

Reduction of photobleaching effects in photoacoustic imaging using noise agnostic, platform-flexible deep-learning methods

Avijit Paul,^a Christopher Nguyen,^a Tayyaba Hasan^{a, b} and Srivalleesha Mallidi^{a, *}

^aTufts University, Department of Biomedical Engineering, Medford, Massachusetts, United States

^bMassachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, United States

ABSTRACT. **Significance:** Molecular photoacoustic (PA) imaging with exogenous dyes faces a significant challenge due to the photobleaching of the dye that can compromise tissue visualization, particularly in 3D imaging. Addressing this limitation can revolutionize the field by enabling safer, more reliable imaging and improve real-time visualization, quantitative analysis, and clinical decision-making in various molecular PA imaging applications such as image-guided surgeries.

Aim: We tackle photobleaching in molecular PA imaging by introducing a platform-flexible deep learning framework that enhances SNR from single-laser pulse data, preserving contrast and signal integrity without requiring averaging of signals from multiple laser pulses.

Approach: The generative deep learning network was trained with an LED-illuminated PA image dataset and tested on acoustic resolution PA microscopy images obtained with single-laser pulse illumination. *In vitro* and *ex vivo* samples were first tested for demonstrating SNR improvement, and then, a 3D-scanning experiment with an ICG-filled tube was conducted to depict the usability of the technique in reducing the impact of photobleaching during PA imaging.

Results: Our generative deep learning model outperformed traditional nonlearning, filter-based algorithms and the U-Net deep learning network when tested with *in vitro* and *ex vivo* single pulse-illuminated images, showing superior performance in terms of signal-to-noise ratio (93.54 ± 6.07 , and 92.77 ± 10.74 compared with 86.35 ± 3.97 , and 84.52 ± 11.82 with U-Net for kidney, and tumor, respectively) and contrast-to-noise ratio (11.82 ± 4.42 , and 9.9 ± 4.41 compared with 7.59 ± 0.82 , and 6.82 ± 2.12 with U-Net for kidney, and tumor respectively). The use of cGAN with single-pulse rapid imaging has the potential to prevent photobleaching ($9.51 \pm 3.69\%$ with cGAN, and $35.14 \pm 5.38\%$ with long-time laser exposure by averaging 30 pulses), enabling accurate, quantitative imaging suitable for real-time implementation, and improved clinical decision support.

Conclusions: We demonstrate the potential of a platform-flexible generative deep learning-based approach to mitigate the effects of photobleaching in PA imaging by enhancing signal-to-noise ratio from single pulse-illuminated data, thereby improving image quality and preserving contrast in real time.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JBO.30.S3.S34102](https://doi.org/10.1117/1.JBO.30.S3.S34102)]

Keywords: photobleaching; platform-flexible generative deep learning; photoacoustic imaging; acoustic resolution photoacoustic microscopy

Paper 240379SSRR received Dec. 16, 2024; revised Apr. 1, 2025; accepted Apr. 3, 2025; published May 28, 2025.

*Address all correspondence to Srivalleesha Mallidi, srivalleesha.mallidi@tufts.edu

1 Introduction

Photobleaching is the irreversible degradation of a dye molecule, triggered by photon absorption from a rare short-lived excited singlet state or a more common long-lived triplet state after inter-system crossing.^{1–3} This phenomenon is influenced by intrinsic molecular properties as well as environmental factors such as oxygen concentration, pH, and temperature. As photobleaching occurs with prolonged light exposure, it limits the time window for obtaining quantitative images in optical imaging.^{4–6} For example, in contrast-enhanced optical image-guided surgical interventions, surgeons may need to frequently readminister or reapply imaging agents, disrupting the procedure. The signal intensity loss due to photobleaching compromises real-time visualization of critical structures during surgery, such as tissues, vessels, or tumors, ultimately leading to diminished sensitivity in detecting subtle anatomical or pathological features. Quantitative analysis, such as measuring the concentration of a dye or tissue metabolic activity, relies on consistent signal intensity. Photobleaching introduces variability in optical signal intensity over time, leading to inaccurate measurements and potentially flawed decisions. Alternatively, to compensate for reduced signal intensity, surgeons might increase laser intensity that can cause tissue heating and phototoxic effects on sensitive structures or develop photostable dyes.^{7–10} Photobleaching is thus a common challenge in optical imaging, particularly photoacoustic (PA) imaging (PAI) due to its high power delivered within the short nanosecond laser pulse, and its impact is particularly significant in applications demanding high spatial and temporal resolution, where consistent image quality and signal reliability are critical.^{11–14}

PAI, an established hybrid biomedical imaging modality combining optical excitation and acoustic detection,^{13,15} offers unique advantages over traditional clinical imaging techniques such as X-ray,¹⁶ MRI,¹⁷ PET,¹⁸ and OCT¹⁹ in terms of radiation exposure or spatial and depth resolution for molecular image-guided surgery. By leveraging optical absorption contrast and providing superior spatial and depth resolution (acoustic), PAI has been successfully applied to both endogenous (label-free) and exogenous chromophores.²⁰ Among its configurations, acoustic-resolution PA microscopy (AR-PAM²¹) and optical-resolution PA microscopy (OR-PAM²²) have been widely used for visualizing microscopic structures. AR-PAM achieves resolution by focusing the acoustic transducer, whereas OR-PAM relies on tight optical focusing for superior lateral resolution.^{22–24} However, AR-PAM systems require 2D raster scanning for image acquisition, during which multiple laser pulses (20 to 30 per scan point) are typically averaged to enhance the signal-to-noise ratio (SNR).²⁵ As the laser pulse repetition frequency for a typical Nd-YAG pumped optical parametric oscillator laser is typically low (5 to 20 Hz), this averaging decreases the AR-PAM scan speed and eventually increases the plausibility of prolonged laser exposure which might impose a burden of photobleaching when using exogenous contrast agents such as indocyanine green (ICG). ICG, a commonly used, clinically approved contrast agent,^{26,27} absorbs photons and undergoes excitation to a singlet state, followed by nonradiative relaxation or interaction with molecular oxygen to generate reactive oxygen species (ROS).^{28,29} ROS-induced oxidative damage causes structural degradation of ICG, leading to reduced optical absorption and diminished PA signals. This degradation complicates quantitative analyses, such as concentration measurements and monitoring biodistribution, and limits the effective duration for which contrast agents can be used in imaging.

Mitigating photobleaching in PAI involves a trade-off between reducing illumination intensity or exposure time and maintaining adequate SNR for image reconstruction. Low-intensity laser pulses below the photobleaching threshold can be employed, but this results in noisy images that compromise contrast and resolution. Alternative approaches, such as modifying or encapsulating ICG with nanoparticles to enhance photostability^{7–10} or reducing ROS generation by altering environmental factors,^{30–33} are technically challenging, lack clinical approval, and require extensive safety and biological feasibility studies before clinical implementation. A promising alternative is to minimize the number of laser pulses per scan point, ideally using single-pulse illumination, to reduce cumulative energy deposition and photobleaching risk. However, single-pulse illumination produces inherently noisy data due to low energy delivery, posing challenges for high-quality image reconstruction.

In this study, we address these challenges by leveraging a platform-flexible generative deep learning (DL) model³⁴ to enhance SNR in single-pulse PAI before photobleaching occurs. Traditional algorithms such as Savitzky–Golay (SG),³⁵ Wiener,³⁶ and BM3D³⁷ are not efficient in generating high SNR images, and they require prior estimation of the noise distribution. DL, with its ability to generalize across unseen test data and automatically learn features from training data, has shown significant promise in biomedical imaging.^{38–40} Specifically, convolutional neural network (CNN)-based architectures,^{41–44} such as U-Net,^{45–48} have been widely applied for denoising and image enhancement tasks. Recent work has demonstrated the use of U-Net variants for improving lateral resolution in AR-PAM and denoising low-fluence OR-PAM images.⁴⁹ However, most existing studies focused on enhancing A-line signals before image reconstruction, which requires access to raw data—a limitation in certain PAI systems. To overcome this, our DL framework directly processes noisy images obtained from single-pulse illumination, bypassing the need for A-line access. Although GAN-based⁵⁰ approaches have been explored in various imaging modalities,^{51–54} their applications to PAI,^{48,55–58} particularly for mitigating photobleaching, remain underexplored. It is worth noting that the current work presents several unique advancements specifically designed for our problem domain, demonstrating noise-invariant platform flexibility through an optimized generative model. To validate our approach, we conducted phantom experiments with ICG solutions, where we observed a decline in PA amplitude with multiple pulse illuminations. Our cGAN model successfully reconstructed high-quality SNR images from single-pulse noisy data, offering a robust solution to the dual challenges of photobleaching and noisy imaging in real-time PAI. Furthermore, developing platform-independent DL architectures is vital for robust biomedical imaging. Such models can handle diverse input data and ensure consistent performance across systems. In PAI, platform flexibility minimizes reliance on specific hardware and maintains high SNR across configurations. Despite its significance, platform-flexible DL networks for PAI remain underexplored. Platform flexibility in this work refers to the model’s ability to generalize across different noise characteristics inherent to LED-based and AR-PAM imaging systems. This flexibility ensures robust denoising performance without requiring retraining or fine-tuning for each platform. Using a custom loss function-based cGAN trained on AcousticX LED-illuminated data and tested on AR-PAM data, we demonstrate platform flexibility for SNR improvement, paving the way for broader DL adoption in diverse PAI applications.

