reliable validation for data-driven methods in PAI?

This thesis addresses RQ3 by presenting an initial, reproducible workflow for the design and fabrication of anatomically informed phantoms, tailored for the validation of quantitative PAI methods. Built on established phantom materials and guided by patient-specific imaging, the approach enables the generation of complex structures with defined sO_2 levels for voxel-level annotation. The resulting SPECTRE dataset offers a first-of-its-kind platform for benchmarking data-driven models for functional tissue property estimation. This thesis provides a foundational contribution toward systematic validation in PAI, offering a practical and extensible approach on which future work can build.

6 Summary

6.1 Summary of contributions

As in many areas of modern research, data availability increasingly drives methodological progress and PAI is no exception. While PAI holds great promise for functional and molecular imaging, the scarcity of annotated reference data remains a major barrier to developing and validating robust data-driven methods. The primary objective of this thesis is to address this central challenge in PAI. Given the inherent complexity involved, the overarching problem is subdivided into three interrelated key subchallenges (C.1–3). Addressing these subchallenges, this thesis provides three contributions, each building sequentially upon the preceding one:

Contribution 1: Simulation framework

The toolkit for *simulation and image processing for photonics and acoustics*, or short, SIMPA, is a modularly designed framework that allows efficient simulation of spectral images while being both user-friendly and highly customizable. SIMPA is the first open-source PA simulation pipeline enabling users to generate large-scale, state-of-the-art simulations with minimal setup effort. It is easily extendable, as shown by the straightforward addition of diffuse reflectance imaging simulations. Further possible extensions are learning-based modules for the synthesis of realistic tissue geometries and a surrogate model for the model-based photon transport simulator. Both were published in peer-reviewed journals (VI and VII) [91, 124]. Despite the comparatively small PAI and optical community, the SIMPA GitHub repository has more than 100 stars with users and contributors from around the world. The pypi package has been downloaded in 45 countries within 2024. Additionally, it has been used by multiple projects of the IPASC, proving that it has a significant impact on the community.

Contribution 2: Domain transfer method

The domain transfer method presented in this thesis is the first cINN- and GAN-based method, combining their advantages of exact maximum likelihood training and high image quality, respectively. It could be shown that the method can synthesize data that can help to improve the performance of data-driven methods for various tasks compared to model-based simulations. Even though it was originally designed for PA images, it could be easily adapted to reflectance imaging. This generalizability was also acknowledged by MICCAI 23 poster chairs who picked it as a poster highlight for having ”the potential to impact other applications of AI and computer vision”.

Contribution 3: Validation platform

The presented anthropomorphic phantoms, combined with the corresponding data from SPECTRE that will be open-sourced as soon as possible, represent the first multispectral validation platform using experimental data with optical and functional reference annotations in PAI. They have the potential to pave the way for rigorous performance assessments of data-driven methods. Clinical translation of PAI can only happen if both a clear use case and a corresponding scientifically sound validation are provided. This thesis makes a substantial step towards this goal.

6.2 Conclusion

The PAI research community is in desperate need of experimental annotated data for algorithm development and validation. To address this critical need, this thesis has introduced a comprehensive approach for data-driven PA image synthesis. Beginning with the development of the modular and efficient simulation framework SIMPA, it advanced towards enhancing data realism through development of a domain transfer method and concluded with establishing a broadly applicable validation strategy using anthropomorphic phantoms. Although initially tailored for PAI applications, the methods developed, including SIMPA, the cINN-based domain transfer, and fabricated validation phantoms, are versatile and can be generalized to other imaging modalities, such as HSI. Specifically, the cINN-based domain transfer approach demonstrated significant potential for unsupervised adaptation across diverse imaging domains beyond its original context.
All tools, code, and datasets presented in this thesis have been or will be released open-source to ensure broad accessibility and transparency. Notably, the wide adoption of SIMPA within the research community exemplifies the impact and utility of these contributions, facilitating external quality checks and reproducibility. While ongoing improvements in simulation fidelity, advancements in deep learning techniques, and increasingly complex phantom designs are necessary for further progress, the outcomes of this thesis constitute a substantial advancement in methodological development, thereby supporting the transition of PAI methods towards clinical application.
To further advance PAI, several factors contribute to a successful clinical translation. Improving accuracy and robustness of methods for quantification of tissue parameters, such as blood oxygen saturation, will stay important with data-driven approaches increasingly dominating research efforts. High-quality simulations will continue to be essential for developing these data-driven quantitative methods due to their unique ability to provide ground truth reference. Finally, more complex, stable, and tissue-mimicking phantoms must be developed to ensure reliable validation and benchmarking of methods intended for clinical PAI applications. Ultimately, adoption of PAI in clinical routine warrants demonstration of competitiveness compared to established medical imaging modalities and is thus highly reliant on methodological advancements as presented in this thesis. The path toward clinical validation of PAI will likely involve iterative and mutual advancements in data availability, computational modeling, machine learning, and phantom fabrication, each driving improvements and validation in the others, supported by a continued commitment to open science practices.

List of publications

Publications featured in this thesis

P.I Gröhl, J., Schellenberg, M., **Dreher, K. K.**, & Maier-Hein, L. (2021). Deep learning for biomedical photoacoustic imaging: A review. *Photoacoustics*, 22, 100241.
DOI: https://doi.org/10.1117/1.JBO.30.7.076006

P.II Gröhl, J., **Dreher, K. K.**, Schellenberg, M., Rix, T., Holzwarth, N., Vieten, P., ... & Maier-Hein, L. (2022). SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics. *Journal of biomedical optics*, 27(8), 083010-083010.
DOI: https://doi.org/10.1117/1.JBO.27.8.083010

P.III **Dreher, K. K.**, Ayala, L., Schellenberg, M., Hübner, M., Nölke, J. H., Adler, T. J., ... & Maier-Hein, L. (2023, October). Unsupervised domain transfer with conditional invertible neural networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 770-780). Cham: Springer Nature Switzerland.
DOI: https://doi.org/10.1007/978-3-031-43907-0_73

P.IV **Dreher, K. K.**, Groehl, J., Grace, F., Ayala, L. A., Nölke, J. H., Bender, C. J., ... & Bohndiek, S. E. (2025). Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging. *Journal of biomedical optics*, 30(7), 076006.
DOI: https://doi.org/10.1117/1.JBO.30.7.076006

P.V **Dreher, K. K.**, Nölke, J. H., Bender, C. J., Grace, F., Seitel, A. & Maier-Hein ,L., (2025). SPECTRE: Simulated photoacoustic and experimental data for comparing tissue realism and quantitative estimation.
In peparation

Peer-reviewed publications

I Gröhl, J., Schellenberg, M., **Dreher, K. K.**, & Maier-Hein, L. (2021). Deep learning for biomedical photoacoustic imaging: A review. *Photoacoustics*, 22,

100241.
DOI: https://doi.org/10.1016/j.pacs.2021.100241

II Holzwarth, N., Schellenberg, M., Gröhl, J., **Dreher, K. K.**, Nölke, J. H., Seitel, A., ... & Maier-Hein, L. (2021). Tattoo tomography: Freehand 3D photoacoustic image reconstruction with an optical pattern. *International Journal of Computer Assisted Radiology and Surgery*, 16(7), 1101-1110.
DOI: https://doi.org/10.1007/s11548-021-02399-w

III Gröhl, J., Hacker, L., Cox, B. T., **Dreher, K. K.**, Morscher, S., Rakotondrainibe, A., ... & International Photoacoustic Standardisation Consortium. (2022). The IPASC data format: A consensus data format for photoacoustic imaging. *Photoacoustics*, 26, 100339.
DOI: https://doi.org/10.1016/j.pacs.2022.100339

IV Schellenberg, M., **Dreher, K. K.**, Holzwarth, N., Isensee, F., Reinke, A., Schreck, N., ... & Gröhl, J. (2022). Semantic segmentation of multispectral photoacoustic images using deep learning. *Photoacoustics*, 26, 100341.
DOI: https://doi.org/10.1016/j.pacs.2022.100341

V Gröhl, J., **Dreher, K. K.**, Schellenberg, M., Rix, T., Holzwarth, N., Vieten, P., ... & Maier-Hein, L. (2022). SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics. *Journal of biomedical optics*, 27(8), 083010-083010.
DOI: https://doi.org/10.1117/1.JBO.27.8.083010

VI Schellenberg, M., Gröhl, J., **Dreher, K. K.**, Nölke, J. H., Holzwarth, N., Tizabi, M. D., ... & Maier-Hein, L. (2022). Photoacoustic image synthesis with generative adversarial networks. *Photoacoustics*, 28, 100402.
DOI: https://doi.org/10.1016/j.pacs.2022.100402

VII Rix, T., **Dreher, K. K.**, Nölke, J. H., Schellenberg, M., Tizabi, M. D., Seitel, A., & Maier-Hein, L. (2023). Efficient photoacoustic image synthesis with deep learning. *Sensors*, 23(16), 7085.
DOI: https://doi.org/10.3390/s23167085

VIII **Dreher, K. K.**, Ayala, L., Schellenberg, M., Hübner, M., Nölke, J. H., Adler, T. J., ... & Maier-Hein, L. (2023, October). Unsupervised domain transfer with conditional invertible neural networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 770-780). Cham: Springer Nature Switzerland.
DOI: https://doi.org/10.1007/978-3-031-43907-0_73

IX Nölke, J. H., Adler, T. J., Schellenberg, M., **Dreher, K. K.**, Holzwarth, N., Bender, C. J., ... & Maier-Hein, L. (2024). Photoacoustic quantification

of tissue oxygenation using conditional invertible neural networks. *IEEE Transactions on Medical Imaging*, 43(9), 3366-3376.
DOI: https://doi.org/10.1109/TMI.2024.3403417

X Knopp, M., Bender, C. J., Holzwarth, N., Li, Y., Kempf, J., Caranovic, M., ... & **Dreher, K. K.** (2025). Shortcut learning leads to sex bias in deep learning models for photoacoustic tomography. *International Journal of Computer Assisted Radiology and Surgery*, 1-9.
DOI: https://doi.org/10.1007/s11548-025-03370-9

XI Holzwarth, N., Rachel, Z., Nölke, J. H., Schellenberg, M., Bauer, L., Schreck, N., Bender, C.J., **Dreher K. K.**, ... & Held, T. (2025). Photoacoustic imaging for monitoring radiotherapy treatment response in head and neck tumors. *Scientific reports*, 15(1), 16344.
DOI: https://doi.org/10.1038/s41598-025-95137-0

XII **Dreher, K. K.**, Groehl, J., Grace, F., Ayala, L. A., Nölke, J. H., Bender, C. J., ... & Bohndiek, S. E. (2025). Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging. *Journal of biomedical optics*, 30(7), 076006.
DOI: https://doi.org/10.1117/1.JBO.30.7.076006

Non-peer-reviewed publications

XIII Gröhl, J., **Dreher, K. K.**, Schellenberg, M., Seitel, A., & Maier-Hein, L. (2021, March). SIMPA: an open source toolkit for simulation and processing of photoacoustic images. In *Photons Plus Ultrasound: Imaging and Sensing 2021* (Vol. 11642, p. 116423C). SPIE.
DOI: https://doi.org/10.1117/12.2578172

XIV Holzwarth, N., Schellenberg, M., Gröhl, J., **Dreher, K. K.**, Nölke, J. H., Biegger, P., ... & Maier-Hein, L. (2021, March). Tattoo tomography: an optical pattern approach for context-aware photoacoustics. In *Photons Plus Ultrasound: Imaging and Sensing 2021* (Vol. 11642, p. 1164217). SPIE.
DOI: https://doi.org/10.1117/12.2578177

XV Gröhl, J., Schellenberg, M., **Dreher, K. K.**, Holzwarth, N., Tizabi, M. D., Seitel, A., & Maier-Hein, L. (2021, March). Semantic segmentation of multispectral photoacoustic images using deep learning. In *Photons Plus Ultrasound: Imaging and Sensing 2021* (Vol. 11642, p. 116423F). SPIE.
DOI: https://doi.org/10.1117/12.2578135

XVI Schellenberg, M., Gröhl, J., **Dreher, K. K.**, Holzwarth, N., Tizabi, M. D., Seitel, A., & Maier-Hein, L. (2021, March). Generation of training data for quantitative photoacoustic imaging. In *Photons Plus Ultrasound: Imaging and Sensing 2021* (Vol. 11642, p. 116421J). SPIE.
DOI: https://doi.org/10.1117/12.2578180

XVII Rix, T., Hübner, M., **Dreher, K. K.**, Nölke, J. H., Ayala, L., Schellenberg, M., ... & Maier-Hein, L. (2022, March). Deep learning for spectral image synthesis. In *Multimodal Biomedical Imaging XVII* (p. PC119520I). SPIE.
DOI: https://doi.org/10.1117/12.2608622

XVIII Vieten, P., **Dreher, K. K.**, Holzwarth, N., Schellenberg, M., Nölke, J. H., Seitel, A., ... & Maier-Hein, L. (2022, March). Deep learning-based semantic segmentation of clinically relevant tissue structures leveraging multispectral photoacoustic images. In *Photons Plus Ultrasound: Imaging and Sensing 2022* (p. PC119600P). SPIE.
DOI: https://doi.org/10.1117/12.2608616

XIX Hübner, M., Ayala, L., Rees, M., Adler, T. J., **Dreher, K. K.**, Seidlitz, S., ... & Maier-Hein, L. (2023, March). How to assess the realism of synthetic spectral images. In *Molecular-Guided Surgery: Molecules, Devices, and Applications IX* (p. PC1236104). SPIE.
DOI: https://doi.org/10.1117/12.2648461

XX **Dreher, K. K.**, Groehl, J., Rix, T., Schellenberg, M., Holzwarth, N., Noelke, J. H., ... & Seitel, A. (2023, March). Photoacoustic image synthesis with SIMPA: updates and future directions. In *Photons Plus Ultrasound: Imaging and Sensing 2023* (p. PC1237931). SPIE.
DOI: https://doi.org/10.1117/12.2649935

XXI Holzwarth, N., Schellenberg, M., Gröhl, J., **Dreher, K. K.**, Nölke, J. H., Seitel, A., ... & Maier-Hein, L. (2023, June). Tattoo-Tomographie: Freihand-3D-Photoakustik und multimodale Bildfusion. In *BVM Workshop* (pp. 114-114). Wiesbaden: Springer Fachmedien Wiesbaden.
DOI: https://doi.org/10.1007/978-3-658-41657-7_25

XXII Rix, T., **Dreher, K. K.**, Schellenberg, M., Nölke, J. H., Seitel, A., & Maier-Hein, L. (2024, March). Efficient photoacoustic image synthesis with Fourier neural operator networks. In *Photons Plus Ultrasound: Imaging and Sensing 2024* (p. PC128422U). SPIE.
DOI: https://doi.org/10.1117/12.3001800

XXIII Bender, C. J., **Dreher, K. K.**, Holzwarth, N., Knopp, M., Nölke, J. H., Rix, T., ... & Maier-Hein, L. (2024, March). Hardware-related biases in machine learning

algorithms for photoacoustic image analysis. In *Photons Plus Ultrasound: Imaging and Sensing 2024* (p. PC128422T). SPIE.
DOI: https://doi.org/10.1117/12.3000348

XXIV White, K. L., Hacker, L., Else, T., **Dreher, K. K.**, Watt, M., Bohndiek, S., & Williams, C. (2025, March). Snapshot multispectral cameras targeted for surgical vision (Conference Presentation). In *Advanced Biomedical and Clinical Diagnostic and Surgical Guidance Systems XXIII* (p. PC1330606). SPIE.
DOI: https://doi.org/10.1117/12.3043900

XXV **Dreher, K. K.**, Schellenberg, M., Nölke, J. H., Rix, T., Seitel, A., & Maier-Hein, L. (2024, March). Unsupervised domain transfer for realistic photoacoustic image simulation. In *Photons Plus Ultrasound: Imaging and Sensing 2024* (p. PC128420W). SPIE.
DOI: https://doi.org/10.1117/12.3003392

XXVI **Dreher, K. K.**, Nölke, J. H., Bender, C. J., Grace, F., Seitel, A. & Maier-Hein ,L., (2025). SPECTRE: Simulated photoacoustic and experimental data for comparing tissue realism and quantitative estimation.
In peparation

Patents

Pt.I Holzwarth, N., **Dreher, K. K.**, Schellenberg, M., Nölke, J.H., Gröhl, J. and Maier-Hein, L., Deutsches Krebsforschungszentrum DKFZ, 2023. *Method And System For Context-Aware Photoacoustic Imaging*. U.S. Patent Application 18/004,689.

Honors and awards

H.I **Invited talk**: Department of medical physics, Martin-Luther-University, Halle, Germany: ”Computational modeling of photoacoustic imaging with SIMPA” (2024).

H.II **Invited talk**: BZKF Summer School, Thurnau, Germany: ”Decoding light: Applications of machine learning in advanced biomedical imaging” (2024).

H.III **Poster prize** runner up at the Biophotonics and Imaging Graduate Summer School for the poster entitled ”Deep learning-enabled realistic simulation of spectral images” (2024).

H.IV **Travel award** for the IPASC ”Knowledge Exchange Partnerships in Photoacoustic Imaging” to the University of Cambridge, Cambridge, UK (2023).

H.V **Travel grant** for short-term research visit of the Helmholtz International Graduate School for Cancer Research at the German Cancer Research Center (DKFZ) (2023)

H.VI **Best Ph.D. Talk Award** at the German Cancer Research Center Ph.D. Retreat held at Evangelische Tagungsstätte Löwenstein from 17th-19th July 2023 for the talk entitled "Unsupervised sim2real transfer of spectral images with conditional invertible neural networks" (2023).

H.VII **Three-year Ph.D. Stipend** of the Helmholtz International Graduate School for Cancer Research at the German Cancer Research Center (DKFZ) (2020).

7 Supplementary material

7.1 Open-source simulation toolkit

Supplemental material for: SIMPA: An Open-source Toolkit for Simulation and Image Processing for Photonics and Acoustics

Janek Gröhl[1,†, now at 6,7], Kris K. Dreher[1,2,*,†], Melanie Schellenberg[1,3,4], Tom Rix[1,3], Niklas Holzwarth[1], Patricia Vieten[1,2], Leonardo Ayala[1,5], Sarah E. Bohndiek[6,7], Alexander Seitel[1,‡], and Lena Maier-Hein[1,3,5,‡]

[1]German Cancer Research Center, Computer Assisted Medical Interventions, Heidelberg, Germany
[2]Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany
[3]Heidelberg University, Faculty of Mathematics and Computer Science, Heidelberg, Germany
[4]HIDSS4Health - Helmholtz Information and Data Science School for Health, Heidelberg, Germany
[5]Heidelberg University, Medical Faculty, Heidelberg, Germany
[6]Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way, Cambridge, CB2 0RE, U.K.
[7]Department of Physics, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, U.K.

