

Review article

## Deep learning for biomedical photoacoustic imaging: A review

Janek Gröhl<sup>a,b</sup>, Melanie Schellenberg<sup>a</sup>, Kris Dreher<sup>a,c</sup>, Lena Maier-Hein<sup>a,b,d,\*</sup>

<sup>a</sup> German Cancer Research Center, Computer Assisted Medical Interventions, Heidelberg, Germany

<sup>b</sup> Heidelberg University, Medical Faculty, Heidelberg, Germany

<sup>c</sup> Heidelberg University, Faculty of Physics and Astronomy, Heidelberg, Germany

<sup>d</sup> 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  $\mu_a$ , the local light fluence  $\phi$ , and the temperature-dependent Grüneisen parameter  $\Gamma$ . 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  $\theta$  (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  $\mu_a$  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](mailto: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

2213-5979/© 2021 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

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  $\mu_a$ ). 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

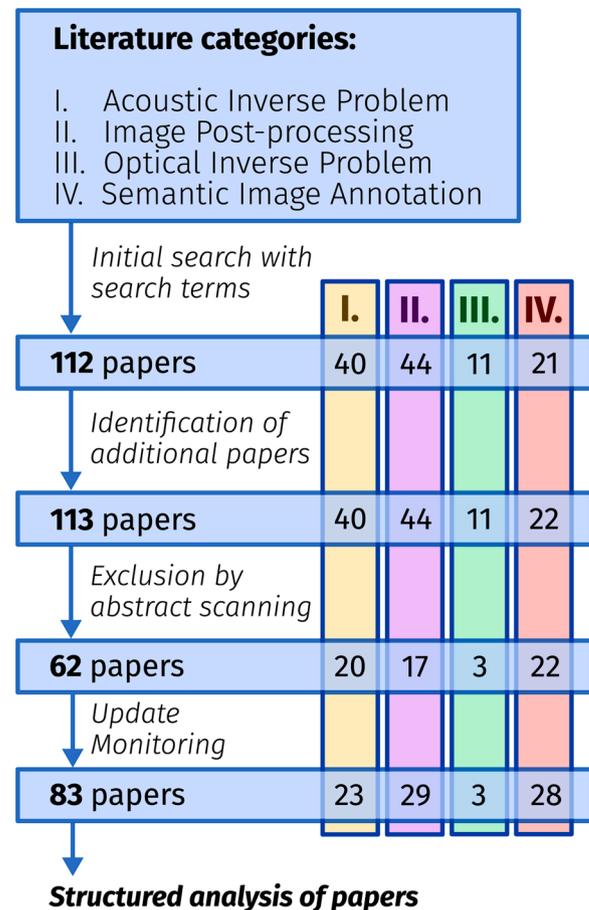


Fig. 1. Overview of the literature review algorithm. First, potentially fitting papers are identified based on an initial search. The search results are complemented by adding additional papers found by other means than the search engines, and finally, non-relevant papers are excluded by removing duplicates 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 pre-determined four areas, the papers were arranged into thematic sub-categories (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

