Navigating challenges and solutions in quantitative photoacoustic imaging 🅕 ⊘

Ruochong Zhang (iD) ; Rabia'tul A'dawiah (iD) ; Tristan Wen Jie Choo (iD) ; Xiuting Li (iD) ;
Ghayathri Balasundaram (iD) ; Yi Qi (iD) ; Yonggeng Goh (iD) ; Renzhe Bi ✉ (iD) ; Malini Olivo ✉ (iD)

🔴 Check for updates

🌐 View Online ↗ Export Citation

Articles You May Be Interested In

A prototype hand-held tri-modal instrument for *in vivo* ultrasound, photoacoustic, and fluorescence imaging

Rev. Sci. Instrum. (March 2015)

A review of methods for solving the optical molecular tomography

J. Appl. Phys. (April 2023)

Applied Physics Reviews

AIP Publishing

Navigating challenges and solutions in quantitative photoacoustic imaging Ⓕ

View Online   Export Citation   CrossMark

Ruochong Zhang,[1,2] ⓘ Rabia'tul A'dawiah,[1] ⓘ Tristan Wen Jie Choo,[1] ⓘ Xiuting Li,[1] ⓘ Ghayathri Balasundaram,[1] ⓘ Yi Qi,[1] ⓘ Yonggeng Goh,[3] ⓘ Renzhe Bi,[1,2,a)] ⓘ and Malini Olivo[1,a)] ⓘ

AFFILIATIONS

[1]A*STAR Skin Research Labs (A*SRL), Agency for Science, Technology and Research (A*STAR), 31 Biopolis Way, Singapore 138669
[2]National Semiconductor Translation and Innovation Center, Singapore
[3]Department of Diagnostic Imaging, National University Hospital, 5 Lower Kent Ridge Road, Singapore 119074

[a)]Authors to whom correspondence should be addressed: Bi_renzhe@asrl.a-star.edu.sg and Malini_olivo@asrl.a-star.edu.sg

ABSTRACT

Photoacoustic imaging, an emerging modality that seamlessly combines advantages of optical absorption contrast and ultrasound resolution, holds great promise for noninvasive imaging of biological tissues. Its applications span across diverse fields, such as dermatology, oncology, cardiology, and neurology. However, achieving accurate image reconstruction and physiological parameters quantification from raw photoacoustic signals presents a significant challenge. This challenge primarily arises from the inherent heterogeneity of tissues, encompassing variations in optical fluence and acoustic properties. In addition, incomplete information acquired from a limited view also leads to artifacts, image distortions, and reduced spatial resolution. Furthermore, robust spectral unmixing approach is another key step to restore the initial biochemical components' distribution with complex or unknown background absorption. To overcome these hurdles, researchers have proposed numerous state-of-the-art techniques, aiming to improve the accuracy and reliability of quantitative photoacoustic imaging (qPAI) in heterogeneous tissue. This review aims to comprehensively overview recent developments over the past decade, for addressing four main challenges frequently encountered in qPAI: limited-view reconstruction, acoustic heterogeneity, optical fluence fluctuations, and robust spectral unmixing, which serves as a reference for readers seeking to understand the specific challenges and corresponding solutions in this field.

TABLE OF CONTENTS

I. INTRODUCTION

Photoacoustic (PA) effect, or optoacoustic (OA) effect, is a physical phenomenon that was initially discovered by Alexander Graham Bell in 1880.[1] This phenomenon involves absorption of light energy by a material, resulting in temporary thermal expansion and the subsequent generation and transmission of acoustic waves. The PA effect forms the basis of photoacoustic imaging (PAI), a cutting-edge biomedical imaging technique that seamlessly merges the strengths of optical and ultrasound imaging. Leveraging the distinctive ability of tissues to absorb light as endogenous contrast and subsequently emit ultrasonic signals, PAI provides non-invasive and real-time insights into biological samples at various scales—from microscopic cellular details[2,3] to macroscopic whole-organ imaging.[4–6] The harmonization of optical and acoustic modalities not only offers exceptional spatial resolution and penetration depths unattainable by conventional optical methods but also possesses comprehensive functional information, with the potential to revolutionize fields such as cancer diagnostics, neuroimaging, and vascular visualization. Utilizing mathematical models and signal processing techniques on measured PA signals enables quantitative analysis of physiological parameters such as tissue oxygenation, blood flow, and concentrations of specific molecules or contrast agents, which is known as quantitative PAI (qPAI). There are mainly three variants of PAI catered for different applications: PA microscopy (PAM), provides detailed visualization of biological samples at the microscale level up with millimeter depth, making it particularly useful for studying cellular structures, microvasculature, and superficial tissues;[7] PA tomography (PAT), which is tailored to provide rapid imaging at greater depths up to a few centimeters, at the cost of coarser resolution than PAM. It is particularly useful in tumor, cardiovascular, and brain imaging;[8] PA endoscopy (PAE), which is a specialized form of PAI designed for internal organ examination, employing miniaturized probes to capture PA signals within body cavities. It facilitates applications in esophageal, gastrointestinal, urogenital, and intravascular imaging.[9]

PA technique has gained significant attention in the past decade, as evidenced by the rapidly growing number of publications (Fig. 1). According to the search results from Google Scholar, there was a notable surge in research studies related to PA techniques in the past decade, showing a substantial increase in 187% from 2013 to 2023. PA techniques can be categorized into two groups. One is PA spectroscopy (PAS), which involves the use of a PA cell for sensing purpose. This is a well-established field and has been studied for nearly a century. The other is PAI, which has emerged and attracted more attention than PAS recently. In 2023, the majority of PA studies, accounting for 83%, were focused on imaging, with the remaining percentage dedicated to PAS. Specifically for PAI, the upward trend is more pronounced, with an increase in 245%. When narrowing the focus to clinical PAI, the increase becomes even more conspicuous, reaching approximately 329%. Among various PAI modalities, PAM has received the most extensive attention, accounting for 57%, followed by PAT at 36%. PAE is comparatively less explored, constituting only 7% of the PAI studies. In the realm of clinical applications of PAI in 2023, the primary focus lies in breast and brain studies, comprising 30% and 28% of the research, respectively. Skin-related investigations represent the third most significant area, constituting 24%. Other areas, including thyroid-, prostate-, and liver-related studies, collectively account for 18% of the research landscape. This trend clearly indicates the evolving role of PAI, transitioning from a primarily research-oriented domain to an emerging clinical imaging tool. This shift highlights its increasing importance in medical diagnostics, suggesting it could significantly influence healthcare practices.

Numerous review papers on PAI have been published, covering various aspects. For instance, Attia et al.,[10] Karlas et al.,[11] Manohar et al.,[12] and others[13–15] delved into clinical applications of PAI. Jeon et al.,[16] Yao et al.,[7] and Omar et al.[17] reviewed state-of-the-art PAM
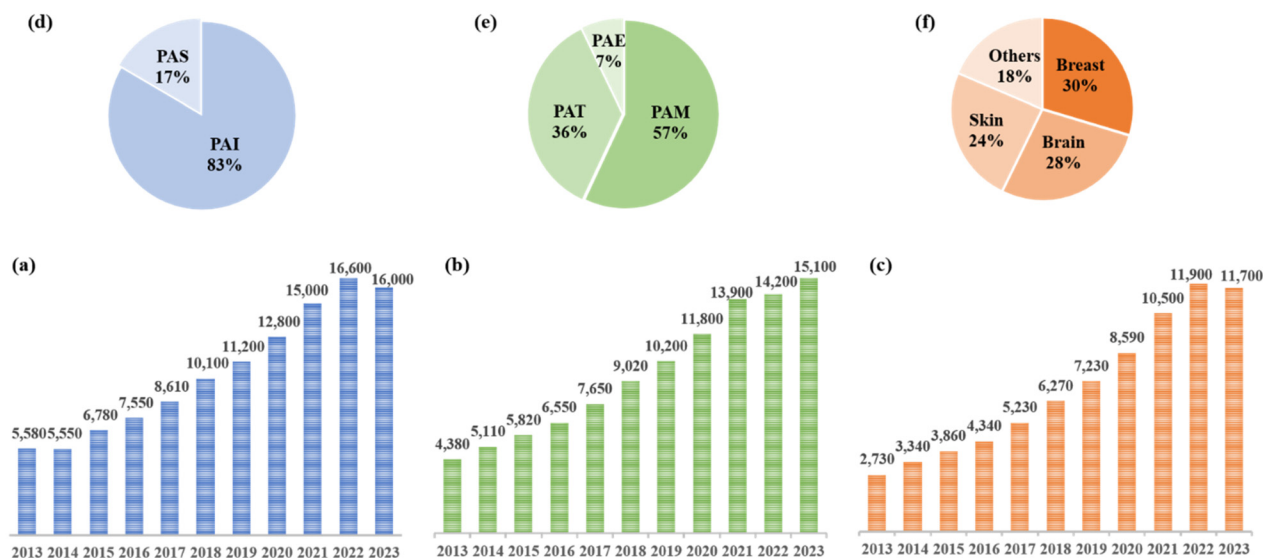
FIG. 1. Research landscaping for PA from 2013 to 2023 based on the search results in Google Scholar. Histograms of the number of publications related to PA, PAI, and clinical PAI by searching (a) "photoacoustic OR optoacoustic." (b) "(photoacoustic OR optoacoustic) AND (imaging OR microscopy OR mesoscopy OR tomography OR endoscopy)." (c) "(photoacoustic OR optoacoustic) AND (Imaging OR microscopy OR mesoscopy OR tomography OR endoscopy) AND (clinical OR human OR patient)." (d) Pie chart representing the percentage of PAI and PA spectroscopy (PAS). (e) Pie chart of different PAI modalities. (f) Pie chart of various applications of clinical PAI.
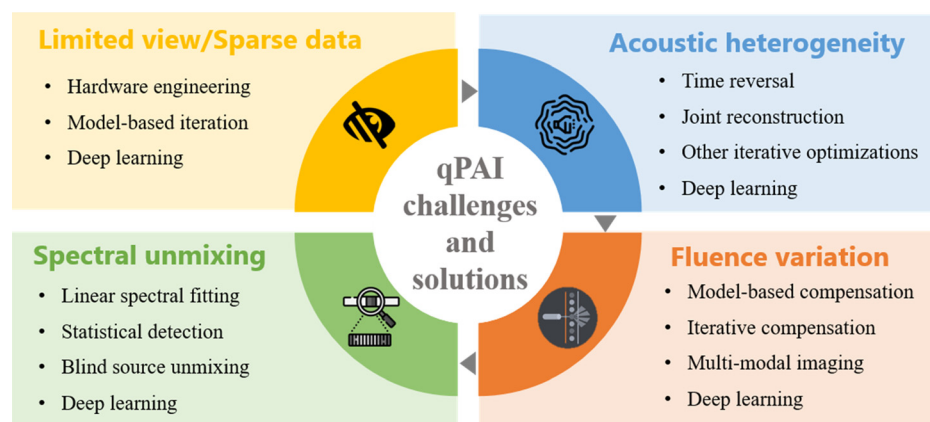
FIG. 2. Illustrative overview outlining the scope of this review.

systems. Wang *et al.*,[18] Taruttis and Ntziachristos,[19] and others[20–22] provided comprehensive reviews of different PAT approaches. Some reviews also explored the synergy between PAI and machine learning,[23–25] along with other general reviews.[26,27] However, few of them focused on the challenges associated with qPAI in complex biological tissues and the corresponding strategies to address them, apart from Lutzweiler and Razansky[28] and Cox *et al.*[29] 10 years ago. In fact, the process of acquiring vital physiological or pathological information from biological tissues through qPAI involves the integration of spatial and spectral information, addressing both acoustic and optical inversion problems. The intermediate goal of qPAI is to reconstruct initial acoustic pressure from the measured PA signals, informing the location and distribution of optical absorbers within the tissue. The ultimate goal is to derive optical absorption coefficients across different wavelengths, decoding the composition, distribution, and concentration of these absorbers. By simultaneously considering both aspects, qPAI offers more comprehensive and detailed information about the structure-functional information of tissues. Quantitative PAI presents several engineering and algorithmic hurdles, spanning from signal acquisition to the interpretation of PA data. This current review primarily focuses on recent advancements (2013–2023) aimed at overcoming the specific challenges encountered in qPAI from four main aspects, as shown in Fig. 2. To begin, we delve into the limited-view problem during data acquisition, reviewing the corresponding solutions that have been developed to compensate for incomplete signals. Afterward, we investigate the acoustic heterogeneity matter in acoustic inversion and the innovative mitigation strategies. Subsequently, fluence compensation methods employed in optical inversion are reviewed. We then review the recent exploration in direct reconstruction techniques. Furthermore, we thoroughly examine recent PA spectral unmixing algorithms, which play a critical role in qPAI to restore chromophores mapping. At the end of each section, a table is appended, summarizing the corresponding solutions in the literature to each specific challenge. Finally, a conclusion is drawn, highlighting future perspectives and the need for further advancements and refinements to fully realize the potential of qPAI in clinical diagnostics.

II. RECONSTRUCTION WITH LIMITED-VIEW OR SPARSE DATA

The principles of PAI, especially PAT, have been explained in details by many scholars.[8,28,30] Briefly, visible or near-infrared light is used to generate localized heat within tissues. Various chromophores absorb energy at certain wavelengths, creating tissue expansion and pressure waves. Subsequently, the pressure waves are measured by broadband acoustic detectors and reconstructed at each wavelength to retrieve the initial pressure distribution. After that, spectral unmixing is done to decipher contributions of individual chromophores, enabling the key capability of qPAI for noninvasive diagnosis. When taking absorber as a point source, the resultant pressure wave spreads spherically. Hence, in ideal tomographic imaging situations, the object should be surrounded by detectors placed along a closed surface, establishing a full-view condition. In practice, this condition is not easy to be satisfied, considering system complexity and practicability. In most cases, detectors only cover certain angle of tissues and the data are under-sampled. When the tomographic coverage is $<\pi$, the reconstructed PA images can be distorted and incomplete.[31,32] This situation is often referred to as the "limited-view" problem. Researchers have proposed various solutions, from both engineering and algorithmic perspectives.

A. Hardware engineering approaches

Limited-view issues with conventional ultrasound linear array probe can be mitigated by adding a rotational scanning platform.[33] Either the probe or the object can be rotated to increase the viewing angle. Li *et al.* introduced a multiview Hilbert transform specifically designed for this rotational approach to address the PA signal bipolarity problem across multiple angles.[34] Shu's group utilized two physical linear array probes with adjustable positions in the same imaging plane to fit target geometry, where the relative positions of the two transducers are pre-calibrated by ultrasound, one probe transmitting and the other probe receiving.[35] Huang *et al.* employed single linear array transducer with an acoustic reflector made of glass, positioned at a 45° angle, as a virtual detector.[36] This arrangement results in a system comprising two linear arrays oriented perpendicular to each other, effectively doubling the detection angle without additional transducers or scanning. They further extended the system by incorporating two acoustic reflectors positioned 120° apart and tripled the detection view.[37]

While hardware engineering approaches have made notable strides in mitigating the limited-view problem in PAI, they still exhibit certain limitations. For instance, the rotational scanning scheme can

be time-consuming and not suitable for real-time or dynamic imaging. Additionally, the physical arrangement of additional transducers or reflectors not always be feasible, particularly in clinical settings. Moreover, these engineering solutions often involve more complex hardware configurations, leading to increased overall cost of the system.

### B. Model-based iterative reconstructions

Beyond hardware improvement, image reconstruction algorithms also play an important role in addressing limited-view issues. Model-based iterative reconstruction is one of the computational techniques, which involves iteratively refining an initial image using mathematical models that describe how light and sound interact within biological tissues, aiming to improve the accuracy and resolution of reconstructed images. In general, the model-based iterative algorithms follow the format as shown in Eq. (1), with various regularizers and optimization techniques. For instance, Liu et al. derived an iterative adaptive weighted filtered back projection (IAWFBP) approach for single-element transducer circular scanning scheme.[38] By employing an imaged-based adaptive weighting factor, the discrepancy between measured and computed signals gradually decreased with each successive iteration, until the image converged. Normalized mean absolute error (NMAE), defined as $\delta_i = \|x_i - x\|_2 / \|x\|_2$, was used as an evaluation index in simulation study and iterative stop criteria in ex vivo experiment. The improvement was more than 50% compared to conventional filtered back projection (FBP). Han et al. proposed to combine L1 norm and total variation (TV) cost functions as regularization,[39] which surpassed conventional Tikhonov (Tik-Lap), L1, or TV regularization by having sharper features and less artifacts, indicated by lower root mean square difference (RMSD), higher signal-to-noise ratio (SNR), and contrast-to-noise ratio (CNR). Another group proposed an interlaced sparse sampling (ISS) method for ring-shaped multispectral PAT image reconstruction,[40,41] as shown in Fig. 3. This method involves a unique scanning-based image acquisition scheme, where sparse signal sets were acquired from n non-overlapping angles at n different illumination wavelengths to form a synthetic dense dataset. A high-quality prior image was then reconstructed using synthetic data. Next, a model-based iterative reconstruction was applied, incorporating a regularization term based on prior anatomy,

$$x = \arg\min\left\{\frac{1}{2}\|p - \mathbf{W}x\|_2^2 + \beta J(x,a)\right\}, \quad (1)$$

where x is the PA image and p represents raw data. $\|\ \|_2$ is the L2 norm. $\mathbf{W}$ is the forward model matrix. $\beta \geq 0$ and $J(x,a)$ is the regularization term, obtained by a non-local means (NLM) regularization.[40] This special term only reflects structural similarity between priori (a) and target (x), while minimizing intensity effect on PA images of individual wavelength,

$$J(x,a) = \sum_{j=1}^{N} \sum_{k \in n_j} w_{kj} \sqrt{1 + \left(\frac{x_j - x_k}{\delta}\right)^2 + \left(\frac{a_j - a_k}{\eta}\right)^2} - 1, \quad (2)$$

where w_{kj} measures the similarity between pixels k and j. δ and η are the experimental parameters. Adaptive gradient decent algorithm was then employed to solve for x. Quantitative performance metrics include mean of structural similarity (MSSIM), mean square error
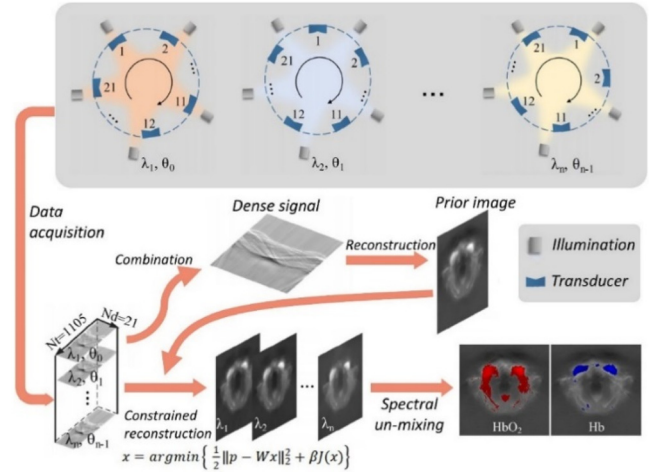
FIG. 3. Process illustration of ISS-PAT. λ_n: nth wavelength; θ_i: rotation angle; Nd: the number of transducer elements; Nt: the number of acquired photoacoustic samples in the time domain for each transducer channel; HbO$_2$: oxyhemoglobin; Hb: deoxyhemoglobin. Reproduced with permission from Li et al., Comput. Methods Programs Biomed. **214**, 106562 (2022). Copyright 2021 Elsevier.

