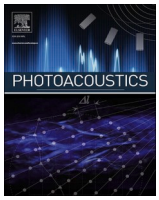
Photoacoustics 44 (2025) 100729

Contents lists available at ScienceDirect

Photoacoustics

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



Spectroscopic photoacoustic denoising framework using hybrid analytical and data-free learning method

Fangzhou Lin [a,1], Shang Gao [a,1], Yichuan Tang [a], Xihan Ma [a], Ryo Murakami [a], Ziming Zhang [b], John D. Obayemi [c,d], Winston O. Soboyejo [d], Haichong K. Zhang [a,c,e,*]

[a] Department of Robotics Engineering, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA
[b] Department of Electrical & Computer Engineering, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA
[c] Department of Biomedical Engineering, Gateway Park Life Sciences Center, Worcester Polytechnic Institute (WPI), 60 Prescott Street, Worcester, MA 01605, USA
[d] Department of Mechanical & Materials Engineering, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA
[e] Department of Computer Science, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA

ARTICLE INFO

Keywords:
Photoacoustic imaging
Quantitative Imaging
Denoising
Data-free
Multiwavelength
In vivo demonstration

ABSTRACT

Spectroscopic photoacoustic (sPA) imaging uses multiple wavelengths to differentiate and quantify chromophores based on their unique optical absorption spectra. This technique has been widely applied in areas such as vascular mapping, tumor detection, and therapeutic monitoring. However, PA imaging is highly susceptible to noise, leading to a low signal-to-noise ratio (SNR) and compromised image quality. Furthermore, low SNR in spectral data adversely affects spectral unmixing outcomes, hindering accurate quantitative PA imaging. Traditional denoising techniques like frame averaging, though effective in improving SNR, can be impractical for dynamic imaging scenarios due to reduced frame rates. Advanced methods, including learning-based approaches and analytical algorithms, have demonstrated promise but often require extensive training data and parameter tuning. Moreover, spectral information preservation is unclear, limiting their adaptability for clinical usage. Additionally, training data is not always accessible for learning-based methods. In this work, we propose a Spectroscopic Photoacoustic Denoising (SPADE) framework using hybrid analytical and data-free learning method. This framework integrates a data-free learning-based method with an efficient BM3D-based analytical approach while preserving spectral integrity, providing noise reduction, and ensuring that functional information is maintained. The SPADE framework was validated through simulation, phantom, in vivo, and ex vivo studies. These studies demonstrated that SPADE improved image SNR by over 15 *dB* in high noise cases and preserved spectral information ($R > 0.8$), outperforming conventional methods, especially in low SNR conditions. SPADE presents a promising solution for preserving the accuracy of quantitative PA imaging in clinical applications where noise reduction and spectral preservation are critical.

1. Introduction

Photoacoustic (PA) imaging is an emerging functional imaging modality that relies on laser-generated ultrasound (US). Chromophores absorb a short light pulse and generate thermoelastic waves through adiabatic expansion, revealing the optical absorption of the material [1,2]. Different materials possess unique spectroscopic characteristics (i.e. PA spectrum). Spectroscopic PA (sPA) imaging, also known as multispectral PA imaging, can be used to characterize different types of chromophores using multiwavelength PA excitation [3]. Quantitative PA imaging leverages this spectral information to estimate the concentration of specific chromophores, allowing for more precise assessments in applications such as neurovascular mapping [4,5], contrast agent quantification for tumor detection [6–8], oxygenation mapping [9,10], and therapeutic monitoring [11–15], thereby enhancing diagnostic accuracy and treatment evaluation.

While it has demonstrated significant promise in many applications, PA imaging is susceptible to noise and artifacts. Noises arise from various sources including acoustic background noise, as well as electronic noise from detectors and amplifiers [16–19]. These different

* Corresponding author at: Department of Robotics Engineering, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA.
E-mail address: hzhang10@wpi.edu (H.K. Zhang).
[1] The two authors contributed equally to this work.

https://doi.org/10.1016/j.pacs.2025.100729
Received 15 November 2024; Received in revised form 24 March 2025; Accepted 27 April 2025
Available online 1 May 2025

noises in the PA signal led to a low signal-to-noise ratio (SNR), resulting in poor PA imaging quality. In order to manage noise signals and provide higher contrast PA imaging, denoising is performed to enhance the quality of this imaging modality. Averaging is a simple and effective technique for enhancing SNR by suppressing random noise and persevering spectrum information. It usually results in a lower frame rate and is not always feasible in dynamic scenarios with motion. Moreover, in an in vivo experimental setting, data acquisition time plays a crucial role in determining the amount of information that can be collected and analyzed. Even without real-time operation, reducing the number of averaging steps—and thereby shortening acquisition time—can minimize the burden on the animal during an imaging session. This also allows for the collection of multiple conditional datasets within the limited available acquisition window. Therefore, we believe that achieving high SNR with limited sPA data is of significant importance.

Several advanced denoising methods that require fewer imaging frames have been reported. Most of these methods can be categorized as data-driven learning-based methods and analytical methods. On the one hand, data-driven learning-based methods achieve remarkable progress in denoising various medical imaging [18,20]. He et al. designed an Attentive GAN for denoising PA microscopy imaging [21]. Although this method is promising for different datasets, it highly relies on the collection and annotation of datasets, which are not widely available in PA imaging [22]. Cheng et al. extend learning-based denoising to an unsupervised method based on the Noise2Noise network, which learns noise patterns from noisy image datasets [20]. However, ground truth images are not always available for supervised learning, and training data may be inaccessible for unsupervised learning, particularly in specific quantitative PA imaging tasks for clinical applications. Additionally, networks in data-driven learning may require additional training for new tasks to avoid performance drops due to the data distribution gap [23–25]. This limits the learning-based methods to work only in some specific tasks and hinders their genericity. On the other hand, numerous analytical algorithms have been introduced for denoising PA signals in single and multiple image frames [26]. Awasthi et al. proposed image-guided filtering for improving PA tomographic image reconstruction [27]. This guided filtering approach was proposed by combining the best features in the input and support images. Although it can improve the reconstructed image quality substantially, the supported images for feature enhancement cannot always be obtained. Kong et al. proposed an empirical mode decomposition for image reconstruction and quality improvement of PA imaging [28]. Despite the fact that both artificial and stochastic noises can be reduced, it needs parameter tuning with an optimized solution. Shi et al. designed a spatiotemporal singular value decomposition (SVD) for low fluency LED-based PA denoising [29]. Although this method shows great promise in enhancing imaging quality, they only demonstrated the effectiveness on single wavelength PA images without demonstration on sPA images. Kazakeviciute et al. provided effective solutions for artifact removal and denoising in multispectral PA images [30]. Their approach significantly improved the image clarity, yet focused more on spatial noise reduction without prioritizing spectral information preservation. The spectral integrity of the output multiwavelength PA image is unclear, and this hinders their applications in the sPA imaging denoise task. Having an advanced annotation-free and data-free (zero-shot fashion) denoising algorithm that could preserve spectral information is critical for sPA imaging.

In this work, we proposed a sPA denoising framework using a hybrid analytical and data-free learning method. A data-free learning-based method has recently been proposed which only consumes input data for imaging denoising [31–33]. The proposed method does not require additional images for network training or extensive hyperparameter tuning. For data preprocessing, we introduce the Spectral Domain Data Re-assembly (SDDR) module, which is designed to process multi-dimensional data, such as multispectral PA imaging, by merging multiple frames into a single entity, while preserving the high-dimensional structure and maintains spectral integrity, including spectroscopic information. In the learning-based component, we adapt residual learning with redesigned Zero-Shot Noise2Noise (ZS-N2N). This model focuses on predicting noise rather than mapping the entire image. Doing so simplifies the learning task and helps preserve functional information. In the analytical component, we design a simple but effective BM3D-based method, which can provide stable denoising performance in sPA imaging with low computational cost. Together, we form a framework that effectively and efficiently denoises sPA images while preserving their spectrum information.

We summarize our contributions as follows: (1) We propose a SPADE framework, to denoise sPA images without training data and parameter tuning. Working in a zero-shot, unsupervised fashion, our framework offers broad applicability and adaptability to various quantitative PA imaging scenarios. (2) The framework preserves the spectral integrity of the pixel during the denoising allowing its usage for functional imaging. (3) We demonstrate notable denoising performance on simulation, phantom, ex vivo, and in vivo setup. The remainder of the paper was structured as follows. Section 2 details the proposed denoising method and algorithm. Section 3 outlines the experiment setup to validate the proposed denoising method in simulation, phantom, ex vivo, and in vivo setup. Section 4 presents the results of the validation experiments. The advantages and limitations of the work are discussed in Section 5, followed by the conclusion in Section 6.

2. Methods

The proposed SPADE framework consists of two components: an SDDR module to directly multispectral sPA imaging for denoising, and Zero-Shot Hybrid (ZS-Hybrid) module leverage redesigned BM3D [34] and ZS-N2N [31] for two-stage sPA imaging denoising. The framework and image processing pipeline are shown in Fig. 1(a). The processing pipeline is design for any image with dimensions of $H \times W \times \lambda$ (axial--lateral-spectral), without any dimensional or scale limitation. Notably, the proposed framework consumes noise-included sPA only and the ZS-Hybrid network is trained in a zero-shot learning fashion. We don't implement any registration or co-registration process in the whole process.

2.1. Spectral domain data re-assembly (SDDR) for multispectral sPA imaging

In multispectral PA imaging, each pixel in a PA image collected at a specific wavelength has a unique spectral intensity while maintaining similar spatial information. This allows the use of multiple images for the same region. To format these regions, we designed a novel SDDR module as the preprocessing backbone for our SPADE. The noisy sPA image data is first processed in the SDDR module before being fed into the denoising model. The design groups pixels from the same position to ensure they share spatial information but differ in spectral information, which helps generate pairs of downsampled images in the following processing steps and enhances residual learning. The goal is to group pixels with different spectra horizontally, and then stack these groups hierarchically, preserving spectral information without compromising spatial integrity.

The SDDR pipeline is illustrated in Fig. 1(b). In each sPA image, there are λ spectral frames with dimensions $W \times$ H. The SDDR module consists of three steps: First, each axial line of pixels is split within each wavelength. Next, these axial lines are recursively concatenated horizontally to form a new 2D image with a size of Hλ. This process is repeated for every axial line across the lateral axis, resulting in grouped axial-spectral plane pseudo-images. Finally, all these images are stacked laterally to create a new 2D image y with dimensions H $\times$ λW.

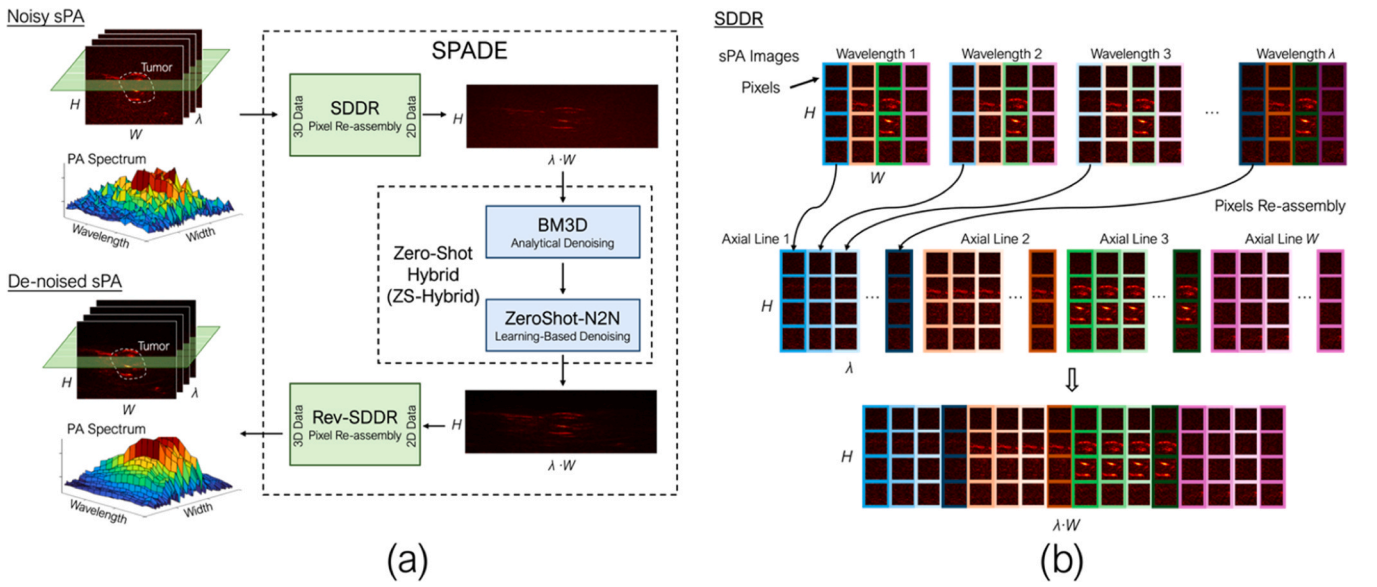


Fig. 1. (a) Spectroscopic Photoacoustic Denoising (SPADE) Framework. Data processing pipeline of the proposed SPADE method: The noisy spectroscopic photoacoustic (sPA) image is first reassembled using Spectral Domain Data Re-assembly (SDDR) before being denoised as a 2D image using Zero-Shot Hybrid (ZS-Hybrid). After denoising, the reverse SDDR is applied to restore the original dimensions of the sPA image. (b) Illustration of sPA image pixel reassembly during the SDDR process. Box color denotes the pixel wavelength, and brightness indicates the axial location.