ABSTRACT

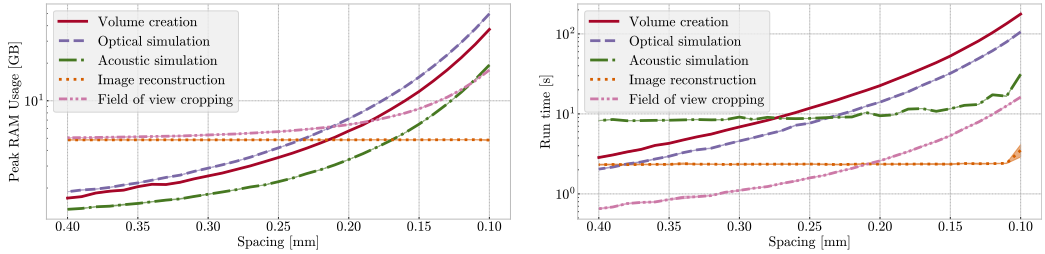
This document contains materials and plots supplementing the paper: SIMPA: An open-source toolkit for simulation and image processing for Photonics and Acoustics. The experiments conducted in this paper do not require any external data. The latest release of the SIMPA code can be downloaded from GitHub (`https://github.com/CAMI-DKFZ/simpa`, last visited 14$^{\text{th}}$ December 2021). The code used to generate the results and figures is available in a GitHub repository (`https://github.com/CAMI-DKFZ/simpa_paper_experiments`) and via zenodo (made available upon publication).

*Send correspondence to K.K.D. k.dreher@dkfz-heidelberg.de
†These authors contributed equally
‡Shared last authorship

1. SIMPA MEMORY AND TIME SCALING

One major use case of SIMPA is to simulate large training datasets for machine learning algorithms. As such, the computational costs for the SIMPA simulations is of special interest. All experiments in this section were conducted using a workstation with an *AMD(R) Ryzen 3900x 12-core* central processing unit (CPU), 64 GB of RAM, and *NVIDIA RTX 3090* GPU running Ubuntu 20.04.

Generally, a higher resolution in medical imaging is advantageous for a physician to be able to differentiate between small structures maybe even in the submillimeter regime. For simulation of PA images, a finer computational grid can simulate higher frequencies of the acoustic waves*. Unfortunately, with a higher resolution, the computational costs can be extreme, such that it is not possible even for advanced scientific computers to simulate PA images arbitrarily highly resolved. The resolution of a SIMPA simulation is given by the isotropic voxel size or *spacing*. A simulation pipeline with the same parameters as the default simulation in section 3.2 of the main paper has been executed ten times each with decreasing spacing from 0.4 mm to 0.1 mm and the RAM usage, as well as the run time, for each pipeline element has been recorded. These simulations indicate what the SIMPA user can expect of the scaling of *Random Access Memory* (RAM) usage, as well as the run time of an example simulation pipeline of SIMPA with decreasing spacing (Figure S1). It is important to note that the absolute numbers of the run times of the simulation runs can vary vastly from device to device, however, the scaling of both RAM usage and run time with decreasing spacing should be consistent.



(a) Scaling of peak RAM usage with spacing.

(b) Scaling of simulation run time with spacing.

Figure S1: Analysis of the random access memory (RAM) and run time requirements of SIMPA. S1a shows the scaling of the peak RAM usage and S1b shows the scaling of the simulation run time with decreasing spacing from 0.4 mm to 0.1 mm. Both the peak RAM consumption in GigaBytes (GB) (S1a) and the run time scaling in seconds (s) (S1b) for each of the pipeline elements are shown on a logarithmic scale. The pipeline elements are colour-coded as follows: volume creation (red), optical modelling (purple), acoustic modelling (green), image reconstruction (orange), and field of view cropping (pink).

It should be noted that the **first** initialisation of any tensor on the GPU using PyTorch takes a lot longer (around two seconds) than the actual reconstruction algorithm. Because of this, the image reconstruction time including tensor initialisation is constantly around two seconds. As the actual reconstruction algorithm only takes about 20 ms, a reconstruction frame rate of 50 Hz can be achieved if the user initialises the PyTorch framework prior to executing the reconstruction algorithm.

*`http://www.k-wave.org/manual/k-wave_user_manual_1.1.pdf`, last visited 14$^{\text{th}}$ December 2021

2. CUSTOMISING SIMULATION PARAMETERS - RUN TIMES

In section 3.2 of the main paper, a simulation pipeline has been customised with various hyperparameters with the aim of reaching the most similar reconstruction result compared to the underlying initial pressure. In Tables S1, S2, S3, and S4 the run times for different parameter combinations can be found. The reconstructions in the different pipelines were performed with the default settings (delay-and-sum), with an applied bandpass filter, with a "differential mode" (delay-and-sum of the first derivative of the time signal), and finally, a customised set of hyperparameters. These times are reported for the spacings 0.15 mm, 0.35 mm, 0.55 mm.

Spacing [mm]	Optical modelling time [s]	Acoustic modelling time [s]	Image reconstruction time [s]
0.55	1.16	7.92	2.16
0.35	2.77	8.47	2.20
0.15	27.89	11.64	2.20

Table S1: Run times of the optical and acoustic forward modules and image reconstruction for the default simulation pipeline in seconds (s). The times are reported for three different spacings: 0.15 mm, 0.35 mm, 0.55 mm.

Spacing [mm]	Optical modelling time [s]	Acoustic modelling time [s]	Image reconstruction time [s]
0.55	1.17	7.91	2.28
0.35	2.77	8.15	2.20
0.15	28.05	11.70	2.25

Table S2: Run times of the optical and acoustic forward modules and image reconstruction in seconds (s) for a simulation pipeline with an applied bandpass filter (Tukey window with an alpha value of 0.5 and 1 kHz as high-pass and 8 MHz as low-pass frequencies). The times are reported for three different spacings: 0.15 mm, 0.35 mm, 0.55 mm.

Spacing [mm]	Optical modelling time [s]	Acoustic modelling time [s]	Image reconstruction time [s]
0.55	1.19	7.99	2.18
0.35	2.80	8.24	2.21
0.15	27.87	11.63	2.21

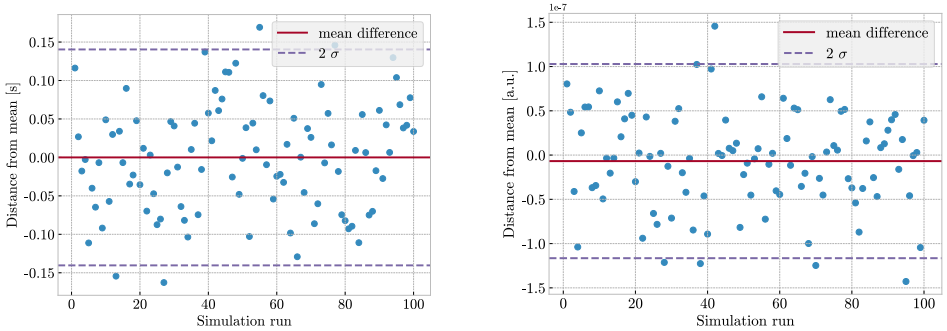
Table S3: Run times of the optical and acoustic forward modules and image reconstruction in seconds (s) for a simulation pipeline with delay-and-sum reconstruction with the first derivative of the time-series data . The times are reported for three different spacings: 0.15 mm, 0.35 mm, 0.55 mm.

Spacing [mm]	Optical modelling time [s]	Acoustic modelling time [s]	Image reconstruction time [s]
0.55	1.15	7.89	2.25
0.35	2.86	8.27	2.24
0.15	27.27	11.51	2.23

Table S4: Run times of the optical and acoustic forward modules and image reconstruction in seconds (s) for a simulation pipeline with delay-and-sum reconstruction with a bandpass filter (Tukey window with an alpha value of 0.5 and 1 kHz as high-pass and 8 MHz as low-pass frequencies), the first derivative of the time-series data and envelope detection. The times are reported for three different spacings: 0.15 mm, 0.35 mm, 0.55 mm.

3. INVESTIGATION OF ADVERSE PROGRAMMING EFFECTS

To simulate a large data set consisting of multiple photoacoustic images such as in section 3.6 of the main paper, the simulation of each image should run as fast as possible without influencing the simulation of a succeeding image of the dataset. This is relevant because the existence of dependencies within sequential runs would potentially lead to unusable or heavily biased data. The risk that this happens is high, as such dependencies could be introduced by minor software bugs (e.g. naming conventions are not kept, file access is not handled properly, instance variables are not reset properly, etc). It is therefore important to examine the behaviour of the toolkit in subsequent executions.



(a) Simulation run time differences.

(b) Simulation result differences.

Figure S2: Analysis of the independence of subsequently executed simulations. Simulation run times of 100 sequential simulation executions are shown, where the pipeline was newly initialised every run. S2a shows the difference of the run times for each simulation in seconds (s). S2b shows the absolute differences between the simulation results, in this case the reconstructed images, where each image has been normalised between 0 and 1 in arbitrary units (a.u.). The dashed purple lines denote the $2\,\sigma$ interval around the mean and the red line denotes the mean difference.

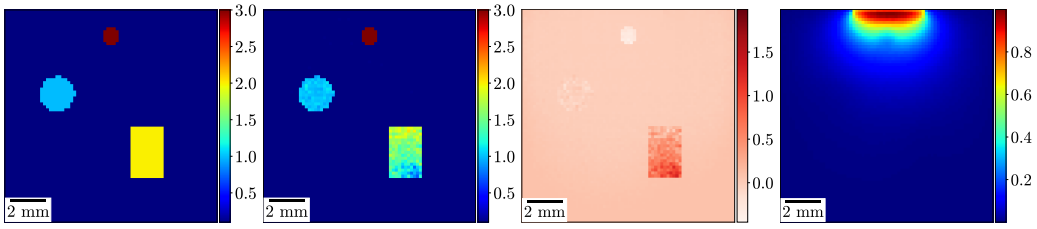
To investigate a possible existence of adverse effects from subsequent simulations runs, 100 simulations with the same parameters are sequentially executed (with a spacing of 1 mm). The simulation run times are highly comparable (Figure S2); the simulation of the largest outlier took 0.17 s longer which is about 2% of the mean simulation time ($\approx$ 7.95 s). As for the resulting simulated PA images, the difference between the mean image and largest outlier is about 1.5e-7 a.u. which is negligible since the images are normalized between 0 and 1 a.u.

4. PHOTOACOUSTIC IMAGE PROCESSING

Currently, there are two widely used image processing algorithms implemented in SIMPA and are demonstrated below.

4.1 Iterative quantitative photoacoustic imaging

The qPAI method introduced by Cox et al. 2006 uses a simple iterative scheme to recover optical absorption coefficients of a known initial pressure distribution by repeatedly simulating the optical fluence based on previous assumptions of the absorption coefficients and updating them accordingly until the algorithm converges. In comparison to the 2D method presented in the original paper, in SIMPA, a slightly advanced 3D method was implemented with MCX as the optical forward model. To reproduce the results obtained by Cox et al., one of the example volumes was recreated and the optical model was simulated. The initial pressure and the optical scattering were then used as input for the iterative qPAI algorithm. The ground truth and estimated absorption coefficient distributions, a difference image and the estimated fluence are shown in figure S3 where the same colour map has been chosen to make a comparison to the original paper easier. The results are surprisingly similar to the original paper even though we use a different optical forward model.



(a) Ground truth absorption [cm^{-1}]

(b) Recovered absorption [cm^{-1}]

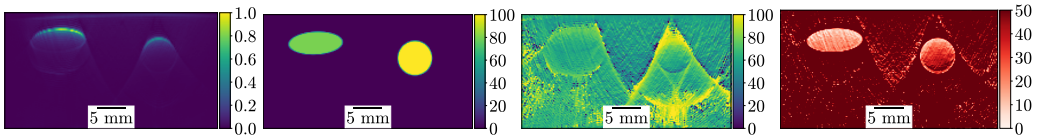
(c) Difference of ground truth and recovered absorption [cm^{-1}]

(d) Recovered fluence [a.u.]

Figure S3: Recreation of the Cox et al. 2006 quantitative PA imaging experiment within SIMPA. From left to right, the panels show S3a the ground truth absorption coefficients in cm^{-1}; S3b the absorption coefficients recovered by the algorithm in cm^{-1}; S3c the absolute difference of S3a and S3b; and S3d the recovered light fluence of the last iteration step normalised between 0 and 1 in arbitrary units (a.u.). It has to be noted that while the original study was conducted using a 2D photon propagation model, a 3D model (MCX) was used in this case.

4.2 Linear unmixing

"Spectral unmixing is the decomposition of a mixed pixel into a collection of distinct spectra, or endmembers, and a set of fractional abundances that indicate the proportion of each endmember." (Keshava et al., 2002) Linear spectral unmixing (the word *spectral* is often omitted) assumes that the pixel spectrum is a linear combination of the spectra of the endmembers. A version of linear spectral unmixing using singular value decomposition to reduce computational cost is implemented in SIMPA. To show the capabilities of this processing component, an example simulation was performed including linear unmixing using the wavelengths 750 nm, 800 nm, and 850 nm to provide an estimation of blood oxygen saturation (Figure S4). Notably, the errors of the estimated blood oxygen saturation are comparably low within the two ellipses especially in the upper area of the ellipses. The error is higher in the background because the signal of the highly absorbing structures dominates the spectrum of the reconstructed image.



(a) Reconstructed image at 750 nm [a.u.]

(b) Ground truth blood oxygen saturation at 750 nm [%]

(c) Estimated blood oxygen saturation [%]

(d) Absolute error [%]

Figure S4: An example application of linear unmixing on the reconstructed images within SIMPA. From left to right, the panels show: S4a the normalised reconstructed PA image at a wavelength of 750 nm which serves as input for the algorithm; S4b the ground truth blood oxygenation, where the background has 0%, the left structure 50%, and the right structure 100% oxygen saturation; S4c the estimated blood oxygenation after linear unmixing using 750 nm, 800 nm and 850 nm; and S4d the absolute error of the estimation when comparing S4b and S4c.

7.2 Unsupervised domain transfer

1

Supplementary Material for: Unsupervised Domain Transfer with Conditional Invertible Neural Networks

Table S1: **Simulated ranges of physiological parameters of a three-layer tissue model for HSI**. $v_{\mathrm{Hb}}[\%]$ represents the blood volume fraction, sO_2 the blood oxygenation, a_{mie} the reduced scattering coefficient at 500 nm, b_{mie} the scattering power, g the scattering anisotropy, n the refractive index, and d the layer thickness. Parameters were uniformly sampled within the specified range.

	$v_{\mathrm{Hb}}[\%]$	$sO_2[\%]$	$a_{\mathrm{mie}}[\mathrm{cm}^{-1}]$	b_{mie}[a.u.]	g[a.u.]	n[a.u.]	d[cm]
layers 1 to 3:	$0-30$	$0-100$	$5-50$	$0.3-3$	$0.80-0.95$	$1.33-1.54$	$0.002-0.2$
simulation framework: GPU-MCML, 10^6 photons per simulation							
simulated samples: 5.5×10^5 in wavelength range: 500 nm-1000 nm, step size 2 nm							

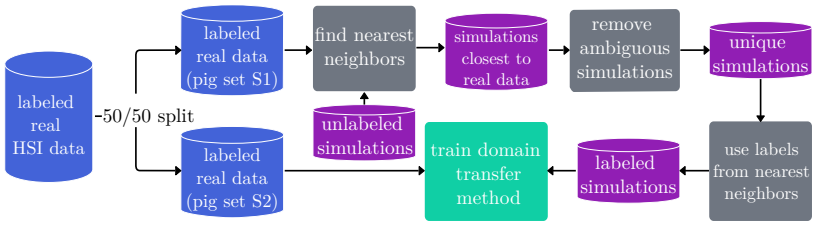


Fig. S1: **HSI data generation**. Real data is leveraged to reduce a large set of simulations to a set of plausible spectra with unambiguous labels.

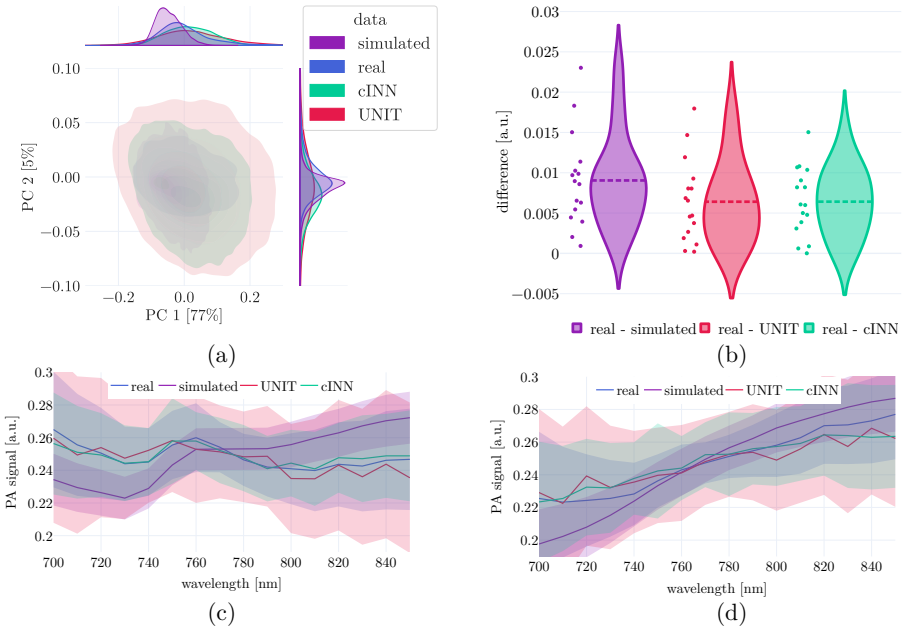


Fig. S2: **Our domain transfer approach yields realistic spectra**. Figures a) and b) are the artery-equivalent of veins in Fig. 5 of the main paper. c) and d) represent the real, simulated, and domain-transferred vein and artery spectra.

Table S2: **Hyperparameters of cINN models and discriminators (Dis) for PAT and HSI.** The hyperparameters have been optimized to yield best spectral consistency and classification performance. The networks have been implemented in PyTorch. Here, "lr" represents the learning rate and "WD" the weight decay. The code and pre-trained models are available at: GITHUB_LINK.