(MSE), and Dice-coefficient (Dice). Compared to the conventional sparse sampling (SS) scheme, ISS achieved much closer performance to the reference obtained by the dense sampling method, indicating its outstanding capability to reduce system cost and complexity. To mitigate high computational burden, the same group proposed a directional total variation regularization (ISS-dTV) method where $J(x,a) = \sum_{j=1}^{N} |(I - \xi_j \xi_j^*)\nabla x_j|$ and ξ_j encodes the structural information of anatomical priori.[41] The reconstruction time reduced by 96% compared to ISS-NLM, and the performance was also significantly improved compared to ISS without anatomical priori (ISS-un).

Apart from above-mentioned methods for circular geometry, several other groups also explored widely used linear detection geometry. Shang et al. proposed a sparsity-based PAI reconstruction (SPAIR),[42] implemented by a two-step iterative shrinkage/thresholding (TwIST) algorithm with actual measurement of impulse responses.[43] The algorithm outperformed conventional back projection (BP), frequency-domain (FD) and time-reversal (TR), and L$_2$ norm minimization (LSQR) reconstructions in phantom test, indicated by averaged mean square error (AMSE) and SNR. Ai et al. evaluated variance-reduced stochastic gradient descent (VR-SGD) iterative model-based algorithm in various detection geometries formed by linear array transducers. The performance was compared to delay and sum (DAS) and evaluated by SNR, CNR and generalized CNR (gCNR), demonstrating its effectiveness in recovering absent structures and minimizing artifacts due to limited-view problem.[44] In addition, for planar detection geometry, Zhu and coworkers recently evaluated various non-iterative and iterative reconstruction algorithms, including time reversal (TR), non-negativity constrained TR (TR+), TR with TV denoising (TR+TV), non-negativity constrained iterative least squares (iLS+), non-negativity constrained TV regularization (TV+), and Bregman iteration enhanced TV+ (TV+Breg).[45] SNR and signal-to-artifact ratio (SAR) were used as evaluation metrics. They concluded that iterative reconstruction methods (iLS+, TV+, TV+Breg) outperformed non-iterative ones (TR, TR+, TR+TV) in general, but the selection of

iterative reconstruction needs to consider specific requirement and constraints. Different iterative reconstruction methods offer varying trade-offs between factors, such as computational complexity, accuracy, and robustness.

C. Machine-learning-based methods

Machine learning approaches in PAI reconstruction have seen significant development in recent years. They have emerged as a versatile and powerful tool with a multitude of applications. They play a pivotal role and have shown promise in recovering details of reconstructed images, particularly in situations where limited-view or sparse data pose challenges. Recently, Wang and coworkers conducted an extensive review on different machine-learning approaches to mitigate limited-view problems,[46] which includes three main aspects: training data preparation, learning strategy, and different network architectures. In this sub-section, we will recap some representative works and highlight their contributions in PAI reconstruction, followed by a brief summary.

Davoudi and team proposed a deep convolutional neural network (CNN) based framework to enhance image quality using limited-view data.[47] A U-Net network was trained to record streak-type artifacts or residual originated from sparse data, and the refined image was obtained by subtracting the learned residual from artifactual image. They also underlined the importance of training with real *in vivo* data obtained from a full-view scanner, instead of phantom or simulated data. To achieve this, they developed a full-view PAT system to acquire high-quality images as ground truth, which consists of 512 elements arranged in a 360° circle around the target. Data obtained from 85 and 192 elements, corresponding to 60° and 135° coverage, were used to evaluate the network architecture's performance under limited-view

conditions. *In vivo* mice imaging demonstrated its capability to rectify distortions and recover structures even with sparse data, which outperformed conventional universal back-projection (UBP). Two years later, the same group further improved the network by incorporating both reconstructed images and time-domain sinograms into a combined U-Net CNN with two encoders.[48] Cross-sectional images on human fingers were evaluated by peak SNR (PSNR), SSIM, MSE, and normalized root mean square error (NRMSE), showing that combined U-Net can better define the vascular structures compared to previous single-input U-Net. Vu *et al.* proposed a Wasserstein generative adversarial network with gradient penalty (WGAN-GP) for the linear-array transducer to reduce limited-view artifacts.[49] It showed improvements in resolution and noise reduction, compared to time-reversal and U-Net. In addition to the post-learning methods, Hauptmann *et al.* presented a learned iterative reconstruction method.[50] Instead of learned post-processing after initial reconstruction, this method took into account the physical forward model in the learning process, where regularizer was learned from the data by deep gradient descent (DGD). ELCAP Public Lung Image Database was used as a training dataset and applied to *in vivo* human palm imaging. The results were compared to U-Net, TV, and non-negative least square (NNLS) reconstructions and demonstrated that learned iterative reconstruction provided better image quality than learned post-processing, but at the cost of longer training time. Guan *et al.* proposed another direct learned method named pixel-wise deep learning (Pixel-DL),[51] which interpolates pixel-wise wave propagation data first, before sending it to CNN as an input for subsequent reconstruction. They illustrated different learning approaches, including Pixel-DL, Post-DL, and Direct-DL in Fig. 4. Moreover, they quantified the performance of time-reversal (TR), iterative reconstruction with TV constraint, Post-DL, and Pixel-DL by
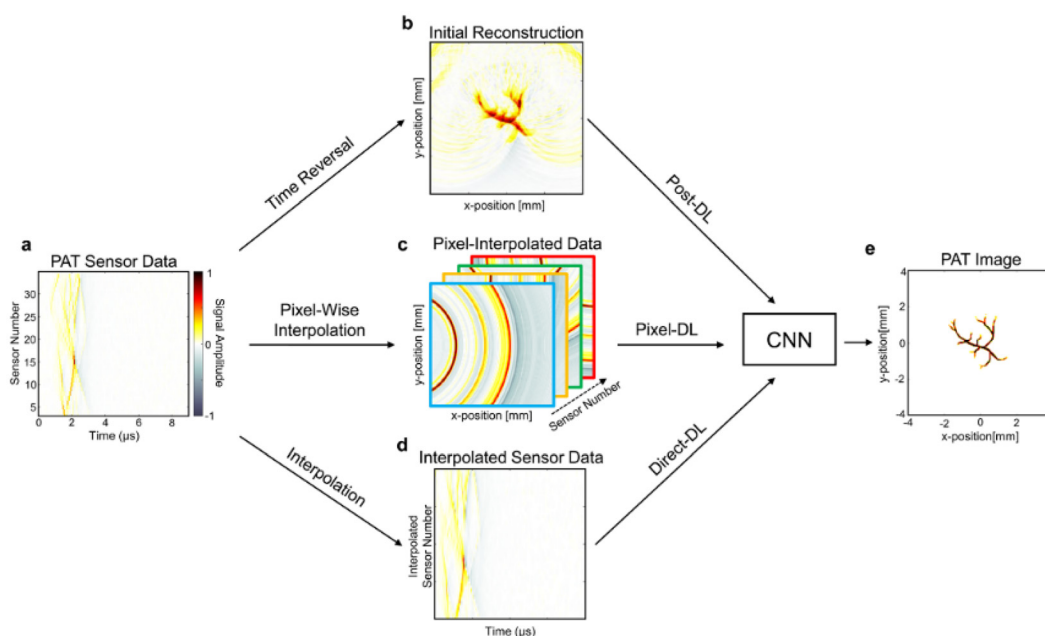
FIG. 4. Comparison of different learning strategies for PAT image reconstruction. Reproduced with permission from Guan *et al.*, Sci. Rep. **10**(1), 8510 (2020). Copyright 2020 Springer Nature.

PSNR and SSIM. The simulation and phantom results indicated that Pixel-DL achieved superior performance to TR and Post-DL methods, albeit slightly inferior to iterative reconstruction. Nevertheless, Pixel-DL exhibited significantly shorter computational time, being faster than iterative reconstruction by a factor of 1000.

The status of machine learning in PAI reconstruction is evolving rapidly. The above-mentioned studies have demonstrated the potential of these approaches to produce high-quality images with reduced artifacts. However, challenges remain, including the need for large and diverse training datasets, lack of high-quality *in vivo* ground truth, optimization of network architectures, and the integration of prior information to guide the reconstruction process effectively.

Sometimes, limited-view reconstruction aiming to improve spatial resolution may alter spectral information at certain image pixel to achieve better spatial definition. Therefore, preservation of spectral information needs to be considered to guarantee the accuracy of qPAI. Hardware engineering approaches provide more comprehensive PA signals and reduce limited-view artifacts by increasing the coverage angles. Model-based reconstructions tend to minimize the discrepancies between the measured and computed signals through a data fidelity term consisting of a forward operator and a regularization term. Zhu *et al.*[45] indicate that retaining feature visibility, minimizing noise, and maximizing amplitude accuracy need to be balanced via careful selection of regularization parameter λ. If amplitude accuracy is priority in the case of qPAI, a smaller λ is beneficial. ISS method[40,41]

maximizes structural similarity between priori and target and has a constraint to minimize intensity effect on PA images of individual wavelengths. In machine learning approaches, different strategies are adopted. For example, *in vivo* data from full-view scanner were used for training and validation to remove artifacts.[47,48] The high-quality ground truth serves as reliable benchmark, potentially enhancing spectral preservation at different locations. Hauptmann *et al.*[50] learn suitable priori constraints of forward operator during training to restore features and intensity, although the performance depends on the quality of training data. Despite some works not explicitly mentioning spectral information preservation, most of the novel reconstruction methods in Sec. II have successfully improved the image quality in terms of SSIM, PSNR, and MSE, indicating faithful spatial representation of spectral information.

Table I provides a summary of recent strategies addressing the challenges posed by limited-view or sparse data in PAI reconstruction. Further elaboration on model-based and machine-learning based reconstructions can be found in Tables II and III, respectively.

III. ACOUSTIC COMPENSATION IN HETEROGENEOUS TISSUES

The propagation of ultrasound waves, $p(r, t)$ at any location $r \in \mathbb{R}^3$ and $t \geq 0$, resulting from thermal expansion triggered by laser pulse excitation, necessitates their traversal through biological tissue before being detected by transducers. The significance of sound speed

TABLE I. Summary of different methods to address limited-view PAI reconstruction.

Reference	Methods		Transducer, detection geometry	Study type
33	Engineering	Rotational scanning	Linear array, circular	*Ex vivo* (mice)
34		Rotational scanning, Multiview Hilbert transform	Linear array, circular	*In vivo* (zebra fish)
35		Two probes at adjustable angles	Linear array	*In vivo* (human hand)
36		One acoustic reflector at 45°	Linear array	*Ex vivo* (mouse ear)
37		Two acoustic reflectors at 120°	Linear array	Phantom (leaf skeleton)
38	Model-based iterative algorithms	IAWFBP	Single-element, circular scan	*Ex vivo* (pork liver)
39		TV-L1	270° circular array	*In vivo* (mice)
40		ISS-NLM	270° circular array	*In vivo* (mice)
41		ISS-dTV	270° circular array	*In vivo* (mice)
42		SPAIR+TwIST	Linear array	Phantom (graphite)
44		VR-SGD	Linear array, L-shape, parallel	Phantoms
45		iLS+, TV+, TV+Breg	Fabry–Pérot planar scanner	*In vivo* (human hand)
47	Machine-learning	Post-processing, U-Net	Full-ring array	Training: full-view *in vivo* (mice) Validation: limited-view *in vivo* (mice)
48		Post-processing, combined U-Net	Full-ring array	Training: full-view *in vivo* (human finger) Validation: limited-view *in vivo* (human finger)
49		Post-processing, WGAN-GP	Linear array	Training: *in silico* (brain vessels) Validation: *in vivo* (mouse brain)
50		Learned iterative reconstruction, DGD	Fabry–Pérot planar scanner	Training: *in silico* (lung vessels) Validation: *in vivo* (human palm)
51		Pixel-DL, FD-UNet	Semi-circular array	Training: *in silico* (vasculature) Validation: *in silico* (vasculature)

TABLE II. Model-based iterative reconstructions with a limited view. Φ is a three-level Daubechies four-wavelet transform. See Eq. (2) and Sec. II B for $J(x,a)$. A is the sparse operator: identity matrix. $\mathfrak{R}$ encodes a priori knowledge.

Reference	Regularizations	Evaluation metrics	Improvement	Compared to
38	$\cdots$	NMAE	>50%	FBP
39	$\beta_1\lvert x\rvert_{TV} + \beta_2\lVert\Phi x\rVert_1,$ $\beta_i \geq 0$	RMSD	3%–15%	Tik-Lap, TV, L1
		SNR	0.4–0.6 dB	
		CNR	5%–34%	
40	$\beta J(x,a), \beta \geq 0$	MSSIM	27%	Sparse sampling
		MSE	64%	
		Dice	117%	
41			46%	ISS-un
			94%	
			430%	
		Time	96%	ISS-NLM
42	$\beta\lVert Ax\rVert_1, \beta \geq 0$	AMSE	13%–43%	BP, FD, TR, LSQR
		SNR	~30 dB	
44	$\beta\lvert x\rvert_{TV}, \beta \geq 0$	CNR	4–20 dB	DAS
		SNR	4–28 dB	
		gCNR	3%–40%	
45	$\beta\mathcal{R}(p_0)$	SNR	0–200 dB	TR
		SAR	10–20 dB	

in accurate PA image reconstruction cannot be overstated, as the time-of-flight (TOF) directly influences the calculated position of the object to be reconstructed. Most image reconstruction algorithms available operate under the assumption of acoustically homogeneous biological tissues, wherein the speed of sound (SoS) remains constant. In most image reconstruction algorithms, the manual selection of the SoS is a common practice, leading to deviations from its actual value. These deviations can, in turn, introduce image distortions and artifacts in the

reconstructed images. The heterogeneity of biological tissues plays a crucial role in these variations of the SoS, challenging the accuracy of the constant speed assumption often employed in such algorithms.

Acoustic attenuation should also be considered for acoustic compensation methods. Some of the acoustic energy generated from PA is lost through acoustic absorption while propagating through the tissue to the surface. Acoustic absorption can occur at both the molecular and cellular levels through chemical and molecular relaxations or through thermal conduction and viscous relative motion, respectively.[52] Acoustic energy dispersion can also occur through scattering. Depending on the frequency of the acoustic waves, scattering is more significant at higher frequencies. Acoustic attenuation, in general, abides by a frequency power law, $\alpha(\mathbf{r}, f) = \alpha_0(\mathbf{r})\omega^y$, affecting imaging sensitivity.[53] The attenuation coefficient α is affected by the temporal angular frequency (ω) and the frequency-independent attenuation coefficient term α_0 in dB MHz^{-y} cm^{-1}. The power law exponent y is usually in the range of 0–3.[53] This section reviews numerous investigations that have probed the impact of tissue heterogeneity on acoustic signals and suggested potential solutions in the past decade.

A. Iterative optimization

1. Time-reversal methods

The time-reversal (TR) method presents a compelling approach to PA image reconstruction, offering distinct advantages and considerations. The essence of TR algorithms is based on the unchanged wave propagation physical process even when time is reversed (i.e., wave equation invariance under the time transformation of $t \rightarrow -t$). In this method, the numerical acoustic propagation model is typically employed to back-propagate the forward-propagating field or its normal derivative in a temporal manner.[54] TR algorithm methods have proven to be highly versatile, capable of mitigating effects from acoustic heterogeneities while requiring fewer assumptions. This flexibility allows for its application across various detector distributions and detection geometries. One notable strength of the TR approach lies in

TABLE III. Machine-learning-based approaches for limited-view reconstructions.

Reference	Purpose of network	Input	Output	Evaluation metrics	Improvement	Compared to
47	Eliminate streak-type artifacts	Primary reconstructed images	Reconstructed images with reduced artifacts	SSIM	~2%–250%	UBP
48	Enhance structures' visibility	Primary reconstructed images, time-domain sinograms		PSNR	~5 dB	U-Net
				SSIM	~10%	
				MSE	~70%	
				NRMSE	~67%	
49	Remove artifact	Artifactual images		SSIM	0%–5%	U-Net
				PSNR	−0.1–0.8 dB	
50	Learn physical forward model	Primary reconstructed images		Time	~20–160 s slower	U-Net, TV, NNLS
				PSNR	~1%–5%	
				SSIM	~0.6–3.7 dB	
				Rel l^2-error	~6%–35%	
51	Correct artifacts	Pixel-interpolated data		PSNR	−4–11 dB	TR, iterative (TV), post-DL
				SSIM	−1%–378%	
				Time	0–489 s faster	

its ability to effectively correct for SoS variations within the imaging medium. Leveraging the TR property of waves, the method exhibits the capacity to refocus signals to their original emission locations, thereby enhancing image accuracy in the presence of moderate acoustic heterogeneities. The method's efficacy may diminish in the presence of pronounced acoustic heterogeneities, compromising its compensation capabilities and leading to artifacts and reduced accuracy. Furthermore, the TR approach is sensitive to model inaccuracies, necessitating precise knowledge of acoustic wave propagation and medium properties for optimal results.

A study conducted by Treeby et al.[55] showed that TR method can regulate resolution, SNR, and image degradation. They proposed a time-domain reversal method to account for acoustic attenuation. The iterative reconstruction algorithm employs a backward-in-time solving model for the following equations:

$$\frac{\partial}{\partial t}\boldsymbol{u}(\boldsymbol{r},t) = -\frac{1}{\rho_0(\boldsymbol{r})}\nabla p(\boldsymbol{r},\mathrm{t}), \qquad (3)$$

$$\frac{\partial}{\partial t}\rho(\boldsymbol{r},t) = -\rho_0(\boldsymbol{r}).\boldsymbol{u}(\boldsymbol{r},t), \qquad (4)$$

$$p(\boldsymbol{r},t) = c_0(\boldsymbol{r})^2\left\{1 - \tau(\boldsymbol{r})\frac{\partial}{\partial t}(-\nabla^2)^{\frac{y}{2}-1} - \eta(\boldsymbol{r})(-\nabla^2)^{\frac{y+1}{2}-1}\right\}\rho(\boldsymbol{r},t), \qquad (5)$$

where $\boldsymbol{u}(\boldsymbol{r},t)$: acoustic particle velocity; $c_0(\boldsymbol{r})$: speed of sound; $\rho(\boldsymbol{r},t)$: acoustic density; $\rho_0(\boldsymbol{r})$: ambient density; and $p_m(t)$: pressure data measured at $\boldsymbol{r}_m$ with maximum time recorded indicated as T.

The variables $\tau(\boldsymbol{r}) = -2\alpha_0 c_0(\boldsymbol{r})^{y-1}$ and $\eta(\boldsymbol{r}) = 2\alpha_0 c_0(\boldsymbol{r})^y \tan\left(\frac{\pi y}{2}\right)$ describe the acoustic absorption and dispersion, respectively. The acoustic equations (3)–(5) abide by the following initial and boundary conditions:

$$p(\boldsymbol{r},t)|_{t=T} = 0, \quad \boldsymbol{u}(\boldsymbol{r},t)|_{t=T} = 0, \quad p(\boldsymbol{r}_m,t) = p_m(t). \qquad (6)$$

Equation (5) incorporates distinct provisions for acoustic attenuation and dispersion, achieved through the utilization of two lossy derivative operators rooted in the fractional Laplacian. This approach aligns seamlessly with the power law equation. During image reconstruction, the sign of the absorption proportionality coefficient was reversed, while the dispersion remained unchanged. The solution to the equations yields an attenuation-compensated image at $t = 0$ by using the fast Fourier transform (FFT) k-space pseudospectral method.[56] This reconstruction algorithm is primarily assessed within media characterized by a constant acoustic attenuation coefficient (α_0) and the power law exponent y.