2.2. Zero shot-denoising (ZED) for sPA denoising

Our proposed denoising method combines both analytical and data-free learning-based approaches. The whole denoising pipeline is illustrated in Fig. 2. In the analytical component, we adapted the concept of the BM3D algorithm. Vanilla BM3D has become a benchmark in the medical imaging field due to its reasonable computational cost and consistent performance. It works by constructing a 3D array by aligning similar blocks extracted from noisy blocks within an image and then reintegrating these processed patches to reduce noise. The method leverages the inherent redundancies in images, following a two-step process: first, it applies hard thresholding during collaborative

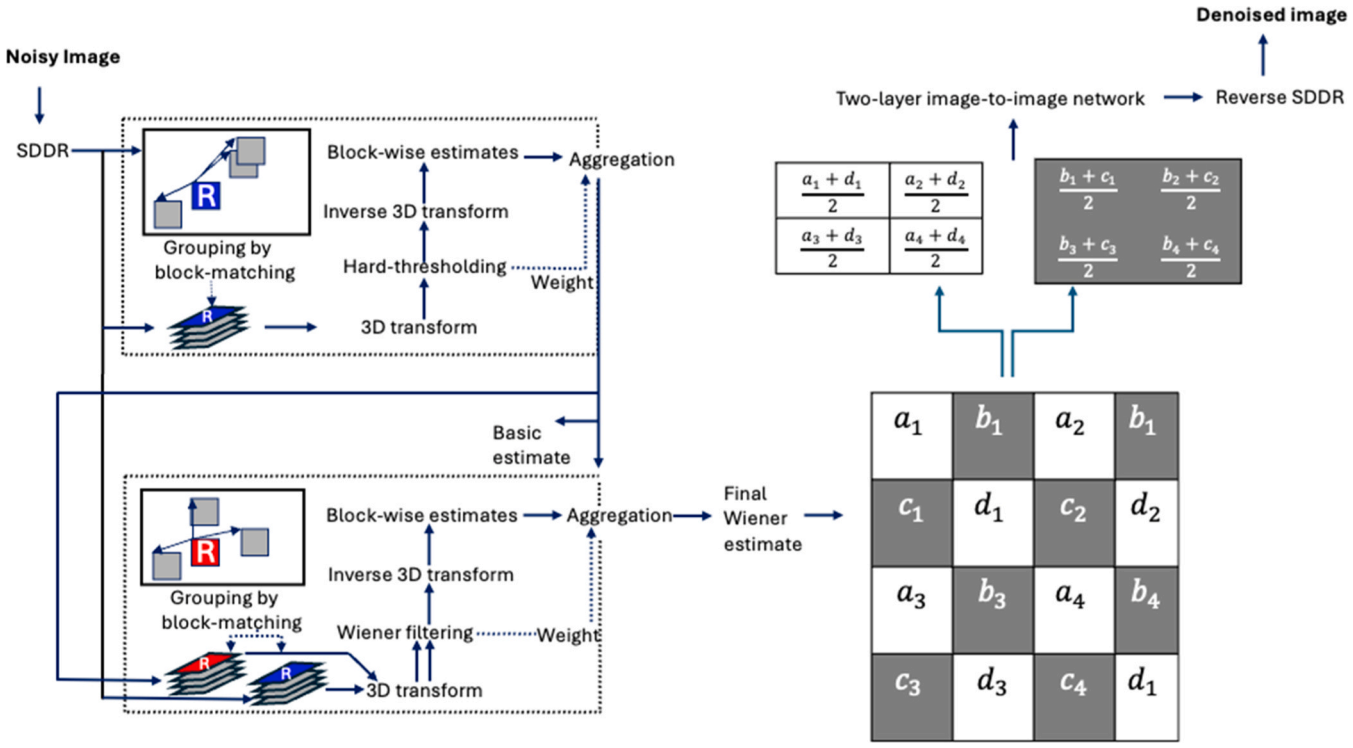


Fig. 2. The denoising process starts with feeding the raw noisy image into SDDR, followed by the analytical component of vanilla BM3D: in the first step, the image is grouped by block matching, then a 3D transform is applied for hard thresholding, followed by an inverse 3D transform and aggregation of block-wise estimates; this is then repeated in a similar second step with further aggregation for the final Wiener estimate. After the analytical component, the image is downsampled by two fixed filters into two downsampled images, which are then processed through a two-layer image-to-image network for further denoising. Finally, the denoised image is recovered using reverse SDDR.

filtering to obtain a basic estimate (hard parameters), then applies a Wiener filter on the original noisy image and the basic estimate to generate the final denoised image [34] (as detailed illustrated in the left of Fig. 2). In our method, we extend vanilla BM3D to spectral-BM3D, which accommodates the multiwavelength PA images processed by the SDDR module. Using vanilla BM3D with the SDDR module directly would result in high computational costs and disrupt spectral information, while spectral-BM3D aligns similar blocks in each grouped image, reducing processing time and preserving spectral information.

For the data-free learning-based component, we adapted ZS-N2N, a zero-shot image denoising algorithm that requires no training examples or prior knowledge of the noise model or level. ZS-N2N builds on Noise2Noise (NB2NB) [32] and Neighbour2Neighbour [33], where Noise2Noise trains on pairs of noisy images, and NB2NB generates such pairs from a single noisy image. ZS-N2N denoises by downsampling a single image into two pairs of images, then training a neural network to map one downsampled image to the other [31]. The downsampled images are generated by applying two convolution masks to the noisy image, as shown in Eq. 1.

$$D_1(y) = yk_1, D_2(y) = yk_2 \tag{1}$$

where $k_1 = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix}$, $k_2 = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$. D_1 and D_2 are downsampled original image y, using fixed filter k_1, and k_2, respectively (This process can be found in the right of Fig. 2). The neural network is trained by minimizing two types of losses, residual loss ($\mathscr{L}_{\text{res.}}(\theta)$) and consistency loss ($\mathscr{L}_{\text{con.}}(\theta)$), using stochastic gradient descent as shown in Eq. 2. θ represents the parameters of the network.

$$\mathscr{L}(\theta) = \mathscr{L}_{\text{res.}}(\theta) + \mathscr{L}_{\text{con.}}(\theta) \tag{2}$$

The residual loss and consistency loss are calculated using Eq. 3 and Eq. 4, where f_θ is the network. ZS-N2N focuses on residual learning, where the network learns only the noise rather than recovering the clean image, simplifying training and better preserving spectral information [35–37]. The consistency loss we are using can enforce consistency by comparing two processes. It ensures that denoising the image first and then downsampling it produces results similar to downsampling first and then denoising it [31]. We have not tried other loss functions because, typically, the chosen loss function is designed to coordinate with the network architecture. We plan to explore this further in our future work.

$$L_{res.}(\theta) = \frac{1}{2}\Big(\|D_1(y) - f_\theta(D_1(y)) - D_2(y)\|_2^2 + \|D_2(y) - f_\theta(D_2(y)) - D_1(y)\|_2^2\Big) \tag{3}$$

$$L_{cons.}(\theta) = \frac{1}{2}\Big(\|D_1(y) - f_\theta(D_1(y)) - D_1(y - f_\theta(y))\|_2^2 + \|D_2(y) - f_\theta(D_2(y)) - D_2(y - f_\theta(y))\|_2^2\Big) \tag{4}$$

The network (f_θ) architecture is a simple two-layer image-to-image network, consisting of two convolutional operators with a 3 × 3 kernel size, followed by a 1 × 1 convolution. While deep learning models often require kernel size adjustments based on input image size, we avoid additional hyperparameter tuning by using this simple structure, maintaining stable performance. The small kernel size also allows future modifications to input image size. In summary, our proposed denoising method involves first processing noisy multi-spectral PA images through the SDDR module to generate a single noisy image y. This image is then fed into the SPADE framework for two-step denoising (BM3D and ZS-N2N). Finally, the denoised image y' is fed into the reverse SDDR module to recover the denoised multi-spectral PA imaging, as shown in Fig. 2.

2.3. Implementation detail and evaluation metric

In the proposed SPADE framework, all modules are integrated and packaged together for an end-to-end implementation. Data processing is conducted on a server equipped with two Intel Xeon Silver 4314 CPUs (2.40 GHz), two NVIDIA A100 80 G GPUs, and 256 GB of RAM, with operating system of Ubuntu 20.04.6 LTS (GNU/Linux 5.4.0–128-generic x86_64) and software with Python 3.8.19 and PyTorch 1.13.0 [38]. To evaluate the performance of our method, we introduced several quantitative metrics. One of the key measures is the signal-to-noise ratio (SNR) [39], which is used to assess the denoising capability. SNR is defined by Eq. 5:

$$SNR = 20log_{10}\frac{|P_{peak}|}{\sigma_{background}} \tag{5}$$

where P_{peak} is the peak PA signal amplitude and $\sigma_{\text{background}}$ is the standard deviation of background noise. The noise region was selected at the same depth as the peak signal, positioned laterally more than 1 *mm* away from each peak.

We also used peak signal-to-noise ratio (PSNR) [35] and structural similarity (SSIM) index [21,40] to quantify the similarity between the denoised image and the reference image. PSNR is calculated based on the image (I) and reference images (I_{ref}), as defined in Eq. 6. For image (I), it refers as the noisy image, while reference images (I_{ref}) refer as ground truth image. Both images (I and I_{ref}) have $m \times n$ dimension where m, n is the pixel coordinates.

$$PSNR = 10log_{10}\frac{(P_{peak})^2}{\sqrt{\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(I_{ref} - I)^2}} \tag{6}$$

SSIM is another metric to evaluate imaging quality highlighting structural similarity. SSIM is defined in Eq. 7. μ_{ref} and μ_I are the mean of the reference image and denoised image, respectively. And σ_{ref} and σ_I are the standard deviation of the reference image and denoised. $\sigma_{ref,I}$ is the covariance between two images, and c_1 and c_2 are constant values to avoid instability when the sum square of means or variances are close to zero.

$$SSIM = \frac{\left(2\mu_{ref}\mu_I + c_1\right)\left(2\sigma_{ref,I} + c_2\right)}{\sqrt{\left(\mu_{ref}^2 + \mu_I^2 + c_1\right)\left(\sigma_{ref}^2 + \sigma_I^2 + c_2\right)}} \tag{7}$$

Additionally, we quantitatively evaluated spectrum information preservation by calculating Pearson correlation coefficients between spectra to assess changes in pixel spectrum. It is the most commonly used linear correlation coefficient. For column X_a in matrix X and column Y_b in matrix Y, having means $\overline{X_a} = \sum_{i=1}^{n}\frac{X_{a,i}}{n}$, and $\overline{Y_b} = \sum_{j=1}^{n}\frac{Y_{b,j}}{n}$, then Pearson linear correlation $rho(a,b)$ is defined as:

$$rho(a,b) = \frac{\sum_{i=1}^{n}\left(X_{a,i} - \overline{X_a}\right)\left(Y_{b,i} - \overline{Y_b}\right)}{\left\{\sum_{i=1}^{n}\left(X_{a,i} - \overline{X_a}\right)^2\sum_{j=1}^{n}\left(Y_{b,j} - \overline{Y_b}\right)^2\right\}^{1/2}} \tag{8}$$

3. Experimental setup

3.1. Simulation implementation

We first conducted a simulation study to statistically understand the performance of our proposed method. This study has two main objectives. First, to analyze the image SNR improvement performance of the proposed method across different noise levels while preserving spectral information. The second is to assess the coefficient of the spectral similarity after denoising to validate that spectral information is preserved.

The simulation runs on a Matlab (Matlab 2023a, MathWorks, USA) platform with K-Wave [41] and ValoMC [42] for the time-domain PA simulation [43]. A 2D simulation environment with a space of 41.6 × 41.6 mm^2 was defined. 128 transducer elements were placed apart from each other at a distance of 0.195 *mm*. We used the default omni-directional setting for the sensor and the sensor is a broadband point detector. Light source propagation simulation (optical simulation) is based on the finite element method (FEM) and the Monte Carlo method. A single Gaussian light source aligned with the transducer has been set up with a standard deviation of 0.1. One point target (diameter: 650 μm) was placed at a depth of 3.9 *mm* with respect to the transducer plane.

In order to obtain multiple frame images, we independently run the light source propagation simulation several times. For multi-spectrum image simulation, we run the Light source propagation simulation several times with different absorption coefficients to mimic different spectra. The simulated PA signal first underwent beamforming using Delay-and-Sum (DAS) algorithms. Then, the different levels of noise are added at each simulation to simulate acoustic noise. The simulated single-frame noisy PA image was processed with vanilla BM3D and our proposed methods. The performance matrices with different methods were compared.