Hyperparameter	**PAT cINN**	**HSI cINN**
Epochs	300	300
Batch size	2	10,000
Optimizer	Adam (lr=0.001)	Adam (lr=0.0001)
Optimizer parameters	β_1=0.4,β_2=0.999,WD=0.001	β_1, β_2, WD=0.9, 0.95, 0.0001
Scales	5	1
Blocks per scale	4, 2, 1, 1, 2	40
Conditions per scale	DY, D, D, D, D	30×DY + 10×None
Exponential clamping	1	1
Subnetwork	3×(2D Conv + ReLU) 256 hidden features	2×(Linear + ReLU) 512 hidden features
Hyperparameter	**PAT Dis**	**HSI Dis**
Scales	3	1
Layers per scale	4×(2D Conv + LeakyReLU)	3×(Linear + LeakyReLU)
Hidden features	64	256
Dropout prob.	0.2	0.2

7.3 Anthropomorphic phantoms

Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging: Supplementary material

Kris K. Dreher[a,b,*], Janek Gröhl[c,d], Friso Grace[a,e], Leonardo Ayala[a,f], Jan-Hinrich Nölke[a,g], Christoph J. Bender[a,g], Melissa J. Watt[c,d], Catherine-Louise White[c,d], Ran Tao[c,d], Wibke Johnen[h,i], Minu D. Tizabi[a], Alexander Seitel[a], Lena Maier-Hein[a,f,g,j,†,*], Sarah E. Bohndiek[c,d,†,*]

[a]German Cancer Research Center (DKFZ), Division of Intelligent Medical Systems (IMSY), Heidelberg, Germany
[b]Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany
[c]University of Cambridge, CRUK Cambridge Institute, Cambridge, United Kingdom
[d]University of Cambridge, Department of Physics, Cambridge, United Kingdom
[e]University of St Andrews, School of Physics and Astronomy, St Andrews, United Kingdom
[f]NCT Heidelberg, a partnership between DKFZ and University Hospital Heidelberg, National Center for Tumor Diseases (NCT), Heidelberg, Germany
[g]Heidelberg University, Faculty of Mathematics and Computer Science, Heidelberg, Germany
[h]DKFZ, Division of Medical Physics in Radiation Oncology, Heidelberg, Germany
[i]Heidelberg Institute for Radiation Oncology (HIRO), National Center for Radiation Research in Oncology (NCRO), Heidelberg, Germany
[j]Heidelberg University, Medical Faculty, Heidelberg, Germany

*Corresponding authors: KKD, LMH {k.dreher,l.maier-hein}@dkfz-heidelberg.de, SEB seb53@cam.ac.uk
†Shared last authorship

S1 Supplementary Notes

The phantom manufacturing protocol used for quality assurance during phantom production is embedded below.

Measurement Protocol for Multispectral Photoacoustic Inclusion Materials

Date: _____

Operator: _____

Experiment ID: _____

Colour: _____

Phantom preparation protocol
Please check the boxes when a step has been finished!

☐ **Step 1: Prepare 30 ml stock solution with 2 mg/ml**

Check?	Instruction	Comment	
☐	Weigh 30 ml (25.14 g) of mineral oil in a skirted tube	$M_{measured}$:	g
☐	Weigh 60 mg of the dye powder	$M_{measured}$:	mg
☐	Mix the mineral oil and the dye in the vortexer (30 s)	$T_{vortexed}$:	s
☐	Sonicate in the water bath sonicator for 10 s and then shake the tube by hand. (repeat ~10 times)	$N_{repetitions}$:	
☐	Crush the little bits and clumps of the dye with a spatula.		
☐	Repeat the previous 3 steps until there are no clumps left that are bigger than a granule of sugar.		
☐	Label the tube with the dye ID and the date of production		

☐ **Step 2: Prepare phantoms according to multicentre study**

The following steps are meant to briefly summarise and amend the IPASC multicenter phantom material verification study 2022 Protocol. This means that the steps are adjusted to a 50 ml volume of mineral oil (conversion ratio of 0.5 from 100 ml to 50 ml). In doubt, please refer to the full study protocol!

Check?	Instruction	Comment	
☐	Prepare a silicone oil bath, set the temperature to 160°C		
☐	Weigh 50 ml (41.90 g) of mineral oil in a beaker (beaker 1)	$M_{measured}$:	g

☐	Vortex the stock solution for 1 min	$T_{vortexed}$:	min
☐	Pipette 12.8 ml of the stock solution	Stock solution:	
☐	Add the stock solution to beaker 1		
☐	Weigh 76.5 mg of TiO2 and add to beaker 1	$M_{measured}$:	mg
☐	Sonicate and sway beaker 1 until no TiO2 is settled at the bottom. This will take roughly 20 - 30 min. In the meantime, the following 3 steps can be performed.	$T_{sonicated}$:	min
☐	Weigh 1.0 g of Butylated Hydroxytoluene	$M_{measured}$:	g
☐	Weigh 12.57 g of SEBS	$M_{measured}$:	g
☐	Add SEBS and Butylated Hydroxytoluene with a magnet stirrer in a new beaker (beaker 2)		
☐	When there is no settled TiO2 left in beaker 1, pour the content of beaker 1 into beaker 2		
☐	Heat beaker 2 in the oil bath for about 45 min with aluminium foil on top	$T_{in\ oil}$:	min
☐	After 10 min, take off the aluminium cover and stir the mixture with a metallic spoon/spatula.		
☐	When the mixture has become liquid, put in vacuum chamber for at least 1 min	$T_{in\ vacuum}$:	min
☐	Scratch residual bubbles from surface and repeat the previous step if necessary		

☐ **Step 3: Casting the phantom material into the mold for optical samples**

This step is for fabricating the optical samples used to measure absorption and scattering of the material. For this, take four microscopy glass slides, put a metallic frame on two of the slides and pour the material in it. Try to be quick when covering the material and the metallic frame so that the material is still fluid and can be compressed into shape.

S2 Supplementary Figures

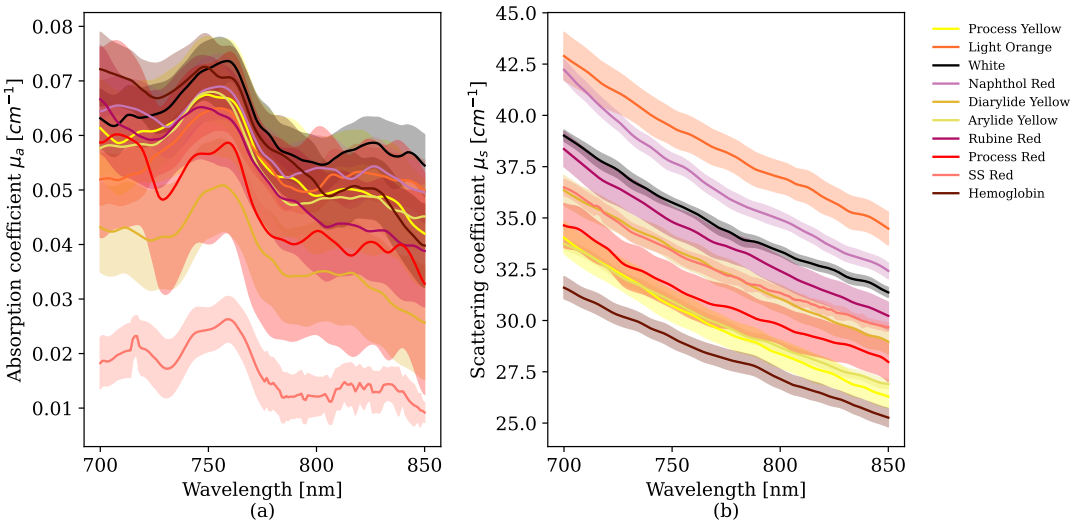


Fig S1: **Optical properties of investigated dyes with low absorption.** In total, 26 dyes were examined for their optical absorption (a) and scattering (b) in a copolymer-in-oil solution using a double-integrating sphere setup. 10 dyes were identified with relatively low absorption coefficient: Process Yellow, Light Orange, White, Naphthol Red, Diarylide Yellow, Rubine Red, Process Red, SS Red, Hemoglobin. For manufacturers and product names, refer to Table S1. Bands around the measured spectra indicate the standard deviation across the 12 measurement points for each optical sample slab.

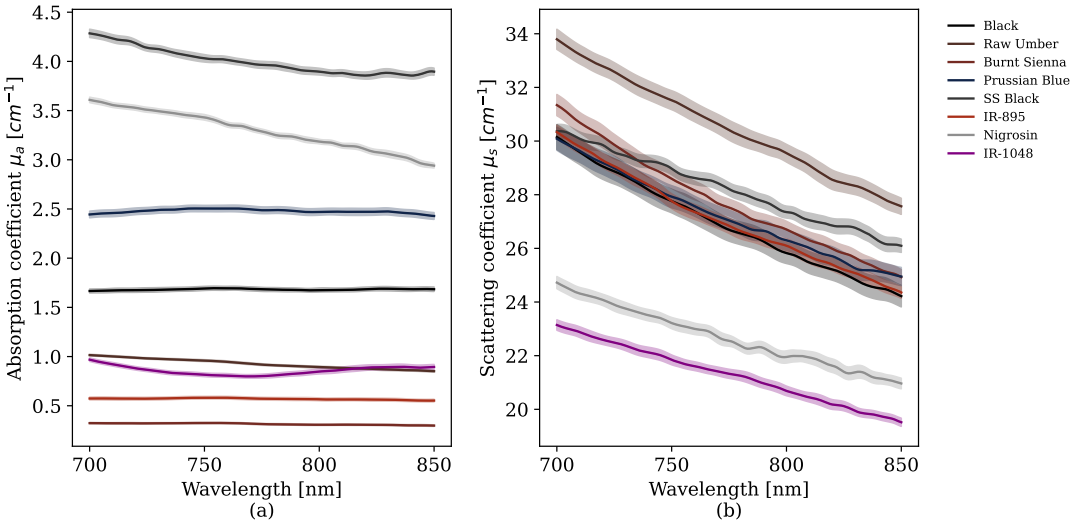


Fig S2: **Optical properties of investigated dyes with high absorption.** In total, 26 dyes were examined for their optical absorption (a) and scattering (b) in a copolymer-in-oil solution using a double-integrating sphere setup. 8 dyes were identified with relatively high absorption coefficient: Black, Raw Umber, Burnt Sienna, Prussian Blue, SS Black, IR-895, Nigrosin, IR-1048. For manufacturers and product names, refer to Table S1. Bands around the measured spectra indicate the standard deviation across the 12 measurement points for each optical sample slab.

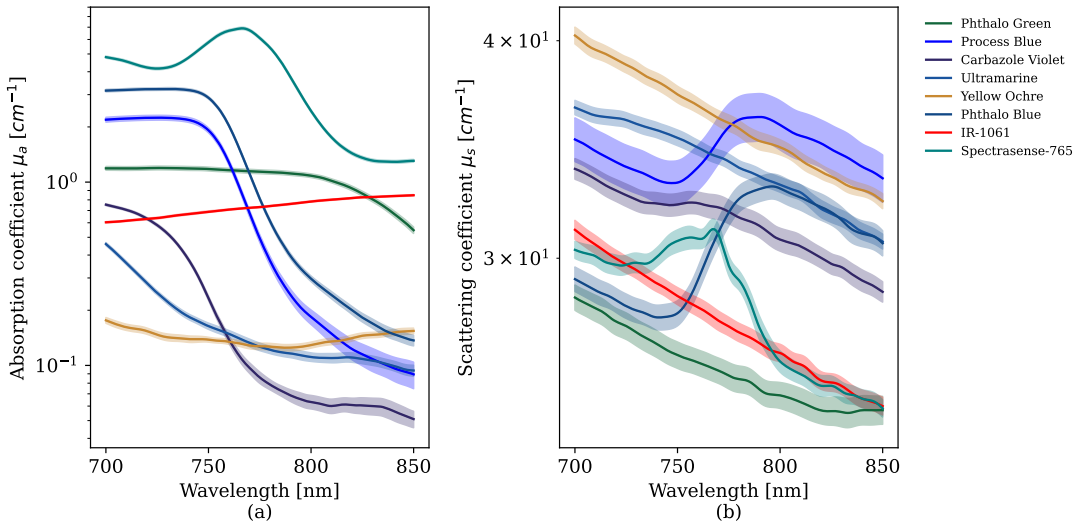


Fig S3: **Optical properties of investigated dyes with high range.** In total, 26 dyes were examined for their optical absorption (a) and scattering (b) in a copolymer-in-oil solution using a double-integrating sphere setup. 8 dyes were identified with a relatively high range of absorption coefficient: Phthalo Green, Process Blue, Carbazole Violet, Ultramarine, Yellow Ochre, Phthalo Blue, IR-1061, Spectrasense-765. For manufacturers and product names, refer to Table S1. Bands around the measured spectra indicate the standard deviation across the 12 measurement points for each optical sample slab.

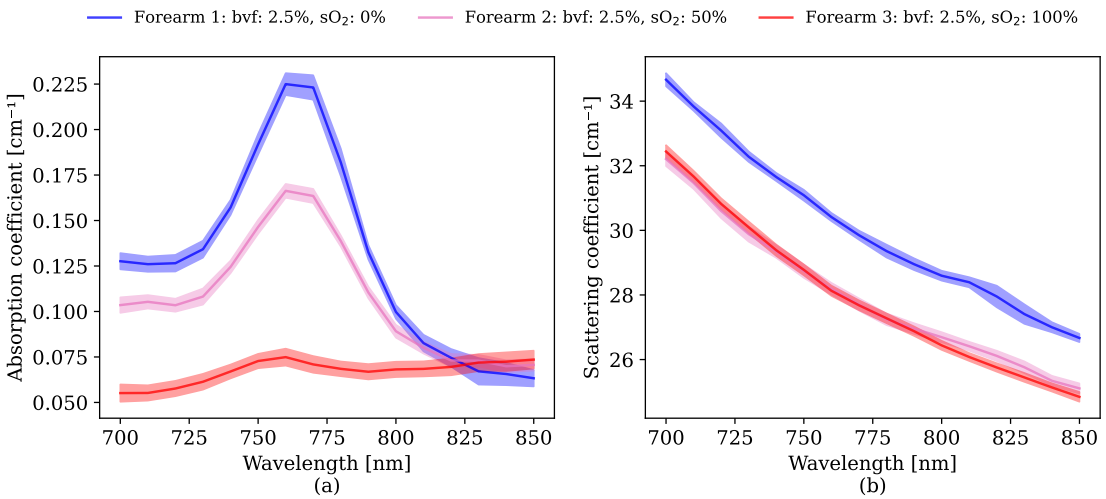


Fig S4: **Optical absorption (a) and scattering (a) coefficients for forearms 1-3.** For scattering anisotropy (g), a value of g=0.7 was assumed.

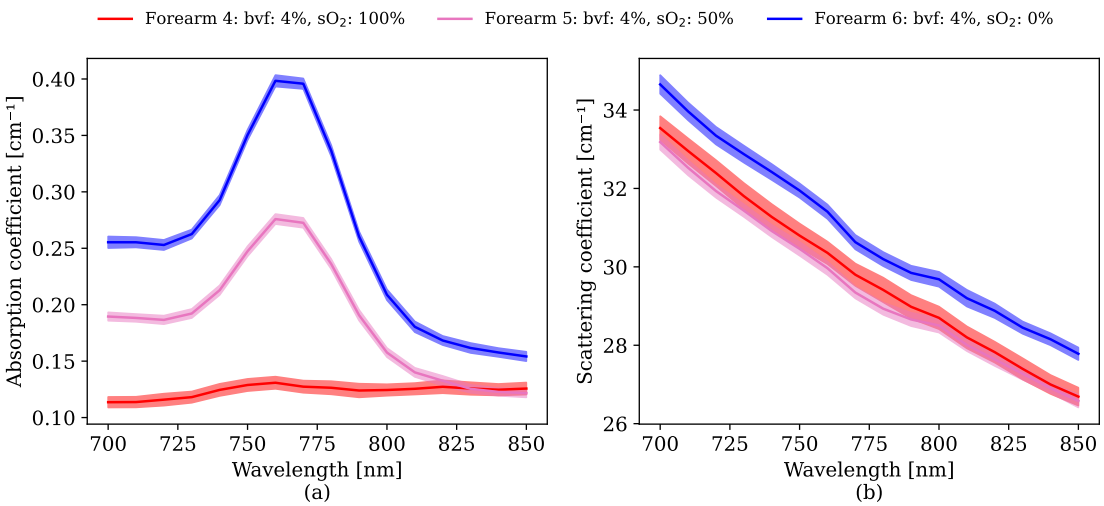


Fig S5: **Optical absorption (a) and scattering (a) coefficients for forearms 4-6.** For scattering anisotropy (g), a value of g=0.7 was assumed.

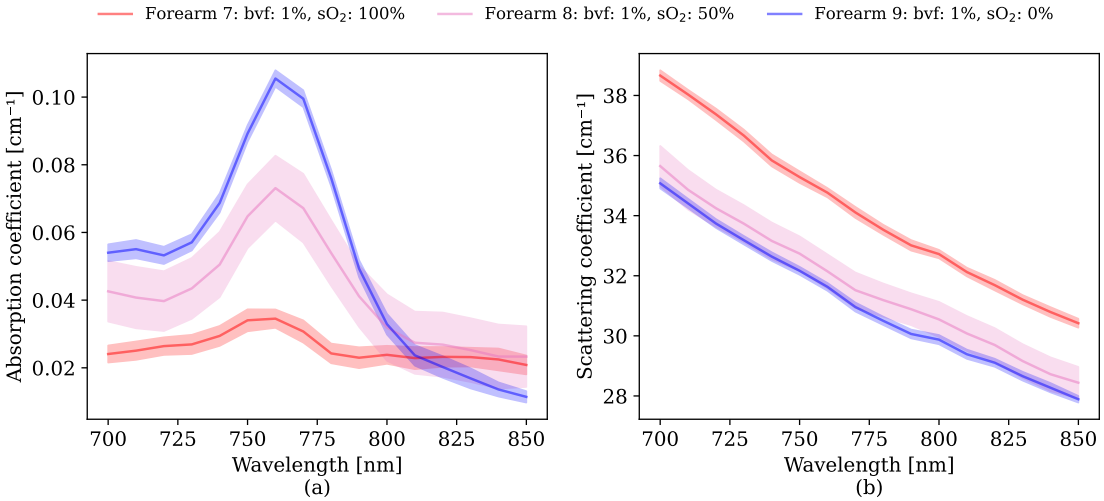


Fig S6: **Optical absorption (a) and scattering (a) coefficients for forearms 7-9.** For scattering anisotropy (g), a value of g=0.7 was assumed.