Huang et al.[57] refined the k-space pseudospectral algorithm developed by Treeby and Cox[56] to allow its use with heterogeneous media by removing the k-space adjustment parameter κ in Eq. (15) in Ref. 56 and setting the acoustic attenuation coefficient α_0 to be a spatially dependent quantity, $\alpha_0(\boldsymbol{r})$. The power law exponent must be assumed a constant value for the k-space TR method. The power law exponent must be assumed a constant value for the k-space TR method. The assumption is only valid if the object is made of soft tissue. The reconstruction method must be modified in cases where objects are composed of bones and soft tissues, and the power law exponents are significantly distinct to avoid image distortions. In their study, Huang et al.[57] demonstrated that acoustic attenuation effects

can be mitigated by considering only the attenuation coefficients of the most attenuating component (e.g., bone), while assuming the attenuation coefficients of other media are negligible.

In addressing complex compensation scenarios, an iterative TR approach, such as the Neumann-series method, has been proposed.[58,59] This TR variant iteratively refines the reconstruction, gradually improving compensation for stronger heterogeneities. While offering enhanced compensation capabilities, the iterative approach introduces increased computational complexity, which may affect real-time applications. In general, although suitable for certain compensation scenarios, TR method's implementation complexity and potential limitations in addressing depth-dependent variations should be carefully considered in the context of specific imaging setups and desired compensation levels.

In 2020, Lai et al.[60] developed a novel image reconstruction approach for imaging heterogeneous media using the TR algorithm. The finite difference time domain (FDTD) method was integrated into the TR algorithm. The iterative reconstruction was done when varimax norm reached minimum. Numerical simulations and one ex vivo experiment on prostate tumor were conducted to test the efficacy of the algorithm. The proposed FDTD-based TR algorithm was able to reconstruct the tumor image and was able to distinguish between malignant from fibrous tissue. This method has shown to overcome the fuzzy image problem brought on by SoS changes in non-homogeneous tissues that is faced by conventional image reconstruction algorithms. However, this method has a slightly longer reconstruction compared to traditional algorithms. Despite the longer reconstruction time, the performance of this algorithm encourages further exploration of its potential for tumor detection in clinical applications.

2. Joint reconstruction of initial pressure and speed-of-sound

In PA image reconstruction, the SoS distribution $c(\mathbf{r})$ usually can be estimated concurrently with the absorbed optical energy density A $(\mathbf{r})$ or initial pressure $p(\mathbf{r},0)$ in tissues, where $p(\mathbf{r},0) = \Gamma A(\mathbf{r})$ and Γ is known Grueneisen coefficient. This technique is referred to as joint reconstruction (JR). Huang et al.[61] developed an iterative algorithm that alternatively updated A$(\mathbf{r})$ and SoS distribution, $c(\mathbf{r})$. There are two sub-problems: (1) reconstruct $\mathbf{A}$ with given $\mathbf{c}$ and (2) reconstruct $\mathbf{c}$ with given $\mathbf{A}$, which are concurrently solved with the initial estimates of $\mathbf{A}^{(0)}$ and $\mathbf{c}^{(0)}$. They conducted a series of numerical experiments and systematically analyzed each sub-problem independently. They indicated that sub-problem (2) is generally non-convex and needs to avoid pitfalls of local minima. The study investigated the effects of spatial supports, relative spatial bandwidths of A$(\mathbf{r})$ and c$(\mathbf{r})$, as well as perturbation of A$(\mathbf{r})$, and offered insights into when accurate JR is feasible. They also emphasized the importance of appropriate regularization. This investigation illustrated that achieving precise JR remains challenging even with impeccable measurement data. Yet, it could enhance the precision of reconstructed $\mathbf{A}$, when compared to assuming a uniform SoS. Recommendations for future research are provided, emphasizing the importance of incorporating experimental data.

Cai et al.[62] introduced a novel feature coupling (FC) method for the JR problem. The method employed finite-dimensional parameterization and utilized a correlation-based optimization approach to iteratively refine the SOS distribution with an initial coarse estimation based on temperature. A clearer initial pressure (IP) image and a more

precise SOS map were achieved by maximizing the similarity of half-ring reconstructions, mitigating feature splits in the full-ring image. The study demonstrated the efficacy of FC through numerical simulations, phantom experiments, and *in vivo* studies, showcasing improved accuracy and sharpness in reconstructed images that were quantified by mean relative error (MRE). However, the study acknowledges certain limitations. The computational burden of FC is relatively high, requiring further exploration of GPU acceleration for efficiency. Additionally, the optimization problem for FC is non-concave, necessitating the use of the gradient ascent with momentum algorithm to avoid local optima. Challenges arise with large structures, impacting the sensitivity of the similarity-based method. The study also notes potential issues with the determination of the initial value for the iteration, especially in multi-SoS problems. Furthermore, the blind segmentation approach is computationally demanding and may require additional techniques for reliable estimation. The impact of the study lies in its significant contributions to photoacoustic imaging, offering a robust method for simultaneous reconstruction of SoS and IP.

Addressing the challenges posed by SoS heterogeneity, Deng *et al.*[63] reviewed the development of four IP-SoS JR methods for ring-array photoacoustic tomography systems (PACT)—namely, multisegmented feature coupling (MSFC), signal compensating (SC), feature coupling (FC), and adaptive PACT (APACT). These methods were purposefully crafted to tackle the challenges posed by SoS heterogeneity in biological tissues. Similar to FC, MSFC entails the separation of the ring arrays into subarrays, with FC dividing them into two subarrays and MSFC requiring eight subarrays. Image reconstruction involved utilizing the independent signals detected by each subarray, resulting in two images for FC and eight images for MSFC. The correlation coefficients between these images were calculated to measure their similarities, with the aim of maximizing these values. Both FC and MSFC exhibited high computational requirements along with

superior IP images and SoS distributions. However, MSFC produced more uniform IP reconstructions and smoother SoS distributions compared to FC. The SC method in this context harnesses the features of electronic impulse response (EIR). By superimposing the negative and positive peaks of the EIR, the lowest intensity of the reconstructed image is achieved. Minimizing the intensity of the reconstructed image allows for the estimation of inhomogeneity-induced wavefront distortion. Fourier transform is subsequently applied to the reconstructed image to analyze the relationship between image intensity and delay time. The computational requirement of SC is moderate, but the resulting SoS map lacks the desired smoothness when compared to the gold standard. Finally, the APACT method employs adaptive imaging techniques. It estimates the distortion of the wavefront due to inhomogeneity in the frequency domain, and the SoS distribution is then estimated from the wavefront distortion. APACT holds an advantage over the FC approach in scenarios with limited *priori* information on the SoS distribution, as seen in *in vivo* experiments. However, it has a very high computational complexity. Their findings were summarized in a table in Ref. 63.

3. Other iterative methods

Haltmeier and Nguyen[64] delved into various iterative methods as an alternative to TR algorithms for solving the inverse problem in PAT in free space. Their approach involved modifying adjoint operator formulations to improve convergence speed and error estimates. The stability analysis demonstrated that their chosen adjoint operator facilitated linear convergence rates in both L2-norm and H1-norm. The study provided insights into convergence speed and appropriate weight selection for mapping spaces from a microlocal perspective. The numerical results showcased the superiority of Nesterov's fast gradient and conjugate gradient methods over Landweber's and iterative TR methods in both visible and invisible cases. The investigation
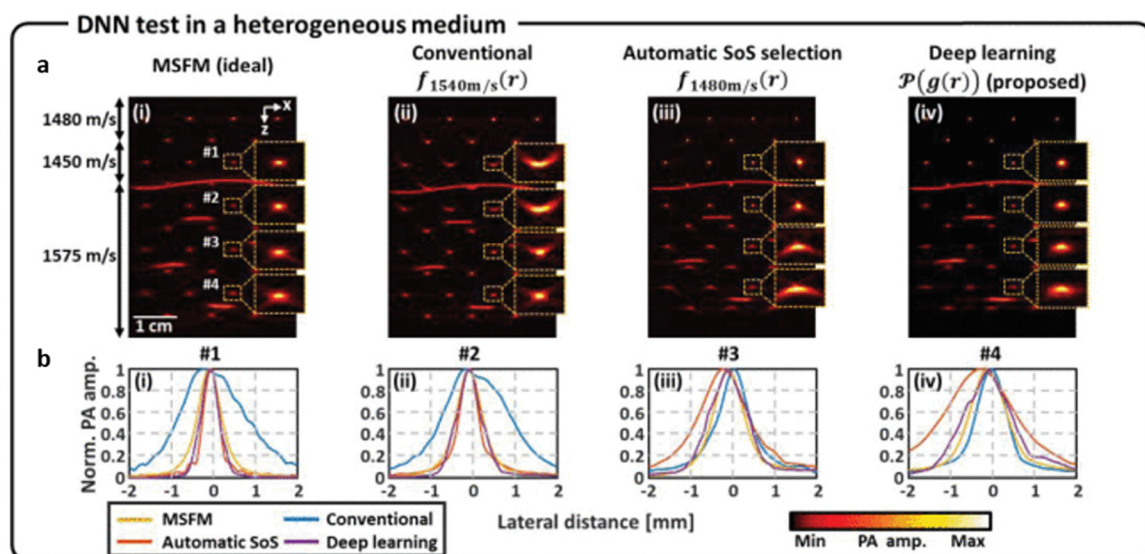
FIG. 5. Performance comparison of different PA reconstruction algorithms in three-layer heterogeneous digital phantom. (a) Imaging results of multi-stencil fast marching with known SoS map (i), conventional beamforming (ii), automatic SoS selection (iii), and proposed SegU-net methods (iv). (b) Lateral resolutions comparison of imaging results in (a). Reprinted with permission from Jeon *et al.*, IEEE Trans. Image Process. **30**, 8773–8784 (2021), licensed under Creative Commons License.

expanded to PAT with variable sound speed, considering full and partial data scenarios. In most cases, Landweber's method performed comparably to iterative TR, while Nesterov's and conjugate gradient methods exhibited faster convergence. The conjugate gradient method displayed semi-convergence for all noisy data, with Landweber's and iterative TR methods being asymptotically slower, particularly for full data. The authors recommended Landweber's and Nesterov's methods for regularization, presenting them as alternatives to TR methods.

Poudel et al.[65] proposed variable temporal data truncation (VTDT) approach as an update to their previous half-time reconstruction method, aiming to reduce distortions.[66] However, a drawback of this algorithm is its dependency on prior information about acoustic heterogeneities to guide the reconstruction process. The adaptive nature of temporal truncation effectively suppresses artifacts arising from acoustic heterogeneity. The results from their study demonstrated superior image quality compared to the previous half-time method. Nevertheless, one limitation of the VTDT-based reconstruction algorithm is the potential for poor image quality if the size of heterogeneity is substantial or if numerous heterogeneities are present. Despite this limitation, the VTDT approach remains valuable for clinical image reconstruction, particularly in samples with a single prominent acoustic heterogeneity.

In another study led by Huang et al.,[67] a pioneering image reconstruction method for PACT in acoustically heterogeneous media effectively overcame limitations inherent in current TR approaches. The study utilized a discrete imaging model based on the exact photoacoustic wave equation, incorporating matched forward and back projection operators to address challenges, such as data incompleteness and noise. A significant advancement was observed with the full-wave iterative (FWI) reconstruction approach in PACT for media with inhomogeneous speed of sound and density distributions. Notably, the study demonstrated the methodology's superiority to traditional TR methods, highlighting reduced root mean square error (RMSE) in simulations and generating cleaner images, particularly in the presence of errors in the speed of sound and density maps. These findings suggest promising potential for enhancing medical imaging quality in preclinical and clinical PACT applications.

Ranjbaran et al.[68] developed a holistic quantitative photoacoustic tomography (qPAT) framework that compensates for SoS inhomogeneity, acoustic compensation, and optical fluence. The study introduces a novel algorithm designed for the enhancement of PAT in environments characterized by heterogeneous optical scattering media, encircled by a configuration of optical fibers and ultrasonic transducers arranged in a ring formation. The proposed iteratively refined wavefield reconstruction inversion (IR-WRI) employed an extended formulation of full-waveform inversion utilizing the alternating direction method of multipliers, coupled with fluence compensation. Integral to this approach is the incorporation of the SoS map as prior information, facilitating the classification and identification of regions exhibiting uniform optical absorption coefficients on the SOS map. The SOS map is acquired through an additional FWI procedure applied to acoustic data generated and recorded by transducers. The method was verified by simulation using a neonatal head consisting of extra-cranial tissue (ECT), cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). Comparative analysis reveals the algorithm's significantly enhanced performance compared to conventional time-reversal algorithms.

The hybrid transmission-reflection optoacoustic ultrasound (TROPUS) was introduced by Merčep et al.[69] TROPUS enabled non-invasive mapping of various tissue parameters, including absorbed optical energy, acoustic reflectivity, SoS, and acoustic attenuation (AA), with submillimeter resolution in live mice. SoS obtained by transmission ultrasound tomography was used to enhance PA image reconstruction. Real-time rendering, facilitated by GPU-based spatial compounding and bent-ray-tracing iterative reconstruction algorithms, provided dynamic insights into organ dynamics and molecular processes. The synergistic combination of optoacoustic and ultrasound techniques enhanced imaging performance, facilitating a comprehensive exploration of anatomical, functional, and molecular properties in health and disease. In vivo mouse imaging experiments revealed fine details on organ parenchyma, vascularization, tissue reflectivity, density, and stiffness. The study also emphasized the importance of addressing challenges including spatial resolution trade-offs, potential artifacts, and the need for algorithmic improvements, such as incorporating full reflectivity information for successful clinical integration.

B. Deep-learning-based methods

In addition to the established techniques mentioned above, recent advancements have also explored the integration of machine learning or deep learning methods for acoustic compensation, offering new possibilities for further enhancing image quality and overcoming some of the limitations discussed here. In this section, we delve into the potential to revolutionize acoustic compensation in photoacoustic imaging.

The use of machine learning for photoacoustic imaging, although still in its nascent stage, has been gaining traction in recent years. Data-driven approaches, particularly deep learning (DL), excel at recognizing intricate patterns within extensive datasets to establish the most effective solutions within the parameter space. DL networks are instrumental in post-processing tasks, such as the removal of artifacts and noise from reconstructed images, which ultimately results in cleaner and more precise images. Beyond these tasks, DL networks are proficient in image segmentation, facilitating the delineation of regions of interest within images. Furthermore, they exhibit remarkable prowess in image classification and labeling, enabling the automated categorization of images or specific regions within images, thus contributing to the overall efficiency and accuracy of photoacoustic imaging. However, in this section, we will focus on using DL as an acoustic compensation technique.

Jeon and Kim introduced a DL algorithm specifically to correct for SoS aberrations.[70,71] They trained a hybrid deep neural network (DNN) model named SegU-net with simulation datasets and successfully applied it to three-layer digital phantom and in vivo human forearm PA images. They benchmarked the performance of DNN with conventional beamformer, multi-stencil fast marching (MSFM), and automatic SoS selection. The results are shown in Fig. 5. Note that the MSFM was applied with known SoS distribution, which is impractical. Among the methods, SegU-net was able to correct for the SoS aberrations in a heterogeneous medium, even though it was trained with a homogeneous media dataset. In addition, the trained DNN effectively reduced the presence of side lobes and background noise, leading to a substantial enhancement in the contrast of the PA image. Their work serves as a compelling indicator of the profound potential of DL in addressing critical challenges within photoacoustic imaging, with

implications extending to various facets of this field. Anas *et al.* proposed a five densely connected layers deep convolutional neural network (CNN) that showed statistically significant improvements in the reconstructed images of simulated circular two-dimensional (2D) targets.[72] The CNN model showed robustness against SoS variations.

The study reported by Dehner *et al.*[73] in August 2023 represents one of the latest advancements in the speed of sound correction in photoacoustic imaging through deep neural networks. The DL framework proposed, DeepMB, achieved real-time photoacoustic imaging that supports dynamic adjustments to the speed of sound in the range of 1475–1525 m/s. This allows for real-time acoustic compensation, increasing the reconstructed image focus. DeepMB was trained on photoacoustic signals created from real images. DeepMB was able to generate accurate image reconstructions by comparing the image it produced with the model-based reconstructions of the corresponding signals. However, it is to be noted that knowledge on the speed of sound is required beforehand and the obtention of sound speed requires manual tuning. Despite this limitation, DeepMB boasts image reconstruction at a speed that is approximately 1000 times faster than a typical generic iterative model-based method.

Shi *et al.*[74] introduced a learning-based method to compensate for SoS variations in a PA-ultrasound (PA-US) hybrid system. The investigation was conducted as a simulation study and tested with *ex vivo* data using chicken breast tissue. The DL model was based on a fully convolutional neural network and was trained on 2000 samples to obtain the speed of sound distributions. The training dataset was modeled based on a clinical ultrasound probe and generated from ultrasound simulations in k-Wave. Acoustic attenuation was considered a constant value of 0.5 dB/MHz/cm. The proposed framework effectively mitigated ultrasound aberrations, demonstrating a high similarity index of 0.8267. The success of the DL framework with the chicken breast *ex vivo* data implies its significant potential for application to human samples or even in *in vivo* clinical applications.

The integration of deep learning techniques into acoustic compensation within the realm of photoacoustic imaging is transforming the capabilities in this field. It promises to reduce the reliance on manual adjustments and iterative methods, offering real-time, automated, and highly accurate solutions to correct for the speed of sound variations. As the field continues to evolve, the integration of deep learning techniques promises to revolutionize our approach and usher in a new era of enhanced imaging capabilities, ultimately contributing to a more profound understanding of the complex world beneath the surface.

C. Other acoustic compensation methods

In addition to previously mentioned works, several other methods have been proposed to account for blurring caused by acoustic propagations. For example, in 2014, Deán-Ben and coworkers studied the effects of tissue acoustic heterogeneities on the reconstructed image.[75] This study investigated the errors associated with heterogeneities and how to mitigate these effects. Their study indicated that TOF errors induced from the assumption of straight acoustic rays are insignificant. However, errors stemming from the uniform speed of sound are substantial. They implemented a modified filtered back-projection (FBP) reconstruction by considering the small SoS variation and compared its performance to that assuming uniform SoS. Their study proved the importance of and the need to have acoustic compensation for the SoS in PA image reconstruction.

Pattyn *et al.*[76] developed a novel straight ray-based (SRB) algorithm designed to overcome the challenges posed by SoS heterogeneity. The algorithm calculates iso-TOF contours between each pixel in the imaging grid and among the transducers. This capability, coupled with fluence compensation, offers a comprehensive approach to image reconstruction in PAT. To assess the algorithm's robustness, the experiments were conducted using a tissue-mimicking polyvinyl chloride plastisol (PVCP) phantom containing cysts and blood inclusions. The signals were back-projected onto iso-TOF contours derived from the straight ray approximation. The time-of-flight (TOF) for each pixel on the imaging grid was computed by integrating the slowness function ($\frac{1}{SoS}$) along the ray connecting pixel locations and transducers. The calculated SoS values were then used to correct for time delays. The results demonstrated that without the SoS correction, and relying solely on optical fluence, the image appeared distorted, with the central artifact being amplified. Therefore, the synergistic use of both the SoS correction algorithm developed and the fluence compensation paved the way for more intricate *in vitro* and *ex vivo* studies.