3.2. Phantom study

We then conducted a wire phantom study to assess the performance of the proposed method and verify its consistency with the simulation study. A phantom study used Nylon fish wire as the imaging target, which was affixed onto a 3D-printed wire holder. The phantom comprised four layers of vertically aligned wires and three-column layouts, facilitating the assessment of PA imaging quality along both axial and lateral axes. Vertical wires were positioned at depths of 8, 13, 18, and 23 *mm* from the transducer surface, while three horizontal wires were situated at depths of 13 *mm* and 18 *mm*, all with a diameter of 0.2 *mm*.

For imaging, the Verasonics Vantage system (Vantage 128, Verasonics, USA) was employed for data acquisition. A laser system (Phocus MOBILE, OPOTEK, USA) capable of emitting wavelength-tunable laser light (690–950 *nm*) at a repetition rate of 20 Hz with a 5 *ns* pulse duration served as the light source. This enables the acquisition of multiwavelength PA images within a short period, allowing SPADE to denoise without averaging. A linear transducer (L12–5, Philips, Netherlands) facilitated both US and PA imaging. In the phantom study, the wires were immersed in a water bath for acoustic coupling, 700–850 *nm* wavelength range was utilized for PA excitation, with US images acquired at the same location to validate the target structure. The acquired PA signal was beamformed using DAS algorithm. The acquired noisy PA image was processed with vanilla BM3D, averaging filter, and our proposed methods. The performance matrices with different methods were compared.

3.3. Ex vivo evaluation with swine cardiac tissue

We also conducted an ex vivo study to analyze the performance of the proposed method using ex vivo tissue samples. The proposed SPADE framework was evaluated in clinically relevant scenarios, specifically for denoising sPA images in cardiac ablation monitoring. Intraoperative sPA imaging for cardiac scans presents challenges due to the dynamic cardiac motion, which limits the use of averaging filters. SPADE's single-frame denoising method is well-suited to address these challenges. The same experiment devices in the phantom study were used in this ex vivo study. For the imaging target, swine cardiac tissue was selected as the experimental substrate, with a total of 35 samples prepared and subjected to scanning procedures. The tissue samples were ablated with various durations to create ablation-induced necrosis lesions. The spectral information of the tissue was analyzed for functional guidance to identify ablated tissue regions under various denoising methods. The tissue samples were immersed in a saline bath for acoustic coupling, and both US and PA images were acquired at the same location using the same imaging setup. PA image was acquired with a wavelength range between 700 and 850 *nm* covering the spectral signature of normal tissue and ablated necrosis [44]. The performance of the methods in denoising and spectrum preservation was evaluated. The image denoised with 64 frames averaging filter was used as a ground truth, and single-frame noisy images were processed with vanilla BM3D and our proposed SPADE method.

3.4. In vivo demonstration

We also conducted an in vivo study to verify the performance of the proposed method on in vivo animal model. The objective of this study is verifying the performance of the proposed method in an in vivo environment. Respiratory motion and myoclonus (involuntary muscle contractions) during anesthesia can reduce the effectiveness of averaging filters. Additionally, using fewer frames for denoising allows for an increased imaging frame rate, which is crucial for intraoperative guidance. The in vivo investigation was conducted utilizing a mouse tumor model employing sublines of a human prostate cancer cell line, originally derived from an advanced androgen-independent bone metastasis. Male non-obese diabetic (NOD)/severe combined immunodeficient (SCID) mice aged six to eight weeks were subcutaneously implanted near the posterior flanks. A total number of 11 mice were utilized in this study. A gold nanoparticle (AuNR) contrast agent, designed to attach to the tumor and enhance tumor detection in PA imaging, was synthesized and administered to the animals. The targeted contrast exhibited a PA spectrum peak at 780 *nm*. This study was approved by the Institutional Animal Care and Use Committee (IACUC) at Worcester Polytechnic Institute (Protocol number: 21–127).

Ultrasound was utilized for tumor localization at baseline, utilizing the same imaging setup as in the phantom study. The mice were anesthetized with 3 % isoflurane in oxygen via a nose cone and positioned dorsally on a fixation stage. Ultrasound gel was applied to ensure proper coupling between the probe and the tumors, as illustrated in Fig. 3. Ultrasound B-mode images aided in transducer positioning and optimizing the interrogated tumor cross-section in PA imaging. After establishing the ultrasound slice position, consecutive PA images within the same image field were acquired, with a wavelength of 780 *nm* selected to highlight the contrast agent's peak spectrum and visualize the targeted tumor.

An identical imaging setup was used to scan the animals. The wavelength between 700 and 850 *nm* was used with an intermediate step of 10 *nm*. The acquired image was processed with vanilla BM3D, averaging filter, and our proposed methods. In addition, the processed sPA images were spectrally unmixed to reveal the distribution of the contrast agent in tissue and around the tumor. Linear spectral unmixing [48] was used to highlight the contrast agent distribution around tumor region [6,9]. Animals were imaged before, and 24 hours post injection to visualize the changes in contrast agent distribution.

4. Results

4.1. Simulation implementation

The proposed SPADE denoising algorithm was evaluated through simulations using a grid of point targets with varying levels of random noise. Its performance was compared to the vanilla BM3D method, with results shown in Fig. 4(a). The raw input image reveals significant noise around the point targets. After processing with BM3D, noise is reduced, but residuals remain, particularly in deeper regions. In contrast, the SPADE-processed image reveals more uniform noise suppression, with clearer, more consistent point targets at all depths.

Quantitative evaluations used four metrics: SNR, spectrum similarity

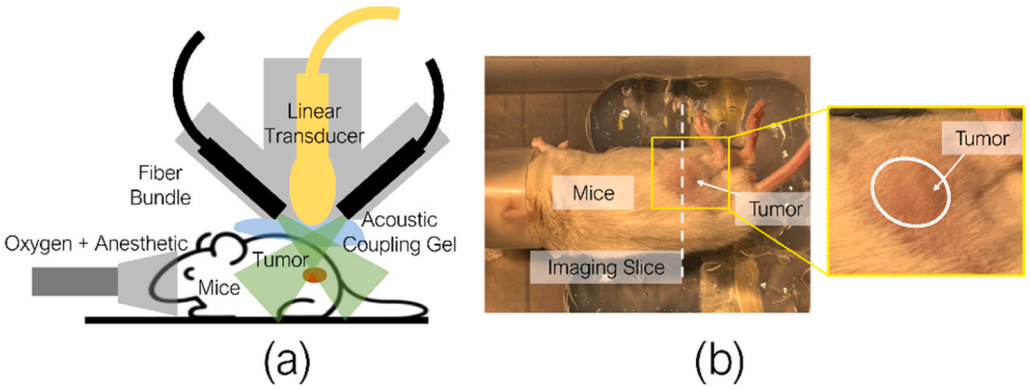


Fig. 3. In vivo tumor scanning: (a) Animal setup sketch, (b) Photograph of prepared animal with tumor region zoomed in.

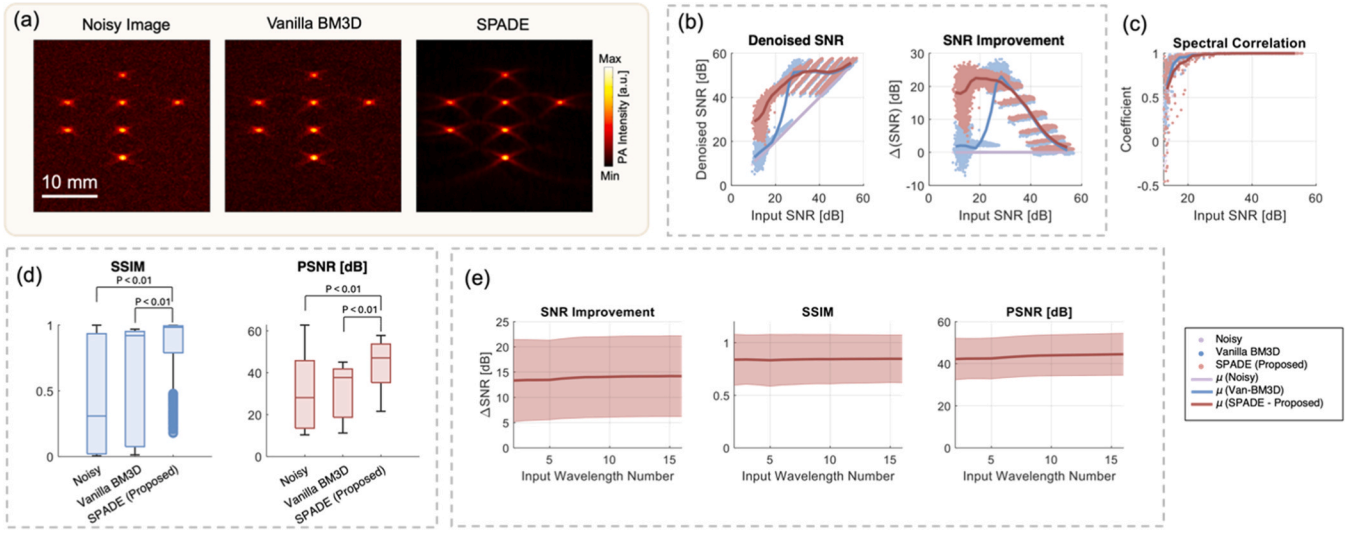


Fig. 4. Simulated spectroscopic photoacoustic (sPA) image at 720 *mm*. (a) PA images (left to right): noisy single-frame PA image, vanilla BM3D denoised image, and image denoised by the proposed SPADE method. Quantitative evaluation of the SPADE method compared to vanilla BM3D and the noisy single-frame PA image: (b) Output signal-to-noise ratio (SNR) at different noise levels and SNR improvement. (c) Spectral correlation coefficient of vanilla BM3D and SPADE at various noise levels. (d) Structural similarity (SSIM) index and Peak signal-to-noise ratio (PSNR), computed with respect to the ground truth simulation image without noise, along with paired wise P value from T test. (e) Denoising performance across varying numbers of input wavelengths.

measured by correlation coefficient, SSIM, and PSNR. SPADE consistently achieved higher denoised SNR across a wide range of input SNRs as shown in Fig. 4(b), excelling when the input SNR was below 25 *dB*. BM3D also improved SNR but was less effective in low-SNR conditions, only matching SPADE at higher input SNRs. The Δ SNR results suggest that SPADE reaches superior performance in high-noise environments, with an average SNR improvement exceeding 15 *dB* for input SNRs below 25 *dB*, far surpassing BM3D. Both methods reached a performance plateau as input SNR increased, with fewer differences at higher SNRs.

Spectrum similarity, measured by the correlation coefficient between the denoised and ground truth spectra, was also evaluated in Fig. 4(c). SPADE maintained a higher correlation across noise levels, especially at lower input SNRs, indicating better spectral preservation. Both methods performed well ($R > 0.8$) at higher SNRs, but SPADE demonstrated greater stability, with higher average coefficients and lower variability compared to BM3D. We also provide the root mean squared error (RMSE) between each method compared with reference spectral: RMSE between noisy raw data and reference spectral is 0.1718 $\pm$ 0.2548, AVG64 and reference spectral is 0.0957 $\pm$ 0.1741, vanilla BM3D and reference spectral is 0.2159 $\pm$ 0.2177, SPADE and reference spectral is 0.1159 $\pm$ 0.0971. Finally, the statistical analysis of SSIM and PSNR in Fig. 4(d) further emphasizes SPADE's advantages. The box plots indicate that SPADE achieved higher scores on both metrics with less variability, demonstrating its consistent performance across different noise levels.

In addition to comparing the denoising performance of the proposed SPADE method, we evaluated its effectiveness across different input wavelength numbers to validate its generality. PA images with 2–15 wavelengths were derived from the original 16-wavelength simulated sPA data and input into the SPADE method. The same evaluation metrics, SNR improvement, SSIM, and PSNR, were used to assess effectiveness at various input dimensions (see Fig. 4(e)). The quantitative data demonstrate stability across the number of input wavelengths. Average SNR improvement increases from 13.32 *dB* at 2 wavelengths to 14.19 *dB* at 16 wavelengths. SSIM improves from 0.838 to 0.846, and PSNR rises from 42.21 *dB* to 44.49 *dB*. Moreover, the paired *t*-test results shown in Fig. 4(d) indicate statistically significant differences between SPADE and Vanilla BM3D, as well as between SPADE and the noisy image, in terms of both SSIM and PSNR.