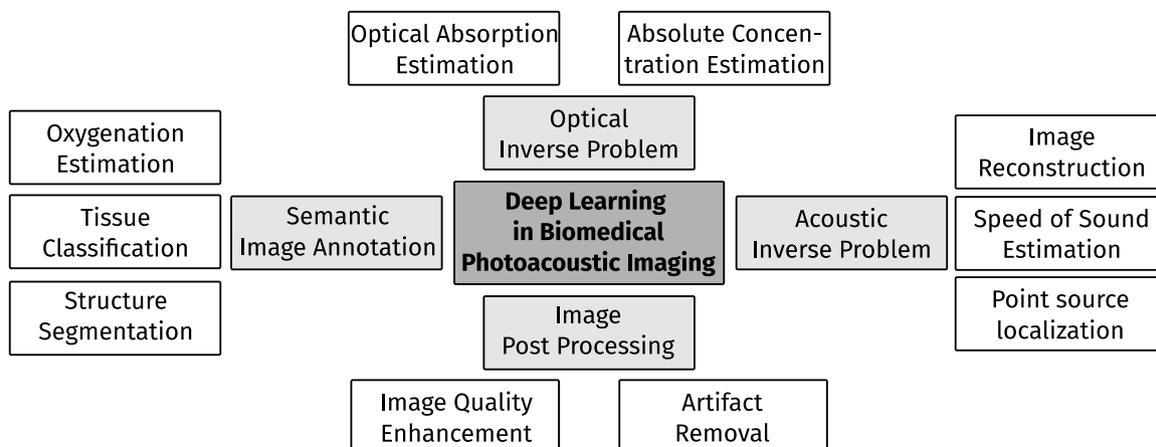
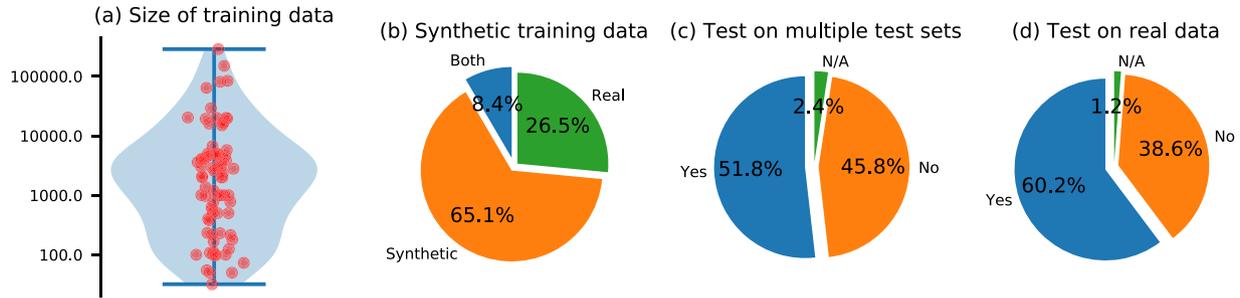


Fig. 2. Overview over the topical foci of current research towards applying deep learning algorithms to problems in biomedical PAI.

### Training and Testing Practices



**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

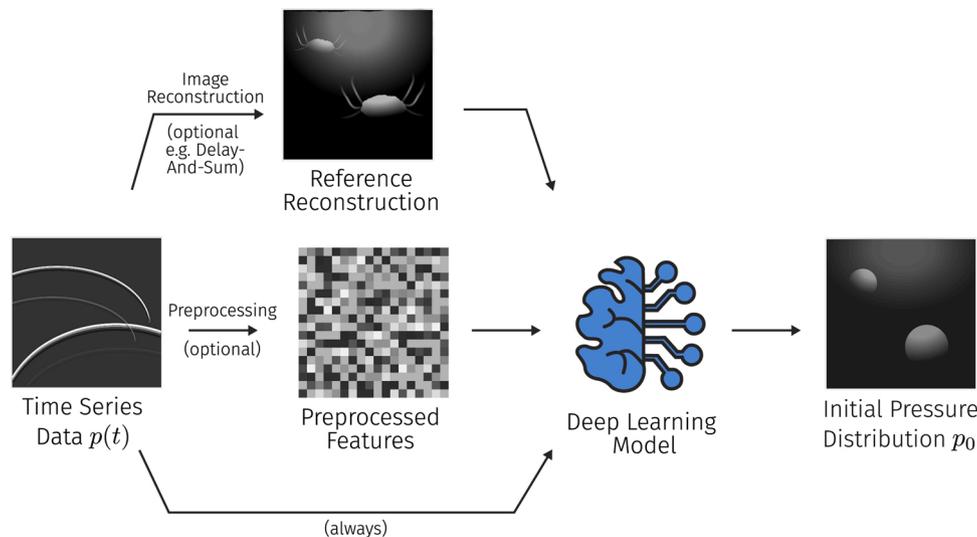
et al. [77] demonstrated the possibility to share the network weights across iterations by using recurrent inference machines for image reconstruction.