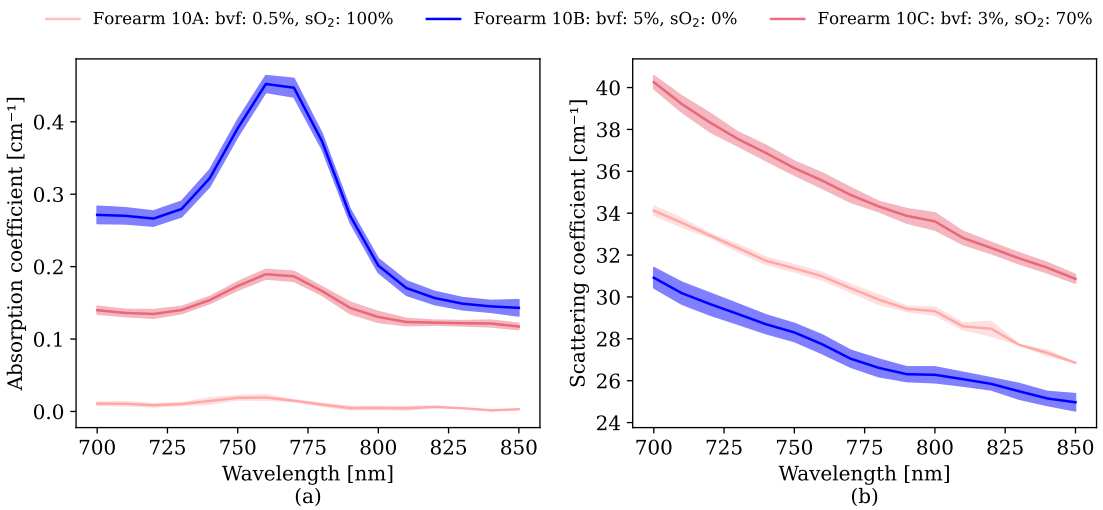
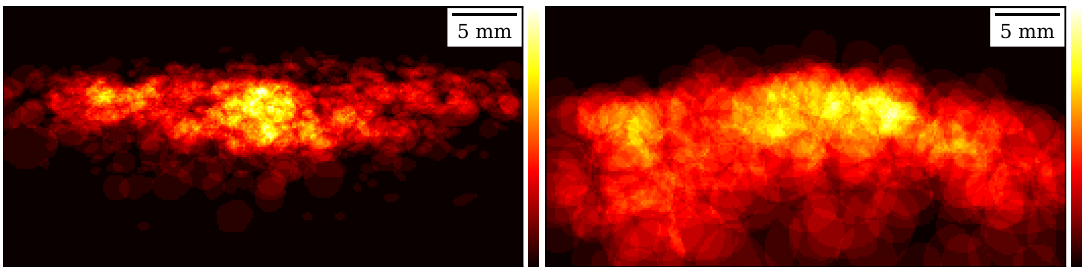


Fig S7: **Optical absorption (a) and scattering (b) coefficients for all three background compositions of forearm 10 (10A, 10B, and 10C).** For scattering anisotropy (g), a value of g=0.7 was assumed.



(a) Human forearm vessel locations (b) Phantom forearm vessel locations

Fig S8: **Heatmap comparison of vessel locations in an in-house human forearm dataset and our forearm phantoms.** The vessels in our phantoms cover the broad area that the human forearm vessels cover with similar hotspot locations. They also extend out into depth more to provide more versatility as we expect more vessels in depth but they were not annotated in the human forearms due to low signal.

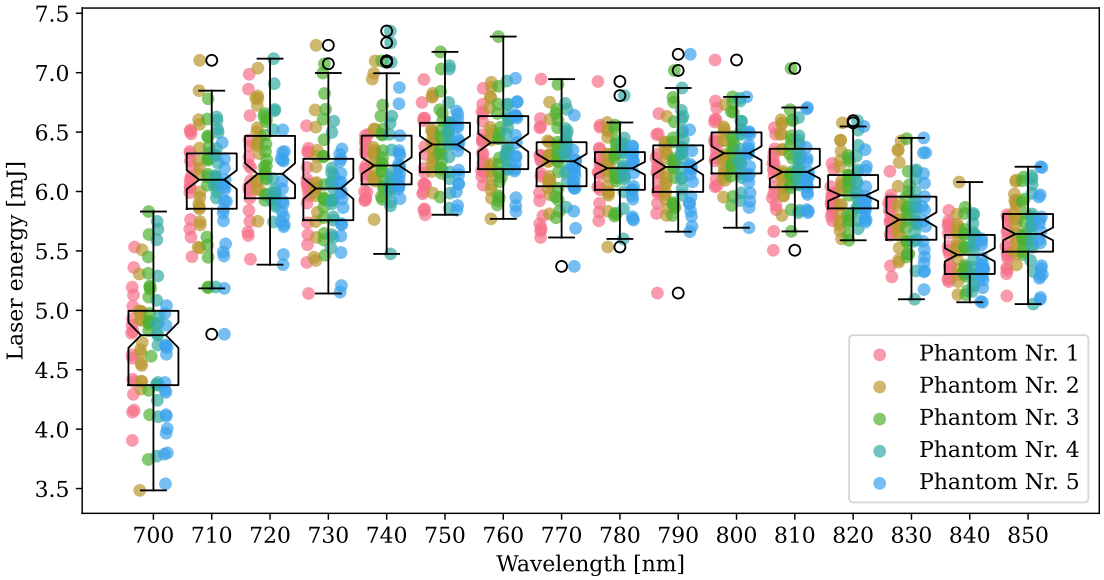


Fig S9: **Measured laser energies in mJ for each of the example phantoms' 27 photoacoustic images.**

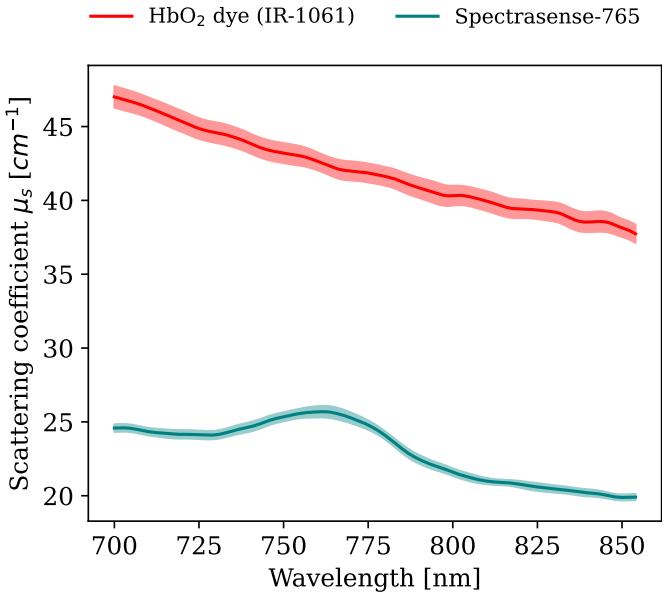


Fig S10: **Scattering spectra of IR-1061 and Spectrasense-765.** Solid lines indicate the measured scattering μ_s spectra of the proxy dyes. Bands around the measured spectra indicate the standard deviation across the 12 measurement points for each optical sample slab.

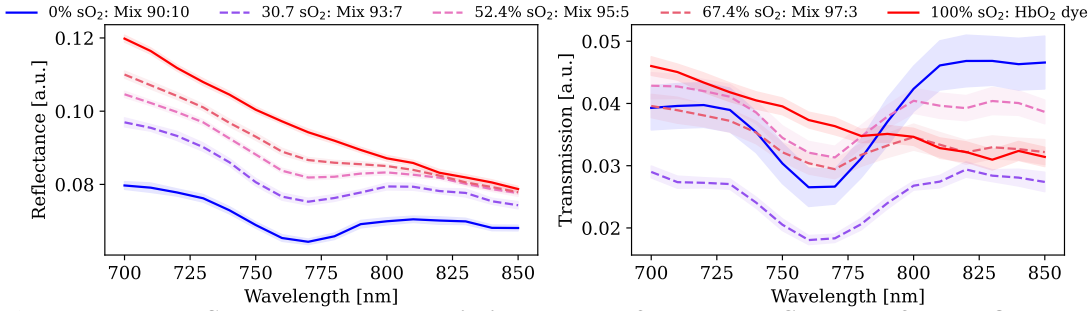
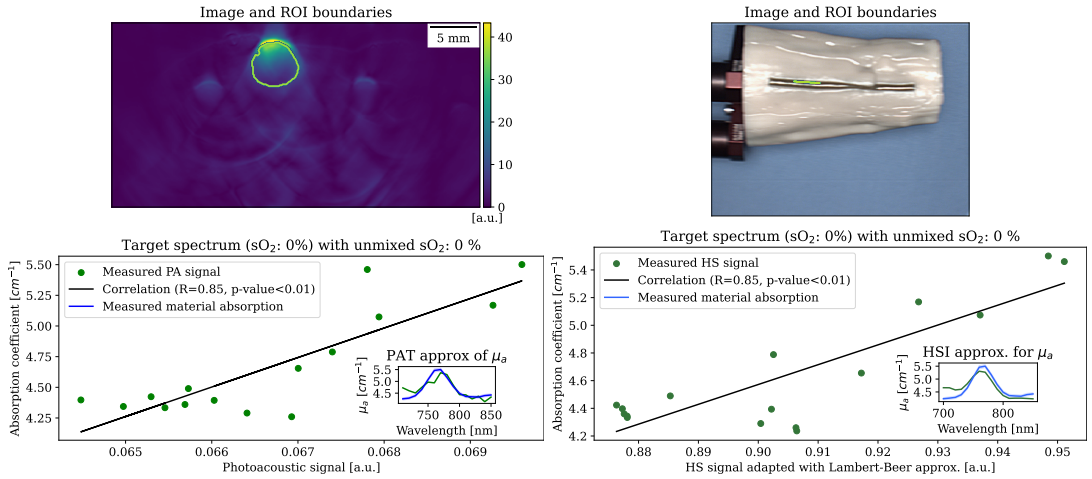
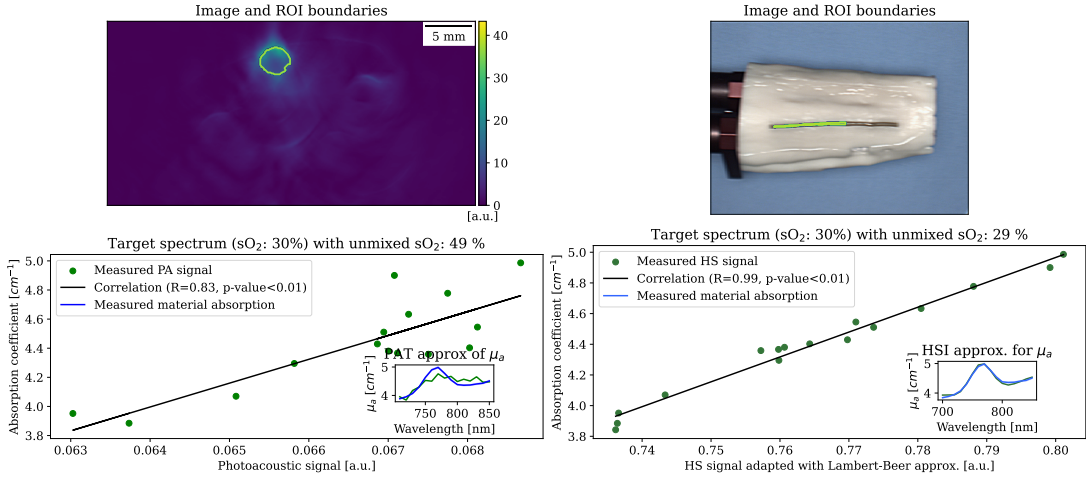


Fig S11: **Raw reflectance and transmission spectra from the DIS system for the five oxygen saturation levels were used for forearm phantom fabrication.** Based on IR-1061 and Spectrasense-765, five oxygen saturation levels (in %) were derived with the respective mixture ratios of 100:0, 97:3, 95:5, 93:7, 90:10. (a) and (b) represent the double integrating sphere measurements of reflectance and transmission, respectively. Solid lines are the spectra that are used as endmembers (0% oxygen saturation and 100% oxygen saturation) for linear spectral unmixing (LSU). Dashed lines represent the intermediate levels and the corresponding percentages in the legend are the LSU results when using the solid lines as endmembers. Bands around the spectra indicate the standard deviation across the 12 measurement points for each optical sample slab.



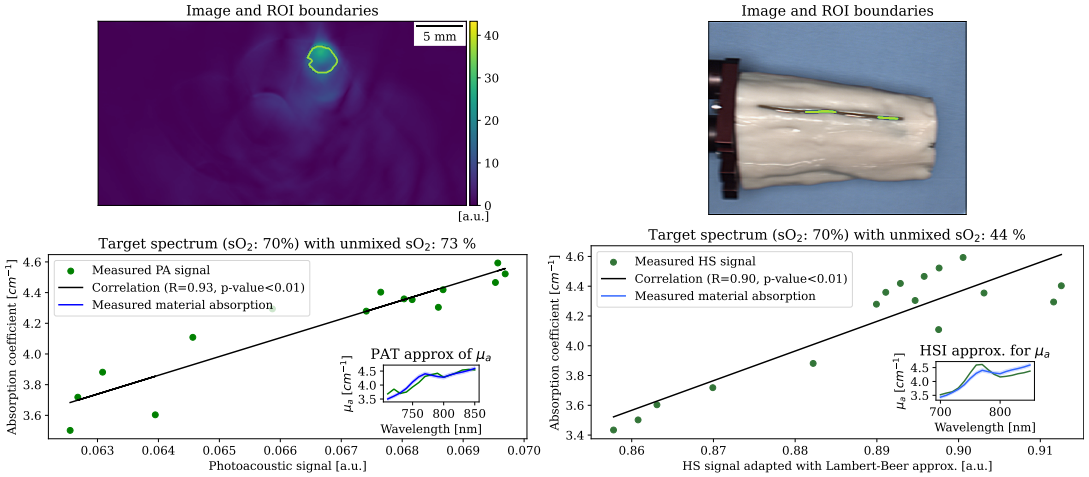
(a) (b)

Fig S12: **Examples of signal correlation of the 0% oxygen saturation superficial vessel for both (a) photoacoustic (PA) tomography (PAT) and (b) hyperspectral (HS) imaging (HSI).** Top pictures show (a) PA image with region of interest (ROI) boundary (green outline) and (b) RGB- (red blue green) rendered HS image with its ROI boundary. Lower pictures show the correlation of measured spectra (green dots as explained in sec. 3 of the main paper) with the measured absorption coefficients. Black solid lines represent the resulting linear regression function with corresponding R values. The inset plots show the measured absorptions (blue) and estimated absorptions (green, using the correlation function) as qualitative confirmation.



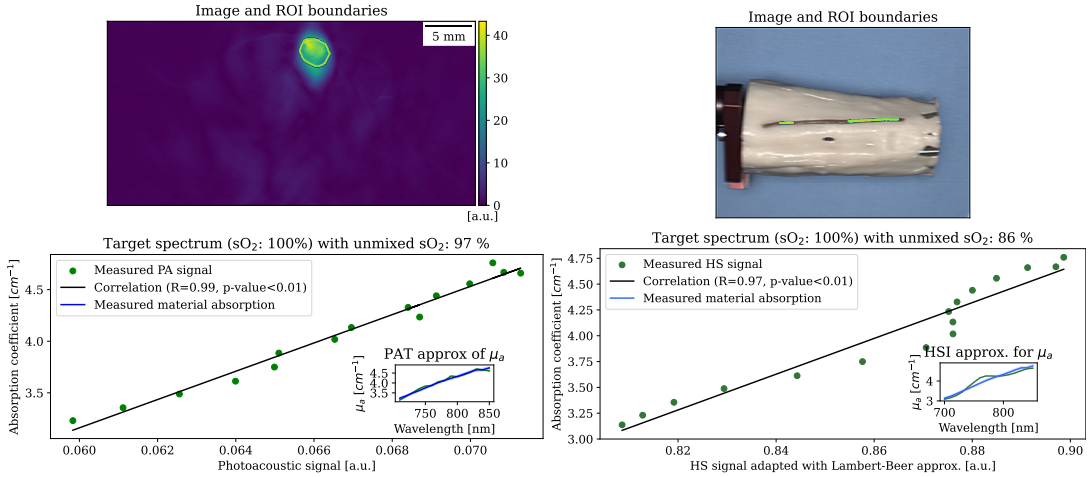
(a) (b)

Fig S13: **Examples of signal correlation of the 30% oxygen saturation superficial vessel for both (a) photoacoustic (PA) tomography (PAT) and (b) hyperspectral (HS) imaging (HSI).** Top pictures show (a) PA image with region of interest (ROI) boundary (green outline) and (b) RGB- (red blue green) rendered HS image with its ROI boundary. Lower pictures show the correlation of measured spectra (green dots as explained in sec. 3 of the main paper) with the measured absorption coefficients. Black solid lines represent the resulting linear regression function with corresponding R values. The inset plots show the measured absorptions (blue) and estimated absorptions (green, using the correlation function) as qualitative confirmation.



(a) (b)

Fig S14: **Examples of signal correlation of the 70% oxygen saturation superficial vessel for both (a) photoacoustic (PA) tomography (PAT) and (b) hyperspectral (HS) imaging (HSI).** Top pictures show (a) PA image with region of interest (ROI) boundary (green outline) and (b) RGB- (red blue green) rendered HS image with its ROI boundary. Lower pictures show the correlation of measured spectra (green dots as explained in sec. 3 of the main paper) with the measured absorption coefficients. Black solid lines represent the resulting linear regression function with corresponding R values. The inset plots show the measured absorptions (blue) and estimated absorptions (green, using the correlation function) as qualitative confirmation.



(a) (b)

Fig S15: **Examples of signal correlation of the 100% oxygen saturation superficial vessel for both (a) photoacoustic (PA) tomography (PAT) and (b) hyperspectral (HS) imaging (HSI).** Top pictures show (a) PA image with region of interest (ROI) boundary (green outline) and (b) RGB- (red blue green) rendered HS image with its ROI boundary. Lower pictures show the correlation of measured spectra (green dots as explained in sec. 3 of the main paper) with the measured absorption coefficients. Black solid lines represent the resulting linear regression function with corresponding R values. The inset plots show the measured absorptions (blue) and estimated absorptions (green, using the correlation function) as qualitative confirmation.