4.2. Phantom study

A phantom study using a multi-point target, similar to the simulation, was conducted to evaluate the real-world effectiveness of the proposed SPADE denoising method. The denoised images are presented

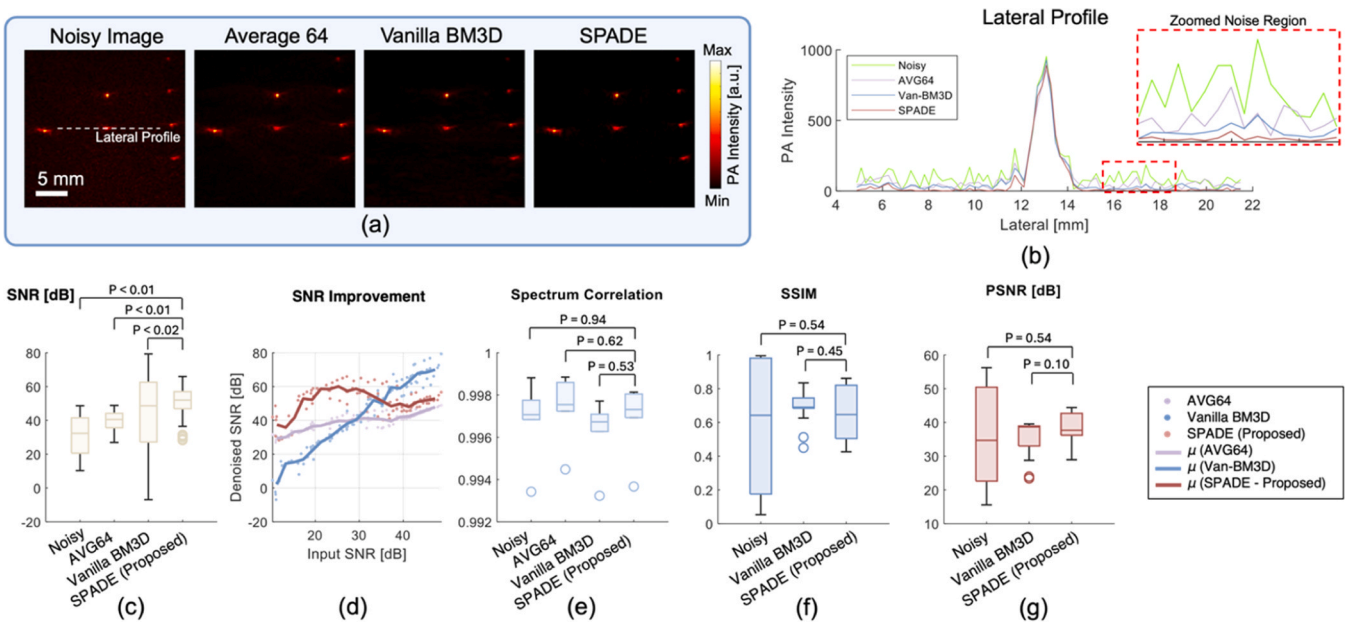


Fig. 5. (a) Photoacoustic (PA) imaging denoising results of the point phantom at 810 *nm*. From left to right: Raw noisy data, 64-frame averaging filter, vanilla BM3D, and the proposed SPADE algorithm. (b) Signal lateral profile across one of the targets with zoomed noise region. The white dash line in (a) indicates where we plot the lateral profile. (c) SNR statistics of point targets for each method, along with paired wise P value from T test. (d) Output SNR from the three denoising methods at various noise levels. (e) Spectral correlation coefficient of each method compared to the ground truth spectrum, along with paired wise P value from T test. (f) Structural similarity (SSIM) index and (g) Peak signal-to-noise ratio (PSNR) computed relative to the averaging filter image, along with paired wise P value from T test.

in Fig. 5, where the noisy image was processed using a 64-frame averaging filter, vanilla BM3D, and the proposed SPADE method.

Fig. 5(a) illustrates the visual results of the denoising process. The noisy image exhibits reduced visibility of the point targets, particularly at lower depths. After applying the 64-frame averaging filter, some noise suppression is achieved, but the reduction is uneven across depths. The vanilla BM3D method further improves noise reduction, though some residual noise remains, especially at the deepest points. In contrast, the SPADE method suppresses the noise more uniformly, with high contrast point targets of preserved shape across the entire image. The lateral signal profile was displayed in Fig. 5(b) showing the same trend of noise suppression across different methods.

Fig. 5(c) and (d) quantitatively evaluate the denoising performance. The box plot shows the distribution of denoised SNR values for each method, with SPADE consistently yielding higher median SNR values and a tighter interquartile range, indicating more reliable performance across varying noise levels. The central plot examines denoised SNR as a function of input SNR, where SPADE demonstrates substantial improvement, particularly at lower input SNRs where other methods struggle. SPADE's ability to maintain a higher SNR across all noise levels underscores its effectiveness in enhancing image quality under challenging conditions. Fig. 5(e) assesses spectral similarity, with all methods showing a strong correlation ($R > 0.99$). We also provide RMSE between each method compared with reference spectral: RMSE between noisy raw data and reference spectral is 0.0325 ± 0.0071, AVG64 and reference spectral is 0.02717 ± 0.0065, vanilla BM3D and reference spectral is 0.0337 ± 0.0058, SPADE and reference spectral is 0.0309 ± 0.0071.

Fig. 5(f) and (g) further support these findings through additional statistical analysis. The box plots compare SSIM across methods, where the proposed SPADE method reaches a lower mean SSIM compared to the averaging filter and vanilla BM3D. Its standard deviation is larger than BM3D's but smaller than the averaging filter's. A higher PSNR is observed with the proposed method compared to vanilla BM3D, with a comparable standard deviation, indicating consistent performance relative to the averaging filter. Moreover, the paired *t*-test results shown in Fig. 5(c, e, f, g) indicate statistically significant differences only in SNR between SPADE and Vanilla BM3D, between SPADE and the noisy image, and between SPADE and AVG64.

4.3. *Ex vivo evaluation with swine cardiac tissue*

The proposed methods were then evaluated using ex vivo swine cardiac tissue. Sixteen wavelengths of PA scanning were performed on 35 ablated cardiac tissue samples. The single-frame noisy image was processed using vanilla BM3D and the proposed SPADE method, with the 64-frame averaging filter result used as the ground truth for comparison. Fig. 6(a) shows the denoised image results from each method. In the single-frame noisy image, the tissue boundary is barely visible, with its location confirmed via ultrasound. All three denoising methods successfully suppressed noise, highlighting both the tissue boundary and the ablated region. The SPADE method revealed more intensity from deeper regions compared to vanilla BM3D. Linear spectral unmixing was performed on the PA images using spectra from ablated and non-ablated tissue [44], displaying the ablation-induced necrotic boundary. The noisy data failed to identify tissue boundaries due to electrical noise above the tissue surface, and the boundary of the ablated tissue was inaccurate. In contrast, all three denoising methods provided a clear boundary, matching the gross pathology results acquired post-scanning. The tissue and lesion shapes obtained by the SPADE method align better with vanilla BM3D.

Statistical analysis of all ex vivo scans was presented. SNR values (Fig. 6(b)) directly reflect the denoising effectiveness of each method. The proposed SPADE method achieved the output (43.05 *dB*median) from the original 28.01 *dB* noisy image, close to the 64-frame averaging filter (43.22 *dB* median). Both PSNR (Fig. 6(c)) and SSIM (Fig. 6(d)) values suggest that all denoising methods improved image structural similarity, with the SPADE method yielding a higher PSNR (36.173 *dB*) compared to vanilla BM3D (32.632 *dB*) and a comparable SSIM (SPADE: 0.697, vanilla BM3D: 0.721). The spectral correlation from the SPADE denoised image was higher, indicating strong agreement with the

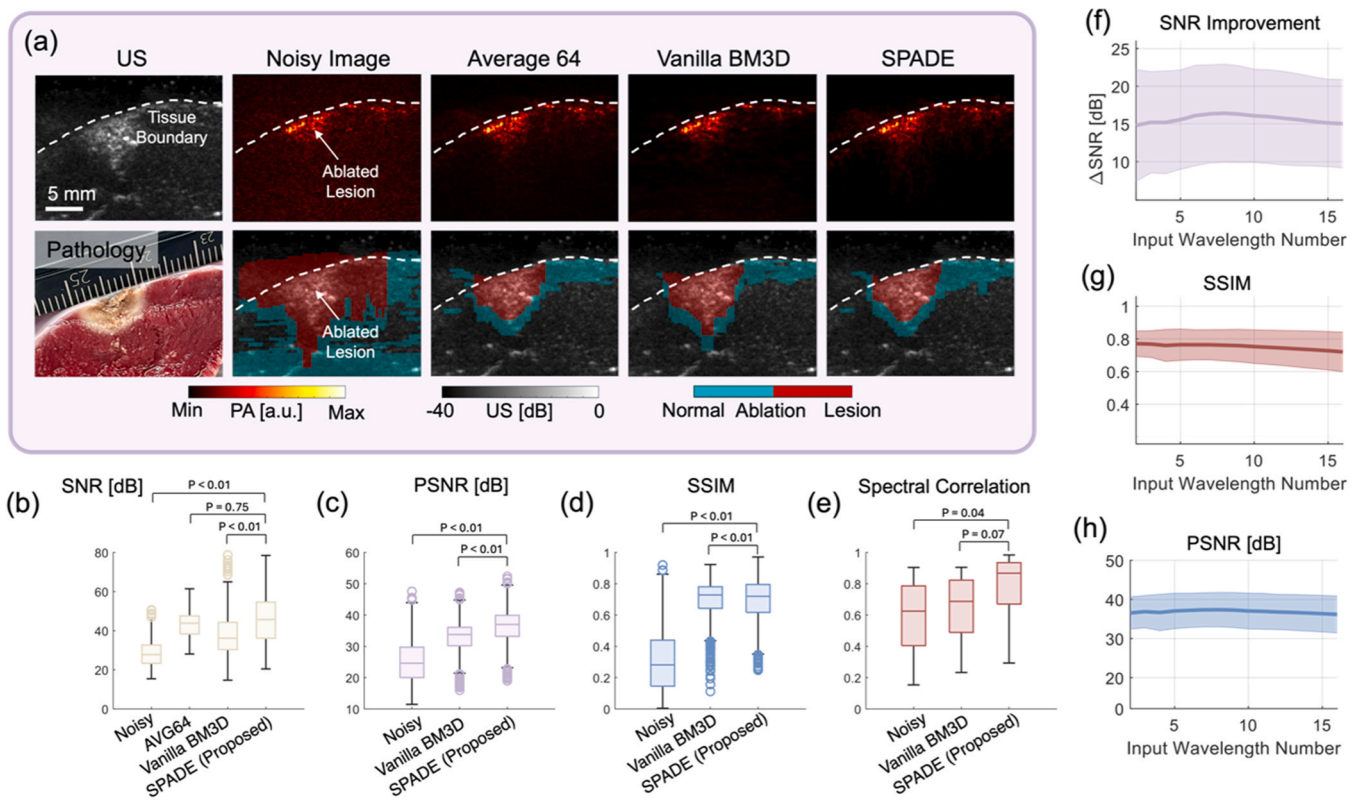


Fig. 6. Ex vivo evaluation of spectroscopic PA (sPA) analysis for cardiac ablation mapping. (a) PA image at 780 *nm* and PA-mapped lesion boundary using different denoising methods, alongside the corresponding US image and post-scan gross-pathology highlighting the lesion. Quantitative analysis statistics: (b) Signal-to-noise ratio (SNR), along with paired wise P value from T test. (c) Peak signal-to-noise ratio (PSNR), along with paired wise P value from T test. (d) Structural similarity (SSIM) index, and (e) spectral correlation coefficient relative to the averaging filter, along with paired wise P value from T test. Denoising performance: SNR improvement (f), SSIM (g), PSNR (h) across varying numbers of input wavelengths based on ex vivo evaluation.

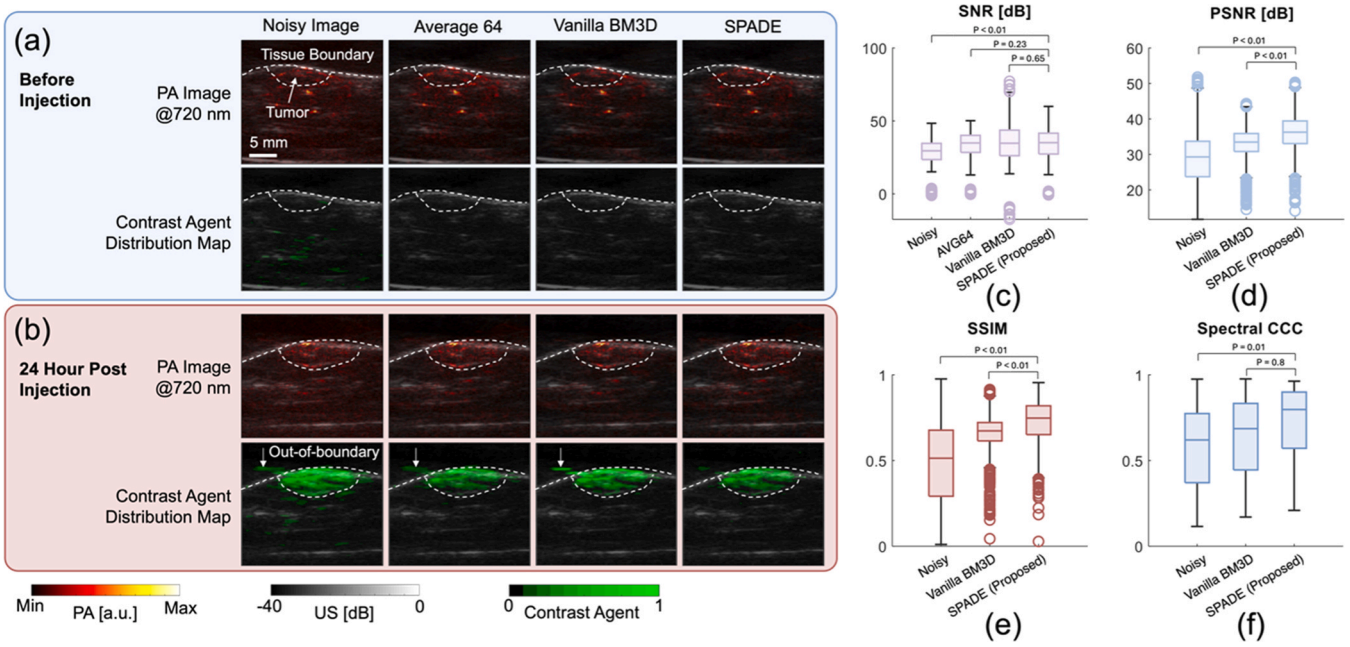


Fig. 7. In vivo mice tumor scanning: (a) before and (b) 24 hours after contrast agent injection. From left to right: single frames, PA image averaged with 64 frames (Reference), vanilla-BM3D, and SPADE (proposed), overlaid with the US image. The same color range is applied within each method before and after injection. Quantitative analysis statistics: (c) Signal-to-noise ratio (SNR), along with paired wise P value from T test. (d) Peak signal-to-noise ratio (PSNR), along with paired wise P value from T test. (e) Structural similarity (SSIM) index, and (f) spectral correlation coefficient relative to the averaging filter, along with paired wise P value from T test.

ground truth spectra from the averaging filter image, as shown in Fig. 6 (e). We also provide RMSE between each method compared with AVG64: RMSE between noisy raw data and AVG64 is 0.7627 ± 0.1763, vanilla BM3D and AVG64 is 0.7164 ± 0.1705, SPADE and AVG64 is 0.6070 ± 0.2166.