2 Methodology

2.1 Microscopic Imaging System

To perform PA imaging, the AR-PAM system employed an OPO (Phocus HE Benchtop, OPOTEK, Carlsbad, California, United States) to generate pulsed laser energy. The subsequent acoustic signal generated was collected using a 1-in. focused transducer (V324-SU, Olympus, Tokyo, Japan) with a center frequency of 25 MHz. This signal was amplified by a pulser/receiver (DPR500, JSR Ultrasonics, Pittsford, New York, United States), which also facilitated signal digitization by a data acquisition card (CSE161G2, GaGe, Ferndale, Michigan, United States).

To transmit the laser energy to the sample, the system utilized an achromatic doublet (AC254-030-B-ML, Thorlabs, Newton, New Jersey, United States) and an aspheric condenser (ACL2520U-B, Thorlabs) to couple the laser beam from the OPO fiber into a secondary fiber bundle. This secondary fiber bundle consisted of seven multimode fibers (FT1000EMT, Thorlabs) epoxied together toward the proximal end for coupling. The distal end of the fiber bundle was designed in a fanned configuration and connected to a 3D-printed transducer-and-fiber mount, which oriented the fibers to mimic an optical condenser.

2.2 Nonlearning Traditional Models

We implemented several traditional noise removal algorithms such as SG, Wiener, and BM3D to compare with the cGAN model. For the SG noise removal algorithm, local polynomial fitting ensures smoothness without significant loss of signal shape. Wiener filter aims at minimizing the mean square error between the original and estimated signals. It adapts based on local signal statistics, resulting in a trade-off between smoothness and signal preservation. However, this filter is ideal when we have prior knowledge about both the signal and noise characteristics,

such as their power spectra. BM3D is another state-of-the-art denoising algorithm that groups similar patches from an image and performs 3D transformation to filter the noise. The key advantage of this filter is the ability to group similar image blocks and process them together in 3D space, allowing it to outperform simpler denoising techniques in terms of detail preservation and noise reduction. In the SG algorithm, we employed a third-order SG filter (`sgolayfilt` in MATLAB[®]) with a frame length of 9 for convolution-based smoothing.⁵⁹ After several hits and trials, we observed that a third-order SG filter with a frame length of 9 was used to preserve signal characteristics while reducing noise. The third-order polynomial enables the filter to preserve both linear and quadratic signal trends, as well as subtle curvature variations, which are common in PA signals (e.g., vascular structures or temporal dynamics). Higher-order filters might be overfitted to noise, introducing artifacts, whereas lower-order filters (e.g., first or second order) might over-smooth, erasing critical features like sharp edges or peaks. We observed that the frame length of 9 balances noise suppression and feature preservation. It is to be noted that larger frame lengths might provide stronger noise suppression but risk smoothing out high-frequency details, whereas smaller frame lengths retain more detail but may inadequately suppress noise.^{60,61} The Wiener algorithm^{36,62} was implemented using a pixel-wise adaptive low-pass filter (`wiener2` function in MATLAB[®]) with a neighborhood size of 5×5 to estimate the local image mean and standard deviation. Multiple trials with several neighborhood window compositions indicate that the choice of a 5×5 neighborhood for the Wiener filter in PA imaging is grounded in its ability to provide a balanced trade-off between noise suppression, feature preservation, computational efficiency, and compatibility with the spatial resolution of PA systems. In BM3D, we initially estimated the standard deviation (σ) using a multichannel approach and subsequently applied Python's `bm3d` function with the estimated σ and `BM3DStages.ALL STAGES` configuration. This configuration, which includes Wiener filtering in addition to hard thresholding, provides slightly improved performance for our system-specific noise compared to other `BM3DStages`.

2.3 Deep Learning Architectures

2.3.1 U-Net

Within our implementation, we employed two DL networks, one of which was the original U-Net architecture. U-Net, a convolutional neural network (CNN)-based architecture, follows an encoder-decoder structure, where the encoder (contracting path) captures context and reduces spatial dimensions through successive layers of convolution, ReLU activation, and max-pooling operations.⁶³ This is followed by a decoder (expanding path) that uses up-convolutions (transposed convolutions) to restore spatial resolution. At each step of up-sampling, skip connections from the corresponding layers of the encoder are concatenated to the decoder, enabling the model to retain fine-grained information from earlier stages. These skip connections help U-Net preserve spatial details while improving localization, making it highly effective for tasks like biomedical imaging. The network's symmetry and ability to work with limited training data are key strengths. We utilized the "mean absolute error" loss function and the "Adam" optimizer with an initial learning rate of 1×10^{-4} . The U-Net is distinguished by its inclusion of residual skip connections, which contribute to the network's notable effectiveness. For a more comprehensive understanding of the U-Net's impact on PA imaging, please refer to Ref. 64.

2.3.2 Conditional GAN

In our implementation, we devised a custom loss function based Pix2Pix network (cGAN denoising model); the process schematic is illustrated in Fig. 1. The general workflow of our cGAN training with AcousticX LED data and testing with AR-PAM data is shown in the figure. G and D showcase the architecture of the generator and the PatchGAN discriminator, respectively. The extended inset portrays the functioning of the cGAN with our custom loss function.

Generator. We employed a U-Net-based model as the generator. The down-sampling and up-sampling operations are accompanied by corresponding residual connections which help recover local information lost during the down-sampling process. Each block in the down-

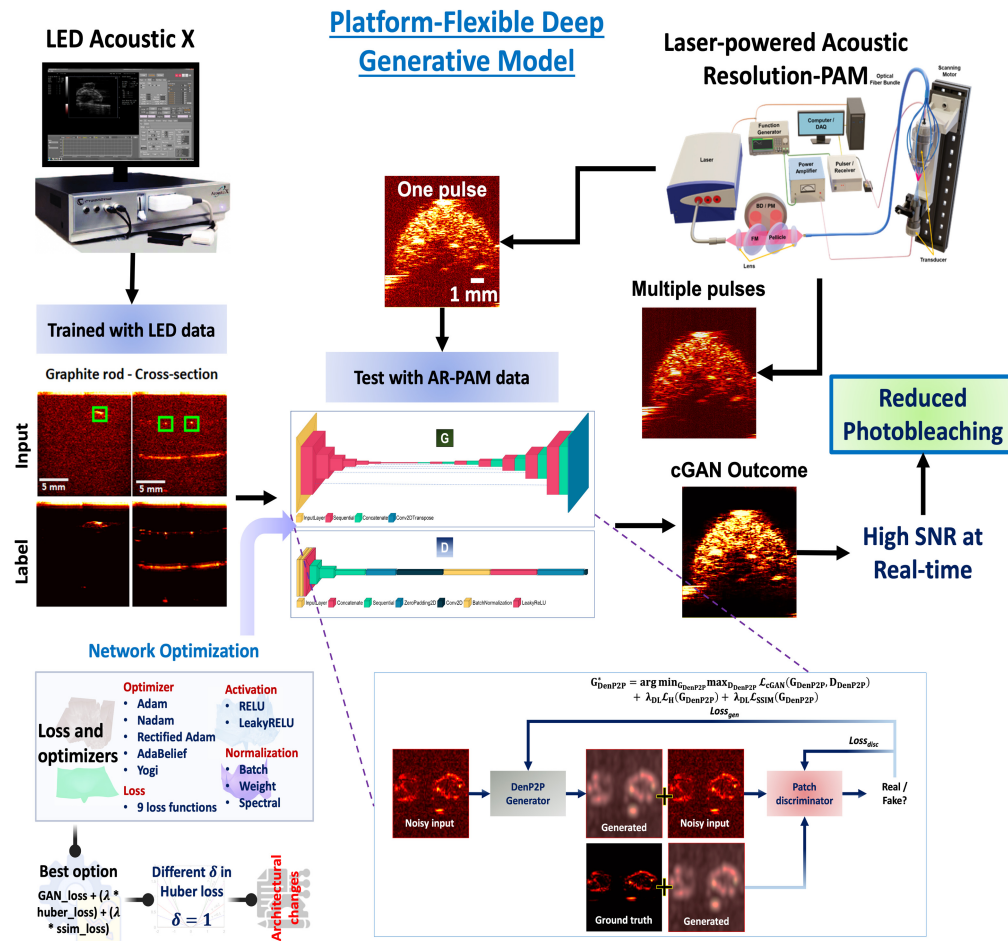


Fig. 1 Platform-flexible cGAN workflow where training was done with LED-illuminated PAI data and tested with AR-PAM data. G, generator; D, PatchGAN discriminator where layer depictions were generated by the visuallkeras library. The network optimization modules are shown at the left bottom portion where we implemented several loss functions, optimizers, activation layer, and architectural changes.

sampling pathway consists of a series of convolution, batch normalization, and rectified linear unit (RELU) activation layers. The first block does not include a batch normalization layer, and the first three blocks do not have a dropout layer. Similarly, each block in the decoder section is comprised of a sequence of transposed convolution, batch normalization, and RELU layers, with the last three blocks excluding a dropout layer.