Lin *et al.*[77] ambitiously proposed a hybrid reconstruction method to overcome SoS heterogeneity, combining the traditional delay and sum (DAS) with multi-stencils fast marching (MSFM) methods. This hybrid approach aimed to reduce reconstruction distortions, allowing for precise measurement of the deep structure and position of the sample. The time delay map in conventional DAS reconstruction was determined using the MSFM method, originally designed for tracking expanding wave fronts. This adaptation computed the time delay map at predefined grid spaces by solving the non-linear wave propagation equation along multiple stencils. Selecting the solution with the shortest path enabled accurate time delay map generation, compensating for variable acoustic speed across different structures within the imaging field of view. Emphasizing the critical role of specifying the SoS in PA imaging reconstruction, this study significantly influences image quality and target positioning. For detecting brain tumors with an intact skull, numerical models, phantoms, and sequential *in vivo* PA experiments were introduced. While acknowledging reliance on prior knowledge of SoS distribution, the study recognized spatial variations in real biological tissue, complicating accurate delay time map calculations. Consequently, SoS mapping during *in vivo* PA experiments was roughly estimated through MRI and US imaging.

One of the most efficient methods for PA image reconstruction involves the use of wavefield methods. The plane wave model has been employed to decompose measured signals and reconstruct images through integral transformations or wavefield extrapolations. A groundbreaking algorithm, reported by Jin *et al.*,[78] introduced a plane wave model designed to account for the variation in SoS, particularly in irregularly layered heterogeneous media. A wavefield extrapolation for irregularly layered media was proposed by introducing three approximations—narrow beam, one-way wave, and wave—into the standard plane wave model. The SoS was determined by passive ultrasound (PU). This method has been shown to accurately reconstruct the photoacoustic image for layered media, requiring only 10% of the computational time compared to the TR method (i.e., reconstructing a $50 \times 60 \times 600$ pixels 3D image in only 5.5 s). In the same year, they proposed an attenuation compensation method for high-frequency acoustic-resolution PA imaging.[79] Their method upgrades the wavenumber term in Fourier-domain (FD) image reconstruction by including acoustic dispersion and attenuation during wavefield

24 October 2024 08:47:13

extrapolations. Their study demonstrated improvements in image resolution, edge sharpening, and handling space-variant attenuation effects with computational efficiency in wavefield extrapolation methods. The method involves Fourier transformation of signals and compensation of the focused transducer, converted into a virtual detector, by including a phase shift of spatial and time. Subsequently, wavefield extrapolation is executed in a layer-wise manner to reconstruct an image across the entire depth with an acoustic attenuation term using a power-law prefactor, α_0, to obtain the image at the specific depth.

Table IV offers a concise report summarizing significant scientific inquiries concerning acoustic compensation from 2013 to 2023.

IV. OPTICAL FLUENCE COMPENSATION

Since PA effect involves the interaction between light and sound, another important factor affecting the accuracy and clinical effectiveness of qPAI is the spatial distribution of optical fluence in tissues. Given the equation for the initial pressure rise from an irradiated chromophore

$$p(\mathbf{r}, \lambda) = \Gamma \mu_a(\mathbf{r}, \lambda)\Phi(\mathbf{r}, \lambda), \qquad (7)$$

where p represents the photoacoustic pressure, Γ is the Gruneisen parameter representing the efficiency of converting thermal energy to mechanical energy, μ_a is the optical absorption coefficient, Φ is the optical fluence, $\mathbf{r}$ is a positional vector, and λ is the wavelength of the irradiating light,[80] compensation for the spatial heterogeneities in optical fluence must be performed to recover absorption maps, which accurately reflect the distribution of the various tissues and chromophores being imaged.

The amount of fluence at any point depends on depth and wavelength. Depth-dependent optical fluence variation exists because the optical absorption and scattering by more superficial tissues causes deeper tissues to receive less energy, while wavelength-dependent optical fluence variation exists because chromophores have different absorption and scattering properties at different wavelengths. The different degrees of absorption and scattering by various tissues at various light wavelengths cause the amount of fluence at any particular point to vary with the wavelength used, particularly at greater depths after the illuminating light has undergone higher degrees of absorption and scattering.[81] Absorption spectra recovered from PA images are thus often distorted as a result of wavelength-dependent fluence, a phenomenon often dubbed as "spectral coloring."[29]

Several methods have been developed to correct the errors that arise as a result of depth- and wavelength-dependent optical fluence variation. In this section, we describe different types of compensation methods, including model-based methods, iterative compensations, deep learning methods, multi-modal imaging approaches, and others.

A. Model-based methods

Several models of light transport exist to determine optical fluence at various spatial locations. These include the photon diffusion equation (PDE), Monte Carlo (MC) model based on the radiative transfer equation (RTE), and several other models such as the diffusion dipole model. These models can be used with various degrees of complexity (2D vs 3D modeling, assumption of homogeneity/heterogeneity, etc.) to generate fluence distribution maps with varying degrees of accuracy often dependent on the level of complexity and accuracy of the input parameters. Based on Eq. (7), reconstructed PA images can

be subsequently divided by the obtained fluence maps and the Gruneisen parameter to obtain fluence-compensated optical absorption maps. In Subsections IV A 1–IV A 3, we describe various studies employing these models for fluence correction.

1. Photon diffusion equation

The photon diffusion equation (PDE) is a partial differential equation that describes the diffusion of photons in a medium, typically employed in situations where a macroscopic description of light transport is sufficient. In a study involving a handheld probe integrating photoacoustic (PA) and ultrasound (US) modalities for dual-mode imaging, Zhao et al. used the PDE in Eq. (8) and finite element analysis (FEA) in COMSOL to generate an optical fluence map that was used to correct raw PA images,[80]

$$\frac{1}{3\mu_a(\mu_a + \mu_s')}\nabla\Phi \cdot \mathbf{n} = S, \qquad (8)$$

where μ_a and μ_s' are absorption and reduced scattering coefficients, respectively. Φ is the fluence rate, $\mathbf{n}$ is the unit normal vector, and S represents the optical source. A phantom and in vivo experiment were conducted to test the effectiveness of the correction technique. In the phantom experiment, a phantom consisting of a gel pad and three layers of agar phantoms with different concentrations of intralipid (3.3%, 1.8%, and 0.0%, respectively) was imaged, while a human breast sample was imaged in the in vivo experiment. Known values of the optical absorption and reduced scattering coefficients for each layer of agar/tissue as well as parameters for the illuminating laser and transport mean free path were used to produce an optical fluence map in COMSOL. After PA images were obtained, the raw values were divided by local fluence values obtained from the optical fluence maps. This correction allowed deeper structures of interest to be better visualized—in the human breast sample, a blood vessel at a depth of 14.4 mm was not well-identified in a PA image; however, after the image was divided by the optical fluence map, it was more clearly visualized. The authors noted, however, that this was at the cost of a lower signal-to-noise ratio at the greater depths of tissue in the images. They also stated the simple sandwiched structure assumed in COMSOL simulations was still far from the complex breast sample in practice and a more realistic model must be established.

A common limitation of model-based methods, which is present in the above example, is the need for image segmentation and knowledge of the optical properties of the various tissues being imaged. In addition, 2D segmentation of images may introduce artifacts when 3D visualization of tissues is required, and light fluence estimation in 2D may also be problematic when fluence correction in 3D is required. To address this, Liang et al. used a modified 3D graph search (GS) algorithm to segment PA images and obtain a smooth and accurate 3D model of a mouse model.[82] The 3D model was transformed into a finite element mesh (FEM) model and the PDE was used to generate a 3D light fluence map. The raw 3D model was then corrected by dividing by the obtained fluence model. This method proved to be superior over manual segmentation, and segmentation using the active contour and conventional 3D GS algorithms. Manual segmentation, which was performed on individual 2D images, did not produce as smooth a structure when a 3D volume was reconstructed as compared to the modified 3D GS algorithm, while the active contour and conventional

TABLE IV. Summary of significant research studies conducted on various acoustic compensation methods from the years 2013 to 2023.

Reference	Methods		Acoustic parameter to compensate	Imaging target	Priori	FOV (mm²)	Study type	Computation requirement	Accuracy
60	Iterative optimization	FDTD-based TR	SoS	Prostate tumor, agar (hemoglobin, melanin, lipid)	N	30×30	*Ex vivo*	High	⋯
61		Alternating optimization scheme-based JR	SoS	Numerical phantoms	N	128×128	*In silico*	High	Depend on regularization, spatial support and relative spatial bandwidth
62		FC JR	SoS	Mice liver and kidney	N	28.5 mm diameter	*In vivo*	High	*MRE = 0.13%–0.3%*
63		MSFC APACT SC based JRs	SoS	Mice with liver tumor	N	28.5 mm diameter	*In vivo*	High High Medium	High High Medium
64		Modified iterative	SoS	Numerical phantoms	N	⋯	*In silico*	High	⋯
65		VTDT	SoS	Water with air void	Y	76×76	*In silico*	Low	⋯
67		FWI	SoS, AA	Pencil lead, agar, acrylic cylinder	N	256×256	Phantom	High	*RMSE: 0.002–0.034*
68		IR-WRI	SoS, AA	Neonatal head (ECT, CSF, GM, WM, water, vessel)	Y	120×120	*In silico*	High	Error rate: <5%
69		TROPUS	SoS, AA	Mice whole body (hemoglobin)	N	20×20	*In vivo*	Medium	⋯
72	Deep-learning	CNN-based beamforming	SoS	Circular targets with different sizes and contrasts	N	19×19	Training: *in silico* Testing: *in silico*	High for training, low-medium for actual reconstruction	PSNR: 38.7 ± 4.3 dB
70, 71		SegU-net	SoS	Three-layer phantom, human limbs	N	⋯	Training: *in silico* Testing: *in vivo*		SSIM: >0.95
74		US SoS, DL	SoS	Two-layer phantom, human wrist and finger	N	38.4×38.4	Training: *in silico* Testing: *in vivo*		SSIM: 0.8267
73		DeepMB	SoS	Human carotid, biceps, breast, abdomen	N	416×416	Training: *in silico* Testing: *in vivo*		RMSE: 1.34%–1.41%, SSIM: 0.98
75	Others	Modified FBP	SoS	Mouse head and liver regions	Y	20×20	*Ex vivo*	Medium	⋯
76		US SoS map, SRB	SoS	Super Soft + TiO_2, water, blood inclusions	N	200 mm diameter	Phantom	High	Error: 0.1–0.9 mm
77		MRI/US, DAS-MSFM	SoS	Mouse brain (scalp, skull, brain), nonlinear wave propagation equation (Eikonal equation)	Y	⋯	*In vivo*	Medium	⋯
78, 79		PU SoS, wavefield extrapolation in FD	SoS	Chicken breast, shell, carbon rod	N	15×15	*Ex vivo*	Medium	Error rate: <11.1%

3D GS algorithms produced errors that were reduced by the modified GS algorithm. Quantitatively, the RMSE of the segmentation produced by the modified GS algorithm compared to the active contour and conventional GS algorithms was 43.35% and 37.28% lower, while the mean absolute difference (MAD) was 64.70% and 56.89% lower, respectively. For light fluence correction, application of the obtained 3D light fluence model improved the signal intensity for the deeper tissues in the reconstructed images. Compared to correction using 2D light fluence models, fluence correction using the 3D fluence model produced lower variance of the signal intensity, coefficient of variance, higher contrast to noise ratios, and effective number of looks. This resulted in PA images with a higher contrast and smoother appearance.

2. Monte Carlo simulations

Monte Carlo (MC) simulation is a computational technique that uses random sampling to solve the RTE for modeling the behavior of individual photons as they travel through a medium. This can also be used to develop maps of optical fluence. In a study involving the use of a sophisticated phantom model of the breast to measure sO_2, Dantuma et al. used MC simulations to develop a 3D fluence map for the Twente Photoacoustic Mammoscope 2 (PAM 2) system.[83] The phantom contained two bulk layers of tissue mimicking the fat and fibrogranular layers of the breast and channels within each layer through which blood-mimicking liquids could be pumped through. Two fluence maps were generated—one considering the phantom having homogeneous optical properties that were determined by taking the average of the optical property values for the fat and fibrograndular layers, and the other taking into account the specific optical property values for each layer. Comparing PA-based sO_2 measurements calculated without fluence compensation with those calculated with fluence compensation using the two fluence maps, the authors found that fluence compensation significantly decreased the mean absolute error from the ground truth sO_2 values obtained using an oximeter for two sets of sO_2 values (8.7% and 57.6%).

Kirillin et al. also used the MC model in a study on raster scanning optoacoustic mesoscopy (RSOM)-based angiography.[84] Having a RSOM system that could focus at different depths, they performed several MC simulations to develop a range of fluence maps for subsequent fluence compensation. The aim was to observe the effectiveness of the simulations in performing fluence compensations and study the effect of varying the parameters of the simulations, including the optical parameters of the tissues, the optical beam geometry, the radius of the acoustic antenna receiving diagram, and the focus position. Through phantom and in vivo studies, the authors found that using parameters for the MC simulation that matched the real values was beneficial and at times critical to successful fluence compensation. In phantom studies, they found that fluence maps produced using MC simulations performed with variations in the chosen optical properties enabled the recovery of PA signals that were attenuated at larger depths. However, using values of the absorption and reduced scattering coefficients in the simulations, which were not equal to the true values, could dampen PA signals from superficial structures and were thus not ideal. Notably, the values of the absorption and reduced scattering coefficients that were greater than the actual values produced larger errors than values that were smaller than the actual values. In in vivo studies of human skin, they also found that using incorrect values focus

positions of the RSOM system in the MC simulations resulted in larger amounts of noise and impaired visualization of the superficial vessels.

Hirasawa et al. also used MC simulations to calculate the fluence in a study to validate the effectiveness of a custom-developed PA probe in a phantom model and when injected into a subcutaneous tumor within a mouse model.[85] US imaging was used to identify skin thickness, and literature values of the tissue optical properties at each wavelength used were input into the simulations. Of note, they also implemented a subsequent spectral fitting method to improve the visualization of the PA signal from the PA probe by suppressing the background PA signals generated by strong optical absorbers, such as hemoglobin. In the phantom study, fluence compensation was found to mitigate the spectral coloring effect, with the peak of the corrected PA absorption spectrum of the probe matching the reference spectrum. For the in vivo study, fluence compensation and spectral fitting allowed the probe to be detected in the presence of hemoglobin.

3. Other model-based methods

In another approach, Zhou et al. used a diffusion dipole model (DDM) to perform fluence corrections on raw PA images and compared its performance to the 1D diffusion approximation (DA) and MC simulations.[86] They found the DDM's performance in facilitating the recovery of the spectra of various chromophores and at various depths to be better than the 1D DA algorithm and comparable to that of MC simulations. A similar finding was observed for the measurement of sO_2. The DDM was achievable this with the added advantages of being computationally more cost-effective and easier to implement. However, the authors noted that as with other model-based fluence correction methods, prior knowledge of tissue optical properties is necessary and may not be practically observable in in vivo cases. In addition, the DDM requires the assumption of a homogeneous medium.

Separately, Kim et al. used an extension of the PDE to correct for the wavelength-related fluence errors obtained using a kilohertz-rate wavelength tunable laser, which illuminates media with narrow laser beams at a high frequency.[87] The linear US probe had ten fibers along each of its long edges, and laser beams were sequentially transmitted through the fibers using a motor-controlled collimating lens. By extending the PDE to analytically calculate the fluence profile from the partial PA image illuminated by different fibers, they developed two fluence models, one a simplification of the other to extract μ_{eff}. Through conducting numerical simulations with these, they found that both models closely approximated the fluence profiles obtained through MC simulations beyond a certain depth. In phantom studies, they found that the proposed models were able to facilitate accurate estimation of optical parameters, which enabled good spectral corrections. Jeng et al. developed a PAUS imaging system using a similar technique to achieve real-time fluence estimation in macro-homogenous media and demonstrated it in vivo.[81]

In a slightly similar manner, Ranasinghesagara et al. used multi-illumination PA microscopy (MI-PAM) to measure the optical properties of the imaged medium and subsequently used these optical properties as inputs in the MC model to calculate fluence.[88] For an optical absorber in a phantom, they illuminated the absorber at different lateral locations by translocating the illuminating laser and obtained the photoacoustic signals. They then performed MC simulations and PDE calculations for various reduced scattering coefficients (at a fixed absorption coefficient) and determined the RMSE between

the experimental and simulation data over all the lateral positions studied for each reduced scattering coefficient μ_s' used. The μ_s' at which the lowest RMSE was achieved was taken to be the estimated μ_s' for the medium. Using this method, the authors found that for a true μ_s' of 4.4 cm^{-1}, the error between the estimated and true μ_s' values was 5.7%. The accuracy of the estimated μ_s' generally decreased when the phantom had larger μ_s' values. However, when the estimated μ_s' was used in MC simulations with fixed values of optical absorption coefficients, fluence maps that facilitated accurate depth-dependent fluence corrections were obtained, and the authors concluded that the method could improve the quantitation of PA imaging. A drawback of this approach is its computational intensity and time-consuming nature, as numerous MC simulations are required to identify the global minimum for the RMSE between the experimental and simulated data for each examined optical property.

The amount of information and degree of detail required almost invariably affects the accuracy of the fluence-corrected image or volume. While higher levels of detail and accuracy may be provided by advanced models and algorithms such as the MC model and the modified 3D GS algorithm highlighted above, these are unfortunately associated with higher computational requirements and tend to require more time to be applied. Stacks of 2D images, which form 3D volumes, in particular, require multiple-fold more computation and time to be fluence-corrected compared to single images. To address this, less typical models such as the DDM and others that have lower computational requirements can be utilized to provide similar performances. Different models, as well as modifications to the conventional algorithms or models, can also be successfully applied to adapt to the characteristics of the instrumentation used and address the specific difficulties of the tissues being imaged. However, a critical limitation of model-based compensation methods that remains is the need for accurate input parameters to be used. These include the optical parameters of the imaged tissue and the experimental characteristics, such as the focus depth and optical beam geometry. To tackle this, we now turn to iterative compensation methods that represent a possible solution.

B. Iterative compensation methods

Iterative fluence compensation involves an iterative optimization process to refine the fluence compensation. It typically starts with an initial estimation of fluence distribution and updates the fluence distribution in each iteration, adjusting it to better match the observed photoacoustic signals. This iterative refinement continues until convergence to a solution. Liu *et al.* used an iterative algorithm that incorporated fluence compensation to obtain optical absorption maps from raw PA images.[89] In their algorithm, 3D MC simulations were performed to obtain fluence maps, and by taking the product of these with optical absorption coefficient maps, absorption energy distributions were calculated. The absorption energy distributions were then compared with the absorption energy distribution obtained from the initial PA image, and if the difference exceeded a certain threshold, the optical absorption coefficient maps were updated and the process was repeated iteratively. The advantage of this algorithm was that knowledge of the optical properties and geometric information of the tissues or models being imaged was not required. Through testing their algorithm on a simulated model, the authors found that the algorithm was able to recover the absorption coefficient for various targets in the simulated model for up to an accuracy of 99.5% and produce significant

improvements in the structural similarity, RMSE, and peak SNR parameters after ten iterations. Application of the algorithm on a phantom model also allowed the optical absorption coefficient maps to be recovered with decent accuracy.