We evaluated the effectiveness of SPADE denoising across different input wavelengths based on ex vivo results. Wavelengths ranging from 2–15 were derived from the original images and input into the SPADE method. In Fig. 6(f-h), the same evaluation metrics were used to assess effectiveness at various input dimensions. The results indicate that average SNR improvement varies between 14.80 *dB* and 16.42 *dB* across different input wavelengths. SSIM remains relatively stable with a slight decline from 0.771 to 0.721, while PSNR remains steady between 36.17 *dB* and 37.37 *dB*. Based on the paired *t*-test results shown in Fig. 6(b–e), with the exception of the SNR between SPADE and AVG64 and the spectral correlation between SPADE and Vanilla BM3D, all other metrics (SNR, PSNR, and SSIM) showed statistically significant differences between SPADE and the other methods (Noisy, AVG64, and Vanilla BM3D).

4.4. In vivo demonstration

In vivo demonstration was conducted using a contrast agent-injected mouse model to enhance tumor detection. The same imaging and spectral analysis pipeline from the previous ex vivo study was applied. The spectrum of the AuNR agent acquired from the earlier study was used as the library spectra, and the scanned sPA images were unmixed with AuNR, oxygenated hemoglobin (HbO2), and deoxygenated hemoglobin (HbR) spectra. The denoised PA images and contrast agent distribution maps, both before and 24 hours after injection, are shown in Fig. 7. The contrast agent distribution was quantified by calculating the ratio of its decomposed pixel intensity to the total hemoglobin intensity (HbR + HbO2), accounting for local illumination variation.

As in the in vivo results, all three denoising methods effectively suppressed noise and enhanced tissue boundaries that were otherwise invisible in the noisy image. Before injection, no contrast agent signals were detected across all denoising cases, as depicted in Fig. 7(a). In the noisy image, some false contrast agent intensity values were observed in the deeper regions due to mislabeled noisy spectra. Moreover, there are some noisy pixels illuminated outside the boundary outliers as well. 24 hours after injection, a strong contrast agent signal was observed around the tumor region as presented in Fig. 7(b). In the noisy image, the contrast agent was mainly distributed in the tumor region, with some illuminated pixels in the deeper regions. In the denoised results, the contrast agent intensity value was predominantly retained within the tumor region. The distribution of AuNR in the proposed method closely matched the averaging filter results. Notably, the SPADE method was the only result without noisy pixels illuminated outside the tissue boundary.

Statistical analysis is presented in Fig. 7(c-f). The median SNR of the SPADE method is 35.09 *dB*, comparable to the averaging filter at 34.86 *dB*, and higher than both vanilla BM3D at 34.71 *dB* and the original noisy image at 29.50 *dB*. The PSNR between the SPADE method and the averaging filter is 36.26 *dB*, an improvement over the noisy image (29.28 *dB*) and vanilla BM3D (33.47 *dB*). A similar trend is observed in SSIM, with SPADE achieving 0.748 compared to 0.514 from the noisy image and 0.673 from BM3D. The spectral correlation of the denoised image with the ground truth spectrum from the averaging filter was also evaluated. The SPADE method achieved a median coefficient of 0.798, higher than 0.620 before denoising, and 0.686 with vanilla BM3D. We also provide RMSE between each method compared with AVG64: RMSE between noisy raw data and AVG64 is 0.7460 ± 0.2051, vanilla BM3D and AVG64 is 0.7121 ± 0.1973, SPADE and AVG64 is 0.6167 ± 0.2192. Similarly, the paired *t*-test results shown in Fig. 7(c–f) clearly indicate whether there are statistically significant differences between SPADE and the others (Noisy, AVG64, and Vanilla BM3D).

4.5. Overall performance

As mentioned previously, this study focuses on denoising multi-wavelength PA images and ensuring spectral information preservation, here we include a systematic presentation of SAM and RMSE metrics to better to evaluate spectral fidelity. As shown in Table 1, it summarizes values across simulation, phantom, ex vivo, and in vivo studies, clearly showing that SPADE consistently yields lower RMSE compared to alternative methods. Our experimental findings consistently favor SPADE in terms of spectral fidelity.

Moreover, as shown in Fig. 8, from (a) to (b), we provide simulation, point phantom SAM since SAM requires a consistent reference spectrum available only in simulation and phantom data (Smaller score: Higher spectral similarity. Larger score: Lower spectral similarity). This also clearly showing that SPADE consistently yields lower SAM compared to alternative methods, has proven the consistent superiority of the proposed SPADE method.

We also provide statistically analysis; we added T test in the present study data to show the statistical significance in performance. We observed that for simulation data, we found SNR, PSNR, and SSIM are statistically different ($P < 0.05$) between SPADE and Vanilla BM3D, supporting the claim that our method is significantly better than Vanilla BM3D. While the spectral correlation coefficients between SPADE and Vanilla BM3D, are not statistically significant, both exhibit high correlation, demonstrating spectral integrity. In phantom data, SNR results are statistically different between SPADE and Vanilla BM3D, SPADE is significantly better than Vanilla BM3D ($P < 0.05$). While PSNR, SSIM, and spectral correlation coefficients show no statistically significant differences, we believe this is due to the low noise in the raw input image (average SNR of 32.33 *dB*). This results in minimal changes before and after denoising, yielding high quantification values without statistical significance. For ex vivo data, we found SNR, PSNR, SSIM are statistically different between SPADE and Vanilla BM3D, SPADE is significantly better than baseline ($P < 0.05$), while spectral correlation coefficient shows no statistically significant differences. Finally in the in vivo data, we found PSNR and SSIM are statistically different SPADE and Vanilla BM3D, SPADE is significantly better than baseline ($P < 0.05$), while SNR and spectral correlation coefficient are not statistically significant with no evidence of a difference. In in vivo studies, the raw input data already exhibits a relatively high SNR. Unlike other evaluation metrics that require a comparison between denoised and ground truth data, SNR is calculated solely based on the ratio between the target point and the background. Therefore, no significant difference can be observed.

Moreover, to mitigate the risk of inflated false positives from multiple t-tests across metrics (SNR, PSNR, SSIM, correlation coefficients, SAM, and RMSE), we applied a one-way Analysis of Variance (ANOVA) to compare our SPADE method with BM3D. Our analysis covered all experimental datasets—simulation, phantom, ex vivo, and in vivo. For phantom data, where all six metrics were evaluated, SPADE significantly outperformed BM3D. In the ex vivo dataset, a one-way ANOVA across the five available metrics (excluding SAM, which requires a consistent

Table 1
Summary of RMSE values (mean ± std) across studies comparing various methods with the designated reference. (Noted, in simulation and phantom study, we can obtain reference spectra.).

Study	Reference	Noisy Raw Data	AVG 64	Vanilla BM3D	SPADE
Simulation	Ref	0.1718±	0.0957±	0.2159	**0.1159**
	Spectral	0.2548	0.1741	±0.2177	**± 0.0971**
Phantom	Ref	0.0325±	0.0271±	0.0337±	**0.0309**
	Spectral	0.0071	0.0065	0.0058	**± 0.0071**
Ex vivo	AVG 64	0.7627±	Ref	0.7164	**0.6070**
		0.1763		±0.1705	**± 0.2166**
In vivo	AVG 64	0.7460±	Ref	0.7121±	**0.6167**
		0.2051		0.1973	**± 0.2192**

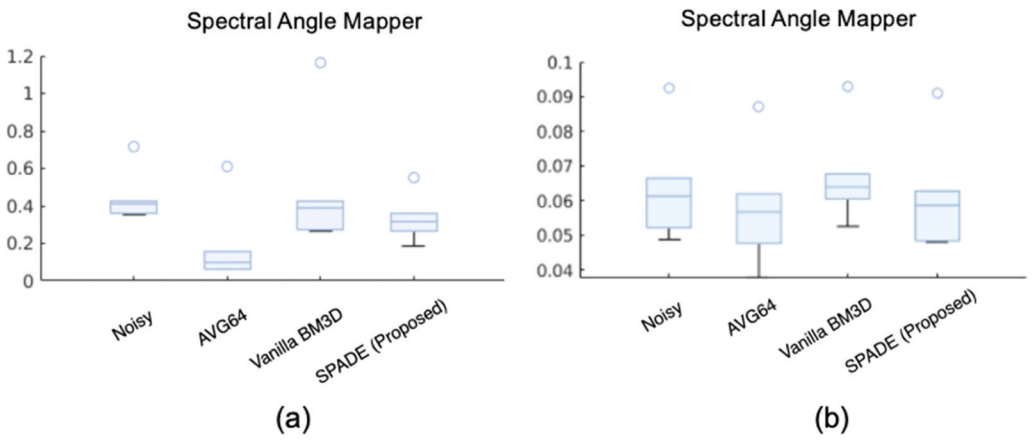


Fig. 8. Spectral angle mapper (SAM). (a) SAM of simulation data. (b) SAM of point phantom data.

reference spectrum available only in simulation and phantom data) demonstrated that SPADE was significantly superior to BM3D. In contrast, for simulation and in vivo datasets, the differences between the methods did not reach statistical significance.

At the same time, we would like to clarify that we do not consider the correlation between multiple metrics to be a major concern in PA imaging. Image quality evaluation metrics inherently share mutual relationships and dependencies by design. SNR, PSNR, and SSIM are widely accepted for assessing PA image quality, particularly in studies introducing denoising methods. Meanwhile, correlation coefficients, SAM, and RMSE serve as effective measures for evaluating spectral information preservation. While we acknowledge the benefits of a conservative evaluation approach, in practical circumstances, conducting t-tests against BM3D for each metric remains a meaningful method to demonstrate the superiority of the SPADE method.

We also acknowledge that while our technical metrics demonstrate statistically significant improvements, the clinical relevance of improvements (e.g., a few dB in SNR or minor differences in SSIM) has not been fully discussed. Future work will explore the correlation between these technical gains (such as an X dB improvement) and potential clinical benefits (for example, a Y% reduction in diagnostic error), along with additional comparisons to established benchmarks in the literature. These aim to enhance both the statistical rigor of our analysis and the practical context of our findings.

4.6. Ablation study

We conducted ablation studies for all our proposed modules and loss functions using simulated sPA images. First, we present the results obtained using individual loss functions. The performance degrades when only a single loss is applied. As shown in Fig. 9(b), the degradation is not severe when only the residual loss is used. However, Fig. 9(c) reveals a significant performance degradation when only the consistency loss is applied.

Next, Fig. 10(b) shows the ablation results obtained by removing the SDDR module. The results indicate a degradation in denoising performance along with a substantial loss of spectral information. In Fig. 10(c), we present the ablation study of the SDDR module combined with BM3D. Without the learning component of ZS-N2N, the noise is visually reduced; however, SPADE consistently outperforms the SDDR + BM3D combination in terms of noise suppression and quantitative metrics. Notably, Fig. 10(c) shows that the spectral information is preserved similarly to SPADE, which underscores the spectral preservation contributed by the SDDR module.

Fig. 10(d) presents the ablation study of the SDDR module with ZS-N2N only. Without the analytical component of BM3D, the noise is only slightly reduced visually, and SPADE again consistently outperforms the SDDR + ZS-N2N combination regarding noise suppression and quantitative metrics. Similarly, Fig. 10 (d) demonstrates that the spectral information is as well preserved as in SPADE, further emphasizing the spectral preservation property of the SDDR module.

Finally, Fig. 10(e) presents the ablation study of the learning method ZS-N2N only. Without the SDDR module and the analytical component of BM3D, the noise is only slightly reduced visually, and SPADE again consistently outperforms the ZS-N2N regarding noise suppression and quantitative metrics by a large margin. Similarly, Fig. 10(e) demonstrates that the spectral information failed to be preserved compared with SPADE, further emphasizing the spectral preservation property of the SDDR module as well.