All dropout layers have a probability of 0.5 for discarding the random nodes. In addition, the weight initialization for all layers was performed using “*he_normal*,” which draws weights from a truncated normal distribution. This initialization strategy works well with RELU activation due to its controlled weight initialization, facilitating the rapid convergence to a global minimum of the loss function.

Discriminator. We employed a convolutional PatchGAN classifier in our implementation. This classifier focuses on classifying a 70×70 portion of the input image based on a 30×30 output image patch. Each block of the discriminator consists of a convolution, batch normalization, and ReLU layer. However, the first block does not include a normalization layer and the last layer has a leakyReLU activation instead of a regular ReLU activation for better gradient flow, avoiding dead neurons, and stabilizing the training while effectively capturing local texture details.

Generator and discriminator objectives. The cGAN computational architecture attempts to find a mapping from the observed image (noisy input) \mathbf{x} and random noise vector

\mathbf{z} to label (ground truth) \mathbf{y} , $G: \{\mathbf{x}, \mathbf{z}\} \rightarrow \mathbf{y}$. The generator of the cGAN learns a structured loss by minimizing the difference between the network outputs and the corresponding targeted images. The cost function of a cGAN is normally expressed as $\mathcal{L}_{\text{cGAN}}(G, D) = \mathbb{E}_{\mathbf{x}, \mathbf{y}}[\log D(\mathbf{x}, \mathbf{y})] + \mathbb{E}_{\mathbf{x}, \mathbf{z}}[\log(1 - D(\mathbf{x}, G(\mathbf{x}, \mathbf{z})))]$, which represents a two-player game between generator G and discriminator D where G tries to maximize the objective against the antagonist D . Hence, the generator's objective function is expressed as $G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D)$. Keeping the loss function of the discriminator unchanged from the Pix2Pix model, we used Huber loss, (instead of $L1$, mean absolute error, loss) and an additional loss, structural similarity (SSIM) loss for the generator. As a result, G attempts to fool D as it tries to produce images that are not distant from the real data following the Huber loss, which considers both L_1 and L_2 norm based on a δ value and are structurally similar to the ground truth based on the SSIM loss. The Huber loss^{65,66} is defined as

$$\mathcal{L}_h(G) = \begin{cases} \frac{1}{2}(y - G(\mathbf{x}, \mathbf{z}))^2 & |y - G(\mathbf{x}, \mathbf{z})| \leq \delta \\ \delta \cdot [\|y - G(\mathbf{x}, \mathbf{z})\|_1 - \frac{1}{2}\delta] & \text{otherwise} \end{cases}$$

and SSIM loss is defined as $\mathcal{L}_{\text{SSIM}}(G) = \frac{1}{N} \sum 1 - \text{SSIM}(\mathbf{y}, G(\mathbf{x}, \mathbf{z}))$ where $\text{SSIM}(A, B)$ is defined in Ref. 67. These additional losses help the cGAN resist the vanishing gradient problem because the loss metrics can measure how far the generated data distribution is from real ground truth, and our generator learns from discriminator losses. SSIM loss⁶⁸ computes localized statistics (mean, variance, etc.) in image patches, ensuring nonzero gradients even in flat regions where pixel-wise losses such as MSE may fail. By emphasizing structural patterns and textures rather than individual pixel intensities, SSIM loss maintains gradient flow and improves the network's ability to capture high-level image features. The combined generator loss functions give us the final objective function of G as follows:

$$G^* = \arg \min_G \max_D \mathcal{L}_{\text{cGAN}}(G, D) + \lambda_{DL} \mathcal{L}_h(G) + \lambda_{DL} \mathcal{L}_{\text{SSIM}}(G),$$

where λ_{DL} is the regularization parameter with a value that was chosen after multiple trials.

The Adam optimizer^{69,70} was utilized for both the generator and discriminator, with an initial learning rate of 0.0001. The parameters β_1 and β_2 were set to 0.5 and 0.999, respectively, and ϵ was set to 10^{-7} . The parameters ϵ , β_1 , and β_2 in the Adam optimizer play distinct roles in controlling how the algorithm updates weights during training. ϵ is a small constant added to the denominator for numerical stability in the learning process, particularly when gradients are very small. β_1 is the exponential decay rate for the first moment estimate (mean of the gradients), which controls how quickly the moving average of the gradients adapts to new gradient information. β_2 is the exponential decay rate for the second moment estimate (variance of the gradients), which controls how the moving average of the squared gradients adapts. During training, a batch size of 1 was employed and the process consisted of 40,000 steps that allowed the network to focus on the unique structural and pixel-level details of each image pair, enhancing fine-grained outputs (instance normalization).^{71,72} This approach is crucial for preserving subtle features in high-resolution tasks. Although larger batch sizes improve convergence, they average gradients across samples, potentially diluting unique contributions and reducing training diversity. Moreover, smaller batch size introduces more stochasticity, helping the discriminator generalize better, avoid overfitting, and stabilize optimization by escaping poor local minima. Within the combined generator objective function, the regularization parameter λ_{DL} was set to 100. In addition, the δ parameter for the Huber loss was set to 1. After systematically evaluating a range of generator loss functions (combinations of $L1$, $L2$, Huber, SSIM, and PSNR losses) and optimization methods (Adam, Nadam, Rectified Adam, AdaBelief, and Yogi),⁷³ Adam emerged as the preferred optimizer when combined with the original GAN, Huber, and SSIM losses, due to its shorter training time (93 s per 1000 steps) and robust generalization, despite no statistically significant differences in image quality metrics across optimizers. In addition, we optimized the delta (δ) value for the Huber loss function, identifying $\delta = 1$ as the optimal setting. Further experimentation with activation functions (ReLU and Leaky ReLU) and normalization layers (batch, weight, and spectral)⁷⁴ led to the selection of ReLU activation and batch normalization as the most effective combination (see bottom left part of Fig. 1). The detailed optimization and tuning of the hyperparameters (nine different loss functions and delta parameter of huber loss) and

architectural modifications (activation layer and normalization strategies) are provided in Tables S1, S2, and S3 in the [Supplementary Material](#).

All DL codes were implemented using TensorFlow (version 2.9.2) with the Keras backend and executed on the Google Colab Pro platform. The computations were performed on a Tesla T4 GPU (CUDA version: 11.2) with 25.45 GB of RAM. The training process took ~93 to 94 s per 1000 steps.

2.4 Database

2.4.1 Training data

Frame averages impact image quality: fewer averages yield low-SNR, noisy images, whereas higher averages provide high-SNR images. Using the LED-based AcousticX PA system, we captured 800 paired images: low-SNR inputs (128 frame averages) and high-SNR labels (25,600 frame averages). There was no spatial misalignment between inputs and labels as the transducer remained stationary, and the noise was system-generated (e.g., shot and thermal noise). The training data came from a different imaging platform, handling macro-scale objects, unlike the AR-PAM system, which focuses on microscopic data.

2.4.2 Test data

We conducted test data collection using our in-house built AR-PAM system. For *in vitro* and *ex vivo* imaging samples, we utilized carbon fiber, mouse kidney, and mouse tumor specimens. To investigate the photobleaching phenomenon, we conducted an experiment using a polyethylene tube filled with a mixture of 0.5- μ M ICG and deionized water, which was securely positioned at both ends of a box. To assess the plausibility of inducing photobleaching of ICG by laser pulses, we illuminated the tube at a single point for 15 min in three separate trials using both deionized water and dimethyl sulfoxide. In all cases, we observed a consistent decrease in the PA amplitude over time, indicating the occurrence of photobleaching. Subsequently, we performed a 2D scanning of the tube in the $x-z$ direction with a step size of 100 μ m. Two laser pulse illuminations were employed: single-pulse illumination (no averaging) and 30-pulse illumination.

2.5 Image Quality Metrics

To check the quality of the cGAN-generated output images, we used three important reference image quality metrics.

2.5.1 SNR

We used SNR as a no-reference quality metric⁷⁵ measured in arbitrary units to accommodate signals with a wide dynamic range. It is defined as $SNR = \frac{\mu_{ROI}}{\sigma_{BG}}$, where μ_{ROI} is the mean signal amplitude of the target region of interest (ROI), and σ_{BG} is the standard deviation of the background noisy region.