*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-rafterd 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.

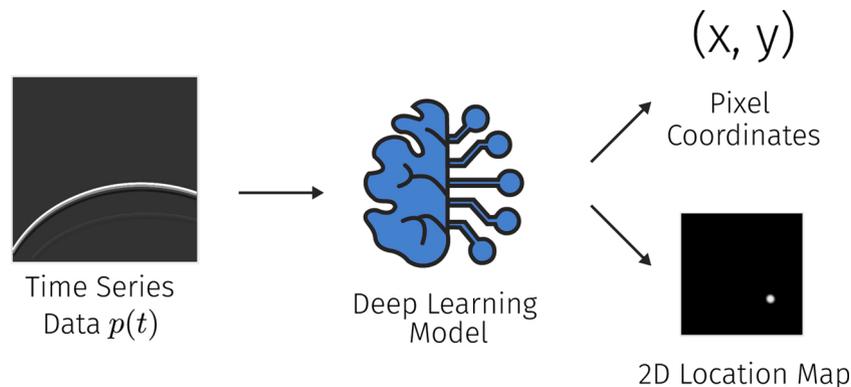
### 5. Image post-processing

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.

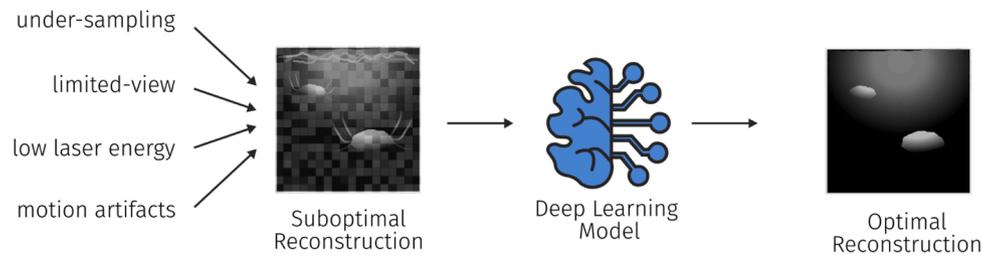
The evaluation metrics that were used to assess the quality of the post-processing results and their relative frequencies are shown in Table 4.

#### 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.

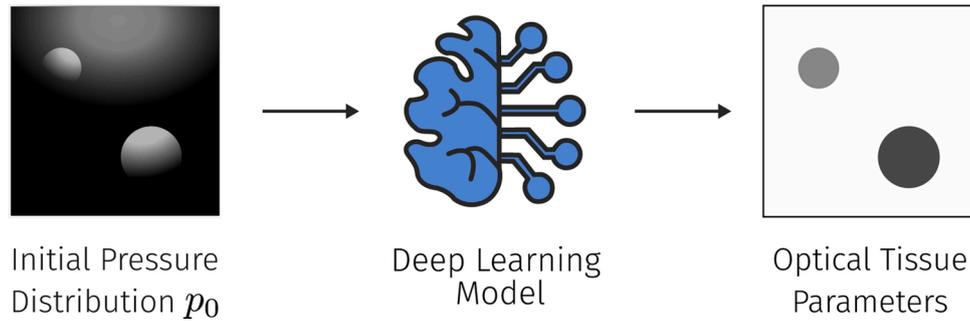
For the identified deep learning-based approaches, the key objective is to estimate the optical absorption coefficients and subsequently the absolute concentrations of chromophores from the initial pressure distribution (see Fig. 7).