5. Discussion

The results from the simulation, phantom, in vivo, and ex vivo studies consistently demonstrate the superior performance of the proposed SPADE denoising algorithm in quantitative PA imaging compared to vanilla BM3D and achieving comparable performance to 64 frames averaging filters which generally considered as ground truth. Across all experiments, SPADE was able to achieve more uniform noise suppression, improved SNR, and enhanced image quality, particularly in the challenging low-SNR environments. These improvements are particularly significant given the limitations of existing denoising techniques, which often struggle in high-noise conditions and result in less reliable performance in deeper regions.

In the simulation study, SPADE consistently outperformed BM3D in terms of noise suppression and quantitative metrics such as SNR, SSIM, and PSNR. Notably, SPADE can better preserve the clarity of the point target across different depths, particularly in low-SNR conditions, whereas BM3D's performance is weakened. This trend was reflected in the statistical evaluations, where SPADE maintained a higher average SNR and correlation coefficient, especially when the input SNR fell below 25 *dB*. This finding underscores the robustness of SPADE in handling noise without sacrificing the fidelity of the spectroscopic information, as evidenced by its performance in spectrum similarity preservation. Regarding the statistical SSIM results for each denoising group. Inspection of the data reveals outliers when the input SNR is between 11.77 *dB* and 15.47 *dB*. In this range, the SNR improvement varies from 9.78 *dB* to 21.08 *dB* — a considerably wider span than that observed at higher input SNR levels. The lower SNR improvements account for the noted outliers. Since the method employs a nonlinear network approach, as opposed to the analytical BM3D method, some nonlinear behavior is expected. Nonetheless, the method outperforms BM3D within this input SNR range. The evaluation of denoising performance across various input wavelengths highlights the stability and robustness of the SPADE framework. The results suggest that SPADE

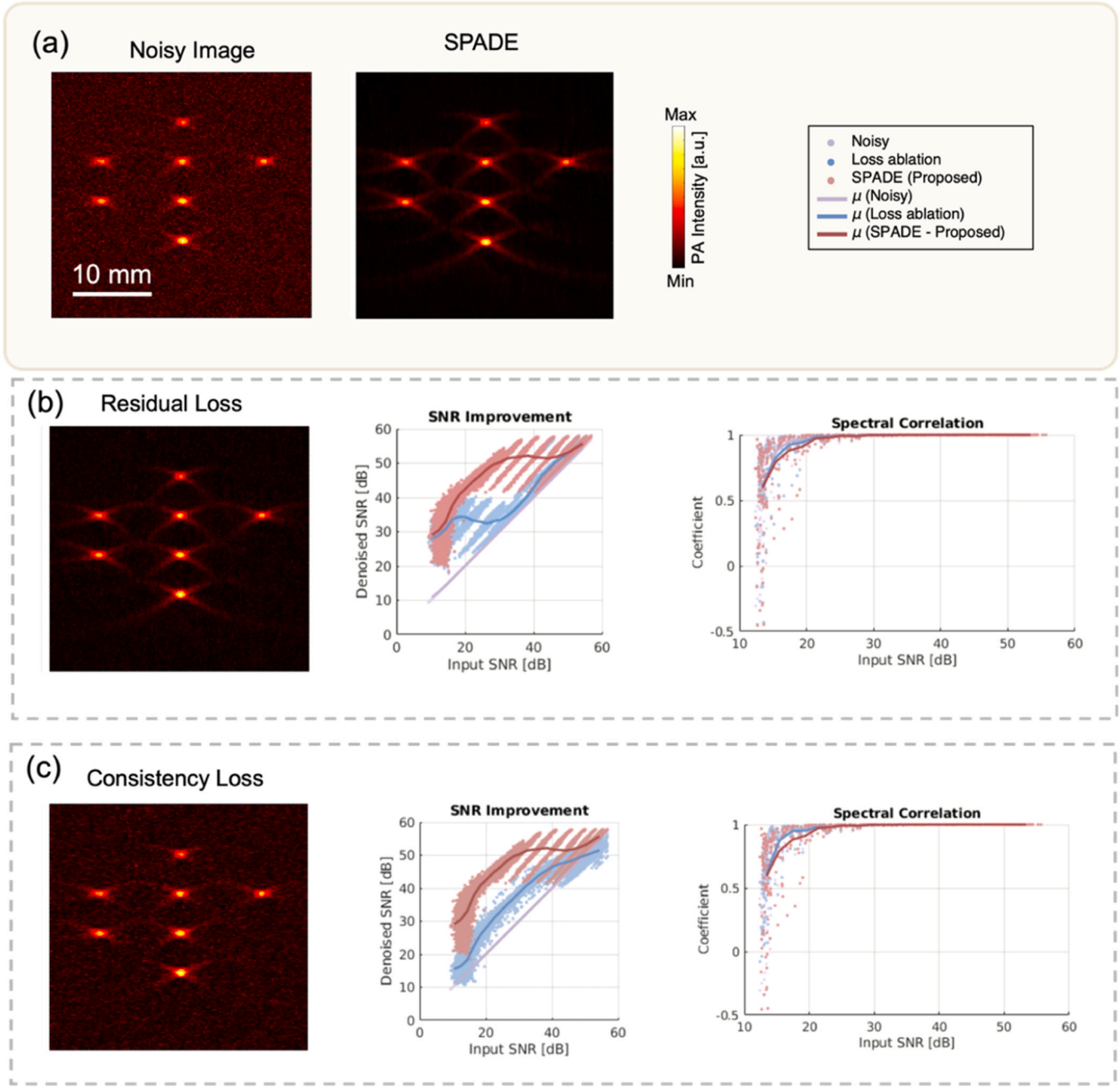
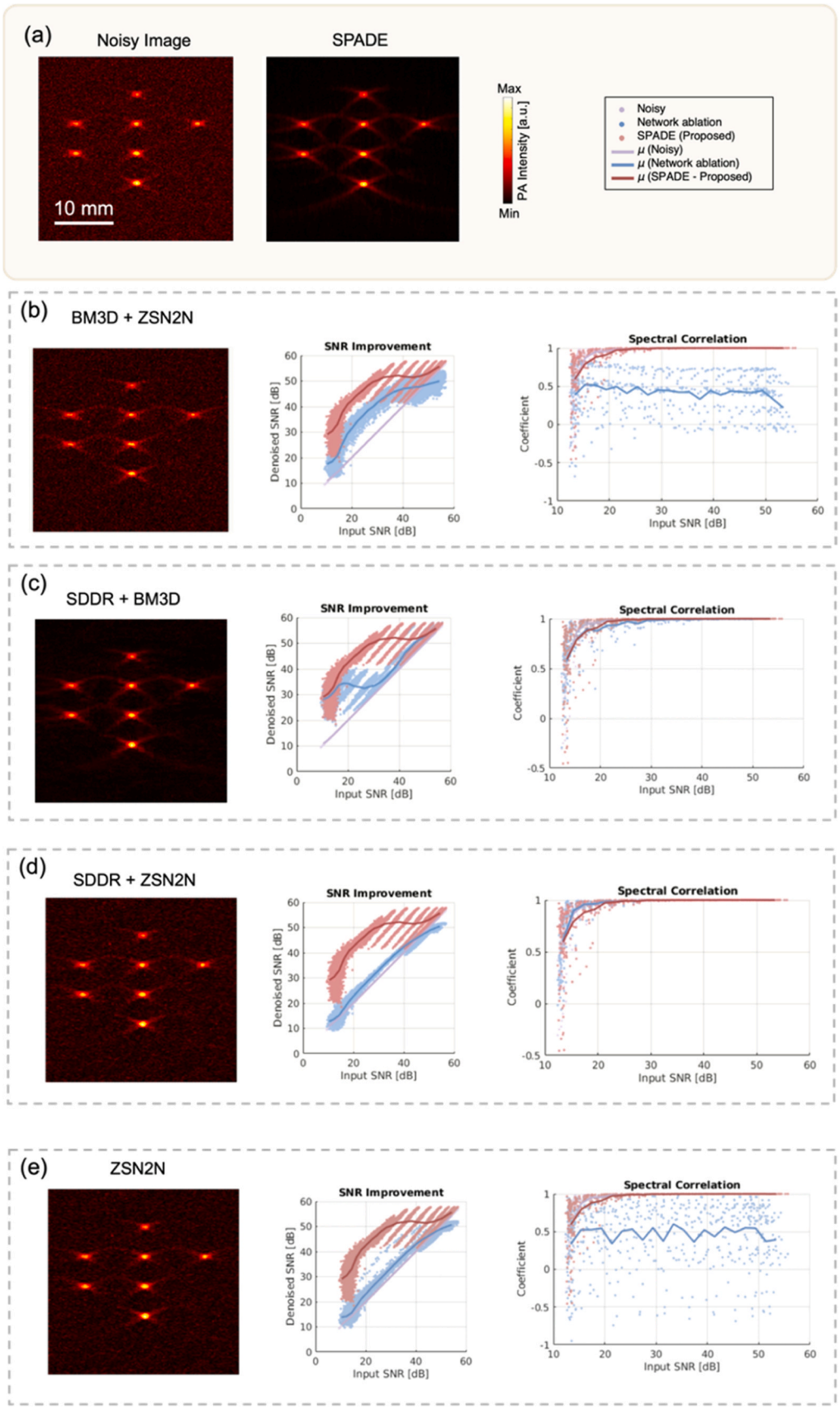


Fig. 9. Simulated spectroscopic photoacoustic (sPA) image with single loss only. (a) PA images (left to right): noisy single-frame PA image, and image denoised by the proposed SPADE method. Quantitative evaluation of the SPADE method compared to two single losses (residual loss, consistency loss): (b) Output signal-to-noise ratio (SNR) at different noise levels. Spectral correlation coefficient of residual loss only and SPADE at various noise levels. (c) Output signal-to-noise ratio (SNR) at different noise levels. Spectral correlation coefficient of consistency loss only and SPADE at various noise levels.

maintains effective denoising even with as few as two input wavelengths, underscoring its versatility and efficiency in handling different input dimensions without compromising performance.

The phantom study further corroborates these findings, demonstrating that SPADE achieves higher SNR values and more effective noise suppression in a physical model. Compared to the 64-frame averaging filter and BM3D, SPADE provided clearer point targets across varying depths. This result indicates the method's usability in scenarios where high frame averaging may be impractical or introduce latency. Additionally, SPADE's performance consistency across different noise levels, as illustrated by the tighter interquartile range in SNR values, reinforces its applicability in diverse imaging conditions. Although the median SSIM value for SPADE (0.647) is lower than that of vanilla BM3D (0.691), and a similar trend is seen with PSNR, SPADE suppressed more noise around the point targets, as confirmed by both SNR measurements and visual inspection. The lower SSIM and PSNR values are likely due to the removal of signal sidelobes in SPADE images, which are present in the averaging filter images due to beamforming.

The ex vivo evaluation using swine cardiac tissue verifies the effectiveness of SPADE in clinical settings. Here, the algorithm demonstrated a strong capacity to enhance tissue boundary visibility, which was challenging the noisy data. The quantitative spectral analysis, which accurately revealed the ablation-induced necrotic boundary, further confirms the spectral preservation of our proposed method. Statistical analysis showed that SPADE achieved the highest median SNR and PSNR values, closely aligning with the ground truth from the 64 frames



(caption on next page)

Fig. 10. Simulated spectroscopic photoacoustic (sPA) image with different network architectures. (a) PA images (left to right): noisy single-frame PA image, and image denoised by the proposed SPADE method. Quantitative evaluation of the SPADE method compared to different network architectures: (b) Output signal-to-noise ratio (SNR) at different noise levels. Spectral correlation coefficient of BM3D + ZSN2N and SPADE at various noise levels. (c) Output signal-to-noise ratio (SNR) at different noise levels. Spectral correlation coefficient of SDDR + BM3D and SPADE at various noise levels. (d) Output signal-to-noise ratio (SNR) at different noise levels. Spectral correlation coefficient of SDDR + ZSN2N and SPADE at various noise levels. (e) Output signal-to-noise ratio (SNR) at different noise levels. Spectral correlation coefficient of ZSN2N and SPADE at various noise levels.

averaging filter. The ability of SPADE to maintain higher spectral correlation coefficients with single-frame noisy data input indicates its potential for increasing PA imaging speed which is critical for future clinical translation. Currently, this study focuses on denoising multi-wavelength PA images using single-frame input per wavelength. With the implementation of the proposed SDDR, we demonstrate effective denoising while preserving spectral information. Achieving single-frame denoising without the need for averaging paves the way for enhanced imaging speed. However, we acknowledge that the current denoising algorithm requires longer processing time and has not yet been optimized for higher frame rates. That said, the core preprocessing module, SDDR, is a plug-and-play component compatible with most learning-based or analytical methods. By replacing zero-shot learning approaches (such as ZS-N2N) with test-time or online training methods, inference time could be significantly reduced. We will revise the draft to ensure a more consistent discussion. The performance evaluation across different input wavelengths aligns with the simulation results, validating the robustness of the SPADE framework.