2.5.2 CNR

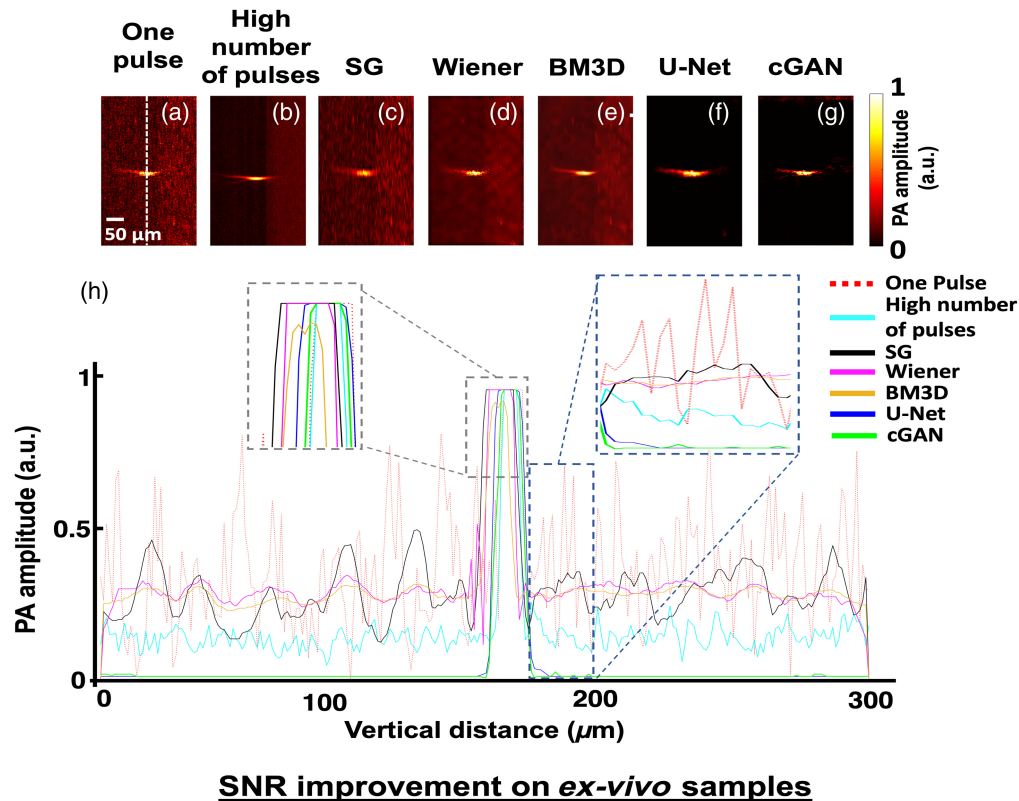
Contrast to noise ratio (CNR)⁷⁶ is a no-reference image quality metric, which compares the standard deviation of the target ROI with respect to the background noisy region. It is defined as $CNR = \frac{\sigma_{ROI}}{\sigma_{BG}}$, where σ_{ROI} is the standard deviation of the target ROI and σ_{BG} is the standard deviation of the background noisy region.

2.5.3 GCNR

Generalized CNR^{77,78} is a quantitative measure of how distinguishable two ROIs are in an image, accounting for overlap in their intensity distributions by combining the effect of contrast and noise. Unlike metrics that rely purely on absolute intensity differences, GCNR evaluates the separability of the distributions, making it robust to noise and intensity variations. GCNR is computed as $GCNR = 1 - \int \min(p_1(I), p_2(I))dI$, where $p_1(I)$ and $p_2(I)$ are the probability density functions of the pixel intensities in the two ROIs, and I is the intensity value.

3 Results and Discussion

To demonstrate the efficacy of our DL model for improving the SNR and CNR, we imaged a carbon fiber with two laser pulse illuminations: one pulse (input) and 20 pulse averaging (ground truth). In Fig. 2, cross-section images of the carbon fiber are shown for one pulse [Fig. 2(a)] and average of 20 pulses [Fig. 2(b)]. Noise removal filters were applied to the one pulse images using traditional nonlearning algorithms [SG: Fig. 2(c), Wiener: Fig. 2(d), and BM3D: Fig. 2(e)], U-Net [Fig. 2(f)] and our Pix2Pix-based cGAN denoising model [Fig. 2(g)]. Figure 2(h) shows the line profile of the one pulse illumination (red line), a high number of pulses illumination (cyan), SG noise removal (black), Wiener filter (magenta), BM3D (yellow), U-Net image (blue), and our cGAN (green).



SNR improvement on *ex-vivo* samples

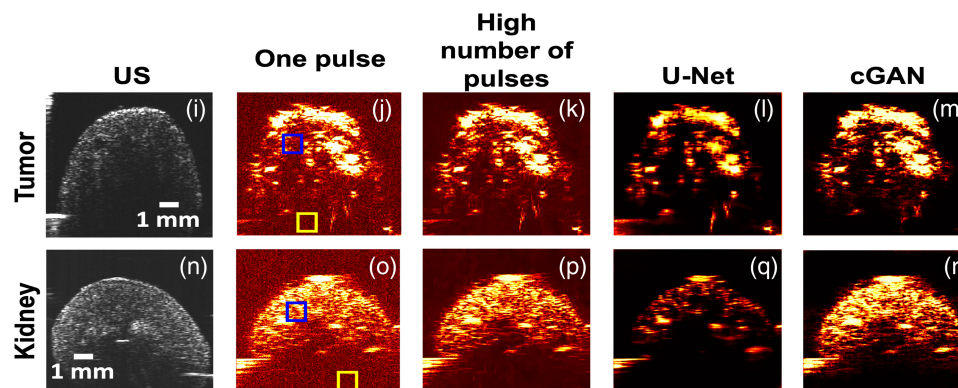


Fig. 2 (a)–(g) Comparison between nonlearning traditional denoising algorithms with our cGAN DL model. The bottom graph (h) shows the PA amplitude along the axial direction depicted by a white dotted line for one pulse illumination (red), high number of pulses (cyan), SG (black), Wiener (magenta), BM3D (yellow), U-Net (blue), and our cGAN (green). The left inset is the enlarged version of the peak signal and the right one shows the zoomed-in version of the background noise. (i)–(r) US and PA images for *ex vivo* samples produced with one laser pulse, average of 20 laser pulses, U-Net, and cGAN methods.

and the cGAN generated image (green) for the carbon fiber along the axial direction. The left inset in Fig. 2(h) shows the enlarged display of the PA peak signal amplitude and the right enlarged inset denotes the background region profile. For *ex vivo* samples, we utilized mouse kidney and tumor data as shown in Figs. 2(i)–2(m) and Figs. 2(n)–2(r), respectively. The ultrasound (US) images accompanying each row provide morphological confirmation of the biological structures. For quantitative analysis, the blue square was selected as the ROI to calculate image quality metrics such as SNR and CNR. The yellow box served as the background to assess how well each method suppressed noise in non-ROI areas. The image quality metrics for *ex vivo* samples are depicted in Table 1.

Traditional nonlearning filters demonstrated a moderate ability to reduce noise, but they fell short in terms of preserving structural details. The SG filter marginally smoothed the background noise but oversimplified the sharp features, whereas the Wiener and BM3D filters introduced varying degrees of blurring. By contrast, the DL-based approaches, especially the Pix2Pix-based cGAN model, achieved superior performance by preserving fine details and maintaining the integrity of the axial line profiles. This can be seen in Fig. 2(h), where the cGAN's line profile closely follows that of the images acquired with a high number of pulses across both the PA peak signal region and background areas. The traditional noise filters' line profiles are less noisy than the single-pulse data but significantly noisy in the background regions compared with the DL methods. The *ex vivo* results reinforced the carbon fiber experiment's observations where our cGAN model (fourth column) outperforms the U-Net in preserving the detailed structure of the tumor while effectively denoising the images, resulting in higher-quality images. Multipulse illumination significantly improved image clarity but required longer acquisition times, which might lead to undesirable effects of photobleaching and phototoxicity. Both DL models performed well in reducing noise, but the U-Net's tendency to blur finer details led to the loss of critical anatomical information. This was evident in the kidney and tumor's more delicate structures, where the Pix2Pix-based cGAN outperformed the U-Net by offering clearer, more detailed visualizations without sacrificing noise suppression, as also evident from Table 1. Our study presents a

Table 1 Image quality metrics for different traditional and DL denoising methods.

Imaging process	SNR	CNR	GCNR
<i>Ex vivo</i> sample: Tumor			
One pulse	73.75 ± 6.32	4.12 ± 2.16	0.734 ± 0.016
High number of pulses	79.81 ± 3.61	7.68 ± 1.29	0.981 ± 0.014
SG	77.01 ± 6.17	5.43 ± 1.47	0.787 ± 0.021
Wiener	75.4 ± 4.74	5.62 ± 2.54	0.791 ± 0.015
BM3D	77.93 ± 8.37	6.32 ± 2.88	0.834 ± 0.014
U-Net	84.52 ± 11.82	6.82 ± 2.12	0.861 ± 0.013
Pix2Pix cGAN	92.77 ± 10.74	9.9 ± 4.41	0.979 ± 0.011
<i>Ex vivo</i> sample: Kidney			
One pulse	72.76 ± 9.25	6.19 ± 4.36	0.739 ± 0.018
High number of pulses	82.11 ± 4.19	7.53 ± 1.96	0.987 ± 0.019
SG	78.09 ± 3.62	7.29 ± 3.79	0.807 ± 0.019
Wiener	81.34 ± 6.49	6.41 ± 2.46	0.793 ± 0.014
BM3D	78.81 ± 7.89	6.98 ± 3.44	0.814 ± 0.012
U-Net	86.35 ± 3.97	7.59 ± 0.82	0.851 ± 0.015
Pix2Pix cGAN	93.54 ± 6.07	11.82 ± 4.42	0.983 ± 0.012

Each of the corresponding table entries is depicted as mean ± standard deviation.

real-time, single-pulse denoising framework designed for applications such as PDT and *in vivo* imaging, where multiframe averaging is impractical due to motion artifacts, increased laser exposure, and photobleaching risks. Although averaging improves GCNR by reducing contrast fluctuations, it is not feasible in real-time settings. Our method enhances image quality at the single-pulse level, achieving significant improvements in SNR and CNR, which are also clinically relevant for preserving fine structural details. Unlike GCNR, which focuses on separability, SNR and CNR impact the visibility of fine structures, making them useful for clinical use. In addition, PA image noise is spatially and spectrally nonuniform, making GCNR sensitive to residual structured noise. Our spatially adaptive noise suppression balances contrast enhancement with feature preservation, ensuring superior perceptual image quality. Unlike post-processing techniques that enhance global contrast at the cost of fine-edge details, our method retains crucial structures essential for clinical interpretation.