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<sub>2</sub>. 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 $\mu_a$ [ $\text{cm}^{-1}$ ]	Background $\mu_a$ [ $\text{cm}^{-1}$ ]	Background $\mu'_s$ [ $\text{cm}^{-1}$ ]	Validation data
Cai et al. [57]	U-Net with residual blocks	N/A	0.2–0.4	5–10	<i>In silico</i>
Gröhl et al. [124]	U-Net	2–10	Const. 0.1	Const. 1.5	<i>In silico</i>
Chen et al. [125]	U-Net	0.1–2	0.1–0.4	Const. 10	<i>In vitro</i>



**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  $p_0$ .

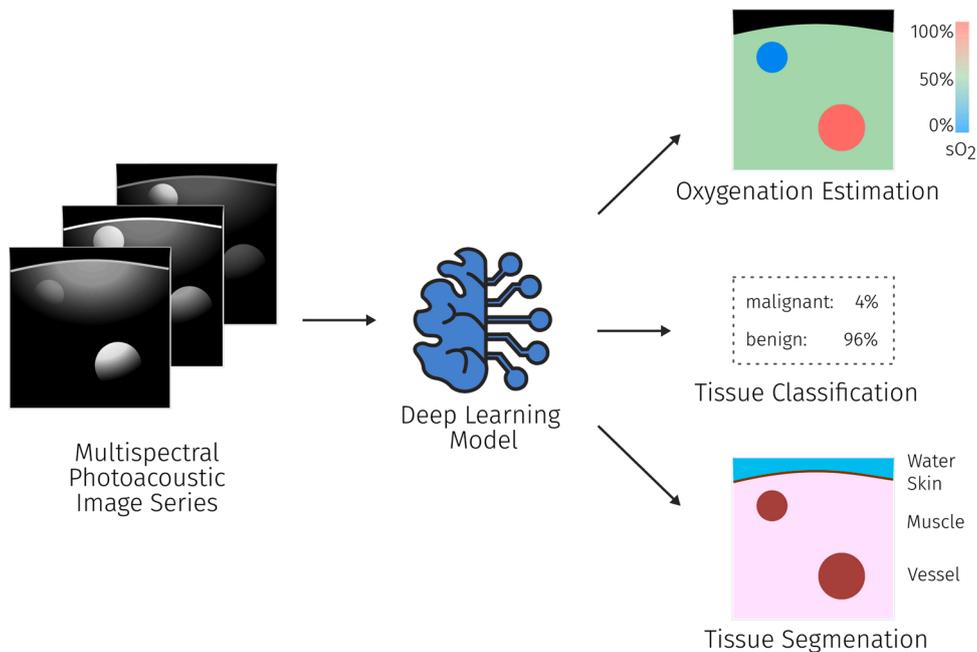
They also presented a method to estimate the expected error of the inversion, yielding an indicator for the model uncertainty. Their approach was trained and tested on simulated data, which contained tubular structures in a homogeneous background. Finally, Chen et al. [125] trained a U-Net to directly 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 ( $\approx 800$  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

**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% $\pm$ 4.5% MAE
Cai et al. [57]	0.8% $\pm$ 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% $\pm$ 0.2% MRE
Yang and Gao [130]	4.8% $\pm$ 0.5% MAE
Nölke et al. [128]	2%, IQR [0%, 4%] to 9%, IQR [3%, 19%] MedAE*

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].

### 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 hyperosteoegeny 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 Program through 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 proof-reading 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.

- [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-Urbe, 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. Karpouk, 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. Yalavarth, 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  $\ell^1$ -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 photoacoustic 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 photoacoustic mesoscopy using artificial intelligence, *Med. Phys.* 46 (9) (2019) 4046–4056.
- [141] B. Lafci, E. Mercep, S. Morscher, et al., Efficient segmentation of multi-modal photoacoustic and ultrasound images using convolutional neural networks, *Photons Plus Ultrasound: Imaging and Sensing 2020*, vol. 11240 (2020) 112402N.
- [142] S. Nitkunanantarajah, K. Haedicke, T.B. Moore, et al., Three-dimensional photoacoustic 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 hyperosteoegony 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 photoacoustic 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. Kennigott, 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. Lings, 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.



**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.



**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.



**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.