In vivo results on contrast agent-enhanced tumor detection provided additional confirmation of SPADE's applicability in living animals. The algorithm successfully suppressed noise while retaining the integrity of the contrast agent quantification, which was observed in both before and 24 hours after injection scanning. Images denoised by SPADE closely matched the reference 64-frame averaged images and outperformed BM3D in terms of all the quantitative matrices. Importantly, SPADE was the only method that substantially eliminated false contrast agent signals outside the tissue boundary. Although some minor signals remain, this result demonstrates SPADE's precision in preserving relevant features while effectively removing artifacts. This capability is particularly valuable in vivo imaging applications where accurate quantification of molecular and structural targets is critical for diagnosis and therapy monitoring. Moreover, we didn't observe that SPADE introducing artifacts when applied to data with extremely low SNR. The reason is that the denoising mechanisms of both analytical and learning-based parts will not introduce extra information to the image. More specifically, the analytical part is based on thresholding and filters to remove noise. The learning part is based on residual learning to estimate noise and remove it. Either method will not introduce artifacts. For the spatially nonuniform noise, as discussed in the previous comments, SPADE can generalize to different types of noise. For example, we have structured noise from the beamforming operation and motion artifacts in an in vivo study, which is demonstrated in our result sections. But, when data with extremely low SNR, the performance of SPADE degraded, as shown in simulation experiment in Fig. 4(b), when the input SNR ranged from 11.77 *dB* to 15.47 *dB*. The image quality improvement varied between 9.78 *dB* and 21.08 *dB* in SNR, a relatively large range compared to cases with higher input SNR. There are certain number of lower SNR improvement denoised data. Given that our method employs a nonlinear network approach, unlike the analytical BM3D approach, some nonlinear behavior is expected. Our inclusion of preclinical ex vivo and in vivo results is specifically intended to demonstrate the framework's effectiveness in realistic noise scenarios, supporting future clinical translation. Major uncertainties, such as subject-specific artifacts, are naturally present. More specifically, the ex vivo evaluation using swine cardiac tissue verifies SPADE's effectiveness in a large-animal setting. Additionally, we conducted an in vivo rodent study to assess the method's performance in a live animal model, where breathing and other motions introduce real-world challenges. The in vivo demonstration used a contrast agent-injected mouse model to enhance tumor detection. This study aims to validate the proposed method in a dynamic in vivo environment, where factors such as respiratory motion and myoclonus (involuntary muscle contractions) during anesthesia can diminish the effectiveness of averaging filters. Furthermore, reducing the number of frames for denoising allows for an increased scanning rate.

Although the results successfully demonstrate the effectiveness of the SPADE method in denoising sPA images, several limitations exist. First, the denoising was performed using PA images collected from the same hardware, and the generalizability of the method across different imaging setups and transducers has not been tested. Therefore, evaluation using various imaging probes is necessary. Secondly, the current denoising process using SPADE is time-consuming, even with high computational power. Denoising a 16-wavelength sPA image takes approximately 70 seconds, comparable to the duration of an 87.5-frame average using a 20 Hz repetition laser. While real-time visualization is not yet feasible, the method eliminates the need for motion-gating during averaging in cases of dynamic cardiac or respiratory motion. Future efforts to optimize the algorithm and network to improve processing speed are needed. For example, we are actively exploring several avenues to accelerate our method in the future, including: (1) Integrating parts of our workflow in C+ + to reduce computational overhead. (2) Adapting (self) supervised trained network for real-time performance. (3) Employing test-time training strategies to further minimize runtime with minimal impact on performance. (4) Optimizing the SDDR module to allow for preliminary real-time denoising after just a few warmup frames. Thirdly, the introduction of a learning-based N2N network relies on stochastic gradient descent, introducing randomness to the output even with identical input [45,46]. This randomness leads to higher standard deviation in performance metrics. Although our network is designed to predict and subtract noise; however, due to its inherent nonlinear characteristics as our method employs a nonlinear network approach, there is a risk that some structural details might be inadvertently classified as noise. This risk is particularly pronounced under low-SNR conditions, where the chance of mistakenly removing important structural information increases, potentially leading to spurious results or performance degradation. This is a challenge with a learning-based method in general. While no serious instability in spectral outcomes was observed in this study, a more robust evaluation is needed to assess potential performance variation due to this randomness. Moreover, the effectiveness of the SPADE framework in suppressing noise arising from electromagnetic interference, such as MRI [47] or radiofrequency ablation energy [48], was not evaluated in this study and requires future investigation.

Moreover, considering that the sPA is obtained by scanning multi-frames, motion artifacts may appear. We did not apply motion artifact processing because it's not an issue in this study. The laser system enables fast continuous wavelength tuning at 20 Hz without delay, allowing multiwavelength PA imaging without significant motion concerns. The primary goal of this work is to propose and validate a denoising framework while preserving PA spectral information. However, future applications may require motion-related processing. Our previous study has demonstrated motion gating based on image similarity [14], which could be incorporated if needed.

Additionally, while this work focuses exclusively on sPA imaging, the SPADE framework holds promising potential for cross-applications in other imaging modalities, such as hyperspectral imaging and

fluorescence imaging. In hyperspectral imaging [49], where the acquisition of spatial and spectral information is critical, SPADE's spectral unmixing capabilities could enhance material classification and detection tasks [50,51]. Similarly, in fluorescence imaging [52], SPADE's ability to manage spectral complexities could aid in resolving overlapping emission spectra, thereby enhancing multiplexed imaging applications [53,54].

6. Conclusion

In conclusion, the proposed SPADE framework effectively addresses the challenges of denoising sPA imaging. By combining a data-free learning-based method with an analytical approach, SPADE achieves higher noise suppression while preserving spectral information, critical for quantitative PA imaging. Through extensive validation in simulation, phantom, ex vivo, and in vivo experiments, the method exhibited significant improvements in SNR compared to noisy images, with comparable outcomes to the averaging filter. The framework consistently provided clearer image visualization and more accurate spectral quantification, confirming its potential for enhancing sPA imaging in various clinical applications. Future work should focus on optimizing processing speed and assess the generality of the method across different imaging systems for real-time clinical implementation.

CRediT authorship contribution statement

Ma Xihan: Writing – review & editing, Methodology. **Tang Yichuan:** Writing – review & editing, Methodology, Data curation. **Murakami Ryo:** Writing – review & editing, Data curation. **Zhang Haichong K.:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Soboyejo Winston O.:** Writing – review & editing, Supervision. **Obayemi John D.:** Writing – review & editing, Data curation. **Zhang Ziming:** Writing – review & editing, Resources, Methodology. **Gao Shang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Data curation, Conceptualization. **LIN FANGZHOU:** Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Haichong K. Zhang reports financial support was provided by National Institutes of Health. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by Worcester Polytechnic Institute Internal Fund and the National Institutes of Health under grants: R01DK133717, DP5OD028162, R01CA134675, R01CA166379 and R01EB030539. We thank the Worcester Polytechnic Institute Animal Care and Use Committee for their assistance with this study.

Data availability

Data will be made available on request.

References

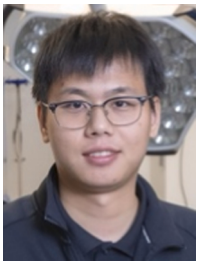
[1] P. Beard, Biomedical photoacoustic imaging, Rev., Interface Focus 1 (2011) 602–631, https://doi.org/10.1098/rsfs.2011.0028.
[2] L.V. Wang, J. Yao, A practical guide to photoacoustic tomography in the life sciences, Nat. Methods 13 (8) (2016) 627–638, https://doi.org/10.1038/nmeth.3925.
[3] J.J.M. Riksen, A.V. Nikolaev, G. van Soest, Photoacoustic imaging on its way toward clinical utility: a tutorial review focusing on practical application in medicine, J. Biomed. Opt. 28 (2023), https://doi.org/10.1117/1.jbo.28.12.121205.
[4] S. Gao, Y. Jiang, M. Li, Y. Wang, Y. Shen, M.C. Flegal, B.C. Nephew, G.S. Fischer, Y. Liu, L. Fichera, H.K. Zhang, Laparoscopic photoacoustic imaging system based on side-illumination diffusing fibers, IEEE Trans. Biomed. Eng. (2023), https://doi.org/10.1109/TBME.2023.3279772.
[5] S. Gao, Y. Wang, X. Ma, H. Zhou, Y. Jiang, K. Yang, L. Lu, S. Wang, B.C. Nephew, L. Fichera, G.S. Fischer, H.K. Zhang, Intraoperative laparoscopic photoacoustic image guidance system in the da Vinci surgical system, Biomed. Opt. Express 14 (2023), https://doi.org/10.1364/boe.498052.
[6] H.K. Zhang, Y. Chen, J. Kang, A. Lisok, I. Minn, M.G. Pomper, E.M. Boctor, Prostate-specific membrane antigen-targeted photoacoustic imaging of prostate cancer in vivo, J. Biophotonics 11 (2018) 1–6, https://doi.org/10.1002/jbio.201800021.
[7] W.G. Lesniak, Y. Wu, J. Kang, S. Boinapally, S. Ray Banerjee, A. Lisok, A. Jablonska, E.M. Boctor, M.G. Pomper, Dual contrast agents for fluorescence and photoacoustic imaging: evaluation in a murine model of prostate cancer, Nanoscale 13 (2021) 9217–9228, https://doi.org/10.1039/D1NR00669J.
[8] Y. Wu, J. Kang, W.G. Lesniak, A. Lisok, H.K. Zhang, R.H. Taylor, M.G. Pomper, E. M. Boctor, System-level optimization in spectroscopic photoacoustic imaging of prostate cancer, Photoacoustics 27 (2022), https://doi.org/10.1016/j.pacs.2022.100378.
[9] M. Li, Y. Tang, J. Yao, Photoacoustic tomography of blood oxygenation: a mini review, Photoacoustics 10 (2018) 65–73, https://doi.org/10.1016/j.pacs.2018.05.001.
[10] S. Gao, X. Ma, H.K. Zhang, Robot-Assisted Wide-Area Photoacoustic System, 2023 IEEE International Ultrasonics Symposium (IUS) (2023) 1–4. ⟨https://doi.org/10.1109/IUS51837.2023.10307874⟩.
[11] L.S. Oey, M. Mitrayana, M.A.J. Wasono, Application of a photoacoustic tomography system: a case study on the monitoring of pig tissue decomposition, Makara J. Sci. 26 (2022), https://doi.org/10.7454/mss.v26i1.1257.
[12] N.E. Pacheco, S. Gao, K. Cleary, R. Shah, H.K. Zhang, L. Fichera, 2024, Photoacoustic image guidance for laser tonsil ablation: approach and initial results, Https://Doi.Org/10.1117/12.3007295 12928 (2024) 278–282. https://doi.org/10.1117/12.3007295.
[13] R. Murakami, S. Gao, M. Chitrapu, A. Sungarian, H.K. Zhang, Preliminary study on characterizing brain-ablation-induced necrosis through spectroscopic photoacoustic imaging, in: 2024. ⟨https://doi.org/10.1117/12.3001173⟩.
[14] S. Gao, H. Ashikaga, M. Suzuki, T. Mansi, Y.-H. Kim, F.-C. Ghesu, J. Kang, E. M. Boctor, H.R. Halperin, H.K. Zhang, Cardiac-gated spectroscopic photoacoustic imaging for ablation-induced necrotic lesion visualization, J. Biophotonics (2024) e202400126, https://doi.org/10.1002/JBIO.202400126.
[15] M. Graham, F. Assis, D. Allman, A. Wiacek, E. Gonzalez, M. Gubbi, J. Dong, H. Hou, S. Beck, J. Chrispin, M.A.L. Bell, In vivo demonstration of photoacoustic image guidance and robotic visual servoing for cardiac catheter-based interventions, IEEE Trans. Med Imaging 39 (2020) 1015–1029, https://doi.org/10.1109/TMI.2019.2939568.
[16] R. Manwar, M. Zafar, Q. Xu, Signal and image processing in biomedical photoacoustic imaging: a review, Optics 2 (2021), https://doi.org/10.3390/opt2010001.
[17] S.V. Mohd Sagheer, S.N. George, A review on medical image denoising algorithms, Biomed. Signal Process Control 61 (2020), https://doi.org/10.1016/j.bspc.2020.102036.
[18] P. Madhumathy, D. Pandey, Deep learning based photo acoustic imaging for non-invasive imaging, Multimed. Tools Appl. 81 (2022), https://doi.org/10.1007/s11042-022-11903-6.
[19] J. Li, B. Yu, W. Zhao, W. Chen, A review of signal enhancement and noise reduction techniques for tunable diode laser absorption spectroscopy, Appl. Spectrosc. Rev. 49 (2014), https://doi.org/10.1080/05704928.2014.903376.
[20] Y. Cheng, W. Zheng, R. Bing, H. Zhang, C. Huang, P. Huang, L. Ying, L. Ying, J. Xia, Unsupervised denoising of photoacoustic images based on the Noise2Noise network, Biomed. Opt. Express 15 (8) (2024) 4390–4405, https://doi.org/10.1364/BOE.529253.
[21] D. He, J. Zhou, X. Shang, X. Tang, J. Luo, S.L. Chen, De-noising of photoacoustic microscopy images by attentive generative adversarial network, IEEE Trans. Med. Imaging 42 (2023), https://doi.org/10.1109/TMI.2022.3227105.
[22] J. Vonk, F. Knieling, S. Kruijff, Collection on clinical photoacoustic imaging, Eur. J. Nucl. Med. Mol. Imaging (2024) 1–2, https://doi.org/10.1007/S00259-024-06780-0/METRICS.
[23] S. Izadi, D. Sutton, G. Hamarneh, Image denoising in the deep learning era, Artif. Intell. Rev. 56 (2023), https://doi.org/10.1007/s10462-022-10305-2.
[24] Z. Yu, K. Tang, X. Song, 2021, Denoising method for image quality improvement in photoacoustic microscopy using deep learning, in: 2021. https://doi.org/10.1117/12.2600759..
[25] N. Ye, K. Li, H. Bai, R. Yu, L. Hong, F. Zhou, Z. Li, J. Zhu, OoD-Bench: quantifying and understanding two dimensions of out-of-distribution generalization, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2022), https://doi.org/10.1109/CVPR52688.2022.00779.
[26] S. Tzoumas, A. Rosenthal, C. Lutzweiler, D. Razansky, V. Ntziachristos, Spatiospectral denoising framework for multispectral optoacoustic imaging based