To substantiate the claim of platform flexibility with respect to noise distribution across different imaging platforms, we first characterized the noise profiles (statistical and spectral properties of noise^{79–81}) of both LED and AR-PAM systems using noise power spectrum and correlation analysis for spatial noise statistics. Figure 3(a) depicts the power spectrum analysis for LED and AR-PAM inputs, which is calculated as the log-transformed of the squared magnitude of the FFT so that the dynamic range is compressed. The spectra are then normalized to their maximum values for comparison. The inset figures [Figs. 3(a,i,ii)] show the spatial frequency content of the noise for the two imaging platforms, highlighting differences between them. The inset in Fig. 3(a,iii) analyzes the cross-power spectrum to study relationships between the noise profiles, which identify frequency regions where the noise profiles are similar or different. Finally, we quantify the similarity between two systems' noise profile distributions by calculating correlation coefficient^{82–84} between the two power spectra flattened to one dimension [Fig. 3(b)]. The LED spectrum has a higher amplitude at low frequencies than AR-PAM, whereas some of the higher frequency zones of AR-PAM data has higher amplitude compared with LED-based outcomes. The cross-power spectrum shows a notable difference in the low-frequency zones. Different amplitude values at the same frequency locations in the power spectra for LED and AR-PAM data suggest that the noise profiles or signal intensities vary between the two platforms. The correlation analysis between the two spectra is significantly <0.8 , which also indicates high dissimilarity in their noise profiles.

For further validation on flexibility, our next approach was to synthetically inject Gaussian ($\sigma = 0.1$),^{85–88} Poisson,^{86,88,89} salt and pepper (S&P) with 10% for pixel destruction^{90,91} and speckle noise (uniform distribution having zero mean and 0.1 as variance)^{86,88,92} into the AR-PAM high number of pulses averaged data to simulate diverse noise conditions. Visual validation from Figs. 4(a)–4(h) corroborates that our DL model consistently achieved high denoising performance, indicating its robustness to noise variations. The three violin plots in Fig. 4(i)–4(k) depicting image quality metrics (SNR, CNR, and GCNR) analysis across datasets with diverse noise profiles evidently suggest our model's noise-invariance capabilities. In addition, we implemented several Gaussian noise profiles with different variances [Figs. S1(a)–S1(f) in the [Supplementary Material](#)] as mentioned in the top right corner of the corresponding input images. We provided a visual depiction of our model's outcomes in the second row [Fig. S1(g–l) in the [Supplementary Material](#)] demonstrating how our model performs under increasing levels of injected noise. This visually confirms the model's robustness up to a certain noise threshold. We also reported quantitative image quality metrics under varying noise levels in the third row [Figs. S1(m)–S1(o) in the [Supplementary Material](#)], showing how performance degrades as noise variance increases. Overall, our results confirm that although extreme noise levels do affect the model, they maintain satisfactory performance within practical noise limits encountered in PAI. Future work will focus on validating the model's noise invariance across additional platforms (e.g., MSOT or handheld devices) and exploring domain adaptation techniques to further enhance its generalizability to systems with varying noise characteristics.

To observe the photobleaching effect of contrast agents used for PAI, we initially illuminated the tube filled with ICG dissolved in deionized water at a specific scanning location for three trials, each for 15 min. The PA signal amplitude consistently reduced by $\sim 35\%$ across all trials, pointing to significant photobleaching. The uniformity of this reduction strongly suggests that water-based ICG solutions are prone to photobleaching under prolonged laser exposure, which

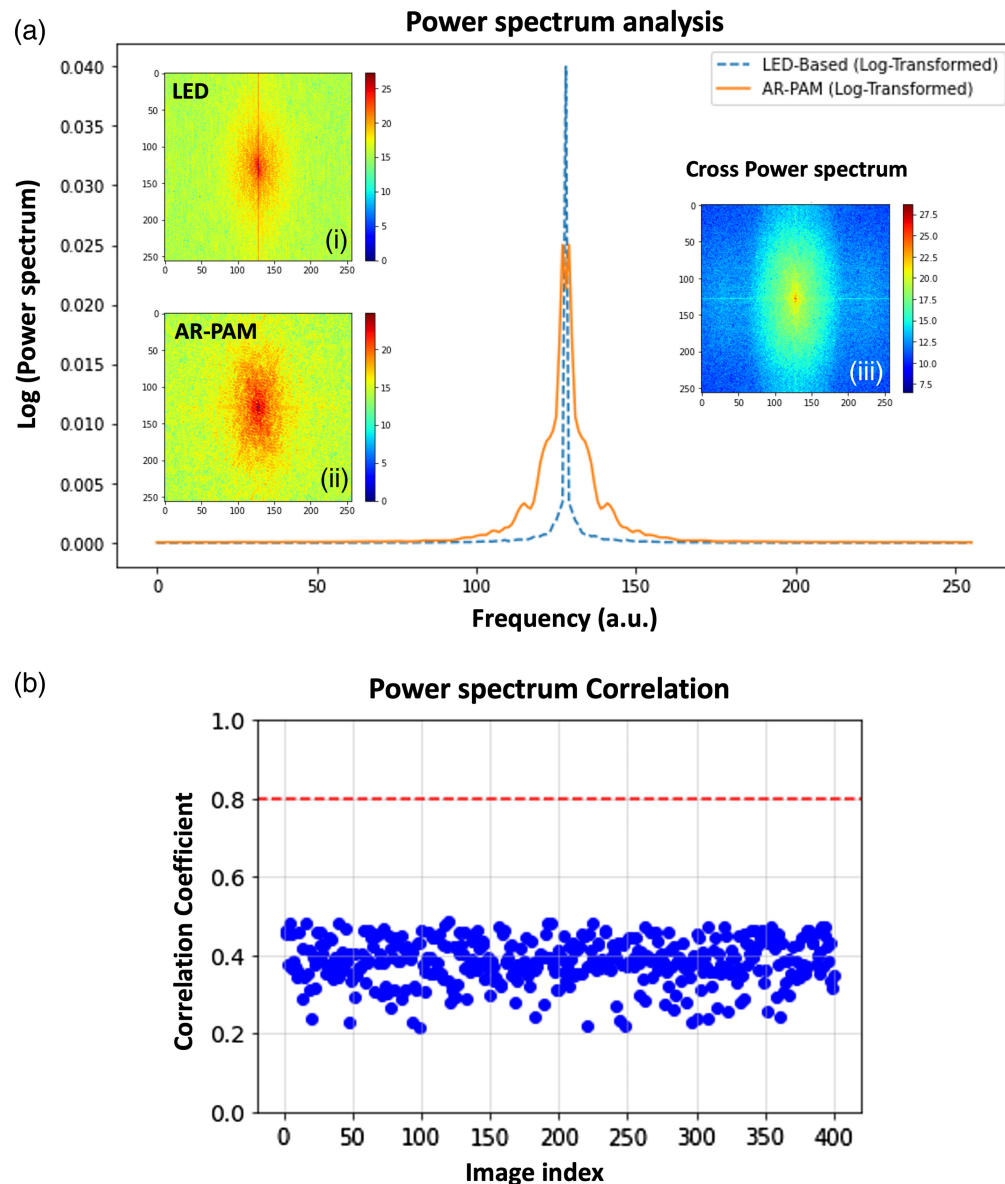


Fig. 3 Characterization of noise profiles across LED and AR-PAM imaging platforms. (a) Normalized power spectrum analysis for LED and AR-PAM systems, calculated as the log-transformed squared magnitude of the FFT to compress the dynamic range. Insets (i) and (ii) show the spatial frequency content of the noise for the two platforms, highlighting differences in noise profiles. Inset (iii) presents the cross-power spectrum, identifying frequency regions where the noise profiles exhibit similarities or differences. (b) The cross-power spectrum highlights notable differences, particularly in the low-frequency regions. Correlation analysis of the flattened power spectra reveals a low correlation coefficient (<0.8), confirming significant dissimilarity in the noise profiles.

can compromise the accuracy and consistency of PA imaging. The fact that photobleaching occurs uniformly indicates that the reduction in signal is not localized but rather affects the entire imaging area. This can lead to a decrease in sensitivity and accuracy in detecting and quantifying contrast-agent-labeled structures, especially in applications such as longitudinal studies where signal consistency is paramount. To visually demonstrate these effects, a 2D raster scan of tubes filled with ICG and deionized water with a step size of $100\ \mu\text{m}$ is shown in Fig. 5. Figure 5(a) illustrates the calculated percentage of photobleaching or photodamage of ICG within the tube at four distinct cross-sections spaced $0.7\ \text{mm}$ apart. The percentage of photobleaching was determined by comparing the PA amplitude values at corresponding locations to the initial starting position. Note that the initial starting position was at $0.1\ \text{mm}$ and therefore shows no