In several model-based methods, optical parameters must be known a *priori*. Although averaged values can be used for *in vivo* studies, estimating these parameters accurately is often difficult. To address this, Wu *et al.* used an iterative fluence compensation and spectral unmixing algorithm to improve PA images and develop fluence maps.[90] First, they initialized the absorption coefficients, scattering coefficients, Gruneisen parameter, chromophore concentrations, and anisotropy values, then used these initialized values in the MC simulation to develop a fluence map. The PA initial pressure was then calculated using the Gruneisen parameter, absorption coefficient, and fluence map and compared with the ground truth PA initial pressure obtained from an image of a prostate tumor in a mouse model. If the calculated PA initial pressure was sufficiently close to the ground truth, the algorithm terminated; otherwise, the optical parameters were recalculated, and a subsequent MC simulation was performed. The algorithm was evaluated on 15 datasets, and its effectiveness was validated using the RMSE and Pearson correlation coefficient (PCC) metrics. The authors found that the algorithm converged after four iterations and was able to improve the RMSE for the quantification of a contrast agent, hemoglobin, and oxyhemoglobin by 337%, 417%, and 337%, respectively. It also improved the PCC for these chromophores by 22%, 11%, and 20%, respectively.

The use of iterative algorithms, however, entails the possibility of overfitting. This tends to occur when the design of the algorithm is highly complex and not optimized. One method to mitigate the consequences of this is to employ a regularization step. In one example, Naser *et al.* used a SNR regulation within the finite element-based local fluence correction (LFC) algorithm with SNR regularization and tested its efficacy in estimating sO$_2$ in *ex vivo* and *in vivo* tissue.[91] The rationale behind using SNR regularization was to prevent iterative models from overfitting to data in PA images that represent noise. Using the algorithm, they found that the average sO$_2$ estimation error decreased from 10.1% to 2.8% in *ex vivo* tissue. In *in vivo* tissue, sO$_2$ estimation increased from 92.9% to 95.9% when rat models were supplied with 100% oxygen, and the estimation error for changes in sO$_2$ decreased from 6.4% to 2.0%.

The advantage of iterative model-based algorithms for fluence compensation is the absence of the requirement for *a priori* knowledge of the optical parameters of the imaged tissue. In this regard, iterative algorithms are more robust and potentially more accurate than model-based methods that are prone to producing biases when inaccurate optical parameters are used. Estimating optical parameters accurately for *in vivo* studies is also often difficult. In addition, iterative algorithms have the ability to better adapt to complex and changing conditions since they allow for continuous refinement based on the actual measured data. However, they suffer from the significantly higher computational requirements, which increase as the number of iterations required increases. This is especially so if MC models are used, and if a higher dimensionality of data (i.e., 3D data instead of 2D data) is required. Ideally, the number of iterations required for convergence should be minimal, which requires the design of the algorithm to be well-crafted and appropriate for the application. Nevertheless, achieving higher accuracies and using lower error thresholds often require

either more iterations or a more complex algorithm, and herein the trade-off between accuracy and computational requirements remains.

C. Multimodal approaches

1. Ultrasound imaging

In many studies, ultrasound (US) imaging is usually combined with PAI due to the complementary nature and the shared utilization of acoustic waves for imaging, to enable dual-modality imaging for both structural and molecular information. US imaging can be used to compensate for fluence variation in PAI through "acoustic feedback," where structural information obtained from US imaging can facilitate the assignment of known optical parameters to the appropriate spatial locations. In one example, Han *et al.* acquired and stacked 2D US B-mode images of human breast to construct a 3D structural model, in order to construct a 3D optical fluence model.[92] Before being fused into a 3D model, each 2D US image was segmented by experienced physicians and each structure was labeled with the appropriate optical properties of the associated tissue (skin, adipose tissue, fibrograndular tissue, tumor tissue, background tissue). Diffusion equations and boundary conditions were then used to calculate the optical fluence throughout the imaged tissue using COMSOL. To validate the effectiveness of their method, the authors studied oxygen saturation (sO_2) in both phantom and *in vivo* models using two wavelengths (750 and 830 nm). Phantom models were constructed with pseudo blood ($NiSO_4$ and $CuSO_4$) that had varying levels of pseudo sO_2 (25%, 50%, and 75%). Raw PA images were compared against fluence-compensated images modified with the constructed optical fluence models, and the authors found that the calculated pseudo sO_2 values in the fluence-compensated PA images were higher and closer to the true pseudo sO_2 values. In *in vivo* models, the authors found that fluence-compensated PA images of the breast tissue in two breast cancer patients indicated sizeable increases in the calculated sO_2 values compared to the uncompensated images, particularly at larger depths (Fig. 6). They also found that the optical fluence maps obtained from their 3D modeling of heterogeneous tissue outperformed maps that were constructed with homogeneous tissue in correcting for the fluence-associated errors. While the modeling of 3D heterogeneous tissue is more accurate and promising, the authors noted that the manual segmentation of 2D images is highly laborious and time-consuming and that more data should be collected to enable machine learning based segmentation.

Bulsink *et al.* used US imaging to obtain images containing structural information, which they subsequently segmented and used with known tissue optical properties in a MC model, instead of PDEs.[93] Fluence maps were obtained and subsequently used to correct raw PA images obtained using two light wavelengths for sO_2 measurement. In phantom studies, uncompensated images resulted in estimated sO_2 values of 85.8%, 83.9%, and 81.7% in three different tubes carrying pseudo-blood with identical oxygenation levels located in the phantom at increasing depths. After fluence compensation, the sO_2 values improved to 99.1%, 99.3%, and 99.6%. The ground truth value was approximately 96.3%. *In vivo* studies on a mouse and human wrist were also performed. The mouse was given two levels of oxygen sequentially—normal air and 100% oxygen. Fluence compensation allowed the PA images to better capture the difference in sO_2 levels between the two oxygenation states. For the human wrist, fluence
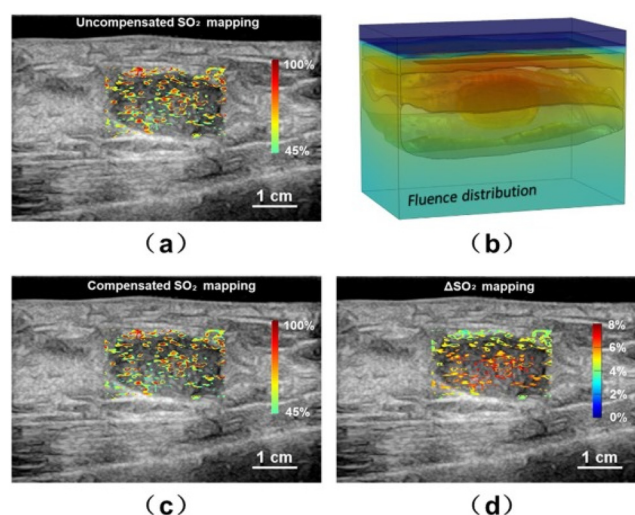
FIG. 6. Breast imaging results on a patient. (a) Uncompensated sO_2 distribution. (b) 3D modeling of fluence map based on US images. (c) Compensated sO_2 distribution. (d) Difference between uncompensated and compensated sO_2. Reprinted with permission from Han *et al.* Photoacoustics **21**, 100222 (2021), licensed under a Creative Commons Attribution (CC BY-NC-ND) license.

compensation also enabled sO_2 differences between an artery and a vein to be measured. Overall, the authors found that the proposed fluence compensation method allowed the accuracy of sO_2 to increase by 12%–15%.

Separately, Pattyn *et al.* obtained a SoS map using US imaging and used it to subsequently construct a fluence map for a phantom study, also with MC simulations.[76] For a heterogeneous phantom with three separate constituent materials [blood, water, and a mix of Super Soft Plastic and titanium dioxide (TiO_2)], they segmented US image and labeled each segment with the appropriate optical properties obtained from the literature using prior knowledge of the constituent material which made up each segment. The labeled map was then used as a numerical phantom to develop a fluence map via MC simulation. The fluence map was then used to correct a processed PAT image to compensate for depth-dependent optical fluence variations.

One disadvantage of using US imaging with PAI is the need for additional instrumentation. To use US imaging with minimal additional instrumentation, Jin *et al.* integrated a passive ultrasound (PU) technique with PA imaging.[94] In PU imaging, passive elements such as polyvinylidene difluoride (PVDF) generate ultrasound waves upon the absorption of photons. A PVDF layer was used in this study to absorb backscattered photons from a sample following laser irradiation, and subsequently generate US waves, which were propagated to a transducer. As such, only a single piezoelectric transducer was required for the detection of US waves from the PA and PU modalities, thus reducing the amount of instrumentation required for a setup capable of acquiring diffuse reflectance information to perform optical fluence compensation. In a phantom setup to image three strips of black tape under a white plastic film at one end and another white plastic film with different scattering properties at another end, the system was able to correct the difference in the calculated optical absorption coefficient in the raw PA image caused by the optical fluence variation almost perfectly. An *ex vivo* experiment using porcine tissues with muscle and

lipids was also conducted. Blue tapes were placed below 1- to 1.5-mm-thick porcine tissues, and the authors showed that the developed system was able to reduce the error in the calculated optical absorption coefficients of the tapes from the raw PA image arising from the differences in scattering properties of muscle and lipid tissue by more than 50% after correction with the diffuse reflectance images.

The use of US imaging to improve the quality of PA images is certainly beneficial as the structural information it provides enables the precise locations of different tissue types to be identified. This facilitates the proper assignment of known optical properties of the various tissue types, which consequently enhances the accuracy of the developed optical fluence maps and hence the quality of the corrected PA images. However, its use necessitates proper segmentation and identification of the different structures, which may require experienced clinicians and domain experts in the case of human and animal images. This also inevitably involves a time cost. However, the possibility of artificial intelligence-based segmentation may alleviate this cost in the future when models have been sufficiently trained to be satisfactorily accurate. Nevertheless, the processing of the US images, particularly when they are stacked to form a 3D volume, will require more computational resources and memory.

2. Diffuse optical tomography

Diffuse optical tomography (DOT) is a non-invasive imaging technique that uses near-infrared light to reconstruct 3D maps of tissue optical properties, which can subsequently be used to develop fluence maps. Consequently, it can be combined with PAI to compensate for the inaccuracies resulting from optical fluence variation in deep tissue. Bauer et al. demonstrated the use of a DOT system to obtain a fluence map of a tissue-mimicking phantom with spatially varying absorption and scattering properties.[95] PA images were first obtained using a modified clinical US system and a dye laser tuned to 650 nm, following which DOT was performed using a 780 nm laser and an electron-multiplying charge coupled device (EMCCD) in a transmission configuration. Reconstructed DOT images were obtained and used with the surface fluence profile of the PAI system to calculate a map of the fluence throughout the phantom. The raw PA image was then divided by the obtained fluence map to produce a corrected image, which resulted in the difference in PA signal magnitude between three optically identical targets placed at a 1.2 cm depth under regions with differing absorption and scattering properties decreasing by approximately sixfold from 33% to 6%. The authors claimed that the developed system was able to correct fluence inhomogeneities from three sources—structure in the illumination pattern, attenuation of the light fluence that arises from bulk optical properties and that is dependent on tissue geometry, and internal absorption and scattering inhomogeneities. For in vivo imaging, the authors noted that a reflection configuration for the DOT system and matching the wavelengths of the PA and DOT lasers would be ideal.

Mahmoodkalayeh et al. also found that fluence maps obtained from DOT imaging were able to improve the accuracy of PA-derived optical absorption maps for a digital phantom consisting of vessel-mimicking and background components.[96] Compared to a PA image that was reconstructed using a fluence map obtained from a model with homogenous optical properties, the DOT-PAT integrated method was able to reduce the error in the optical absorption map by 44%. In addition, the authors also used the structural information obtained

from PA images to improve the reconstruction of DOT images and use these improved reconstructions to obtain a separate set of fluence maps, which were subsequently used to correct PA images. Terming this technique PAT-DOT-PAT, they found that it was able to reduce the error in the optical absorption map obtained from the model-based reconstructed PA image by 48%.

While the use of DOT with PA imaging offers benefits, which are similar to those offered by US imaging, its limitations are also similar. Additional optical equipment such as fibers and sensors are required, which may not be ideal if a small footprint is required, and it also necessitates further processing and analysis, which requires more computational resources.

D. Deep learning-based methods

A relatively new method for optical fluence compensation in PA images is deep learning (DL). Recently, Madasamy et al. studied the effectiveness of several DL methods in correcting raw PA images.[97] They evaluated the U-Net, fully dense U-Net, Y-Net, fully dense Y-Net, Deep ResU-Net, and generative adversarial networks (GANs) on 2D and 3D phantoms and an in vivo dataset, using peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), generalized contrast to noise ratio (GCNR), and contrast to noise ratio (CNR) as the evaluation metrics. Generally, the deep learning models were able to perform satisfactory optical fluence compensation, with the fully dense U-Net architecture generally outperforming the others. For 2D blood vessel phantoms, the authors found that the DL models performed better on images with lower data noise and that they were better able to recover deep structures in images imaged at wavelengths higher than that which the training images were imaged at. Along the same lines, the performance of the models in recovering deep structures was suboptimal in images imaged at wavelengths lower than that which the training images were imaged at. This was likely due to the higher penetration depth of the longer wavelengths in the near-infrared region. With 3D numerical breast phantoms, they trained and tested the deep learning models on 2D slices, then re-stacked the 2D slices into a 3D phantom. Although the models were unable to recover some of the small vasculature, they demonstrated superior performance over model-based reconstructed images. However, the authors also noted that the plane of the images that the models are trained on is of importance. They also found that the DL models were able to perform accurate fluence compensation on in vivo PA images from mice; however, the dataset used only consisted of ten images and thus to further prove the effectiveness of deep learning methods in correcting for optical fluence errors, larger datasets will need to be used.

Yang et al. also used a deep residual recurrent neural network (DR2U-Net) for estimation of sO₂ in PA images.[98] The inputs were optical absorption images obtained from MC simulations, and the output was a quantitative sO₂ image. MC simulations were performed for simulated tissue structures, and the sO₂ of arteries, veins, and the dermis were varied within a certain range according to a Gaussian distribution. The thickness and location of the arteries and veins were also varied to provide wider range of data. Compared to the linear unmixing method, the DR2U-Net was able to achieve a reduction of 47.49% in the mean error and a 1.21% reduction in the standard error for images with arteries only, and 60.96% reduction in the mean error and a 2.02% reduction in the standard deviation for images with arteries and veins. The trained network was also able quantify the distribution

of sO$_2$ in an image in 18.4 ms, a significant improvement for the time taken to do the same using linear unmixing.

The use of DL methods to correct the fluence-based errors is promising, and more studies, which show better results, are expected to be reported in the future. However, their accuracy and generalizability depend on the availability of sufficient PA data. Larger volumes of suitable PA data are desired; however, training DL models with these requires much computational resources and time. Trained DL models may also require longer times to run, especially if they have highly complex architectures. This may be an important consideration when deciding whether to employ them for fluence correction.

E. Other fluence compensation methods

Another method used to compensate for optical fluence variations is the use of wavefront shaping. In a study by Caravaca-Aguirre *et al.*, a spatial light modulator (SLM) was used to modulate the wavefront of a 532 nm laser beam before it was focused onto two capillary tubes behind a glass diffuser.[99] An US transducer was placed behind the tubes to measure the PA signal. Feedback from the PA signal was fed into a genetic algorithm in an optimization process to generate a phase mask, which was projected on the SLM. This enhanced the localized fluence on the capillary tubes, which enhanced the amplitude of the PA signal tenfold. 2D reconstruction was performed by translating the capillary tubes along the x and y axes, and depth information was obtained by extracting temporal information from the PA signal to reconstruct a 3D image.

To improve the measurement of oxygen saturation (sO$_2$) using PA images, Fadhel *et al.* developed a fluence matching algorithm to correct errors in raw PA images arising from wavelength-dependent fluence variations.[100] Power spectra from PA signals obtained using two different wavelengths (720 and 870 nm) were obtained, following which the ratio of the power spectra was calculated. The slope of the linear regression line of the ratio of the power spectra (termed the spectral slope) was then calculated and used, and its antilog value was used to construct a frequency filter. The filter was applied to the PA image obtained at 720 nm, and an inverse fast Fourier transform was used to transform the signal back to the time domain. In phantom studies, the authors showed that their developed method was able to match the fluence profile of the PA image obtained at 720 nm to that at 870 nm and thereby improve sO$_2$ estimates. *In vivo* studies also showed that the fluence matching method could improve sO$_2$ estimates in accordance with the expected trend.

In a short summary, this section reviewed different methods on optical fluence compensation, a critical aspect to enhance the accuracy and clinical relevance of qPAI. Acknowledging the involvement of optical fluence in the photoacoustic pressure equation [Eq. (7)], it emphasizes the necessity of compensating for spatial variations in fluence to derive accurate absorption maps in tissue imaging. Strategies to mitigate inherent challenges such as the "spectral coloring" effect have been summarized in Table V, offering a concise reference for fast takeaways. A key consideration in choosing the appropriate method for fluence compensation is the trade-off between the accuracy of the compensation method with the amount of computational resources required to execute the method. Between model-based methods and iterative compensation algorithms, the amount of computational resources required to execute iterative compensation algorithms is multiple-fold more than those required for model-based methods. The

exact multiplication factor is dependent on the number of iterations required. However, iterative compensation algorithms can potentially produce more accurate fluence maps, which are closer to the ground truth. For DL methods, the amount of computational requirements for DL methods depends on the complexity of their architecture but is likely to be between those of model-based methods and iterative compensation algorithms because their models are typically more complex but are only run once. In addition to these, it is also important to note that the information provided by other imaging modalities in multimodal approaches requires computational resources to be analyzed. The overall computational burden can be segregated into two components, namely, the resources required for image segmentation of the various tissues and those required for developing fluence maps. Ultimately, the accuracy of the developed fluence map is contingent on the accuracy of both the image segmentation and the method used to generate the fluence map, and thus computational resources will have to be balanced and optimized for both processes to produce the most ideal outcome. Many model-based, iterative compensation and DL methods, however, require the prior use of a separate imaging modality to provide structural information to generate a fluence map, and thus the amount of computational resources required for image segmentation is often a necessary factor to account for in the entire process of fluence compensation.

V. DIRECT RECONSTRUCTION OF OPTICAL PROPERTIES IN HETEROGENEOUS TISSUES

The optical absorption coefficients μ_a are typically determined by two-step process of acoustic and optical inversions in qPAI. Some research groups have explored an alternative approach: direct reconstruction of μ_a from measured time-domain PA signals using a composite opto-acoustic model. The JR problems discussed in Sec. III A 2 only involve acoustic inversion to derive initial pressure distribution $p(\mathbf{r},0)$ or absorbed optical energy density A($\mathbf{r}$), which still requires additional optical compensation to further derive μ_a. In contrast, direct reconstruction methods introduced here integrate both acoustic and optical inversions to derive the optical absorption coefficient $\mu_a(\mathbf{r})$ in heterogeneous tissues, bypassing the intermediate step to reconstruct the initial pressure field.

Sun *et al.* concurrently addressed optical and acoustic forward problems in an intravascular photoacoustic tomography (IVPAT) context, by alternately minimizing the calculation errors in numerical simulation.[101] The light transportation was modeled by RTE, and the acoustic simulation was based on FDTD. Reconstruction of $\mu_a(\mathbf{r})$ with given SoS and reconstruction of SoS with given $\mu_a(\mathbf{r})$ were done by iterative optimization with TV and Tikhonov regularizations, respectively. Evaluation of the algorithmic performance was conducted via a phantom experiment, assessing normalized mean square absolute distance (NMSAD). By optimizing iteration step sizes and initial values, the reconstruction quality outperformed those based on a fixed SoS and conventional two-step qPAI by ∼27% and ∼15%, respectively. Nonetheless, the computational load is substantial due to the full-wave inversion scheme utilized. Ding *et al.* proposed a one-step reconstruction technique aimed at directly retrieving both the optical coefficients and the wave speed from the acquired PA data.[102] The authors employed the Levenberg–Marquardt method to solve the nonlinear optimization problem associated with forward wave equations and forward diffusion equations. A numerical experiment was carried out

24 October 2024 08:47:13

TABLE V. Summary of different methods for PA optical fluence variation compensation.