S3 Phantom quality assurance

S3.1 Air bubble investigation within phantoms

Our forearm phantoms are considerably larger than many conventional tissue phantoms, making it difficult to achieve uniform fabrication without entrapping air. Indeed, in our experiments, we frequently observed air bubbles within the phantoms. To assess how these inclusions might affect the photoacoustic signal, we performed a small simulation study. First, we segmented example forearms (including visible air bubbles) from ultrasound images. We then simulated photoacoustic images for two conditions, one with and one without the air bubbles, and qualitatively compared the resulting vessel spectra.

Figure S16 illustrates two simulation outcomes for a digital twin of a forearm phantom, depicting scenarios with and without air bubbles. Two vessels, along with their respective oxygen saturation levels, are highlighted, and the mean spectra within each vessel are plotted for both conditions. In superficial vessels, the spectra agreed closely across both models, whereas deeper vessels showed noticeable distortion in the presence of air bubbles. Similar findings were observed in additional forearm phantoms (see Figures S17, S18, S19, and S20). Overall, vessels with centres located deeper than approximately 1.5 cm from the surface appear susceptible to reverberation artefacts caused by the embedded air bubbles.

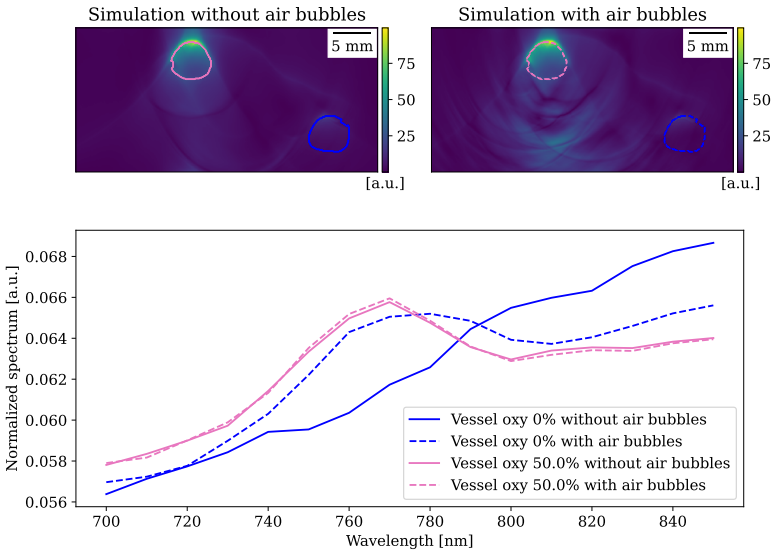


Fig S16: **Image and spectrum comparison of simulations with and without air bubbles.** Photoacoustic images of the forearm phantom 1 were simulated with (top right) and without (top left) air bubbles. Images are shown at 700 nm. Two vessels with different oxygen saturations are marked in the images and their mean spectra are plotted below. The solid lines represent the spectra without air bubbles and the dashed lines represent the spectra with air bubbles.

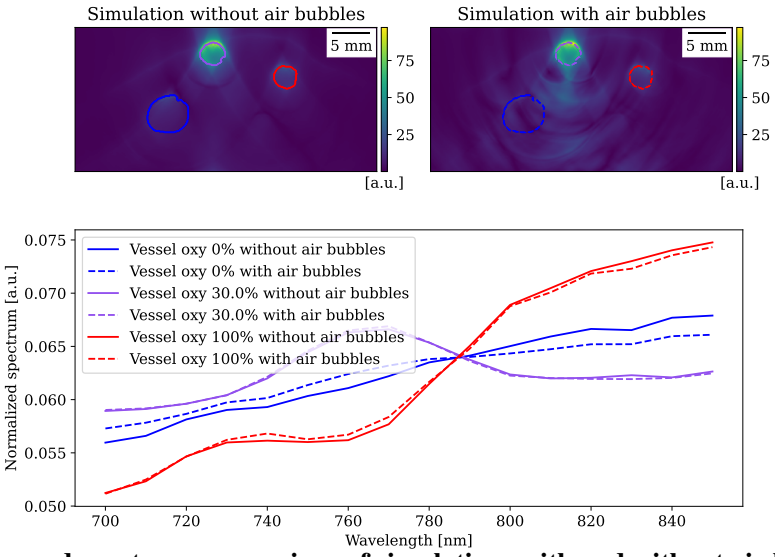


Fig S17: **Image and spectrum comparison of simulations with and without air bubbles.** Photoacoustic images of the forearm phantom 2 were simulated with (top right) and without (top left) air bubbles. Images are shown at 700 nm. Three vessels with different oxygen saturations are marked in the images and their mean spectra are plotted below. The solid lines represent the spectra without air bubbles and the dashed lines represent the spectra with air bubbles.

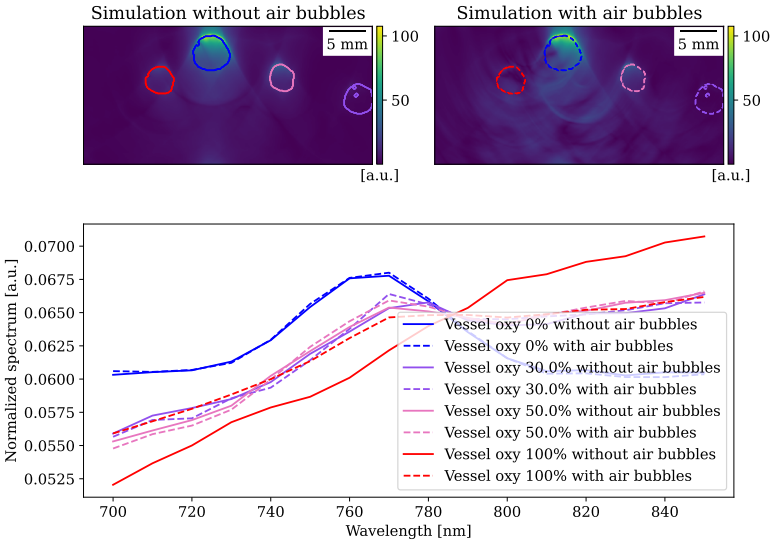


Fig S18: **Image and spectrum comparison of simulations with and without air bubbles.** Photoacoustic images of the forearm phantom 3 were simulated with (top right) and without (top left) air bubbles. Images are shown at 700 nm. Four vessels with different oxygen saturations are marked in the images and their mean spectra are plotted below. The solid lines represent the spectra without air bubbles and the dashed lines represent the spectra with air bubbles. Small purple circles indicate air bubbles.

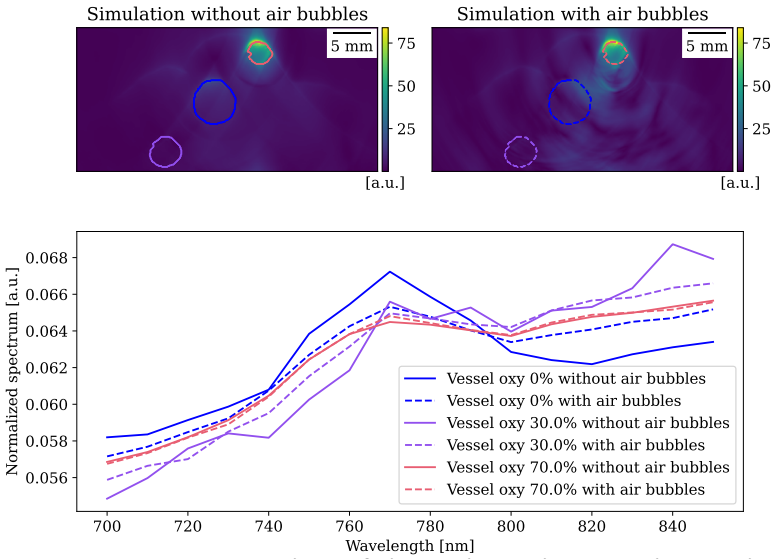


Fig S19: **Image and spectrum comparison of simulations with and without air bubbles.** Photoacoustic images of the forearm phantom 4 were simulated with (top right) and without (top left) air bubbles. Images are shown at 700 nm. Three vessels with different oxygen saturations are marked in the images and their mean spectra are plotted below. The solid lines represent the spectra without air bubbles and the dashed lines represent the spectra with air bubbles.

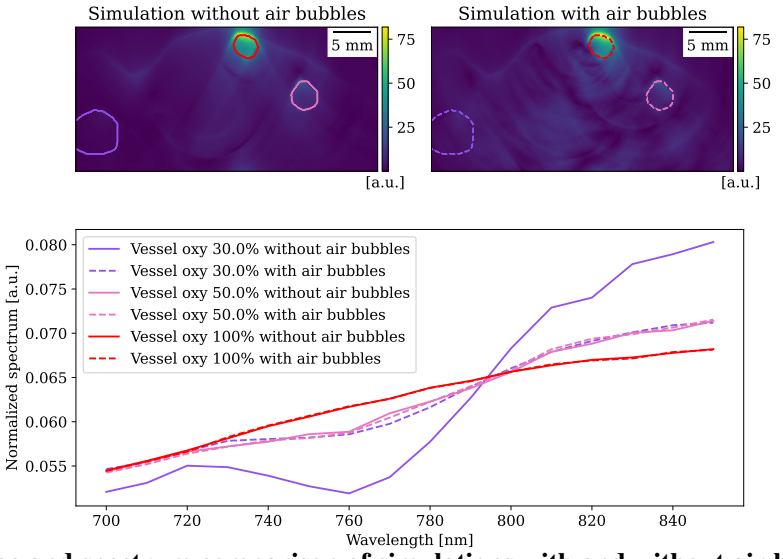


Fig S20: **Image and spectrum comparison of simulations with and without air bubbles.** Photoacoustic images of the forearm phantom 5 were simulated with (top right) and without (top left) air bubbles. Images are shown at 700 nm. Three vessels with different oxygen saturations are marked in the images and their mean spectra are plotted below. The solid lines represent the spectra without air bubbles and the dashed lines represent the spectra with air bubbles.

S3.2 Speed of sound investigation

We assumed a speed of sound of approximately 1470 ms^{-1} in our forearm phantoms. To evaluate whether this approximation is reasonable or significantly off, we conducted a small simulation study using segmentations from the previously described air-bubble investigation. First, we simulated photoacoustic images of the phantoms at 1470 ms^{-1} and then repeated the process with the same segmentations at $\pm$50 ms^{-1} and $\pm$100 ms^{-1} from this baseline value.

Next, we manually segmented superficial vessels in each simulated image according to the previously described protocol, computed the segmented vessel area (A), and derived an approximate vessel diameter (D) from $D = 2\sqrt{A/\pi}$. We compared these diameters to a reference diameter in the image D_{im}, calculated using the fabricated diameter (D_{fab}) and the ratio between the assumed phantom speed of sound (sos_{ref}) and the reconstruction speed of sound (sos_{recon}). Specifically, the reference diameter is given by $D_{\text{im}} = \frac{sos_{\text{recon}}}{sos_{\text{ref}}} D_{\text{fab}}$.

Figure S21 shows the simulation results for a digital twin of a forearm phantom, assuming an sos of 1470 ms^{-1} and variations of $\pm$50 ms^{-1} and $\pm$100 ms^{-1} (for the other phantoms, cf. Figures S22, S23, S24, and S25. In Fig. S21, the fabricated vessel diameter was 5 mm. Under the assumption of 1470 ms^{-1} with a reconstruction at 1497.4 ms^{-1}, we would expect an imaged diameter of 5.09 mm. For the other sos values of -100 ms^{-1}, -50 ms^{-1}, +50 ms^{-1}, and +100 ms^{-1}, the respective imaged diameters were 5.46 mm, 5.27 mm, 4.93 mm, and 4.77 mm. Given the simulated image resolution of 0.1 mm, these findings align closely with theoretical expectations, suggesting that 1470 ms^{-1} is a sufficiently accurate estimate for our phantom material.

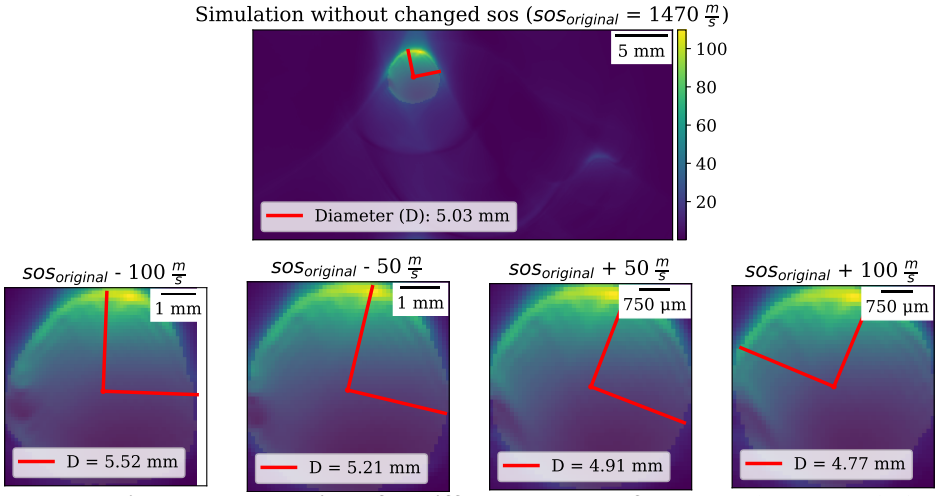


Fig S21: **Vessel diameter comparison for different speeds of sound.** Photoacoustic images of the forearm phantom 1 were simulated with a phantom material speed of sound of 1470 ms^{-1} (top) and then with speeds of sounds that differ from 1470 ms^{-1} by -100 ms^{-1}, -50 ms^{-1}, +50 ms^{-1}, +100 ms^{-1} (from left to right). As the manual segmentations of the superficial vessel were not perfectly circular, the segmentation area A for all images was computed and their diameter D was calculated via $D = 2\sqrt{A/\pi}$. D is indicated in all images by the major axes of the segmented areas.

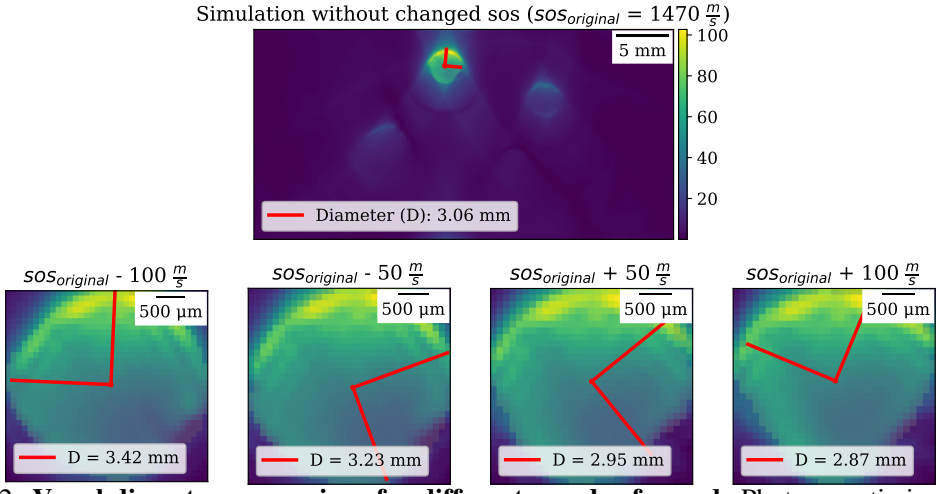


Fig S22: **Vessel diameter comparison for different speeds of sound.** Photoacoustic images of the forearm phantom 2 were simulated with a phantom material speed of sound of 1470 ms^{-1} (top) and then with speeds of sounds that differ from 1470 ms^{-1} by -100 ms^{-1}, -50 ms^{-1}, +50 ms^{-1}, +100 ms^{-1} (from left to right). As the manual segmentations of the superficial vessel were not perfectly circular, the segmentation area A for all images was computed and their diameter D was calculated via $D = 2\sqrt{A/\pi}$. D is indicated in all images by the major axes of the segmented areas.

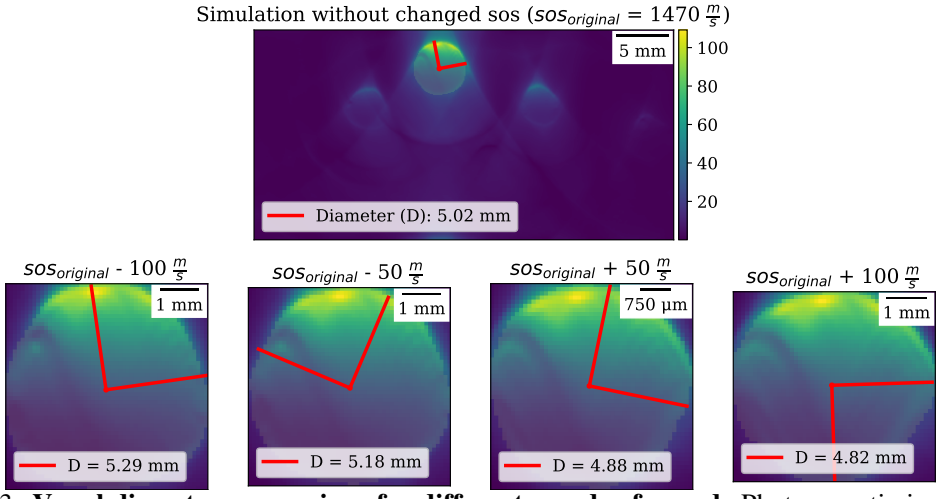


Fig S23: **Vessel diameter comparison for different speeds of sound.** Photoacoustic images of the forearm phantom 3 were simulated with a phantom material speed of sound of 1470 ms^{-1} (top) and then with speeds of sounds that differ from 1470 ms^{-1} by -100 ms^{-1}, -50 ms^{-1}, +50 ms^{-1}, +100 ms^{-1} (from left to right). As the manual segmentations of the superficial vessel were not perfectly circular, the segmentation area A for all images was computed and their diameter D was calculated via $D = 2\sqrt{A/\pi}$. D is indicated in all images by the major axes of the segmented areas.