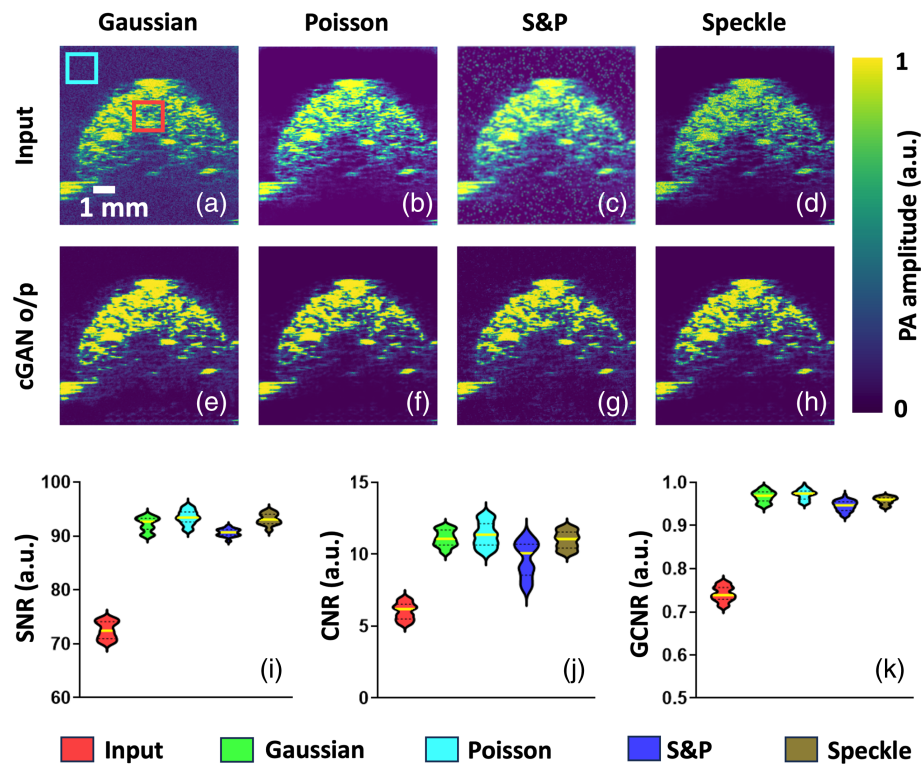


Fig. 4 Validation of the model's flexibility and robustness to diverse noise conditions. (a)–(h) Denoising performance on AR-PAM high-pulse-averaged data with synthetically injected Gaussian, Poisson, S&P, and speckle noise. (i)–(k) Violin plots of image quality metrics (e.g., SNR, CNR, and GCNR) across datasets with varying noise profiles, showcasing the model's noise-invariance capabilities. The two sample regions are denoted by red and cyan colored boxes where one has PA signal contents and another one does not have them.

photobleaching. Thirty-pulse illumination imaging at 2.25-mm scanning distance shows significant photobleaching % ($35.14 \pm 5.38\%$) compared with single pulse imaging ($9.58 \pm 3.74\%$) and our cGAN derived imaging ($9.51 \pm 3.69\%$). Although the current results are derived from retrospective data analysis, the computational speed of the cGAN model (<27 ms per frame) demonstrates its potential for real-time applications as the use of a single-pulse rapid imaging technique minimizes the need for long exposure inherently reducing photobleaching. To further validate this, we plan to conduct experiments integrating the model with a live imaging setup in future work.

Figure 5(b) showcases volumetric images obtained using one pulse and 30 laser pulse illumination. Individual B-scans were obtained in the Y direction, orthogonal to the tube's diameter. A $2 \times 3 \text{ mm}^2$ area of the tube was scanned, consisting of 30 cross-sections spaced $100 \mu\text{m}$ apart. Figure 5(c) presents three of those cross-sections at different locations (start, middle, and end). Each row corresponds to a specific scanning location, and the columns display the outcomes for one-pulse illumination, 30-pulse illumination, and the resulting cGAN denoising. The 2D raster scan experiment further reinforces these findings, with Fig. 5(a) illustrating the photobleaching progression at different cross-sections of the ICG-filled tube. The consistent photobleaching across the tube sections, apart from the initial starting point, shows that even slight shifts in imaging location result in notable reductions in signal amplitude. The uniform step size ($100 \mu\text{m}$) of the scan ensures that the results are comparable across sections, providing a reliable measure of the extent of photobleaching. Interestingly, the results suggest that ICG-water solutions are prone to photodamage even with relatively short intervals between cross-sectional scans. Although the volumetric scans and cross-sectional comparisons in Figs. 5(b) and 5(c) provide deeper insights into how photobleaching impacts 3D PA imaging, they also underscore the limitations posed by multipulse imaging in this context. In the one-pulse illumination, the PA signal was weaker but exhibited less cumulative photobleaching, preserving more structural information. However, in the 30-pulse illumination, although the signal was initially stronger, the cumulative effect of

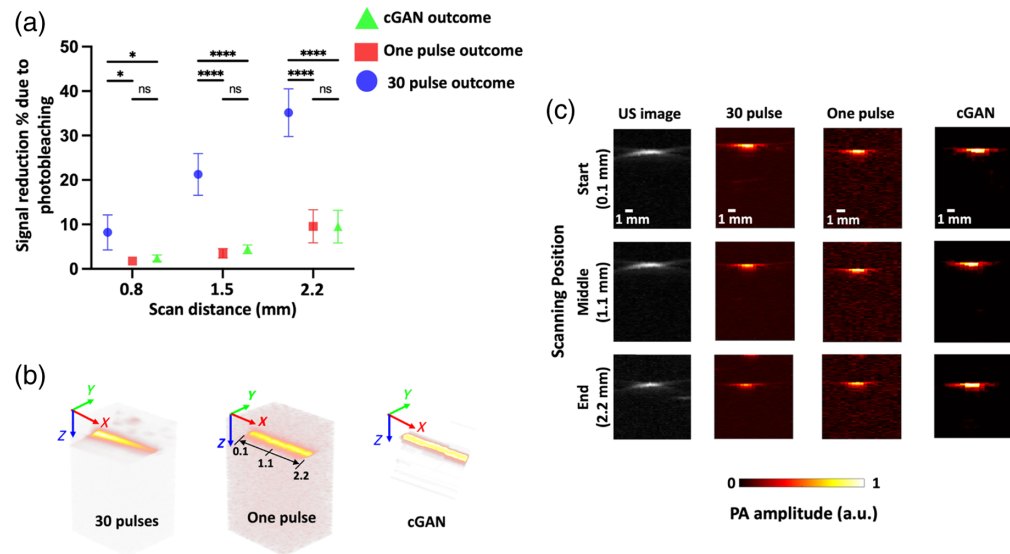


Fig. 5 Comparison of photobleaching experiment of ICG. (a) Percentage of ICG photobleaching at different scanning positions within the tube for one pulse, high number of pulses (30), and cGAN denoising image obtained with single pulse illumination. Photobleaching is considerably more pronounced at the tube's end position compared to the initial scanning position. We performed a two-way ANOVA to compare the results within each method across the different scanning distances. The analysis reveals that photobleaching becomes more pronounced as the scanning distance increases. Comparing the different imaging procedures (single pulse, 30-pulse illumination, and cGAN) across all scanning distances (especially 1.8 and 2.2 mm), we found that the 30-pulse illumination caused significantly more photobleaching than both the single pulse and cGAN methods. ns, not significant, $*p < 0.05$, $****p < 0.0001$. (b) Volumetric 3D images of the tube showing a comparison between one-pulse and 30-pulse illuminated PA images. (c) Comparison of PA amplitude at three scan positions for one pulse, high number of pulses (30), and cGAN denoising. The left most column depicts the ultrasound images of the corresponding scanning positions of the tube.

multiple laser pulses significantly exacerbated photobleaching, further diminishing the signal quality over time. This finding poses a challenge for multipulse imaging strategies: although they may initially offer higher SNR and better contrast, they can induce more pronounced photobleaching, leading to progressively poorer image quality, particularly in water-based ICG solutions.

Our denoising Pix2Pix-based cGAN DL model applied to one-pulse illumination images successfully generated a high SNR image without background noise [Fig. 3(c) third column]. This approach significantly reduced the scanning time by ~ 25 to 30 times when compared with scans with averaging. Consequently, the impact of photobleaching was mitigated, and we obtained an image with satisfactory SNR and CNR. Although endogenous contrast agents are typically not affected by optical bleaching, exogenous contrast agents are prone to it over time. In future work, it will be important to explore strategies to mitigate photobleaching during PA imaging. These might include the development of more photostable contrast agents, optimizing laser parameters to minimize photodamage, or using adaptive imaging strategies that minimize exposure without compromising image quality. In addition, further research into the ability of DL models to correct photobleaching artifacts and their limitations in restoring true quantitative signal strength could offer pathways to improve post-processing techniques.