Reference	Methods	Methods	Wavelengths (nm)	Chromophores	Depth (mm)	Study type	Relative accuracy	Relative computational requirement
80	Model-based	PDE, FEA	1064	Blood	14.4	*In vivo* (human breast)	Low	Low
82		Modified 3D GS, PDE	760	Mammary carcinoma	10	*In vivo* (mice)		
83		MC simulation	755, 1064	Blood (sO_2)	23	Phantom (breast)		
84		MC simulation	532	Blood	4.6	*In vivo* (human palm)		
85		MC simulation	500–600	gGlu-HMDiMeR PA probe	5.5	*In vivo* (mice)		
86		DDM	700–890	India ink, blood, sO_2, copper sulfate	35	Phantom		
87		Fast-sweep PA, PDE extensions	700–900	Human hair, GNR, black ink	9.7	Phantom		
81		Fast-sweep PA, PDE extensions	700–900	GNR, Black ink	8.5–9.7	*In vivo* (mice)		
88		MI-PAM, MC simulation, PDE	532	Hair, crystal violet dye	5	Phantom		
89	Iterative Model-based	Iterative MC simulation	532	India ink	2.7	Phantom	High	High
90		Iterative MC simulation	700–900	Blood, contrast agent	24.2	*In silico*		
91		FEM-based LFC with SNR regularization	710–850	Blood (sO_2)	10	*In vivo* (rat)		
92	Multimode	US imaging segmentation, PDE	750, 830	Blood (sO_2)	26.7	*In vivo* (human breast)	Variable, overall accuracy depends on image quality, image segmentation accuracy, and fluence model used	Variable, image processing and segmentation require additional resources
93		US imaging, MC simulation	750, 850	Blood (sO_2)	6–8	*In vivo* (human wrist)		
76		US imaging, MC simulation	680	Blood	32.5	Phantom		
94		PU	808	Black tape, blue tape	1.5	*Ex vivo* (porcine tissue)		
95		DOT, PDE	650	India ink	12	Phantom		
96		DOT, PDE	800	Blood	25	*In silico*		
97	Deep learning	U-Net, FD U-Net, Y-Net, FD Y-Net, Deep ResU-Net, GAN	600–800	Blood	25	Training: *in silico* Validation: *in vivo*	Medium, though variable and dependent on complexity of model architecture	Medium, though variable and dependent on complexity of model architecture
98		DR2U-Net	700, 800	Blood (sO_2)	· · ·	*In silico*		
99	Others	Wavefront shaping	532	India ink	· · ·	Phantom	· · ·	· · ·
100		Fluence matching	720, 870	Blood (sO_2)	10.5	*In vivo* (mice)		

with different noise level. The maximum relative errors (mRE) of SOS and μ_a are 0.97–1.41 and 0.56–0.99, respectively. The results demonstrated the effectiveness of the one-step reconstruction with *a priori* bounds imposed on the unknown parameters. Javaherian and Holman employed the DA model for the optical problem and three TV-regularized optimizations for the composite opto-acoustic problem.[103] They further developed inexact Newton algorithms and benchmarked with a gradient-based quasi-Newton approach. Numerical evaluation was conducted in various heterogeneous tissue settings. The proposed approach achieved a 12%–21% reduction in relative error (RE) compared to the gradient-based quasi-Newton method, while maintaining a similar level of computational complexity.

The pursuit of direct reconstruction methods in qPAI represents a significant shift from traditional approaches, offering the potential for more streamlined and accurate determination of optical absorption coefficients μ_a in heterogeneous tissues. Unlike the reconstruction methods discussed in Sec. II, which require additional consideration of spectral preservation, direct reconstruction methods bypass the intermediate step of initial pressure reconstruction, eliminating the effect of spatial changes on spectral preservation. By integrating acoustic and optical information simultaneously, direct reconstruction methods capture the spectral characteristics of tissue components directly in the reconstructed images, ensuring that the spatial distribution of μ_a reflects the genuine optical absorption of the imaged tissue. Current direct reconstruction approaches are mainly based on model-based iterative algorithms, which have high computation complexity. Future research may focus on enhancing computational efficiency without compromising accuracy, through the development of novel algorithms or leveraging advancements in hardware acceleration. Additionally, there is potential for the integration of machine learning techniques to avoid iterations, optimize reconstruction processes, and improve the robustness of results in complex tissue environments.

VI. SPECTRAL UNMIXING

PAI stands out for its noninvasive capability to obtain physiological and biochemical information at depth. PA data are usually acquired across several wavelengths, resulting in multispectral images that contain mixed information of various chromophores.[104,105] The reconstructed μ_a images at multiple wavelengths can be further analyzed to derive the concentration distribution of specific chromophores, achieving the ultimate goal of qPAI. Despite the complex composition of tissues, spectral unmixing techniques can effectively separate these chromophores. As implied by its name, spectral unmixing aims to decode the mixed pixel into a set of endmembers by determining the contribution of individual endmember to the observed spectrum at each pixel location.[106]

A. Linear spectral fitting

Conventional approach for addressing the spectral unmixing problem involves the application of a linear regression algorithm,[107] assuming each pixel in the PA image represents a linear combination of absorptions and concentrations of different chromophores, known as linear mixture model (LMM),[108]

$$
\begin{bmatrix} p(\lambda_1) & p(\lambda_2) & \dots & p(\lambda_M) \end{bmatrix}
$$
$$
= k \begin{bmatrix} \varepsilon_1(\lambda_1) & \varepsilon_1(\lambda_2) & \dots & \varepsilon_1(\lambda_M) \\ \varepsilon_2(\lambda_1) & \varepsilon_2(\lambda_2) & \dots & \varepsilon_2(\lambda_M) \\ \dots & \dots & \dots & \dots \\ \varepsilon_N(\lambda_1) & \varepsilon_N(\lambda_2) & \dots & \varepsilon_N(\lambda_M) \end{bmatrix} \times \begin{bmatrix} C_1 & C_2 & \dots & C_N \end{bmatrix},
$$

(9)

where k is a system coefficient assumed to be constant and ε_λ is the known molar extinction coefficient of a certain chromophore at a certain wavelength. C is the corresponding molar concentration. Assuming the fluence is uniform, the concentration matrix can then be obtained by solving

$$
\mathbf{C} = (\boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon})^{-1} \boldsymbol{\varepsilon}^T (\mathbf{p}/k). \tag{10}
$$

Chromophore concentrations at each pixel can be represented by a distinct color and shade. For instance, hemoglobin and oxyhemoglobin are usually represented by red and blue, respectively, with darker shades indicating higher concentrations and lighter shades indicating lower concentrations. The color coded representations allow for a better visualization of the different chromophores' concentrations beneath the skin surface.

Feng *et al.* detected the collagen by multi-wavelength PA (MWPA) to use it as a bio-marker for bone health evaluation.[109] They used 1300–1800 nm wavelengths with 5 nm step size for illumination, falling in the strong absorption region collagen, lipid, and water. The measured spectra were fitted to the major components' known spectra by least square fitting (LSF). In the *ex vivo* study on the human bone sample, a high R^2 value of 0.97 was achieved, indicating the high accuracy of the fitting through linear spectral unmixing. Subsequently, they successfully used relative collagen values to distinguish between healthy and osteoporotic bones, indicating the potential of MWPA and the measured relative collagen value to be the new tool and bio-marker for bone health monitoring.

Ding *et al.* investigated the effect by incorporating non-negativity constraints at reconstruction or unmixing or both stages.[110] They considered seven methods in this study, involving separate and combined constrained or unconstrained reconstruction and unmixing. Different extrinsic contrast agents were measured *in vivo*, including AF750, gold nanorods (GNR), iRFP, and IRdye800CW by using MSOT with near-infrared illumination (680–900 nm). Generally, adding the non-negative constraints is able to reduce image artifacts. More interestingly, by formulating it to a convex optimization problem, the approach of combined reconstruction and unmixing method with a non-negative constraint solely on the contrast agent showed the most robust and accurate performance, even though its computational time is the longest.

Although linear regression is a straightforward method to unmix various chromophores, the absorption spectra of all chromophores must be known as *priori* for spectral fitting, which is usually challenging for *in vivo* applications, especially for complex biological tissues.[105] In addition, the "spectral coloring" effect caused by location-varying fluence $\Phi(\mathbf{r}, \lambda)$ limits its accuracy,[29] unless optical fluence correction is explicitly done as discussed in Sec. IV.

B. Statistical detection methods

Statistical detection methods concentrate solely on a specific target spectrum, without the requirement to know explicit background

absorption spectra. This characteristic makes them highly suitable to detect certain molecules with known spectra, such as exogenous nanoparticles, reporter genes, oxy- and deoxy-hemoglobins. Tzoumas *et al.* introduced eigenspectra multispectral optoacoustic tomography (eMSOT) to estimate oxygen saturation (sO$_2$) level *in vivo*.[111] They used PCA to obtain a mean fluence spectrum (Φ_M) and three fluence eigenspectra (Φ_i), whose superposition can model any arbitrary and complex tissues. Then the measured pressure at each pixel can be represented by

$$p(\mathbf{r}, \lambda) = \left(\Phi_M(\lambda) + \sum_{i=1}^{3} m_i(\mathbf{r})\Phi_i(\lambda) \right) \left(c'_{HbO_2}(\mathbf{r})\varepsilon_{HbO_2}(\lambda) + c'_{Hb}(\mathbf{r})\varepsilon_{Hb}(\lambda) \right), \tag{11}$$

where m_i is the eigenfluence parameter, $c'(\mathbf{r})$ is the relative concentration, and $\varepsilon(\lambda)$ is the known distinct coefficient. It becomes a non-linear constrained optimization problem to be solved for the quantitative sO$_2$. Even though five wavelengths are enough, 21 wavelengths from 700 to 900 nm were employed for high accuracy. *In vivo* experiment on skeletal muscle and 4T1 tumor model showed good correlation with histological results. However, eMSOT's performance can be affected by the noise and artifacts. To address this challenge, Olefir *et al.* proposed a Bayesian implementation of eMSOT.[112] A spectral reliability map (SRM) was developed to evaluate the noise level, allowing the automatic weighting of spectra based on their reliability. The constraints to formulate the Bayesian inversion problem were based on probabilistic graphical models. Simulation results showed that it is more robust than conventional eMSOT.

Tzoumas *et al.* also adapted a different statistical detection framework in the context of MSOT using a generalized likelihood ratio test (GLRT) with adaptive matched filter (AMF), assuming background tissue spectral follows multivariate Gaussian distribution and not affected by sparsely present target molecules.[113,114] The background statistical distribution including mean and covariance matrix can be calculated from the available pixels, based on which to decide whether the specific pixel contains target. To mitigate covariance contamination caused by target size and intensity, a robust statistical detection framework (RSDF) was proposed, using archived MSOT images as training data to estimate covariation matrix.[114] *In vivo* imaging experiment on mice containing fluorophores demonstrated that AMF and RSDF enhanced imaging sensitivity in molecular detection by MSOT. A mini-review was published, evaluating these two classes of molecular detection methods.[108]

C. Blind source unmixing

The exploration of various blind source unmixing algorithms holds promise in mitigating the challenges in linear spectral unmixing and enhancing the precision of qPAI. More importantly, blind source unmixing is a method capable of separating an image without relying on prior knowledge of spectral signature or fluence distribution. This technique possesses the capability to recover unknown source "signals" solely based on the understanding of observed mixtures of these signals within an image, which is more suitable to study complex biological tissues.[115]

Principal component analysis (PCA) and independent component analysis (ICA) are both commonly used blind unmixing approaches for qPAI.[116,117] PCA identifies and emphasizes the most

significant patterns in a dataset by transforming the original variables into a new set of uncorrelated variables, known as principal components. The measured spectra **M** can be decomposed by PCA into a set of principal component vectors **R** and an absorption matrix **U**. Additionally, data dimension reduction and noise removal can be done by discarding the last few components. ICA is another blind source unmixing method aiming to decompose a multivariate signal into additive, statistically independent components. Unlike PCA, which focuses on maximizing variance and capturing linear relationships, ICA seeks to identify non-Gaussian independent components. However, there is no components' significance indicated. PCA and ICA are usually used together to achieve a more ideal unmixing outcomes, where PCA served as a preprocessor to reduce noise and data dimension. Glatz *et al.* investigated linear spectral fitting, PCA and ICA unmixing methods through *in vivo* mice imaging to extract fluorophores (ICG, Cy7) distribution.[116] The results demonstrated that ICA and PCA are superior to spectral fitting to isolate the low-concentration chromophores. Deán-Ben and coworkers proposed to blindly unmix the logarithm representation of measured PA data to address the spectral coloring effect.[117] In this way, the ratio of target (small amount of dyes) and background (blood) absorption can be isolated, which is proportional to the target concentration assuming the background is uniform. Experiment on mouse *postmortem* with embedded AF750 and ICG showed that it can correct the spatial variation of local fluence. Arabul *et al.* employed non-negative ICA and data-driven approach to deal with multi-spectral PAI data on human carotid plaque samples, as shown in Fig. 7.[105] The unmixed results were corroborated by histology images, indicating the potential of non-negative ICA to identify different chromophores in human carotid plaques.

Cano *et al.* also investigated carotid plaque constituents by blind unmixing. They adopted Piecewise Convex Multiple-Model Endmember Detection (PCOMMEND) dealing with nonhomogeneous illumination.[118] PCOMMEND extends the linear unmixing to multiple sets of endmembers and identifies the regions containing the same set of endmembers.[119] The authors used 162 wavelengths swept from 500 to 1300 nm and carefully selected six wavelengths (500, 560, 680, 780, 940, and 1020 nm) to successfully differentiate hemorrhage, lipid, and collagen, which are the main biomarkers of disease progression.

Grasso *et al.* utilized non-negative matrix factorization (NNMF) to explain the measured spectra without any prior knowledge.[120] After matrix factorization with non-negative constrains, endmembers spectra and abundance maps can be obtained. Unlike PCA, the endmember spectra in NNMF are not required to be orthogonal and their non-negativity makes them more interpretable. The authors used 680–900 nm wavelengths in the *in vivo* experiment to detect hemoglobins and injected ICG. It showed promising performance of NNMF to detect contrast agent and major tissue chromophores in qPAI. Two years later, the same group proposed an automatic superpixel spectral unmixing (SPAX) framework to detect weaker endogenous absorbers and compensate spectral coloring.[121] First, ultrasound images were used to segment different structures and tag them with different optical properties. Then, Monte Carlo simulation was run to obtain a fluence map for normalization and compensation, followed by NNMF unmixing. Last but not least, spectral superpixel sub-sampling (SS) was proposed to identify the less prominent components. *In vivo* validation on

mouse model proved SPAX's capability to detect oxy- and deoxy-hemoglobin, as well as less prominent lipid component. A review paper talking more details about PCA, ICA, NNMF, and SPAX was published in 2022.[122]

D. Learning-based spectral unmixing

Kirchner and colleagues proposed a method named voxel-level context-encoding (CE) qPAI to deduce optical fluence and absorption information.[123] PA signal information and system characteristics were encoded into a context image (CI), which was served as a feature vector. To achieve so, voxel-specific fluence contribution maps (FCMs) telling the likelihood of photon's trajectory were first simulated by Monte Carlo tool. Then, neighborhood PA signals were combined with FCMs to form CIs, which were used to train random forests regressor. The trained regressor enabled the prediction of local fluence in unseen images, facilitating the subsequent reconstruction of absorption. Both training and validation were done using *in silico* data with the aim of oxygenation estimation. For simulation, 750, 800, and 850 nm were selected. The authors adopted two approaches: (1) applied linear spectral unmixing after finding the local fluence by CE-qPAI and (2) direct estimation using a functional CE-qPAI where oxygenation was set as the final target during training. The results showed that both methods outperformed the conventional linear spectral unmixing. However, CE-qPAI only took into account the PA signals in neighborhood, neglecting others in the process.

Cai *et al.* introduced an end-to-end learning approach utilizing ResU-net, wherein the entire set of initial pressure images was served as input and the quantitative images were the final output.[124] The proposed ResU-net mainly consists of convolutional layers, batch normalization layers, rectified linear units, and identity mapping shortcut connections. Its performance was evaluated by the simulated data to mimic sO_2 and ICG quantification in deep tissue with 700–900 nm illumination. Three *in silico* experiments all proved the accuracy and robustness of the end-to-end ResU-net. In practice, it is challenging to acquire a significant number of quantitative images as gold standard, which impedes its application for *in vivo* imaging.

Gröhl's team suggested a learned spectral decoloring (LSD) approach enabling more accurate oxygenation estimation.[125] A fully connected feed-forward neural network was employed. It took pixel-wise initial pressure spectra ($S_{p\lambda}$) at different wavelengths as input and oxygenation level (sO_2) as output. The network was trained by *in silico* data and verified by *in silico*, *in vitro*, and *in vivo* dataset, respectively. All the results showed distinct advantages over linear spectral
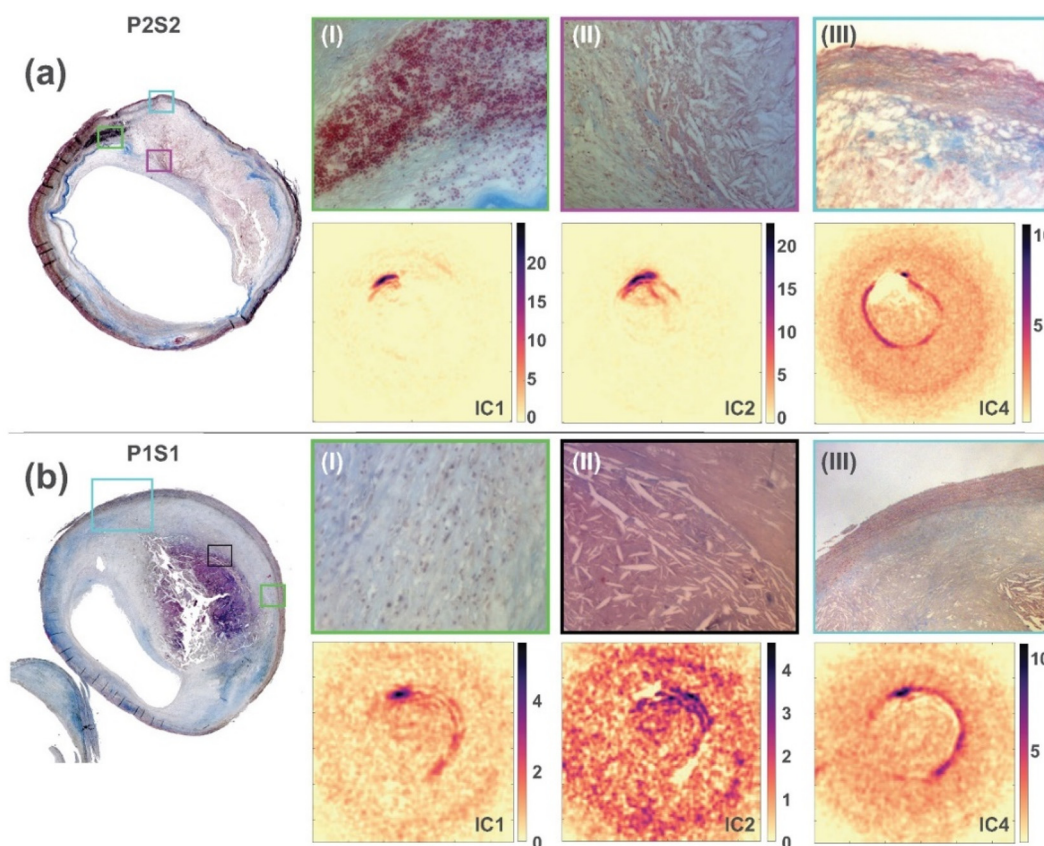
FIG. 7. Correlation of histological details and PAI of plaque samples P1S1 and P2S2. (I)–(III) The detailed view of the corresponding histologic section indicated by squares in different colors. IC1–IC4 are the independent components obtained from spectral unmixing. Reprinted with permission from Arabul *et al.*, Photoacoustics **15**, 100140 (2019), licensed under a Creative Commons Attribution (BY-NC-ND) license.