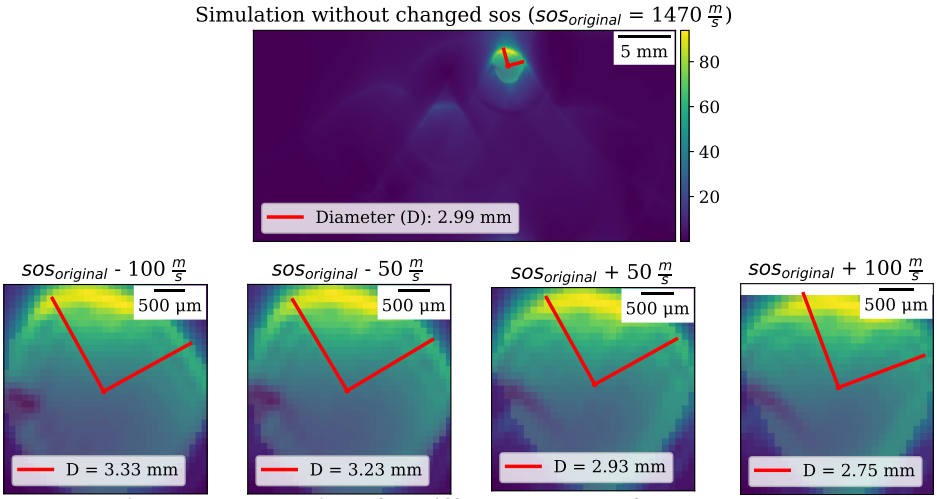


Fig S24: **Vessel diameter comparison for different speeds of sound.** Photoacoustic images of the forearm phantom 4 were simulated with a phantom material speed of sound of 1470 ms^{-1} (top) and then with speeds of sounds that differ from 1470 ms^{-1} by -100 ms^{-1}, -50 ms^{-1}, +50 ms^{-1}, +100 ms^{-1} (from left to right). As the manual segmentations of the superficial vessel were not perfectly circular, the segmentation area A for all images was computed and their diameter D was calculated via $D = 2\sqrt{A/\pi}$. D is indicated in all images by the major axes of the segmented areas.

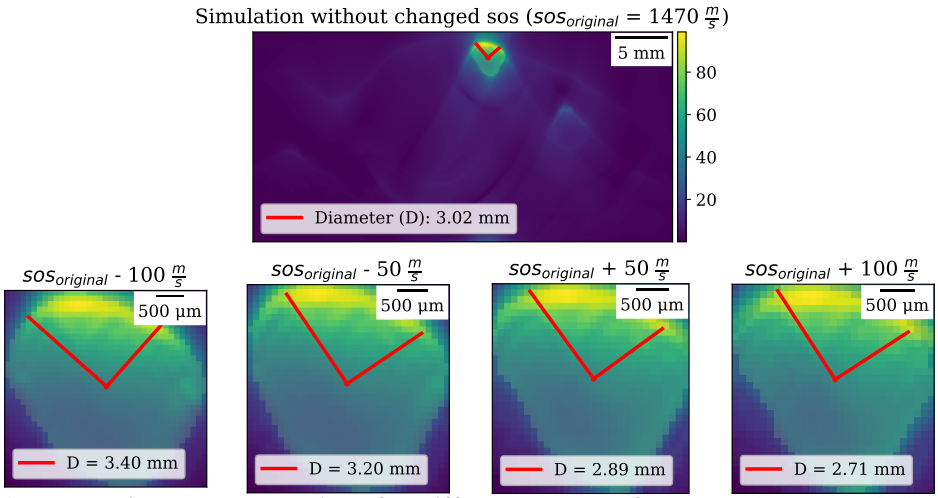


Fig S25: **Vessel diameter comparison for different speeds of sound.** Photoacoustic images of the forearm phantom 5 were simulated with a phantom material speed of sound of 1470 ms^{-1} (top) and then with speeds of sounds that differ from 1470 ms^{-1} by -100 ms^{-1}, -50 ms^{-1}, +50 ms^{-1}, +100 ms^{-1} (from left to right). As the manual segmentations of the superficial vessel were not perfectly circular, the segmentation area A for all images was computed and their diameter D was calculated via $D = 2\sqrt{A/\pi}$. D is indicated in all images by the major axes of the segmented areas.

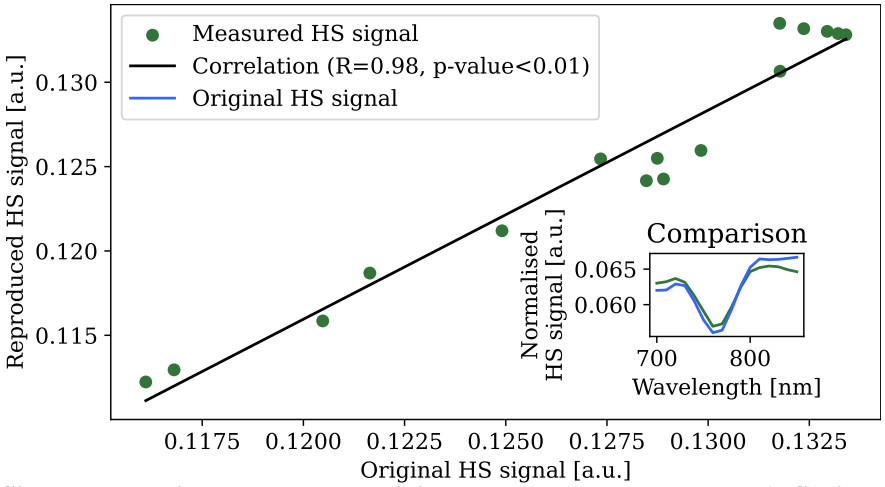


Fig S26: **Signal correlation between the original measured hyperspectral (HS) signal and the re-recorded signal of the 0% oxygen saturation superficial vessel.** Green dots represent the reproduced HS signal and the black solid line represents the resulting linear regression function with corresponding R value. The inset plot shows the normalised re-recorded signal and original signal (blue) as qualitative confirmation.

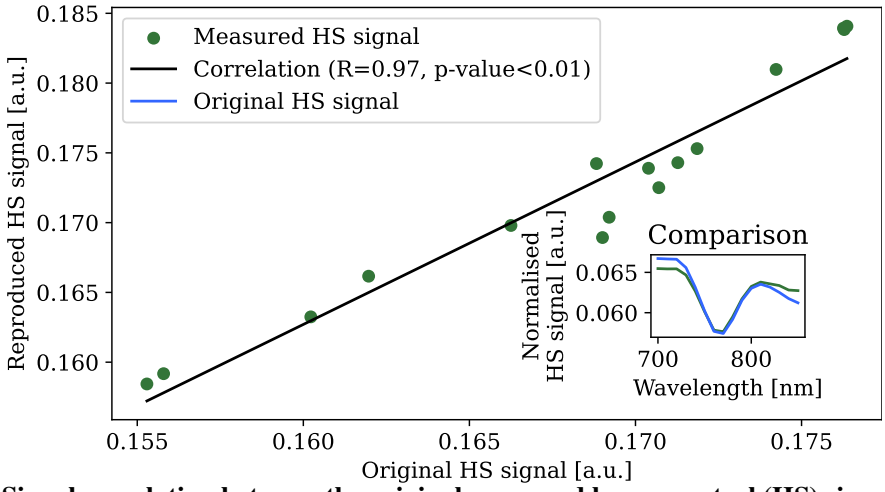


Fig S27: **Signal correlation between the original measured hyperspectral (HS) signal and the re-recorded signal of the 30% oxygen saturation superficial vessel.** Green dots represent the reproduced HS signal and the black solid line represents the resulting linear regression function with corresponding R value. The inset plot shows the normalised re-recorded signal and original signal (blue) as qualitative confirmation.

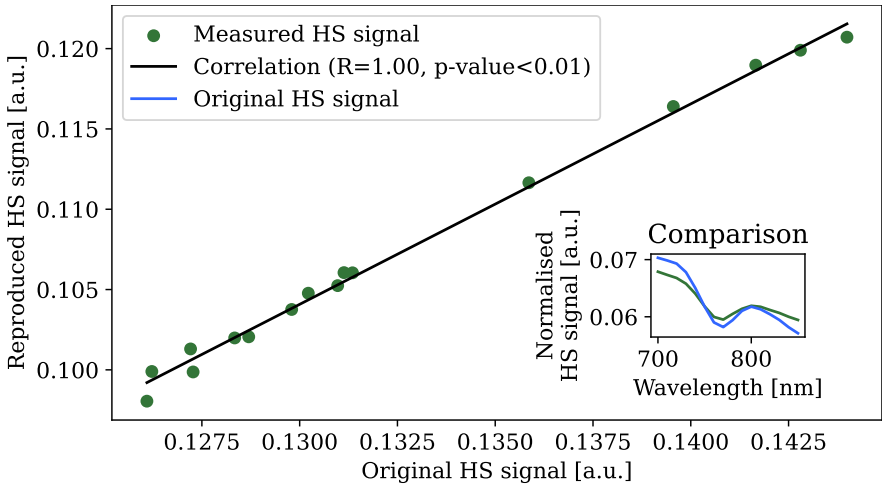


Fig S28: **Signal correlation between the original measured hyperspectral (HS) signal and the re-recorded signal of the 50% oxygen saturation superficial vessel.** Green dots represent the reproduced HS signal and the black solid line represents the resulting linear regression function with corresponding R value. The inset plot shows the normalised re-recorded signal and original signal (blue) as qualitative confirmation.

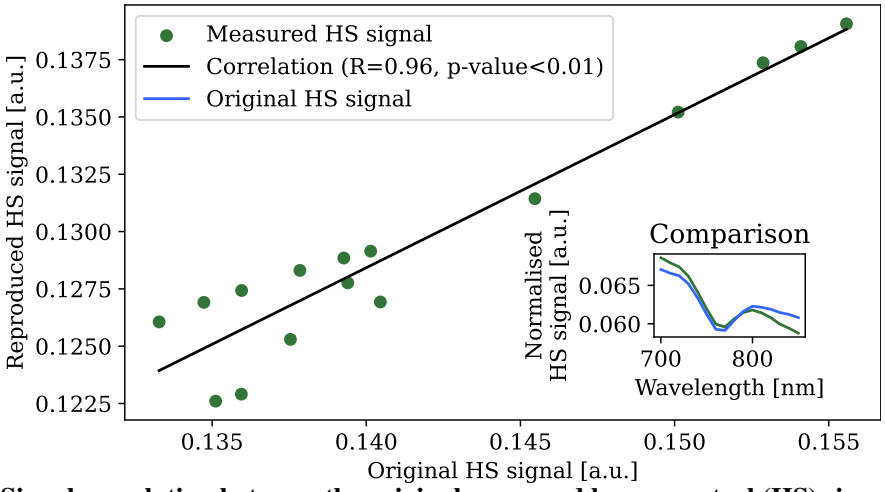


Fig S29: **Signal correlation between the original measured hyperspectral (HS) signal and the re-recorded signal of the 70% oxygen saturation superficial vessel.** Green dots represent the reproduced HS signal and the black solid line represents the resulting linear regression function with corresponding R value. The inset plot shows the normalised re-recorded signal and original signal (blue) as qualitative confirmation.

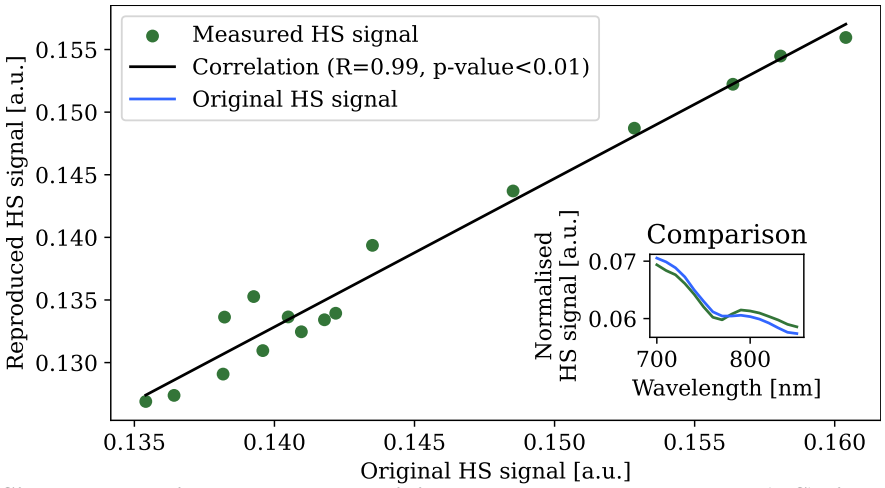


Fig S30: **Signal correlation between the original measured hyperspectral (HS) signal and the re-recorded signal of the 100% oxygen saturation superficial vessel.** Green dots represent the reproduced HS signal and the black solid line represents the resulting linear regression function with corresponding R value. The inset plot shows the normalised re-recorded signal and original signal (blue) as qualitative confirmation.

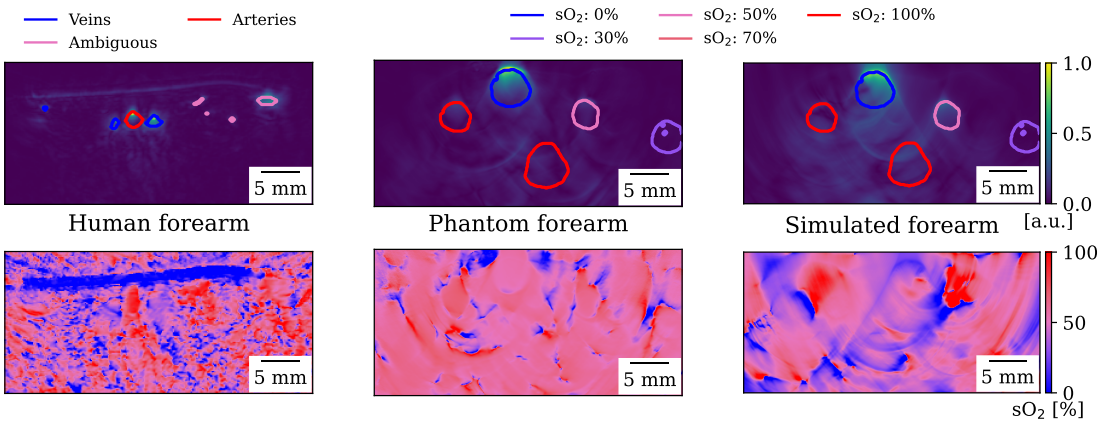


Fig S31: **Side-by-side comparison of a human forearm, an example forearm phantom and its corresponding simulation.** In the upper row, photoacoustic (PA) images with their segmented vessels are presented and in the lower row, the corresponding unmixed images. As we do not have a ground truth for the annotations in the human forearm image, we divided the type of vessels in arteries, veins and ambiguous vessels where we were uncertain. For phantom and simulated forearm, we could annotate with ground truth. We used deoxyhaemoglobin (Hb) and oxyhaemoglobin (HbO_2) as endmembers for the human forearm and our dye proxies as endmembers for the phantom and simulated image.

S4 Supplementary Tables

Dye name	Manufacturer/Vendor	Product name
Black	Cranfield (Jackson's art)	RCR2501860
Process Blue	Cranfield (Jackson's art)	RCR25025291
Process Red	Cranfield (Jackson's art)	RCR25063827
Process Yellow	Cranfield (Jackson's art)	RCR25091884
Arylide Yellow	Cranfield (Jackson's art)	RCR25091630
Diarylide Yellow	Cranfield (Jackson's art)	RCR25091759
Light Orange	Cranfield (Jackson's art)	RCR25091637
Rubine Red	Cranfield (Jackson's art)	RCR25063254
Naphthol Red	Cranfield (Jackson's art)	RCR50063266
Carbazole Violet	Cranfield (Jackson's art)	RCR50071139
Ultramarine	Cranfield (Jackson's art)	RCR50024283
Prussian Blue	Cranfield (Jackson's art)	RCR25024309
Phthalo Blue	Cranfield (Jackson's art)	RCR25024760
Phthalo Green	Cranfield (Jackson's art)	RCR25043104
Yellow Ochre	Cranfield (Jackson's art)	RCR25091737
Burnt Sienna	Cranfield (Jackson's art)	RCR50032371
Raw Umber	Cranfield (Jackson's art)	RCR25032211
White	Cranfield (Jackson's art)	RCR25083391
SS Black	Culture Hustle (Stuart Semple)	BLACK 1.0 Pigment- 50g
SS Red	Culture Hustle (Stuart Semple)	THE WORLD'S PINKEST PINK - 50g
Hemoglobin	Sigma-Aldrich (Merck)	H2625-100G
Nigrosin	Sigma-Aldrich (Merck)	211680-100G
IR-895	Sigma-Aldrich (Merck)	392375
IR-1048	Sigma-Aldrich (Merck)	405175-500MG
IR-1061	Sigma-Aldrich (Merck)	405124-250MG
Spectrasense-765	Sun Chemical Colors & Effects GmbH	Spectrasense™ IR 765

Table S1: Dyes that were investigated for their optical absorption and scattering spectra. The column "Dye name" indicates the name which is used in this work to refer to the specific dye and the "Product name" indicates the name that is used by the manufacturer/vendor.

	1	2	3	4	5	6	7	8	9	10A	10B	10C
sO_2	0%	50%	100%	100%	50%	0%	100%	50%	0%	100%	0%	70%
Bvf	2.5%	2.5%	2.5%	4%	4%	4%	1%	1%	1%	0.5%	5%	3%
sO_2 vessel	50%	30%	0%	70%	100%	100%	/	/	/	/	/	/

Table S2: Oxygen saturation (sO_2) and blood volume fraction (Bvf) levels for the background material of each forearm phantom. Bvf refers to the percentage of the total volume that is made up of the dye mixture. While all the others have as uniform as possible properties, the 10th forearm was separated into three parts, with different properties, named 10A, 10B, and 10C. Six of the forearms have a superficial vessel, the oxygen saturation level of which is also reported.

	Hb dye	HbO_2 dye	30% sO_2	50% sO_2	70% sO_2
R-value	1.00	0.90	1.00	1.00	1.00

Table S3: Pearson correlation coefficients (R-values) of the absorption spectra of proxy dyes and consequent oxygen saturation (sO_2) levels with the target absorption coefficients. The proxy dyes were directly correlated with the absorption spectra of Hb and HbO_2. The sO_2 levels 30%, 50%, and 70% were correlated with their target spectrum generated by $\mu_a^{\text{target}} = sO_2 * \mu_a^{HbO_2\, dye} + (1 - sO_2) * \mu_a^{\text{Hb dye}}$.

BIBLIOGRAPHY

[1] L. V. Wang and H.-i. Wu. *Biomedical Optics: Principles and Imaging*. en. Google-Books-ID: RKukklVYAY4C. John Wiley & Sons, May 2007 (cit. on pp. 3, 11, 13, 14, 18, 23).