on sparse signal representation, Med. Phys. 41 (2014), https://doi.org/10.1118/1.4893530.
[27] N. Awasthi, S.K. Kalva, M. Pramanik, P.K. Yalavarthy, Image-guided filtering for improving photoacoustic tomographic image reconstruction, J. Biomed. Opt. 23 (2018), https://doi.org/10.1117/1.jbo.23.9.091413.
[28] Q. Kong, Q. Song, Y. Hai, R. Gong, J. Liu, X. Shao, Denoising signals for photoacoustic imaging in frequency domain based on empirical mode decomposition, Opt. (Stuttg.) 160 (2018), https://doi.org/10.1016/j.ijleo.2018.02.023.
[29] M. Shi, T. Vercauteren, W. Xia, Spatiotemporal singular value decomposition for denoising in photoacoustic imaging with a low-energy excitation light source, Biomed. Opt. Express 13 (2022), https://doi.org/10.1364/boe.471198.
[30] A. Kazakeviciute, C.J.H. Ho, M. Olivo, Multispectral photoacoustic imaging artifact removal and denoising using time series model-based spectral noise estimation, IEEE Trans. Med. Imaging 35 (2016), https://doi.org/10.1109/TMI.2016.2550624.
[31] Y. Mansour, R. Heckel, Zero-Shot Noise2Noise: efficient image denoising without any data, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2023), https://doi.org/10.1109/CVPR52729.2023.01347.
[32] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, T. Aila, Noise2Noise: Learning image restoration without clean data, in: 35th International Conference on Machine Learning, ICML 2018, 2018.
[33] T. Huang, S. Li, X. Jia, H. Lu, J. Liu, Neighbor2Neighbor: Self-supervised denoising from single noisy images, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2021. ⟨https://doi.org/10.1109/CVPR46437.2021.01454⟩.
[34] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Trans. Image Process. 16 (2007), https://doi.org/10.1109/TIP.2007.901238.
[35] D.R. Bull, F. Zhang, 2021, Intelligent Image and Video Compression: Communicating Pictures, 2021. https://doi.org/10.1016/B978-0-12-820353-8.00002-5..
[36] X.J. Mao, C. Shen, Y.Bin Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections, Adv. Neural Inf. Process Syst. (2016).
[37] W. Jifara, F. Jiang, S. Rho, M. Cheng, S. Liu, Medical image denoising using convolutional neural network: a residual learning approach, J. Supercomput. 75 (2019), https://doi.org/10.1007/s11227-017-2080-0.
[38] F. Lin, Y. Yue, S. Hou, X. Yu, Y. Xu, K.D. Yamada, Z. Zhang, Hyperbolic chamfer distance for point cloud completion, Proc. IEEE Int. Conf. Comput. Vis. (2023), https://doi.org/10.1109/ICCV51070.2023.01342.
[39] K.M. Kempski, M.T. Graham, M.R. Gubbi, T. Palmer, M.A. Lediju Bell, Application of the generalized contrast-to-noise ratio to assess photoacoustic image quality, Biomed. Opt. Express 11 (2020), https://doi.org/10.1364/boe.391026.
[40] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (2004), https://doi.org/10.1109/TIP.2003.819861.
[41] B.E. Treeby, B.T. Cox, k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, J. Biomed. Opt. 15 (2010) 021314, https://doi.org/10.1117/1.3360308.
[42] A.A. Leino, A. Pulkkinen, T. Tarvainen, ValoMC: a Monte Carlo software and MATLAB toolbox for simulating light transport in biological tissue, OSA Contin. 2 (2019) 957, https://doi.org/10.1364/osac.2.000957.
[43] S. Gao, R. Tsumura, D.P. Vang, K. Biland, K. Xu, Y. Tsunoi, H.K. Zhang, Acoustic-resolution photoacoustic microscope based on compact and low-cost delta configuration actuator, Ultrasonics 118 (2021) 106549, https://doi.org/10.1016/j.ultras.2021.106549.
[44] S. Gao, H. Liu, A. Post, L. Jaworski, D. Bernard, M. John, E. Cosgriff-Hernandez, M. Razavi, H.K. Zhang, Enhancing boundary detection of radiofrequency ablation lesions through photoacoustic mapping, Sci. Rep. 14 (2024) 1–11, https://doi.org/10.1038/s41598-024-68046-x.
[45] F. Lin, Y. Yue, Z. Zhang, S. Hou, K.D. Yamada, V.B. Kolachalama, V. Saligrama, InfoCD: a contrastive chamfer distance loss for point cloud completion, : Adv. Neural Inf. Process Syst. (2023).
[46] F. Lin, Y. Xu, Z. Zhang, C. Gao, K.D. Yamada, Cosmos propagation network: deep learning model for point cloud completion, Neurocomputing 507 (2022), https://doi.org/10.1016/j.neucom.2022.08.007.
[47] R. Murakami, Y. Wang, W. Lesniak, R. Tsumura, Y. Tang, Y. Tsunoi, C.J. Nycz, M. G. Pomper, G.S. Fischer, H.K. Zhang, In-bore MRI-compatible transrectal ultrasound and photoacoustic imaging, BioRxiv (2023) 2023.11.27.568947, https://doi.org/10.1101/2023.11.27.568947.
[48] S. Gao, H. Ashikaga, T. Mansi, H.R. Halperin, H.K. Zhang, Photoacoustic Necrotic Region Mapping for Radiofrequency Ablation Guidance, in: IEEE International Ultrasonics Symposium, IUS, IEEE, 2021: pp. 1–4. ⟨https://doi.org/10.1109/IUS52206.2021.9593388⟩.
[49] J. Yoon, Hyperspectral imaging for clinical applications, BioChip J. 16 (1) (2022) 1–12.
[50] L. Bian, Z. Wang, Y. Zhang, L. Li, Y. Zhang, C. Yang, J. Zhang, A broadband hyperspectral image sensor with high spatio-temporal resolution, Nature 635 (8037) (2024) 73–81.
[51] K. Jo, S. Lee, D.H. Lee, H. Jeon, S. Jung, Hyperspectral imaging–based assessment of fresh meat quality: progress and applications, Microchem. J. 197 (2024) 109785.
[52] F. Wang, Y. Zhong, O. Bruns, Y. Liang, H. Dai, In vivo NIR-II fluorescence imaging for biology and medicine, Nat. Photonics (2024) 1–13.
[53] H. Cheng, H. Xu, B. Peng, X. Huang, Y. Hu, C. Zheng, Z. Zhang, Illuminating the future of precision cancer surgery with fluorescence imaging and artificial intelligence convergence, NPJ Precis. Oncol. 8 (1) (2024) 196.
[54] D. Schraivogel, T.M. Kuhn, B. Rauscher, M. Rodríguez-Martínez, M. Paulsen, K. Owsley, L.M. Steinmetz, High-speed fluorescence image–enabled cell sorting, Science 375 (6578) (2022) 315–320.

Fangzhou Lin received his M.S. and Ph.D. in information sciences and data science from Tohoku University, Japan, Sendai, in 2021 and 2024. His research interests include 3D point cloud and medical imaging applications.

Shang Gao, born in Nagoya, Japan, is a Ph.D. candidate at the Medical FUSION (Frontier Ultrasound Imaging and Robotic Instrumentation) Laboratory at Worcester Polytechnic Institute (WPI) in the USA. He earned his M.S. degree in Robotics Engineering from WPI in 2020. In 2018, he received dual B.Eng. degrees: Robotics & Mechatronic Systems Engineering from the University of Detroit Mercy, USA; and Mechanical Engineering from Beijing University of Chemical Technology, China, graduating with honors. His research interests focus on medical robotics, photoacoustic imaging, and image-guided interventions. He is the recipient of the Robert F. Wagner All-Conference Best Student Paper Award from SPIE Medical Imaging 2023.

Yichuan Tang is a PhD candidate in the Department of Robotics Engineering in Worcester Polytechnic Institute, Worcester, MA, USA, since 2019. He received his BEng degrees with honors in mechanical engineering from the University of Nottingham (China campus), Ningbo, 2016. He received his MS degree in robotics engineering from Johns Hopkins University in 2018. His research focuses on ultrasound and photoacoustic imaging instrumentation.



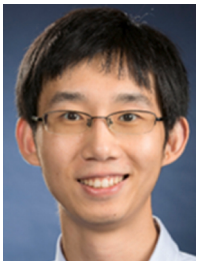
Xihan Ma received dual B.Eng. degrees in mechanical engineering and automation from Beijing University of Chemical Technology, Beijing, China, and mechatronics and robotics systems engineering from the University of Detroit Mercy, Detroit, Michigan, USA, in 2019. He received the M.Sc. degree in robotics engineering from Worcester Polytechnic Institute, Worcester, Massachusetts,USA, in 2021. He is currently pursuing Ph.D. degree in robotics engineering with the Department of Robotics Engineering, Worcester Polytechnic Institute, Worcester, Massachusetts, USA. His research interests include medical imaging and medical robotics, with a focus on robotic ultrasound imaging and ultrasound guided intervention procedures.

Ryo Murakami is a Ph.D. candidate at the Medical FUSION Laboratory, Robotics Engineering at Worcester Polytechnic Institute. His research focuses on the combination of photoacoustic/ultrasound technology and control theory to improve image-guided treatment/diagnosis. Prior to joining the Medical FUSION Laboratory, he worked in clinical sciences at Boston Scientific Japan K.K. and business development at Zimmer Biomet G.K. He earned his bachelor's and master's degrees in Mechanical Engineering at Osaka University, Japan.

Ziming Zhang is an assistant professor at Worcester Polytechnic Institute (WPI). Before joining WPI he was a research scientist at Mitsubishi Electric Research Laboratories (MERL) in 2017–2019. Prior to that, he was a research assistant professor at Boston University in 2016–2017. Dr. Zhang received his PhD in 2013 from Oxford Brookes University, UK, under the supervision of Prof. Philip H. S. Torr. His research interests lie in computer vision and machine learning. He won the R\&D 100 Award 2018.

John D. Obayemi, Ph.D., MBA, is an Assistant Teaching Professor in the Biomedical Engineering Department at Worcester Polytechnic Institute (WPI). He earned his Bachelor of Engineering in Mechanical Engineering from Ahmadu Bello University, Zaria, Nigeria, and both his Master of Science and Ph.D. in Materials Science and Engineering from the African University of Science and Technology (AUST), Abuja, Nigeria. His doctoral research, supervised by a Princeton University professor, focused on "Nano- and Micro-particles for the Detection and Treatment of Cancer." Prior to joining WPI, Dr. Obayemi served as a Lecturer and Postdoctoral Research Associate in the Department of Mechanical and Aerospace Engineering at Princeton University, where he conducted research and taught courses on bioengineering and medical devices. His research interests include drug development, targeted therapies, biomedical devices, and cancer detection and treatment using nanoparticles.

Winston O. Soboyejo, Ph.D., is the President of SUNY Polytechnic Institute. Prior to this, he held significant leadership roles at Worcester Polytechnic Institute (WPI), including Senior Vice President, Provost, and the Bernard M. Gordon Dean of Engineering. He received his Bachelor of Science in Mechanical Engineering from King's College London and a Ph.D. in Materials Science and Metallurgy from Churchill College, Cambridge. Dr. Soboyejo's career encompasses roles as a Research Scientist at McDonnell Douglas Research Labs, faculty member at The Ohio State University, and visiting professor at MIT. He later joined Princeton University as a Professor of Mechanical and Aerospace Engineering, directing the U.S./Africa Materials Institute. His research interests include biomaterials, cancer-targeting nanotechnology, and sustainable solutions for clean water and affordable housing, with over 400 peer-reviewed publications.



Haichong (Kai) Zhang is an Associate Professor in Biomedical Engineering and Robotics Engineering with an appointment in Computer Science at Worcester Polytechnic Institute (WPI). He is the founding director of the Medical Frontier Ultrasound Imaging and Robotic Instrumentation (FUSION) Laboratory. The research in his lab focuses on the interface of medical imaging, sensing, and robotics, developing robotic-assisted imaging systems as well as image-guided robotic interventional platforms, where ultrasound and photoacoustic imaging are two key modalities to be investigated and integrated with robotics. Dr. Zhang received his B.S. and M.S. in Human Health Sciences from Kyoto University, Japan, and subsequently earned his M.S. and Ph.D. in Computer Science from Johns Hopkins University.