24 October 2024 08:47:13

unmixing. Despite that the network was trained by simulated data, the proposed single-pixel training framework bridged the domain gap between simulation and real data.

A deep learning (DL) eMSOT[126] was introduced by Olefir *et al.* to further improve the performance of eMSOT mentioned in Sec. VI B. Previous eMSOT and Bayesian-based eMSOT provided good estimation of sO₂, but still limited by the *ad hoc* constraints and non-convex nature of the optimization problem. In this study, the authors proposed a neural network that consists of recurrent and convolutional layers to learn eigenfluence parameters from the spatially distributed spectral information. Although the DL-eMSOT was trained by synthetic data, it still outperformed the conventional eMSOT in most of the experimental cases.

The above-mentioned spectral unmixing methods were summarized in Table VI. To briefly summarize, the multispectral nature of PA data, acquired across various wavelengths, enables the capability to restore complex chromophores mapping based on optical contrast. Conventional linear spectral fitting partially solves the problem but is

limited by known absorption spectra and spectral coloring effect in complex tissues. The computational requirement can be moderate to high, especially when fitting multiple chromophores' spectra to the measured data using least squares fitting. Adding non-negativity constraints increased not only accuracy but also computational time.[110] Statistical detection techniques prove efficient in detecting specific target spectra including low-probability molecules, without explicit prior knowledge of background absorption spectra. The algorithm effectiveness depends on case-specific experimental and system parameters that affect the ability of unmixing algorithms to accurately represent the target and background spectral variations. Typically, these methods rely on large datasets, thereby increasing the computational load. Additionally, generating such datasets can pose challenges, particularly in the context of *in vivo* measurements. Blind source unmixing methods such as PCA, ICA, and NNMF provide accurate results by separating image components without prior knowledge of spectral signatures. The computational demand varies, depending on the complexity of the algorithms. For instance, using PCA and ICA together for noise

TABLE VI. Summary of different spectral unmixing methods for qPAI.

Ref.	Methods		Wavelengths (nm)	Chromophores	Study type	Accuracy	Computational requirement
109	Linear spectral fitting	LSF	1300–1800	Collagen, lipid, water	*Ex vivo* (human bone)	$R^2 = 0.3$–0.6	Low
110		Constrained inversion	680–900	Extrinsic molecules	*In vivo* (mice)	$R^2 > 0.95$	High
111	Statistical detection	eMSOT	700, 710, 720, …, 900	sO₂	*In vivo* (mice)	Error = 1%–4%	High, need training of Eigenspectra
112		Bayesian eMSOT	700, 710, 720, …, 900	sO₂	*In silico*	High	
113,114		AMF, RSDF	700, 710, 720, …, 900	Extrinsic molecules	*In vivo* (mice)	Variable, depend on case-specific parameters	High
116,117	Blind source unmixing	Non-negative PCA + ICA	808, 915, 940, 980	Blood, thrombus, lipid	*Ex vivo* (carotid endarterectomy)	Medium, crosstalk noise appear in spatially variant background	Medium
118		PCOMMEND	500, 560, 680, 780, 940, 1020	Blood, lipid, collagen	*Ex vivo* (carotid endarterectomy)	Error <10%	High
120		NNMF	680, 685, 690, …, 900	Blood, ICG	*In vivo* (mice)	$R > 0.8$	High
121		SPAX	680, 685, 690, …, 970	Blood, fat	*In vivo* (mice)	$R > 0.9$	High
123	Machine learning	CE-qPAI	750, 800, 850	sO₂	Training: *in silico* Validation: *in silico*	Error <4%	High for training. Low-medium for actual calculation
124		ResU-Net	700, 705, 710, …, 800	sO₂, ICG	Training: *in silico* Validation: *in silico*	Error <10%	
125		LSD	700, 710, 720, …, 950	sO₂	Training: *in silico* Validation: *in vivo* (forearm)	Error <10%	
126		DL-eMSOT	700, 710, 720, …, 900	sO₂	Training: *in silico* Validation: *in vivo* (mice)	Error 1%–20%	

reduction and dimensionality reduction may increase computational demand. Similarly, PCOMMEND extends linear unmixing to multiple sets of endmembers, potentially increasing computational complexity. While machine learning methods demonstrate high accuracy in simulated data, they may encounter significant errors when estimating results from experimental data, due to the lack of labeled experimental data with available ground truth. In certain instances, ensemble models show notable discrepancy, but neural networks usually do not allow for detailed analysis of suboptimal performance.[126] The ML methods often require significant computational resources for training neural networks, especially for end-to-end approaches like ResU-Net and DL-eMSOT. However, once trained, they can provide efficient inference. Although recent advancements in spectral unmixing methods tailored to various application scenarios, yield more robust and accurate outcomes, a quantitative comparison of algorithmic performance on experimentally collected datasets is practically unfeasible, as it relies on exact knowledge of the position and content of the molecular component of interest. Further efforts would be necessary to fine-tune the balance between acquisition time and spectral resolution, as demonstrated by Luke *et al.*[127] The careful selection of relevant wavelengths would facilitate effective unmixing and reduce acquisition time.

VII. CONCLUSION

Photoacoustic imaging stands at the forefront of non-invasive imaging modalities, offering a unique blend of optical absorption contrast and ultrasound resolution. Despite its tremendous potential in biomedical field, several challenges persist in accurately reconstructing images and quantifying physiological parameters from raw PA signals. First, image acquisition is usually restricted by transducer and detection geometry. The sparse data from limited views may cause incomplete information and distorted images. This kind of problem can be mitigated by both hardware engineering and algorithmic approaches, such as rotational scanning platform, model-based iterative reconstructions, and machine learning methods. Second, the complexity also arises from nonhomogeneous acoustic properties in biological tissues, leading to artifacts in image reconstruction if assuming homogeneous sound speed and attenuation. From iterative optimization techniques like time-reversal methods to joint reconstruction approaches and emerging deep learning-based methods, researchers have tirelessly pursued solutions to alleviate the negative effect of acoustic heterogeneity on the PA image quality. Third, one should also address spatial variations in optical fluence to derive accurate absorption maps in tissue imaging. Various approaches have been elucidated, including model-based methods, iterative compensation, multimodal approaches, and deep-learning methods. Finally, robust spectral unmixing is necessary in obtaining physiological and biochemical information accurately at depth. Recent advancements in spectral unmixing methods, including linear spectral fitting, statistical detection methods, blind source unmixing, and learning-based approaches, have been instrumental in decoding mixed pixel information into distinct chromophores.

It is worth noting that the above-mentioned challenges in qPAI are often interconnected due to the hybrid nature of optical absorption and acoustic detection in qPAI. Recognizing their interdependence can be beneficial for developing effective strategies to enhance qPAI's reliability and applicability across various biomedical contexts. Some researchers have successfully tackled multiple challenges collectively.[61,76,81,128] Additionally, from this review, it can be seen that machine learning or deep learning-based methods are emerging as a powerful tool recently, to address the various challenges in qPAI. Moreover, thanks to hardware compatibility, data fusion capability, and versatile contrast mechanism of PAI, multimode imaging integration is another new trend of qPAI development, where the additional imaging modality can complement or compensate qPAI in terms of structural and functional information.[74,76,92–96,129]

By synthesizing recent developments, this review provides a comprehensive overview of state-of-the-art qPAI compensation methods reported over the past decade and serves as a valuable reference for researchers and practitioners aiming to enhance the accuracy and reliability of qPAI in complex tissue environments. Thanks to these hardware and software advancements, there are increasing efforts toward clinical translation and application of qPAI. This includes validation studies in preclinical models and clinical trials in human subjects to evaluate the diagnostic and prognostic utility of qPAI in various medical contexts, such as cancer detection, cardiovascular imaging, and neuroimaging. Understanding the real clinical needs and the associated challenges at the outset is crucial for implementing corresponding techniques effectively. The successful adoption and integration of these techniques hold the potential to augment the clinical utility of PAI and streamline its transition into routine medical practice, thereby fostering the integration of PAI into clinical settings and advancing its utilization in healthcare.

AUTHOR DECLARATIONS
Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Ruochong Zhang, Rabia'tul A'dawiah, and Tristan Wen Jie Choo contributed equally to this work.

Ruochong Zhang: Conceptualization (lead); Investigation (equal); Writing – original draft (equal); Writing – review & editing (lead). **Rabia'tul A'dawiah:** Investigation (equal); Writing – original draft (equal). **Tristan Wen Jie Choo:** Investigation (equal); Writing – original draft (equal). **Xiuting Li:** Writing – original draft (equal). **Ghayathri Balasundaram:** Writing – review & editing (equal). **Yi Qi:** Writing – review & editing (equal). **Yonggeng Goh:** Writing – review & editing (equal). **Renzhe Bi:** Resources (equal); Supervision (equal); Writing – review & editing (equal). **Malini Olivo:** Resources (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article because no new data were created or analyzed in this study.

REFERENCES

[1] A. G. Bell, "On the production and reproduction of sound by light," Am. J. Sci. **s3-20**(118), 305–324 (1880).

[2] B. Dong, H. Li, Z. Zhang, K. Zhang, S. Chen, C. Sun, and H. F. Zhang, "Isometric multimodal photoacoustic microscopy based on optically transparent micro-ring ultrasonic detection," Optica **2**(2), 169–176 (2015).

[3] P. Hai, T. Imai, S. Xu, R. Zhang, R. L. Aft, J. Zou, and L. V. Wang, "High-throughput, label-free, single-cell photoacoustic microscopy of intratumoral metabolic heterogeneity," Nat. Biomed. Eng. **3**(5), 381–391 (2019).

[4] T. T. W. Wong, R. Zhang, C. Zhang, H.-C. Hsu, K. I. Maslov, L. Wang, J. Shi, R. Chen, K. K. Shung, and Q. Zhou, "Label-free automated three-dimensional imaging of whole organs by microtomy-assisted photoacoustic microscopy," Nat. Commun. **8**(1), 1386 (2017).

[5] J. Zhang, D. Peng, W. Qin, W. Qi, X. Liu, Y. Luo, Q. Guo, and L. Xi, "Organ-PAM: Photoacoustic microscopy of whole-organ multiset vessel systems," Laser Photonics Rev. **17**, 2201031 (2023).

[6] L. V. Wang and S. Hu, "Photoacoustic tomography: In vivo imaging from organelles to organs," Science **335**(6075), 1458–1462 (2012).

[7] J. Yao and L. V. Wang, "Photoacoustic microscopy," Laser Photonics Rev **7**(5), 758–778 (2013).

[8] J. Xia, J. Yao, and L. V. Wang, "Photoacoustic tomography: Principles and advances," Electromagn. Waves (Cambridge, MA) **147**, 1–22 (2014).

[9] H. Guo, Y. Li, W. Qi, and L. Xi, "Photoacoustic endoscopy: A progress review," J. Biophotonics **13**(12), e202000217 (2020).

[10] A. B. E. Attia, G. Balasundaram, M. Moothanchery, U. S. Dinish, R. Bi, V. Ntziachristos, and M. Olivo, "A review of clinical photoacoustic imaging: Current and future trends," Photoacoustics **16**, 100144 (2019).

[11] A. Karlas, M. A. Pleitez, J. Aguirre, and V. Ntziachristos, "Optoacoustic imaging in endocrinology and metabolism," Nat. Rev. Endocrinol. **17**(6), 323–335 (2021).

[12] S. Manohar and M. Dantuma, "Current and future trends in photoacoustic breast imaging," Photoacoustics **16**, 100134 (2019).

[13] I. Steinberg, D. M. Huland, O. Vermesh, H. E. Frostig, W. S. Tummers, and S. S. Gambhir, "Photoacoustic clinical imaging," Photoacoustics **14**, 77–98 (2019).

[14] W. Choi, E.-Y. Park, S. Jeon, and C. Kim, "Clinical photoacoustic imaging platforms," Biomed. Eng. Lett. **8**(2), 139–155 (2018).

[15] L. Lin and L. V. Wang, "The emerging role of photoacoustic imaging in clinical oncology," Nat. Rev. Clin. Oncol. **19**(6), 365–384 (2022).

[16] S. Jeon, J. Kim, D. Lee, J. W. Baik, and C. Kim, "Review on practical photoacoustic microscopy," Photoacoustics **15**, 100141 (2019).

[17] M. Omar, J. Aguirre, and V. Ntziachristos, "Optoacoustic mesoscopy for biomedicine," Nat. Biomed. Eng. **3**(5), 354–370 (2019).

[18] L. V. Wang and J. Yao, "A practical guide to photoacoustic tomography in the life sciences," Nat. Methods **13**(8), 627–638 (2016).

[19] A. Taruttis and V. Ntziachristos, "Advances in real-time multispectral optoacoustic imaging and its applications," Nat. Photonics **9**, 219 (2015).

[20] P. K. Upputuri and M. Pramanik, "Recent advances toward preclinical and clinical translation of photoacoustic tomography: A review," J. Biomed. Opt. **22**(4), 41006 (2017).

[21] M. Li, Y. Tang, and J. Yao, "Photoacoustic tomography of blood oxygenation: A mini review," Photoacoustics **10**, 65–73 (2018).

[22] L. Li and L. V. Wang, "Recent advances in photoacoustic tomography," BME Front. **2021**, 9823268.

[23] J. Gröhl, M. Schellenberg, K. Dreher, and L. Maier-Hein, "Deep learning for biomedical photoacoustic imaging: A review," Photoacoustics **22**, 100241 (2021).

[24] C. Yang, H. Lan, F. Gao, and F. Gao, "Review of deep learning for photoacoustic imaging," Photoacoustics **21**, 100215 (2021).

[25] M. Kim, I. Pelivanov, and M. O'Donnell, "Review of deep learning approaches for interleaved photoacoustic and ultrasound (PAUS) imaging," IEEE Trans. Ultrason. Ferroelectr. Freq. Control **70**, 1591–1606 (2023).

[26] D. Das, A. Sharma, P. Rajendran, and M. Pramanik, "Another decade of photoacoustic imaging," Phys. Med. Biol. **66**(5), 05TR01 (2021).

[27] M. W. Schellenberg and H. K. Hunt, "Hand-held optoacoustic imaging: A review," Photoacoustics **11**, 14–27 (2018).

[28] C. Lutzweiler and D. Razansky, "Optoacoustic imaging and tomography: Reconstruction approaches and outstanding challenges in image performance and quantification," Sensors **13**(6), 7345–7384 (2013).

[29] B. T. Cox, J. G. Laufer, P. C. Beard, and S. R. Arridge, "Quantitative spectroscopic photoacoustic imaging: A review," J. Biomed. Opt. **17**(6), 61202 (2012).

[30] L. V. Wang, "Tutorial on photoacoustic microscopy and computed tomography," IEEE J. Sel. Top. Quantum Electron. **14**(1), 171–179 (2008).

[31] Y. Xu, L. V. Wang, G. Ambartsoumian, and P. Kuchment, "Reconstructions in limited-view thermoacoustic tomography," Med. Phys. **31**(4), 724–733 (2004).

[32] X. L. Deán-Ben and D. Razansky, "On the link between the speckle free nature of optoacoustics and visibility of structures in limited-view tomography," Photoacoustics **4**(4), 133–140 (2016).

[33] J. Gateau, M. Á. A. Caballero, A. Dima, and V. Ntziachristos, "Three-dimensional optoacoustic tomography using a conventional ultrasound linear detector array: Whole-body tomographic system for small animals," Med. Phys. **40**(1), 13302 (2013).

[34] G. Li, L. Li, L. Zhu, J. Xia, and L. V. Wang, "Multiview Hilbert transformation for full-view photoacoustic computed tomography using a linear array," J. Biomed. Opt. **20**(6), 66010 (2015).

[35] W. Shu, M. Ai, T. Salcudean, R. Rohling, P. Abolmaesumi, and S. Tang, "Broadening the detection view of 2D photoacoustic tomography using two linear array transducers," Opt. Express **24**(12), 12755–12768 (2016).

[36] B. Huang, J. Xia, K. Maslov, and L. V. Wang, "Improving limited-view photoacoustic tomography with an acoustic reflector," J. Biomed. Opt. **18**(11), 110505 (2013).

[37] G. Li, J. Xia, K. Maslov, M. A. Anastasio, and L. V. Wang, "Tripling the detection view of high-frequency linear-array-based photoacoustic computed tomography by using two planar acoustic reflectors," Quant. Imaging Med. Surg. **5**(1), 57 (2015).

[38] X. Liu, D. Peng, X. Ma, W. Guo, Z. Liu, D. Han, X. Yang, and J. Tian, "Limited-view photoacoustic imaging based on an iterative adaptive weighted filtered backprojection approach," Appl. Opt. **52**(15), 3477–3483 (2013).

[39] Y. Han, S. Tzoumas, A. Nunes, V. Ntziachristos, and A. Rosenthal, "Sparsity-based acoustic inversion in cross-sectional multiscale optoacoustic imaging," Med. Phys. **42**(9), 5444–5452 (2015).

[40] X. Li, S. Zhang, J. Wu, S. Huang, Q. Feng, L. Qi, and W. Chen, "Multispectral interlaced sparse sampling photoacoustic tomography," IEEE Trans. Med. Imaging **39**(11), 3463–3474 (2020).

[41] X. Li, J. Ge, S. Zhang, J. Wu, L. Qi, and W. Chen, "Multispectral interlaced sparse sampling photoacoustic tomography based on directional total variation," Comput. Methods Programs Biomed. **214**, 106562 (2021).

[42] R. Shang, R. Archibald, A. Gelb, and G. P. Luke, "Sparsity-based photoacoustic image reconstruction with a linear array transducer and direct measurement of the forward model," J. Biomed. Opt. **24**(3), 031015 (2019).

[43] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," IEEE Trans. Image Process. **16**(12), 2992–3004 (2007).

[44] M. Ai, J. Cheng, D. Karimi, S. E. Salcudean, R. Rohling, P. Abolmaesumi, and S. Tang, "Investigation of photoacoustic tomography reconstruction with a limited view from linear array," J. Biomed. Opt. **26**(9), 96009 (2021).

[45] J. Zhu, N. Huynh, O. Ogunlade, R. Ansari, F. Lucka, B. Cox, and P. Beard, "Mitigating the limited view problem in photoacoustic tomography for a planar detection geometry by regularised iterative reconstruction," IEEE Trans. Med. Imaging **42**, 2603–2615 (2023).

[46] R. Wang, J. Zhu, J. Xia, J. Yao, J. Shi, and C. Li, "Photoacoustic imaging with limited sampling: A review of machine learning approaches," Biomed. Opt. Express **14**(4), 1777–1799 (2023).