Most studies using DL for denoising PA images rely on numerically simulated data or *in vitro* phantoms, with limited application to *in vivo* datasets. When *in vivo* data are used, training and testing datasets often come from similar samples and the same imaging hardware, limiting generalizability and clinical relevance. In our study, training was conducted exclusively on metal wire phantom images from an LED-based AcousticX system with a 7 MHz transducer, whereas testing used AR-PAM images with a 25-MHz transducer. This introduced morphological and noise-profile differences between the datasets, providing a stringent test of the model's generalization ability. By training on one platform and testing on another, we demonstrate that

the DL model is not overfitted to the training platform's characteristics. Instead, it generalizes across systems, accommodating diverse noise profiles, resolutions, and hardware configurations. This cross-platform validation highlights the model's robustness and versatility, enabling its use across various clinical and research settings without retraining. Such flexibility enhances the utility of PA imaging for diverse experimental and biological applications, saving time and computational resources while maintaining high performance.

4 Conclusion

Photobleaching poses significant challenges in molecular PAI using exogenous dyes. Prolonged laser exposure degrades signal quality, limiting imaging duration and complicating quantitative analyses. Addressing these limitations, we developed a platform-flexible generative DL framework that leverages a custom loss function-based cGAN to enhance SNR in single-pulse acquired AR-PAM data after being trained with LED-illuminated images, effectively mitigating photobleaching risks. Our model is validated across two PA imaging systems, ensuring platform flexibility, where the model has adaptability to varying noise conditions while preserving critical image features. Unlike platform-independent models, our approach is optimized for specific systems while retaining the potential for broader applicability with minimal adaptation. The primary focus of this study was to develop and benchmark a computational framework for image denoising and signal enhancement using well-controlled *in vitro* and *ex vivo* datasets. These datasets, with their precise and reproducible conditions, were critical for establishing the reliability and efficacy of the proposed method, allowing for controlled experimentation and clear benchmarks for validation. We will integrate the proposed framework into real-time *in vivo* imaging workflows, addressing challenges such as physiological and motion-related artifacts that may impact performance evaluation. Our focus will be on validating the method under complex, dynamic conditions to further assess its adaptability and clinical utility. Other future work involves improving the image quality by integrating multimodal DL models combining PA and US data to fuse complementary information such as providing better anatomical context. We will look for self-supervised, few-shot or zero-shot DL techniques for better generalization with minimal labeled data.

Disclosures

The authors declare no conflicts of interest.

Code and Data Availability

Materials are available upon request from the authors.

Acknowledgments

The authors would like to acknowledge support from Tufts School of Engineering, Tufts Data Intensive Science Center Pilot grant, and Dr. Tayyaba Hasan for subcontract funds on a grant from the National Institutes of Health (NIH Grant No. 5R01CA231606). The authors would also like to acknowledge Dr. Marvin Xavierselvan for help with tumor implantation and Ms. Allison Sweeney for handling animal care.

References

1. R. Bonnett and G. Martinez, "Photobleaching of sensitizers used in photodynamic therapy," *Tetrahedron* **57**(47), 9513–9547 (2001).
2. A. Diaspro et al., "Photobleaching," in *Handbook of Biological Confocal Microscopy*, J. Pawley, Ed., pp. 690–702, Springer, Boston, MA (2006).
3. L. Song et al., "Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy," *Biophys. J.* **68**(6), 2588–2600 (1995).
4. M. Caroline Müllenbroich et al., "Strategies to overcome photobleaching in algorithm-based adaptive optics for nonlinear in-vivo imaging," *J. Biomed. Opt.* **19**(1), 016021 (2014).
5. A. P. Demchenko, "Photobleaching of organic fluorophores: quantitative characterization, mechanisms, protection," *Methods Appl. Fluoresc.* **8**(2), 022001 (2020).
6. N. Lorén et al., "Fluorescence recovery after photobleaching in material and life sciences: putting theory into practice," *Quart. Rev. Biophys.* **48**(3), 323–387 (2015).

7. X. Rong et al., "Near-infrared and highly photostable squaraine-based nanoparticles for photoacoustic imaging guided photothermal therapy," *Dyes Pigments* **211**, 111055 (2023).
8. M. A. Saad et al., "Optimizing axial and peripheral substitutions in Si-centered naphthalocyanine dyes for enhancing aqueous solubility and photoacoustic signal intensity," *Int. J. Mol. Sci.* **24**(3), 2241 (2023).
9. K. Yan et al., "Ultra-photostable small-molecule dyes facilitate near-infrared biophotonics," *Nat. Commun.* **15**(1), 2593 (2024).
10. J. Yu et al., "Photostable, hydrophilic, and near infrared quaterrylene-based dyes for photoacoustic imaging," *Mater. Sci. Eng.: C* **93**, 1012–1019 (2018).
11. L. Gao et al., "Photothermal bleaching in time-lapse photoacoustic microscopy," *J. Biophotonics* **6**(6–7), 543–548 (2013).
12. P. Van den Berg, K. Daoudi, and W. Steenbergen, "Review of photoacoustic flow imaging: its current state and its promises," *Photoacoustics* **3**(3), 89–99 (2015).
13. J. Weber, P. C. Beard, and S. E. Bohndiek, "Contrast agents for molecular photoacoustic imaging," *Nat. Methods* **13**(8), 639–650 (2016).
14. X. Zhang et al., "Optical and photoacoustic imaging in vivo: opportunities and challenges," *Chem. Biomed. Imaging* **1**(2), 99–109 (2023).
15. M. Xu and L. V. Wang, "Photoacoustic imaging in biomedicine," *Rev. Sci. Instrum.* **77**(4), 041101 (2006).
16. M. J. Pushie et al., "X-ray fluorescence microscopy methods for biological tissues," *Metalomics* **14**(6), mfac032 (2022).
17. E. J. van Beek and E. A. Hoffman, "Functional imaging: CT and MRI," *Clin. Chest Med.* **29**(1), 195–216 (2008).
18. R. Bar-Shalom, A. Y. Valdivia, and M. D. Blafox, "PET imaging in oncology," *Semin. Nucl. Med.* **30**(3), 150–185 (2000).
19. J. Olsen, J. Holmes, and G. B. Jemec, "Advances in optical coherence tomography in dermatology—a review," *J. Biomed. Opt.* **23**(4), 040901 (2018).
20. L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *Science* **335**(6075), 1458–1462 (2012).
21. S. Park et al., "Acoustic resolution photoacoustic microscopy," *Biomed. Eng. Lett.* **4**(3), 213–222 (2014).
22. S. Hu and L. V. Wang, "Optical-resolution photoacoustic microscopy: auscultation of biological systems at the cellular level," *Biophys. J.* **105**(4), 841–847 (2013).
23. J. Shi, Y. Tang, and J. Yao, "Advances in super-resolution photoacoustic imaging," *Quant. Imaging Med. Surg.* **8**(8), 724 (2018).
24. R. Shintate et al., "Development of optical resolution photoacoustic microscopy with sub-micron lateral resolution for visualization of cells and their structures," *Jpn. J. Appl. Phys.* **59**(SK), SKKE11 (2020).
25. J. Liu et al., "A novel parameter-tuned stochastic resonator for binary PAM signal processing at low SNR," *IEEE Commun. Lett.* **18**(3), 427–430 (2014).
26. E. Cassinotti, L. Boni, and L. Baldari, "Application of indocyanine green (ICG)-guided surgery in clinical practice: lesson to learn from other organs—an overview on clinical applications and future perspectives," *Updates Surg.* **75**(2), 357–365 (2023).
27. C. W. Teng et al., "Applications of indocyanine green in brain tumor surgery: review of clinical evidence and emerging technologies," *Neurosurg. Focus* **50**(1), E4 (2021).
28. L. Greenbaum et al., "Green fluorescent protein photobleaching: a model for protein damage by endogenous and exogenous singlet oxygen," *Biol. Chem.* **381**(12), 1251–1258 (2000).
29. Q. Zheng et al., "The contribution of reactive oxygen species to the photobleaching of organic fluorophores," *Photochem. Photobiol.* **90**(2), 448–454 (2014).
30. K. Das and A. Roychoudhury, "Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants," *Front. Environ. Sci.* **2**, 53 (2014).
31. P. Garcia-Caparrós et al., "Oxidative stress and antioxidant metabolism under adverse environmental conditions: a review," *Botanical Rev.* **87**, 421–466 (2021).
32. M. Hasanuzzaman et al., "Regulation of ROS metabolism in plants under environmental stress: a review of recent experimental evidence," *Int. J. Mol. Sci.* **21**(22), 8695 (2020).
33. S. Sachdev et al., "Abiotic stress and reactive oxygen species: generation, signaling, and defense mechanisms," *Antioxidants* **10**(2), 277 (2021).
34. M. Mirza and S. Osindero, "Conditional generative adversarial nets," arXiv:1411.1784 (2014).
35. W. H. Press and S. A. Teukolsky, "Savitzky-Golay smoothing filters," *Comput. Phys.* **4**(6), 669–672 (1990).
36. J. Chen et al., "New insights into the noise reduction Wiener filter," *IEEE Trans. Audio Speech Lang. Process.* **14**(4), 1218–1234 (2006).
37. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
38. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**(7553), 436–444 (2015).