[2] V. Ntziachristos. "Going deeper than microscopy: the optical imaging frontier in biology". en. In: *Nature Methods* 7.8 (Aug. 2010), pp. 603–614. DOI: 10.1038/nmeth.1483 (cit. on p. 3).

[3] S. Manohar and D. Razansky. "Photoacoustics: a historical review". EN. In: *Advances in Optics and Photonics* 8.4 (Dec. 2016), pp. 586–617. DOI: 10.1364/AOP.8.000586 (cit. on p. 3).

[4] A. Rosencwaig and A. Gersho. "Theory of the photoacoustic effect with solids". In: *Journal of Applied Physics* 47.1 (Jan. 1976), pp. 64–69. DOI: 10.1063/1.322296 (cit. on p. 3).

[5] *Shining a Light on Photoacoustic Imaging*. en. Tech. rep. Radiological Society of North America (RSNA) (cit. on p. 3).

[6] M. Taylor-Williams et al. "Noninvasive hemoglobin sensing and imaging: optical tools for disease diagnosis". In: *Journal of Biomedical Optics* 27.8 (Aug. 2022), p. 080901. DOI: 10.1117/1.JBO.27.8.080901 (cit. on p. 3).

[7] H. J. Knox et al. "A bioreducible N-oxide-based probe for photoacoustic imaging of hypoxia". en. In: *Nature Communications* 8.1 (Nov. 2017), p. 1794. DOI: 10.1038/s41467-017-01951-0 (cit. on p. 4).

[8] J. Park et al. "Clinical translation of photoacoustic imaging". en. In: *Nature Reviews Bioengineering* (Sept. 2024), pp. 1–20. DOI: 10.1038/s44222-024-00240-y (cit. on pp. 4–7).

[9] T. L. Lefebvre et al. "The Potential of Photoacoustic Imaging in Radiation Oncology". English. In: *Frontiers in Oncology* 12 (Mar. 2022). DOI: 10.3389/fonc.2022.803777 (cit. on p. 5).

[10] J. J. M. Riksen, A. V. Nikolaev, and G. v. Soest. "Photoacoustic imaging on its way toward clinical utility: a tutorial review focusing on practical application in medicine". In: *Journal of Biomedical Optics* 28.12 (June 2023), p. 121205. DOI: 10.1117/1.JBO.28.12.121205 (cit. on p. 5).

[11] H. Assi et al. “A review of a strategic roadmapping exercise to advance clinical translation of photoacoustic imaging: From current barriers to future adoption”. In: *Photoacoustics* 32 (Aug. 2023), p. 100539. DOI: 10.1016/j.pacs.2023.100539 (cit. on pp. 5, 37).

[12] J. Gröhl et al. “Deep learning for biomedical photoacoustic imaging: A review”. In: *Photoacoustics* 22 (June 2021), p. 100241. DOI: 10.1016/j.pacs.2021.100241 (cit. on pp. 7, 39).

[13] A. Hauptmann and B. T. Cox. “Deep learning in photoacoustic tomography: current approaches and future directions”. In: *Journal of Biomedical Optics* 25.11 (Oct. 2020), p. 112903. DOI: 10.1117/1.JBO.25.11.112903 (cit. on pp. 7, 127).

[14] J.-Y. Zhu et al. “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. ISSN: 2380-7504. Oct. 2017, pp. 2242–2251. DOI: 10.1109/ICCV.2017.244 (cit. on p. 7).

[15] K. Li et al. “Adversarial Feature Hallucination Networks for Few-Shot Learning”. In: 2020, pp. 13470–13479 (cit. on pp. 7, 32).

[16] L. Hacker and J. Joseph. “Technical Validation of Photoacoustic Imaging Systems Using Phantoms”. en. In: *Biomedical Photoacoustics: Technology and Applications*. Ed. by W. Xia. Cham: Springer Nature Switzerland, 2024, pp. 213–227. DOI: 10.1007/978-3-031-61411-8_7 (cit. on p. 7).

[17] R. Tachet des Combes et al. “Domain Adaptation with Conditional Distribution Matching and Generalized Label Shift”. In: *Advances in Neural Information Processing Systems*. Vol. 33. Curran Associates, Inc., 2020, pp. 19276–19289 (cit. on pp. 9, 131).

[18] Y. Zhang et al. “Exact Feature Distribution Matching for Arbitrary Style Transfer and Domain Generalization”. en. In: 2022, pp. 8035–8045 (cit. on pp. 9, 131).

[19] K. K. Dreher. “Efficient Simulation of Realistic Photoacoustic Images with Unsupervised Domain Adaptation”. MA thesis. Heidelberg, Germany: Ruprecht-Karls-Universität, July 2020 (cit. on p. 11).

[20] G. Keiser. *Biophotonics: Concepts to Applications*. en. Graduate Texts in Physics. Singapore: Springer Nature, 2022. DOI: 10.1007/978-981-19-3482-7 (cit. on pp. 11, 12).

[21] P. Beard. “Biomedical photoacoustic imaging”. In: *Interface Focus* 1.4 (June 2011), pp. 602–631. DOI: 10.1098/rsfs.2011.0028 (cit. on p. 11).

[22] M. Clement, G. Daniel, and M. Trelles. “Optimising the design of a broad-band light source for the treatment of skin”. In: *Journal of Cosmetic and Laser Therapy* 7.3-4 (Jan. 2005), pp. 177–189. DOI: 10.1080/14764170500344575 (cit. on p. 12).

[23] L. Pezzi et al. “Antimicrobial Effects of Chemically Functionalized and/or Photo-Heated Nanoparticles”. en. In: *Materials* 12.7 (Jan. 2019), p. 1078. DOI: 10.3390/ma12071078 (cit. on p. 12).

[24] N. Bloembergen and P. S. Pershan. “Light Waves at the Boundary of Nonlinear Media”. In: *Physical Review* 128.2 (Oct. 1962), pp. 606–622. DOI: 10.1103/PhysRev.128.606 (cit. on p. 12).

[25] A. J. Welch and M. J. Van Gemert, eds. *Optical-Thermal Response of Laser-Irradiated Tissue*. en. Dordrecht: Springer Netherlands, 2011. DOI: 10.1007/978-90-481-8831-4 (cit. on p. 13).

[26] V. V. Tuchin. *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*. en. Society of Photo-Optical Instrumentation Engineers (SPIE), 2015. DOI: 10.1117/3.1003040 (cit. on p. 14).

[27] J. Ripoll. “Derivation of the scalar radiative transfer equation from energy conservation of Maxwell’s equations in the far field”. EN. In: *JOSA A* 28.8 (Aug. 2011), pp. 1765–1775. DOI: 10.1364/JOSAA.28.001765 (cit. on p. 15).

[28] L. G. Henyey and J. L. Greenstein. “Diffuse radiation in the Galaxy.” In: *The Astrophysical Journal* 93 (Jan. 1941). ADS Bibcode: 1941ApJ....93...70H, pp. 70–83. DOI: 10.1086/144246 (cit. on p. 15).

[29] W. Demtröder. *Mechanics and Thermodynamics*. en. Undergraduate Lecture Notes in Physics. Cham: Springer International Publishing, 2017. DOI: 10.1007/978-3-319-27877-3 (cit. on pp. 17, 18).

[30] P. S. Yarmolenko et al. “Thresholds for thermal damage to normal tissues: An update”. In: *International Journal of Hyperthermia* 27.4 (June 2011), pp. 320–343. DOI: 10.3109/02656736.2010.534527 (cit. on p. 19).

[31] J. Debus. “Grundlagen der Physikalischen Ultraschalldiagnostik und -therapie”. de. In: *Medizinische Physik 2: Medizinische Strahlenphysik*. Ed. by W. Schlegel and J. Bille. Berlin, Heidelberg: Springer, 2002, pp. 211–228. DOI: 10.1007/978-3-642-56259-4_8 (cit. on p. 19).

[32] K. Wang and M. A. Anastasio. “Photoacoustic and thermoacoustic tomography: Image formation principles”. In: *Handbook of Mathematical Methods in Imaging*. Springer, Jan. 2015, pp. 1081–1116. DOI: 10.1007/978-1-4939-0790-8_50 (cit. on p. 20).

[33] J. Poudel, Y. Lou, and M. A. Anastasio. “A survey of computational frameworks for solving the acoustic inverse problem in three-dimensional photoacoustic computed tomography”. en. In: *Physics in Medicine & Biology* 64.14 (July 2019), 14TR01. DOI: 10.1088/1361-6560/ab2017 (cit. on p. 20).

[34] B. T. Cox, J. G. Laufer, and P. C. Beard. "The challenges for quantitative photoacoustic imaging". In: *Photons Plus Ultrasound: Imaging and Sensing 2009*. Vol. 7177. SPIE, Feb. 2009, pp. 294–302. DOI: 10.1117/12.806788 (cit. on p. 20).

[35] B. T. Cox et al. "Quantitative spectroscopic photoacoustic imaging: a review". In: *Journal of Biomedical Optics* 17.6 (June 2012), p. 061202. DOI: 10.1117/1.JBO.17.6.061202 (cit. on p. 20).

[36] Z. Wang, W. Tao, and H. Zhao. "The Optical Inverse Problem in Quantitative Photoacoustic Tomography: A Review". en. In: *Photonics* 10.5 (May 2023), p. 487. DOI: 10.3390/photonics10050487 (cit. on p. 20).

[37] T. Tarvainen and B. Cox. "Quantitative photoacoustic tomography: modeling and inverse problems". In: *Journal of Biomedical Optics* 29.S1 (Dec. 2023), S11509. DOI: 10.1117/1.JBO.29.S1.S11509 (cit. on pp. 20, 129).

[38] M. Xu and L. V. Wang. "Universal back-projection algorithm for photoacoustic computed tomography". In: *Physical Review E* 71.1 (Jan. 2005), p. 016706. DOI: 10.1103/PhysRevE.71.016706 (cit. on pp. 20, 28).

[39] B. E. Treeby, E. Z. Zhang, and B. T. Cox. "Photoacoustic tomography in absorbing acoustic media using time reversal". en. In: *Inverse Problems* 26.11 (Sept. 2010), p. 115003. DOI: 10.1088/0266-5611/26/11/115003 (cit. on p. 20).

[40] X. L. Dean-Ben et al. "Accurate Model-Based Reconstruction Algorithm for Three-Dimensional Optoacoustic Tomography". In: *IEEE Transactions on Medical Imaging* 31.10 (Oct. 2012), pp. 1922–1928. DOI: 10.1109/TMI.2012.2208471 (cit. on p. 20).

[41] C. Dehner et al. "A deep neural network for real-time optoacoustic image reconstruction with adjustable speed of sound". en. In: *Nature Machine Intelligence* 5.10 (Oct. 2023), pp. 1130–1141. DOI: 10.1038/s42256-023-00724-3 (cit. on p. 20).

[42] J. Laufer et al. "Quantitative determination of chromophore concentrations from 2D photoacoustic images using a nonlinear model-based inversion scheme". EN. In: *Applied Optics* 49.8 (Mar. 2010), pp. 1219–1233. DOI: 10.1364/AO.49.001219 (cit. on p. 21).

[43] J. Gröhl et al. "Learned spectral decoloring enables photoacoustic oximetry". en. In: *Scientific Reports* 11.1 (Mar. 2021), p. 6565. DOI: 10.1038/s41598-021-83405-8 (cit. on p. 21).

[44] X. Zhou et al. "Evaluation of Fluence Correction Algorithms in Multispectral Photoacoustic Imaging". In: *Photoacoustics* 19 (Sept. 2020), p. 100181. DOI: 10.1016/j.pacs.2020.100181 (cit. on p. 23).

[45] *Monte Carlo Light Scattering Programs* (cit. on p. 23).

[46] Q. Fang and D. A. Boas. “Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units”. EN. In: *Optics Express* 17.22 (Oct. 2009), pp. 20178–20190. DOI: 10.1364/OE.17.020178 (cit. on pp. 23, 129).

[47] B. E. Treeby and B. T. Cox. “k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields”. In: *Journal of Biomedical Optics* 15.2 (Mar. 2010), p. 021314. DOI: 10.1117/1.3360308 (cit. on pp. 26, 129).

[48] H. Nyquist. “Certain topics in telegraph transmission theory”. In: *Proceedings of the IEEE* 90.2 (Feb. 2002), pp. 280–305. DOI: 10.1109/5.989875 (cit. on p. 26).

[49] C. Shannon. “Communication in the Presence of Noise”. In: *Proceedings of the IRE* 37.1 (Jan. 1949), pp. 10–21. DOI: 10.1109/JRPROC.1949.232969 (cit. on p. 26).

[50] B. Stephanian et al. “Additive noise models for photoacoustic spatial coherence theory”. EN. In: *Biomedical Optics Express* 9.11 (Nov. 2018), pp. 5566–5582. DOI: 10.1364/BOE.9.005566 (cit. on p. 26).

[51] C. Bench, A. Hauptmann, and B. T. Cox. “Toward accurate quantitative photoacoustic imaging: learning vascular blood oxygen saturation in three dimensions”. In: *Journal of Biomedical Optics* 25.8 (Aug. 2020), p. 085003. DOI: 10.1117/1.JBO.25.8.085003 (cit. on p. 26).

[52] D. L. Donoho and I. M. Johnstone. “Ideal spatial adaptation by wavelet shrinkage”. In: *Biometrika* 81.3 (Sept. 1994), pp. 425–455. DOI: 10.1093/biomet/81.3.425 (cit. on p. 26).

[53] T. Kirchner et al. “Signed Real-Time Delay Multiply and Sum Beamforming for Multispectral Photoacoustic Imaging”. en. In: *Journal of Imaging* 4.10 (Oct. 2018), p. 121. DOI: 10.3390/jimaging4100121 (cit. on p. 28).

[54] I. K. Holfort, F. Gran, and J. A. Jensen. “Broadband minimum variance beamforming for ultrasound imaging”. In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 56.2 (Feb. 2009), pp. 314–325. DOI: 10.1109/TUFFC.2009.1040 (cit. on p. 28).

[55] L. Alzubaidi et al. “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”. In: *Journal of Big Data* 8.1 (Mar. 2021), p. 53. DOI: 10.1186/s40537-021-00444-8 (cit. on p. 28).

[56] Y. LeCun, Y. Bengio, and G. Hinton. “Deep learning”. en. In: *Nature* 521.7553 (May 2015), pp. 436–444. DOI: 10.1038/nature14539 (cit. on p. 28).

[57] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016 (cit. on p. 28).

[58] M. Schellenberg. *Learning Tissue Geometries for Photoacoustic Image Analysis*. eng. Dissertation. 2024. DOI: 10.11588/heidok.00034187 (cit. on pp. 29, 130).

[59] O. Elharrouss et al. *Loss Functions in Deep Learning: A Comprehensive Review*. arXiv:2504.04242 [cs] version: 1. Apr. 2025. DOI: 10.48550/arXiv.2504.04242 (cit. on p. 30).

[60] Y. Bengio. "Deep Learning of Representations for Unsupervised and Transfer Learning". en. In: *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*. JMLR Workshop and Conference Proceedings, June 2012, pp. 17–36 (cit. on p. 30).

[61] D. Bank, N. Koenigstein, and R. Giryes. "Autoencoders". en. In: *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook*. Ed. by L. Rokach, O. Maimon, and E. Shmueli. Cham: Springer International Publishing, 2023, pp. 353–374 (cit. on p. 30).

[62] J. Pennington, S. Schoenholz, and S. Ganguli. "Resurrecting the sigmoid in deep learning through dynamical isometry: theory and practice". In: *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc., 2017 (cit. on p. 30).

[63] A. F. Agarap. *Deep Learning using Rectified Linear Units (ReLU)*. arXiv:1803.08375 [cs]. Feb. 2019. DOI: 10.48550/arXiv.1803.08375 (cit. on p. 30).

[64] K. Hornik, M. Stinchcombe, and H. White. "Multilayer feedforward networks are universal approximators". In: *Neural Networks* 2.5 (Jan. 1989), pp. 359–366. DOI: 10.1016/0893-6080(89)90020-8 (cit. on p. 30).

[65] D. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. arXiv:1412.6980 [cs] version: 1. Dec. 2014. DOI: 10.48550/arXiv.1412.6980 (cit. on p. 31).

[66] I. J. Goodfellow et al. "Generative Adversarial Nets". In: *Advances in Neural Information Processing Systems*. Vol. 27. Curran Associates, Inc., 2014 (cit. on p. 32).

[67] X. Mao et al. "Least Squares Generative Adversarial Networks". In: *2017 IEEE International Conference on Computer Vision (ICCV)*. ISSN: 2380-7504. Oct. 2017, pp. 2813–2821. DOI: 10.1109/ICCV.2017.304 (cit. on p. 32).

[68] I. Kobyzev, S. J. Prince, and M. A. Brubaker. "Normalizing Flows: An Introduction and Review of Current Methods". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.11 (Nov. 2021), pp. 3964–3979. DOI: 10.1109/TPAMI.2020.2992934 (cit. on p. 33).

[69] L. Dinh, J. Sohl-Dickstein, and S. Bengio. *Density estimation using Real NVP*. arXiv:1605.08803 [cs] version: 2. Nov. 2016. DOI: 10.48550/arXiv.1605.08803 (cit. on p. 33).

[70] L. Dinh, D. Krueger, and Y. Bengio. *NICE: Non-linear Independent Components Estimation*. arXiv:1410.8516 [cs]. Apr. 2015. DOI: 10.48550/arXiv.1410.8516 (cit. on p. 33).

[71] D. P. Kingma and P. Dhariwal. *Glow: Generative Flow with Invertible 1x1 Convolutions*. arXiv:1807.03039 [stat]. July 2018. DOI: 10.48550/arXiv.1807.03039 (cit. on p. 33).

[72] L. Ardizzone. *Conditional Invertible Generative Models for Supervised Problems*. eng. Dissertation. 2023. DOI: 10.11588/heidok.00033932 (cit. on p. 34).

[73] J. Gröhl et al. "SIMPA: an open-source toolkit for simulation and image processing for photonics and acoustics". In: *Journal of Biomedical Optics* 27.8 (Apr. 2022), p. 083010. DOI: 10.1117/1.JBO.27.8.083010 (cit. on p. 39).

[74] K. K. Dreher et al. "Unsupervised Domain Transfer with Conditional Invertible Neural Networks". en. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*. Ed. by H. Greenspan et al. Cham: Springer Nature Switzerland, 2023, pp. 770–780. DOI: 10.1007/978-3-031-43907-0_73 (cit. on p. 40).