[47] N. Davoudi, X. L. Deán-Ben, and D. Razansky, "Deep learning optoacoustic tomography with sparse data," Nat. Mach. Intell. **1**(10), 453–460 (2019).

[48] N. Davoudi, B. Lafci, A. Özbek, X. L. Deán-Ben, and D. Razansky, "Deep learning of image-and time-domain data enhances the visibility of structures in optoacoustic tomography," Opt. Lett. **46**(13), 3029–3032 (2021).

[49] T. Vu, M. Li, H. Humayun, Y. Zhou, and J. Yao, "A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer," Exp. Biol. Med. **245**(7), 597–605 (2020).

[50] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, "Model-based learning for accelerated, limited-view

3-D photoacoustic tomography," IEEE Trans. Med. Imaging **37**(6), 1382–1393 (2018).

[51] S. Guan, A. A. Khan, S. Sikdar, and P. V. Chitnis, "Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning," Sci. Rep. **10**(1), 8510 (2020).

[52] B. E. Treeby, E. Z. Zhang, A. S. Thomas, and B. T. Cox, "Measurement of the ultrasound attenuation and dispersion in whole human blood and its components from 0–70 MHz," Ultrasound Med. Biol. **37**(2), 289–300 (2011).

[53] F. A. Duck, *Physical Properties of Tissues: A Comprehensive Reference Book* (Academic Press, 2013).

[54] Y. Xu and L. V. Wang, "Time reversal and its application to tomography with diffracting sources," Phys. Rev. Lett. **92**(3), 33902 (2004).

[55] B. E. Treeby, E. Z. Zhang, and B. T. Cox, "Photoacoustic tomography in absorbing acoustic media using time reversal," Inverse Probl. **26**(11), 115003 (2010).

[56] B. E. Treeby and B. T. Cox, "k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," J. Biomed. Opt. **15**(2), 21314 (2010).

[57] C. Huang, L. Nie, R. W. Schoonover, L. V. Wang, and M. A. Anastasio, "Photoacoustic computed tomography correcting for heterogeneity and attenuation," J. Biomed. Opt. **17**(6), 61211 (2012).

[58] P. Stefanov and G. Uhlmann, "Thermoacoustic tomography with variable sound speed," Inverse Probl. **25**(7), 75011 (2009).

[59] J. Qian, P. Stefanov, G. Uhlmann, and H. Zhao, "An efficient Neumann series–based algorithm for thermoacoustic and photoacoustic tomography with variable sound speed," SIAM J. Imaging Sci. **4**(3), 850–883 (2011).

[60] D. Lai, L. Mao, Q. Xu, and J. Li, in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biological Society* (IEEE, 2020), pp. 1923–1926.

[61] C. Huang, K. Wang, R. W. Schoonover, L. V. Wang, and M. A. Anastasio, "Joint reconstruction of absorbed optical energy density and sound speed distributions in photoacoustic computed tomography: A numerical investigation," IEEE Trans. Comput. Imaging **2**(2), 136–149 (2016).

[62] C. Cai, X. Wang, K. Si, J. Qian, J. Luo, and C. Ma, "Feature coupling photoacoustic computed tomography for joint reconstruction of initial pressure and sound speed in vivo," Biomed. Opt. Express **10**(7), 3447–3462 (2019).

[63] K. X. Deng, M. X. Cui, H. Z. Zuo, X. H. Wang, C. J. Cai, J. W. Luo, and C. Ma, "Speed-of-sound heterogeneity compensation method in photoacoustic computed tomographic image reconstruction," Chin. J. Lasers **48**, 1507001 (2021).

[64] M. Haltmeier and L. V. Nguyen, "Analysis of iterative methods in photoacoustic tomography with variable sound speed," SIAM J. Imaging Sci. **10**(2), 751–781 (2017).

[65] J. Poudel, T. P. Matthews, M. A. Anastasio, and L. V. Wang, "Compensation for acoustic heterogeneities in photoacoustic computed tomography using a variable temporal data truncation reconstruction method," Proc. SPIE **9708**, 97083W (2016).

[66] M. A. Anastasio, J. Zhang, X. Pan, Y. Zou, G. Ku, and L. V. Wang, "Half-time image reconstruction in thermoacoustic tomography," IEEE Trans. Med. Imaging **24**(2), 199–210 (2005).

[67] C. Huang, K. Wang, L. Nie, L. V. Wang, and M. A. Anastasio, "Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media," IEEE Trans. Med. Imaging **32**(6), 1097–1110 (2013).

[68] S. M. Ranjbaran, H. S. Aghamiry, A. Gholami, S. Operto, and K. Avanaki, "Quantitative photoacoustic tomography using iteratively refined wavefield reconstruction inversion: A simulation study," IEEE Trans. Med. Imaging **43**, 874–885 (2023).

[69] E. Merčep, J. L. Herraiz, X. L. Deán-Ben, and D. Razansky, "Transmission–reflection optoacoustic ultrasound (TROPUS) computed tomography of small animals," Light Sci. Appl. **8**(1), 18 (2019).

[70] S. Jeon and C. Kim, "Deep learning-based speed of sound aberration correction in photoacoustic images," Proc. SPIE **11240**, 112400J (2020).

[71] S. Jeon, W. Choi, B. Park, and C. Kim, "A deep learning-based model that reduces speed of sound aberrations for improved in vivo photoacoustic imaging," IEEE Trans. Image Process. **30**, 8773–8784 (2021).

[72] E. M. A. Anas, H. K. Zhang, C. Audigier, and E. M. Boctor, *Lecture Notes in Computer Science* (Springer Verlag, 2018), pp. 3–11.

[73] C. Dehner, G. Zahnd, V. Ntziachristos, and D. Jüstel, "A deep neural network for real-time optoacoustic image reconstruction with adjustable speed of sound," Nat. Mach. Intell. **5**, 1130–1141 (2023).

[74] M. Shi, T. Vercauteren, and W. Xia, in *2023 IEEE International Ultrasonics Symposium* (IEEE, 2023), pp. 1–4.

[75] X. L. Deán-Ben, V. Ntziachristos, and D. Razansky, "Effects of small variations of speed of sound in optoacoustic tomographic imaging," Med. Phys. **41**(7), 73301 (2020).

[76] A. Pattyn, Z. Mumm, N. Alijabbari, N. Duric, M. A. Anastasio, and M. Mehrmohammadi, "Model-based optical and acoustical compensation for photoacoustic tomography of heterogeneous mediums," Photoacoustics **23**, 100275 (2021).

[77] X. Lin, M. Sun, Y. Liu, Z. Shen, Y. Shen, and N. Feng, "Variable speed of sound compensation in the linear-array photoacoustic tomography using a multi-stencils fast marching method," Biomed. Signal Process. Control **44**, 67–74 (2018).

[78] H. Jin, R. Zhang, S. Liu, and Y. Zheng, "Rapid three-dimensional photoacoustic imaging reconstruction for irregularly layered heterogeneous media," IEEE Trans. Med. Imaging **39**(4), 1041–1050 (2020).

[79] H. Jin, S. Liu, R. Zhang, Z. Zheng, and Y. Zheng, in *2020 IEEE International Symposium on Circuits System* (IEEE, 2020).

[80] L. Zhao, M. Yang, Y. Jiang, and C. Li, "Optical fluence compensation for handheld photoacoustic probe: An in vivo human study case," J. Innov. Opt. Health Sci. **10**(04), 1740002 (2017).

[81] G.-S. Jeng, M.-L. Li, M. Kim, S. J. Yoon, J. J. Pitre, Jr., D. S. Li, I. Pelivanov, and M. O'Donnell, "Real-time interleaved spectroscopic photoacoustic and ultrasound (PAUS) scanning with simultaneous fluence compensation and motion correction," Nat. Commun. **12**(1), 716 (2021).

[82] Z. Liang, S. Zhang, J. Wu, X. Li, Z. Zhuang, Q. Feng, W. Chen, and L. Qi, "Automatic 3-D segmentation and volumetric light fluence correction for photoacoustic tomography based on optimal 3-D graph search," Med. Image Anal. **75**, 102275 (2022).

[83] M. Dantuma, S. Kruitwagen, R. Ortega-Julia, R. P. Pompe van Meerdervoort, and S. Manohar, "Tunable blood oxygenation in the vascular anatomy of a semi-anthropomorphic photoacoustic breast phantom," J. Biomed. Opt. **26**(3), 36003 (2021).

[84] M. Kirillin, V. Perekatova, I. Turchin, and P. Subochev, "Fluence compensation in raster-scan optoacoustic angiography," Photoacoustics **8**, 59–67 (2017).

[85] T. Hirasawa, R. J. Iwatate, M. Kamiya, S. Okawa, Y. Urano, and M. Ishihara, "Multispectral photoacoustic imaging of tumours in mice injected with an enzyme-activatable photoacoustic probe," J. Opt. **19**(1), 14002 (2016).

[86] X. Zhou, N. Akhlaghi, K. A. Wear, B. S. Garra, T. J. Pfefer, and W. C. Vogt, "Evaluation of fluence correction algorithms in multispectral photoacoustic imaging," Photoacoustics **19**, 100181 (2020).

[87] M. Kim, G.-S. Jeng, M. O'Donnell, and I. Pelivanov, "Correction of wavelength-dependent laser fluence in swept-beam spectroscopic photoacoustic imaging with a hand-held probe," Photoacoustics **19**, 100192 (2020).

[88] J. C. Ranasinghesagara, Y. Jiang, and R. J. Zemp, "Reflection-mode multiple-illumination photoacoustic sensing to estimate optical properties," Photoacoustics **2**(1), 33–38 (2014).

[89] Y. Liu, M. Sun, T. Liu, Y. Ma, D. Hu, C. Li, and N. Feng, "Quantitative reconstruction of absorption coefficients for photoacoustic tomography," Appl. Sci. **9**(6), 1187 (2019).

[90] Y. Wu, J. Kang, W. G. Lesniak, M. G. Pomper, and E. M. Boctor, in *2021 IEEE International Ultrasonics Symposium* (IEEE, 2021), pp. 1–4.

[91] M. A. Naser, D. R. T. Sampaio, N. M. Muñoz, C. A. Wood, T. M. Mitcham, W. Stefan, K. V. Sokolov, T. Z. Pavan, R. Avritscher, and R. R. Bouchard, "Improved photoacoustic-based oxygen saturation estimation with SNR-regularized local fluence correction," IEEE Trans. Med. Imaging **38**(2), 561–571 (2018).

[92] T. Han, M. Yang, F. Yang, L. Zhao, Y. Jiang, and C. Li, "A three-dimensional modeling method for quantitative photoacoustic breast imaging with hand-held probe," Photoacoustics **21**, 100222 (2021).

[93] R. Bulsink, M. Kuniyil Ajith Singh, M. Xavierselvan, S. Mallidi, W. Steenbergen, and K. J. Francis, "Oxygen saturation imaging using LED-based photoacoustic system," Sensors **21**(1), 283 (2021).

[94] H. Jin, R. Zhang, S. Liu, Z. Zheng, and Y. Zheng, "A single sensor dual-modality photoacoustic fusion imaging for compensation of light fluence variation," IEEE Trans. Biomed. Eng. **66**(6), 1810–1813 (2019).

[95]A. Q. Bauer, R. E. Nothdurft, T. N. Erpelding, L. V. Wang, and J. P. Culver, "Quantitative photoacoustic imaging: Correcting for heterogeneous light fluence distributions using diffuse optical tomography," J. Biomed. Opt. **16**(9), 96016 (2011).

[96]S. Mahmoodkalayeh, M. Zarei, M. A. Ansari, K. Kratkiewicz, M. Ranjbaran, R. Manwar, and K. Avanaki, "Improving vascular imaging with co-planar mutually guided photoacoustic and diffuse optical tomography: A simulation study," Biomed. Opt. Express **11**(8), 4333–4347 (2020).

[97]A. Madasamy, V. Gujrati, V. Ntziachristos, and J. Prakash, "Deep learning methods hold promise for light fluence compensation in three-dimensional optoacoustic imaging," J. Biomed. Opt. **27**(10), 106004 (2022).

[98]C. Yang, H. Lan, H. Zhong, and F. Gao, in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (IEEE, 2019), pp. 741–744.

[99]A. M. Caravaca-Aguirre, D. B. Conkey, J. D. Dove, H. Ju, T. W. Murray, and R. Piestun, "High contrast three-dimensional photoacoustic imaging through scattering media by localized optical fluence enhancement," Opt. Express **21**(22), 26671–26676 (2013).

[100]M. N. Fadhel, E. Hysi, H. Assi, and M. C. Kolios, "Fluence-matching technique using photoacoustic radiofrequency spectra for improving estimates of oxygen saturation," Photoacoustics **19**, 100182 (2020).

[101]Z. Sun and L. Sun, "Simultaneous reconstruction of optical absorption property and speed of sound in intravascular photoacoustic tomography," Inverse Probl. Sci. Eng. **29**(12), 1764–1788 (2021).

[102]T. Ding, K. Ren, and S. Vallélian, "A one-step reconstruction algorithm for quantitative photoacoustic imaging," Inverse Probl. **31**(9), 95005 (2015).

[103]A. Javaherian and S. Holman, "Direct quantitative photoacoustic tomography for realistic acoustic media," Inverse Probl. **35**(8), 84004 (2019).

[104]A. Dolet, R. Ammanouil, V. Petrilli, C. Richard, P. Tortoli, D. Vray, and F. Varray, "In vitro and in vivo multispectral photoacoustic imaging for the evaluation of chromophore concentration," Sensors **21**(10), 3366 (2021).

[105]M. U. Arabul, M. C. M. Rutten, P. Bruneval, M. R. H. M. van Sambeek, F. N. van de Vosse, and R. G. P. Lopata, "Unmixing multi-spectral photoacoustic sources in human carotid plaques using non-negative independent component analysis," Photoacoustics **15**, 100140 (2019).

[106]C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," Remote Sens. Environ. **149**, 70–87 (2014).

[107]C. Bench and B. Cox, "Quantitative photoacoustic estimates of intervascular blood oxygenation differences using linear unmixing," J. Phys.: Conf. Ser. **1761**, 12001 (2021).

[108]S. Tzoumas and V. Ntziachristos, "Spectral unmixing techniques for optoacoustic imaging of tissue pathophysiology," Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **375**(2107), 20170262 (2017).

[109]T. Feng, Y. Ge, Y. Xie, W. Xie, C. Liu, L. Li, D. Ta, Q. Jiang, and Q. Cheng, "Detection of collagen by multi-wavelength photoacoustic analysis as a biomarker for bone health assessment," Photoacoustics **24**, 100296 (2021).

[110]L. Ding, X. L. Deán-Ben, N. C. Burton, R. W. Sobol, V. Ntziachristos, and D. Razansky, "Constrained inversion and spectral unmixing in multispectral optoacoustic tomography," IEEE Trans. Med. Imaging **36**(8), 1676–1685 (2017).

[111]S. Tzoumas, A. Nunes, I. Olefir, S. Stangl, P. Symvoulidis, S. Glasl, C. Bayer, G. Multhoff, and V. Ntziachristos, "Eigenspectra optoacoustic tomography achieves quantitative blood oxygenation imaging deep in tissues," Nat. Commun. **7**(1), 12121 (2016).

[112]I. Olefir, S. Tzoumas, H. Yang, and V. Ntziachristos, "A Bayesian approach to eigenspectra optoacoustic tomography," IEEE Trans. Med. Imaging **37**(9), 2070–2079 (2018).

[113]S. Tzoumas, N. C. Deliolanis, S. Morscher, and V. Ntziachristos, "Unmixing molecular agents from absorbing tissue in multispectral optoacoustic tomography," IEEE Trans. Med. Imaging **33**(1), 48–60 (2013).

[114]S. Tzoumas, A. Kravtsiv, Y. Gao, A. Buehler, and V. Ntziachristos, "Statistical molecular target detection framework for multispectral optoacoustic tomography," IEEE Trans. Med. Imaging **35**(12), 2534–2545 (2016).

[115]M. S. Karoui, Y. Deville, S. Hosseini, and A. Ouamri, "Blind spatial unmixing of multispectral images: New methods combining sparse component analysis, clustering and non-negativity constraints," Pattern Recognit. **45**(12), 4263–4278 (2012).

[116]J. Glatz, N. C. Deliolanis, A. Buehler, D. Razansky, and V. Ntziachristos, "Blind source unmixing in multi-spectral optoacoustic tomography," Opt. Express **19**(4), 3175–3184 (2011).

[117]X. L. Deán-Ben, A. Buehler, D. Razansky, and V. Ntziachristos, "Estimation of optoacoustic contrast agent concentration with self-calibration blind logarithmic unmixing," Phys. Med. Biol. **59**(17), 4785 (2014).

[118]C. Cano, C. Matos, A. Gholampour, M. van Sambeek, R. Lopata, and M. Wu, "Blind spectral unmixing for characterization of plaque composition based on multispectral photoacoustic imaging," Sci. Rep. **13**(1), 4119 (2023).

[119]A. Zare, P. Gader, O. Bchir, and H. Frigui, "Piecewise convex multiple-model endmember detection and spectral unmixing," IEEE Trans. Geosci. Remote Sens. **51**(5), 2853–2862 (2012).

[120]V. Grasso, J. Holthof, and J. Jose, "An automatic unmixing approach to detect tissue chromophores from multispectral photoacoustic imaging," Sensors **20**(11), 3235 (2020).

[121]V. Grasso, R. Willumeit-Römer, and J. Jose, "Superpixel spectral unmixing framework for the volumetric assessment of tissue chromophores: A photoacoustic data-driven approach," Photoacoustics **26**, 100367 (2022).

[122]V. Grasso, H. W. Hassan, P. Mirtaheri, R. Willumeit-Römer, and J. Jose, "Recent advances in photoacoustic blind source spectral unmixing approaches and the enhanced detection of endogenous tissue chromophores," Front. Signal Process. **2**, 984901 (2022).

[123]T. Kirchner, J. Gröhl, and L. Maier-Hein, "Context encoding enables machine learning-based quantitative photoacoustics," J. Biomed. Opt. **23**(5), 56008 (2018).

[124]C. Cai, K. Deng, C. Ma, and J. Luo, "End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging," Opt. Lett. **43**(12), 2752–2755 (2018).

[125]J. Gröhl, T. Kirchner, T. J. Adler, L. Hacker, N. Holzwarth, A. Hernández-Aguilera, M. A. Herrera, E. Santos, S. E. Bohndiek, and L. Maier-Hein, "Learned spectral decoloring enables photoacoustic oximetry," Sci. Rep. **11**(1), 6565 (2021).

[126]I. Olefir, S. Tzoumas, C. Restivo, P. Mohajerani, L. Xing, and V. Ntziachristos, "Deep learning-based spectral unmixing for optoacoustic imaging of tissue oxygen saturation," IEEE Trans. Med. Imaging **39**(11), 3643–3654 (2020).

[127]G. P. Luke, S. Y. Nam, and S. Y. Emelianov, "Optical wavelength selection for improved spectroscopic photoacoustic imaging," Photoacoustics **1**(2), 36–42 (2013).

[128]K. Johnstonbaugh, S. Agrawal, D. Abhishek, M. Homewood, S. P. K. Karri, and S.-R. Kothapalli, "Novel deep learning architecture for optical fluence dependent photoacoustic target localization," Proc. SPIE **10878**, 108781L (2019).

[129]E. Merčep, X. L. Deán-Ben, and D. Razansky, "Combined pulse-echo ultrasound and multispectral optoacoustic tomography with a multi-segment detector array," IEEE Trans. Med. Imaging **36**(10), 2129–2137 (2017).