39. X. Wang, Y. Zhao, and F. Pourpanah, "Recent advances in deep learning," *Int. J. Mach. Learn. Cybern.* **11**(4), 747–750 (2020).
40. M. R. Minar and J. Naher, "Recent advances in deep learning: an overview," arXiv:1807.08169 (2018).
41. S. Guan et al., "Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal," *IEEE J. Biomed. Health Inf.* **24**(2), 568–576 (2019).
42. D. Allman, A. Reiter, and M. A. L. Bell, "Photoacoustic source detection and reflection artifact removal enabled by deep learning," *IEEE Trans. Med. Imaging* **37**(6), 1464–1477 (2018).
43. S. Antholzer, M. Haltmeier, and J. Schwab, "Deep learning for photoacoustic tomography from sparse data," *Inverse Probl. Sci. Eng.* **27**(7), 987–1005 (2019).
44. S. Guan et al., "Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning," *Sci. Rep.* **10**(1), 1–12 (2020).
45. T. Chen et al., "A deep learning method based on U-Net for quantitative photoacoustic imaging," *Proc. SPIE* **11240**, 112403V (2020).
46. A. DiSpirito et al., "Reconstructing undersampled photoacoustic microscopy images using deep learning," *IEEE Trans. Med. Imaging* **40**(2), 562–570 (2020).
47. M. K. A. Singh et al., "Deep learning-enhanced LED-based photoacoustic imaging," *Proc. SPIE* **11240**, 1124038 (2020).
48. M. Lu et al., "Artifact removal in photoacoustic tomography with an unsupervised method," *Biomed. Opt. Express* **12**(10), 6284–6299 (2021).
49. A. Sharma and M. Pramanik, "Convolutional neural network for resolution enhancement and noise reduction in acoustic resolution photoacoustic microscopy," *Biomed. Opt. Express* **11**(12), 6826–6839 (2020).
50. I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM* **63**(11), 139–144 (2020).
51. J. Chen et al., "Image blind denoising with generative adversarial network based noise modeling," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 3155–3164 (2018).
52. K. Lin et al., "Real photographs denoising with noise domain adaptation and attentive generative adversarial network," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit. Workshops*, pp. 1717–1721 (2019).
53. D. Mishra et al., "Ultrasound image enhancement using structure oriented adversarial network," *IEEE Signal Process. Lett.* **25**(9), 1349–1353 (2018).
54. K. J. Halupka et al., "Retinal optical coherence tomography image enhancement via deep learning," *Biomed. Opt. Express* **9**(12), 6205–6221 (2018).
55. H. Lan et al., "Ki-GAN: knowledge infusion generative adversarial network for photoacoustic image reconstruction in vivo," *Lect. Notes Comput. Sci.* **11764**, 273–281 (2019).
56. R. Cao et al., "Label-free intraoperative histology of bone tissue via deep-learning-assisted ultraviolet photoacoustic microscopy," *Nat. Biomed. Eng.* **7**, 124–134 (2022).
57. Z. Zhang et al., "Deep and domain transfer learning aided photoacoustic microscopy: acoustic resolution to optical resolution," *IEEE Trans. Med. Imaging* **41**(12), 3636–3648 (2022).
58. M. Huang et al., "Unveiling precision: a data-driven approach to enhance photoacoustic imaging with sparse data," *Biomed. Opt. Express* **15**(1), 28–43 (2023).
59. A. Savitzky and M. J. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.* **36**(8), 1627–1639 (1964).
60. S. R. Krishnan and C. S. Seelamantula, "On the selection of optimum Savitzky-Golay filters," *IEEE Trans. Signal Process.* **61**(2), 380–391 (2012).
61. Y. Liu et al., "Applications of Savitzky-Golay filter for seismic random noise reduction," *Acta Geophys.* **64**, 101–124 (2016).
62. J. Benesty et al., "Study of the Wiener filter for noise reduction," in *Speech Enhancement*, J. Benesty, S. Makino, and J. Chen, Eds., pp. 9–41, Springer, Berlin, Heidelberg (2005).
63. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
64. A. Paul and S. Mallidi, "U-Net enhanced real-time LED-based photoacoustic imaging," *J. Biophotonics* **17**(6), e202300465 (2024).
65. P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in Statistics: Methodology and Distribution*, S. Kotz and N. L. Johnson, Eds., pp. 492–518, Springer, New York, NY (1992).
66. G. P. Meyer, "An alternative probabilistic interpretation of the Huber loss," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 5257–5265 (2021).
67. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
68. M. Pekmezci and Y. Genc, "Evaluation of SSIM loss function in RIR generator GANs," *Digit. Signal Process.* **154**, 104685 (2024).
69. D. P. Kingma, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).
70. S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of Adam and beyond," arXiv:1904.09237 (2019).

71. J. Henry, T. Natalie, and D. Madsen, "Pix2pix gan for image-to-image translation," Research Gate Publication, 1-5 (2021).
72. P. Isola et al., "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 5967–5976 (2017).
73. R.-Y. Sun, "Optimization for deep learning: an overview," *J. Oper. Res. Soc. China* **8**(2), 249–294 (2020).
74. E. S. Lubana, R. Dick, and H. Tanaka, "Beyond BatchNorm: towards a unified understanding of normalization in deep learning," in *Adv. in Neural Inf. Process. Syst.*, Vol. 34, pp. 4778–4791 (2021).
75. K. Sivasubramanian and L. Xing, "Deep learning for image processing and reconstruction to enhance led-based photoacoustic imaging," in *LED-Based Photoacoustic Imaging*, M. Kuniyil Ajith Singh, Ed., pp. 203–241, Springer, Singapore (2020).
76. M. Welsaert and Y. Rosseel, "On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data," *PLoS One* **8**(11), e77089 (2013).
77. K. M. Kempinski et al., "Application of the generalized contrast-to-noise ratio to assess photoacoustic image quality," *Biomed. Opt. Express* **11**(7), 3684–3698 (2020).
78. A. Rodriguez-Molares et al., "The generalized contrast-to-noise ratio: a formal definition for lesion detectability," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**(4), 745–759 (2019).
79. C. Bingham, M. Godfrey, and J. Tukey, "Modern techniques of power spectrum estimation," *IEEE Trans. Audio Electroacoust.* **15**(2), 56–66 (1967).
80. M. F. Kijewski and P. F. Judy, "The noise power spectrum of CT images," *Phys. Med. Biol.* **32**(5), 565 (1987).
81. J. Siewerdsen, I. Cunningham, and D. Jaffray, "A framework for noise-power spectrum analysis of multi-dimensional images," *Med. Phys.* **29**(11), 2655–2671 (2002).
82. J. D. Cohn, "Power spectrum and correlation function errors: Poisson vs. Gaussian shot noise," *New Astron.* **11**(4), 226–239 (2006).
83. C. Park et al., "Power spectrum, correlation function, and tests for luminosity bias in the CfA redshift survey," *Astrophys. J.* **431**(2, pt. 1), 569–585 (1994).
84. S. V. Vaseghi, "Power spectrum and correlation," in *Advanced Digital Signal Processing and Noise Reduction*, pp. 263–296, John Wiley & Sons Ltd (2000).
85. A. Balakrishnan and R. R. Mazumdar, "On powers of Gaussian white noise," *IEEE Trans. Inf. Theory* **57**(11), 7629–7634 (2011).
86. A. Jain and V. Bhateja, "A versatile denoising method for images contaminated with Gaussian noise," in *Proc. CUBE Int. Inf. Technol. Conf.*, pp. 65–68 (2012).
87. O. Lepskii, "On a problem of adaptive estimation in Gaussian white noise," *Theory Probab. Appl.* **35**(3), 454–466 (1991).
88. C. Saxena and D. Kourav, "Noises and image denoising techniques: a brief survey," *Int. J. Emerg. Technol. Adv. Eng.* **4**(3), 878–885 (2014).
89. S. W. Hasinoff, "Photon, Poisson noise," in *Computer Vision: A Reference Guide*, K. Ikeuchi, Ed., pp. 980–982, Springer (2021).
90. R. H. Chan, C.-W. Ho, and M. Nikolova, "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization," *IEEE Trans. Image Process.* **14**(10), 1479–1485 (2005).
91. Y. Jiang et al., "Salt and pepper noise removal method based on the edge-adaptive total variation model," *Front. Appl. Math. Stat.* **8**, 918357 (2022).
92. M. Tur, K.-C. Chin, and J.W. Goodman, "When is speckle noise multiplicative?" *Appl. Opt.* **21**(7), 1157–1159 (1982).

Avijit Paul received his Bachelor of Technology degree in computer science and engineering from the Kalyani Government Engineering College, West Bengal University of Technology, India, and his Master of Science degree in intelligent systems from University of Sussex, UK. He has more than 11 years of experience as a software developer and a project management lead in several multinational corporations in India. He is currently pursuing PhD in biomedical engineering from iBIT lab, Tufts University, USA and his research interests include ML/DL applications for photoacoustic imaging and signal processing.

Srivalleesha Mallidi received her master's and PhD degrees in biomedical engineering from the University of Texas at Austin. She was a NIH Ruth L. Kirschstein postdoctoral fellow at Harvard Medical School. Currently, she is a Tiampo Family Assistant Professor in the Department of Biomedical Engineering at Tufts University, Medford, Massachusetts, and she directs the integrated biofunctional imaging and therapeutics (iBIT) lab that focuses on ultrasound and photoacoustic imaging based therapeutic strategies.

Biographies of the other authors are not available.