[75] K. K. Dreher et al. "Anthropomorphic tissue-mimicking phantoms for oximetry validation in multispectral optical imaging". In: *Journal of Biomedical Optics* 30.7 (July 2025), p. 076006. DOI: 10.1117/1.JBO.30.7.076006 (cit. on p. 40).

[76] C. Yang et al. "Review of deep learning for photoacoustic imaging". In: *Photoacoustics* 21 (Mar. 2021), p. 100215. DOI: 10.1016/j.pacs.2020.100215 (cit. on p. 127).

[77] H. Deng et al. "Deep learning in photoacoustic imaging: a review". In: *Journal of Biomedical Optics* 26.4 (Apr. 2021), p. 040901. DOI: 10.1117/1.JBO.26.4.040901 (cit. on p. 127).

[78] J. Yang et al. "Recent advances in deep-learning-enhanced photoacoustic imaging". In: *Advanced Photonics Nexus* 2.5 (July 2023), p. 054001. DOI: 10.1117/1.APN.2.5.054001 (cit. on p. 127).

[79] X. Wei et al. "Deep learning-powered biomedical photoacoustic imaging". In: *Neurocomputing* 573 (Mar. 2024), p. 127207. DOI: 10.1016/j.neucom.2023.127207 (cit. on p. 127).

[80] R. Wang et al. "Photoacoustic imaging with limited sampling: a review of machine learning approaches". EN. In: *Biomedical Optics Express* 14.4 (Apr. 2023), pp. 1777–1799. DOI: 10.1364/BOE.483081 (cit. on p. 127).

[81] M. Kim, I. Pelivanov, and M. O'Donnell. "Review of Deep Learning Approaches for Interleaved Photoacoustic and Ultrasound (PAUS) Imaging". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 70.12 (Dec. 2023), pp. 1591–1606. DOI: 10.1109/TUFFC.2023.3329119 (cit. on p. 127).

[82] S. Huang et al. “Partial Differential Equations Meet Deep Neural Networks: A Survey”. In: *IEEE Transactions on Neural Networks and Learning Systems* 36.8 (Aug. 2025), pp. 13649–13669. DOI: 10.1109/TNNLS.2025.3545967 (cit. on p. 127).

[83] J. Jumper et al. “Highly accurate protein structure prediction with AlphaFold”. en. In: *Nature* 596.7873 (Aug. 2021), pp. 583–589. DOI: 10.1038/s41586-021-03819-2 (cit. on p. 127).

[84] M. Raissi, P. Perdikaris, and G. E. Karniadakis. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. In: *Journal of Computational Physics* 378 (Feb. 2019), pp. 686–707. DOI: 10.1016/j.jcp.2018.10.045 (cit. on p. 127).

[85] G. E. Karniadakis et al. “Physics-informed machine learning”. en. In: *Nature Reviews Physics* 3.6 (June 2021), pp. 422–440. DOI: 10.1038/s42254-021-00314-5 (cit. on p. 127).

[86] S. Cuomo et al. “Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What’s Next”. en. In: *Journal of Scientific Computing* 92.3 (July 2022), p. 88. DOI: 10.1007/s10915-022-01939-z (cit. on p. 127).

[87] M. Tang, Y. Liu, and L. J. Durlofsky. “A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems”. In: *Journal of Computational Physics* 413 (July 2020), p. 109456. DOI: 10.1016/j.jcp.2020.109456 (cit. on p. 127).

[88] L. Ayala et al. “Spectral imaging enables contrast agent–free real-time ischemia monitoring in laparoscopic surgery”. In: *Science Advances* 9.10 (Mar. 2023), eadd6778. DOI: 10.1126/sciadv.add6778 (cit. on p. 127).

[89] J. Zhang et al. “PAFormer: Photoacoustic reconstruction via transformer with mask mechanism”. In: *2022 IEEE International Ultrasonics Symposium (IUS)*. ISSN: 1948-5727. Oct. 2022, pp. 1–4. DOI: 10.1109/IUS54386.2022.9957348 (cit. on p. 127).

[90] X. Song et al. “Sparse-view reconstruction for photoacoustic tomography combining diffusion model with model-based iteration”. In: *Photoacoustics* 33 (Oct. 2023), p. 100558. DOI: 10.1016/j.pacs.2023.100558 (cit. on p. 127).

[91] T. Rix et al. “Efficient Photoacoustic Image Synthesis with Deep Learning”. en. In: *Sensors* 23.16 (Aug. 2023), p. 7085. DOI: 10.3390/s23167085 (cit. on pp. 127, 137).

[92] J. Gröhl et al. “Distribution-informed and wavelength-flexible data-driven photoacoustic oximetry”. In: *Journal of Biomedical Optics* 29.S3 (June 2024), S33303. DOI: 10.1117/1.JBO.29.S3.S33303 (cit. on p. 127).

[93] I. Loc and M. B. Unlu. “Accelerating photoacoustic microscopy by reconstructing undersampled images using diffusion models”. en. In: *Scientific Reports* 14.1 (July 2024), p. 16996. DOI: 10.1038/s41598-024-67957-z (cit. on p. 127).

[94] X. Wang et al. “Residual-conditioned sparse transformer for photoacoustic image artifact reduction”. In: *Photoacoustics* 44 (Aug. 2025), p. 100731. DOI: 10.1016/j.pacs.2025.100731 (cit. on p. 127).

[95] C.-J. Wu et al. “Sustainable AI: Environmental Implications, Challenges and Opportunities”. en. In: *Proceedings of Machine Learning and Systems* 4 (Apr. 2022), pp. 795–813 (cit. on p. 127).

[96] R. Dwivedi et al. “Explainable AI (XAI): Core Ideas, Techniques, and Solutions”. In: *ACM Comput. Surv.* 55.9 (Jan. 2023), 194:1–194:33. DOI: 10.1145/3561048 (cit. on p. 127).

[97] Y. Chen et al. *Reasoning Models Don’t Always Say What They Think*. arXiv:2505.05410 [cs]. May 2025. DOI: 10.48550/arXiv.2505.05410 (cit. on p. 128).

[98] E. Tjoa and C. Guan. “A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI”. In: *IEEE Transactions on Neural Networks and Learning Systems* 32.11 (Nov. 2021), pp. 4793–4813. DOI: 10.1109/TNNLS.2020.3027314 (cit. on p. 128).

[99] H. Dehghani et al. “Near infrared optical tomography using NIRFAST: Algorithm for numerical model and image reconstruction”. en. In: *Communications in Numerical Methods in Engineering* 25.6 (2009), pp. 711–732. DOI: 10.1002/cnm.1162 (cit. on p. 129).

[100] B. Parvitte et al. “Quantitative simulation of photoacoustic signals using finite element modelling software”. en. In: *Applied Physics B* 111.3 (May 2013), pp. 383–389. DOI: 10.1007/s00340-013-5344-2 (cit. on p. 129).

[101] S. L. Jacques. “Optical properties of biological tissues: a review”. en. In: *Physics in Medicine & Biology* 58.11 (May 2013), R37. DOI: 10.1088/0031-9155/58/11/R37 (cit. on p. 129).

[102] N. Rauch and M. Harders. “Methods for User-Controlled Synthesis of Blood Vessel Trees in Medical Applications: A Survey”. In: *IEEE Access* 13 (2025), pp. 22300–22313. DOI: 10.1109/ACCESS.2025.3536998 (cit. on p. 130).

[103] R. Tao et al. “Tutorial on methods for estimation of optical absorption and scattering properties of tissue”. In: *Journal of Biomedical Optics* 29.6 (June 2024), p. 060801. DOI: 10.1117/1.JBO.29.6.060801 (cit. on p. 130).

[104] Baumgartner C et al. *Tissue Properties Database V5.0*. 2025. DOI: 10.13099/VIP21000-05-0 (cit. on p. 130).

[105] IPASC. *PACFISH*. original-date: 2020-02-14T12:22:24Z. Jan. 2025 (cit. on p. 130).

[106] IPASC. *standardised-image-reconstruction*. original-date: 2021-09-02T09:01:23Z. June 2025 (cit. on p. 130).

[107] D. Martin et al. "Effect of skin tone on the accuracy of the estimation of arterial oxygen saturation by pulse oximetry: a systematic review". In: *British Journal of Anaesthesia* 132.5 (May 2024), pp. 945–956. DOI: 10.1016/j.bja.2024.01.023 (cit. on p. 131).

[108] T. R. Else et al. "Effects of skin tone on photoacoustic imaging and oximetry". In: *Journal of Biomedical Optics* 29.S1 (Dec. 2023), S11506. DOI: 10.1117/1.JBO.29.S1.S11506 (cit. on p. 131).

[109] M. Knopp et al. "Shortcut learning leads to sex bias in deep learning models for photoacoustic tomography". en. In: *International Journal of Computer Assisted Radiology and Surgery* 20.7 (July 2025), pp. 1325–1333. DOI: 10.1007/s11548-025-03370-9 (cit. on p. 131).

[110] P. Godau et al. "Navigating prevalence shifts in image analysis algorithm deployment". In: *Medical Image Analysis* 102 (May 2025), p. 103504. DOI: 10.1016/j.media.2025.103504 (cit. on p. 131).

[111] A. Ramesh et al. *Hierarchical Text-Conditional Image Generation with CLIP Latents*. arXiv:2204.06125 [cs]. Apr. 2022. DOI: 10.48550/arXiv.2204.06125 (cit. on p. 132).

[112] L. Zhang, A. Rao, and M. Agrawala. "Adding Conditional Control to Text-to-Image Diffusion Models". en. In: 2023, pp. 3836–3847 (cit. on p. 132).

[113] H. Chang et al. "Muse: Text-To-Image Generation via Masked Generative Transformers". en. In: *Proceedings of the 40th International Conference on Machine Learning*. PMLR, July 2023, pp. 4055–4075 (cit. on p. 132).

[114] V. v. Veldhuizen et al. *Foundation Models in Medical Imaging – A Review and Outlook*. arXiv:2506.09095 [eess]. June 2025. DOI: 10.48550/arXiv.2506.09095 (cit. on p. 132).

[115] L. Zhang et al. "Generative AI enables medical image segmentation in ultra low-data regimes". en. In: *Nature Communications* 16.1 (July 2025), p. 6486. DOI: 10.1038/s41467-025-61754-6 (cit. on p. 132).

[116] L. Hacker et al. "Criteria for the design of tissue-mimicking phantoms for the standardization of biophotonic instrumentation". en. In: *Nature Biomedical Engineering* 6.5 (May 2022), pp. 541–558. DOI: 10.1038/s41551-022-00890-6 (cit. on p. 134).

[117] L. Hacker et al. "A Copolymer-in-Oil Tissue-Mimicking Material With Tuneable Acoustic and Optical Characteristics for Photoacoustic Imaging Phantoms". In: *IEEE Transactions on Medical Imaging* 40.12 (Dec. 2021), pp. 3593–3603. DOI: 10.1109/TMI.2021.3090857 (cit. on p. 134).

[118] V. Grasso et al. “Development of a morphologically realistic mouse phantom for pre-clinical photoacoustic imaging”. en. In: *Medical Physics* 50.9 (2023), pp. 5757–5771. DOI: 10.1002/mp.16651 (cit. on p. 134).

[119] R. Ragunathan et al. “Direct 3-D printing of complex optical phantoms using dynamic filament mixing”. en. In: *Scientific Reports* 15.1 (Mar. 2025), p. 9705. DOI: 10.1038/s41598-025-94390-7 (cit. on p. 134).

[120] L. Maier-Hein et al. “Metrics reloaded: recommendations for image analysis validation”. en. In: *Nature Methods* 21.2 (Feb. 2024), pp. 195–212. DOI: 10.1038/s41592-023-02151-z (cit. on p. 135).

[121] W. Choi et al. “Open-source implementation and systematic comparison of image reconstruction algorithms for linear array transducers”. In: *Photons Plus Ultrasound: Imaging and Sensing 2024*. Vol. PC12842. SPIE, Mar. 2024, PC128421B. DOI: 10.1117/12.3001124 (cit. on p. 135).

[122] J. Gröhl et al. *Digital twins enable full-reference quality assessment of photoacoustic image reconstructions*. May 2025. DOI: 10.48550/arXiv.2505.24514 (cit. on p. 135).

[123] A. Breger et al. *PhotIQA: A photoacoustic image data set with image quality ratings*. en. July 2025 (cit. on p. 135).

[124] M. Schellenberg et al. “Photoacoustic image synthesis with generative adversarial networks”. In: *Photoacoustics* 28 (Dec. 2022), p. 100402. DOI: 10.1016/j.pacs.2022.100402 (cit. on p. 137).

ACKNOWLEDGEMENTS

Die letzten (fast) fünf Jahre, in denen ich meinen Ph.D. gemacht habe, waren eine verrückte Zeit, in denen ich nicht nur fachlich viel über mein Projekt und mein Thema gelernt habe, sondern auch sehr viel über mich. An dieser Stelle möchte ich mich bei allen Menschen bedanken, die mich während dieser Zeit auf verschiedenste Arten unterstützt haben.

Zu aller erst möchte ich mich bei Lena Maier-Hein bedanken, die mir erst die Chance gegeben hat, Teil des wundervollen Teams IMSY sein zu dürfen, das sie über die Jahre aufgebaut hat. Vielen Dank, Lena, für Deine kontinuierliche Unterstützung, für das Vertrauen, das Du mir jederzeit entgegen gebracht hast, für all die tollen Reisen, die ich als Teil meines Ph.D.'s machen durfte, und einfach für die nicht nur fachlich sondern einfach auch menschlich exzellente Betreuung! Dein schierer Enthusiasmus für Forschung hat mich angesteckt und mich für meine bisherige und auch zukünftige Arbeit motiviert.

Vielen Dank auch an Jürgen Hesser, meinen Betreuer der Physik Fakultät. Die TAC Meetings und Feedback Gespräche waren super hilfreich und haben mir immer wieder auch einen neuen Blickwinkel auf einige Themen geöffnet.

Next, I would like to thank Sarah Bohndiek, who hosted me in her amazing lab for about half a year. During that time, I got to learn so much and meet the amazing people in her team, with whom I made new friends.

Ein Riesendank gilt Janek Gröhl, der nicht nur Betreuer meiner Masterarbeit war, sondern auch über die komplette Zeit meines Ph.D.'s immer offen für Fragen war und mir Feedback gegeben hat, wann auch immer ich es gebraucht habe. Danke Dir, Janek, dass Du durch deine entspannte Art und Deine unnachgiebig positive Einstellung mir nicht nur einmal Sorgen nehmen konntest und mir immer wieder auch empathischen Rat gegeben hast.

Ein weiterer Dank gilt natürlich auch Alexander Seitel, dem Group Lead des Photoakustik Teams, der durch seine riesen Erfahrung in der Wissenschaft an allen Ecken und Kanten tatkräfig angepackt hat, auch wenn es mal um nervige Aufgaben ging. Deine Alex-Hours waren immer mal wieder eine tolle Abwechslung, um entweder etwas Dampf abzulassen, Dinge zu fragen, die man sonst vielleicht nicht so direkt fragen würde, oder einfach ein chilliges Gespräch.

Das gesamte Photoakustik Team ist einfach absolute weltklasse und ich wäre in keiner anderen Gruppe lieber gewesen. Das geht an alle aktuellen und vergangenen Mitglieder, insbesondere auch Thomas Kirchner, der mich bei meiner Bachelor Arbeit betreut hat und mich von Anfang an für Photoakustik begeistern konnte. Aus unserem Büro möchte ich mich auch nochmal bedanken bei Niklas Holzwarth als mein (fast von Beginn an) treuer Photoakustik Begleiter, Melanie Schellenberg, die eine unglaubliche Wärme ins Team gebracht hat, Jan-Hinrich Nölke, Tom Rix, Christoph "CJ" Bender, und Marcel Knopp. Ihr seid die besten Office- und random Gesprächsbuddies, die ich mir hätte wünschen können. Vielen lieben Dank für den "alle für einen und einer für alle"-Spirit!

A big thank you goes out to the whole IMSY team! The time I had during lunch and coffee breaks, and all sorts of social events was amazing and I can definitely say that this is a team with an overwhelming team spirit. I love talking to every single one in the team and forcing my stupid puns on everyone is so much fun (mostly for me). I've met many amazing people that I can truly call my friends. Special thank you to Leo for all the decafferonis that were so much fun! At this point, also special thanks to the Mädels in the office who supported me and all of us, even though I'm sure we can be a bit annoying at times, so thank you for keeping up with us!

Auch nochmal lieben Dank an alle, die meine Thesis Korrektur gelesen haben: Alex, Janek, Flo, Leo, Niklas, Jan und Tom.

Vielen lieben Dank an alle meine Freunde, ob in Heidelberg, Boston (Flo), oder aus dem wunderschönen Saarland, die mich überall immer warm und willkommen fühlen lassen. Für all die schöne Tage und Abende, in denen ich mal den Stress zur Seite schieben und einfach entspannen konnte!

Zu guter letzt, möchte ich mich von ganzem Herzen bei meiner Familie bedanken, die immer für mich da war und ohne die ich das alles so auf jeden Fall nicht geschafft hätte. Mama und Papa, vielen Dank dafür, dass ihr mich in allem und jederzeit unterstützt habt, mir den Rücken frei gehalten habt und dass ihr mir mein Studium überhaupt erst ermöglicht habt! Ich weiß das sehr zu schätzen! Kim, vielen Dank dafür, dass du ein großartiger Bruder bist und sicher gehst, dass ich immer geerdet bleibe. Tanja und Oma Hilde, vielen lieben Dank, für die hunderten Kerzen, die ihr angemacht habt und dass ihr buchstäblich jeden Tag an mich denkt!

Vielen Dank!

Disclosure

During the writing of this work, I used the search engines *Google Scholar*, *Semantic Scholar*, and *ChatGPT Deep Research* as research tools in order to identify related existing literature and assess their relevance. I further used the large language models (LLMs) *OpenAI's ChatGPT 4o* and *DeepL Write* for linguistic and grammatical improvements. The prompts I used were variants of:
"For the following text, improve the clarity, conciseness, and scientific tone. Use academic language and avoid overly complex phrasing, but ensure precision and objectivity. Do not change the meaning or content of the original text."
Any generated content by these services has been carefully reviewed and verified.

Data-driven synthesis of photoacoustic images
Ph. D. Thesis

Supervised by Prof. Dr. Lena Maier-Hein
Prof. Dr. Jürgen Hesser

This work has been set using LaTeX and KOMA-Script.