

Evaluation of a pipeline for simulation, reconstruction, and classification in ultrasound-aided diffuse optical tomography of breast tumors

Giuseppe Di Sciacca^{a,b,*}, Giulia Maffei^b, Andrea Farina^{b,c},
Alberto Dalla Mora^b, Antonio Pifferi^{b,c}, Paola Taroni^{b,c}
and Simon Arridge^a

^aUniversity College London, Department of Computer Science, London, United Kingdom

^bPolitecnico di Milano, Dipartimento di Fisica, Milano, Italy

^cIstituto di Fotonica e Nanotecnologie, Consiglio Nazionale delle Ricerche, Milano, Italy

Abstract

Significance: Diffuse optical tomography is an ill-posed problem. Combination with ultrasound can improve the results of diffuse optical tomography applied to the diagnosis of breast cancer and allow for classification of lesions.

Aim: To provide a simulation pipeline for the assessment of reconstruction and classification methods for diffuse optical tomography with concurrent ultrasound information.

Approach: A set of breast digital phantoms with benign and malignant lesions was simulated building on the software VICTRE. Acoustic and optical properties were assigned to the phantoms for the generation of B-mode images and optical data. A reconstruction algorithm based on a two-region nonlinear fitting and incorporating the ultrasound information was tested. Machine learning classification methods were applied to the reconstructed values to discriminate lesions into benign and malignant after reconstruction.

Results: The approach allowed us to generate realistic US and optical data and to test a two-region reconstruction method for a large number of realistic simulations. When information is extracted from ultrasound images, at least 75% of lesions are correctly classified. With ideal two-region separation, the accuracy is higher than 80%.

Conclusions: A pipeline for the generation of realistic ultrasound and diffuse optics data was implemented. Machine learning methods applied to a optical reconstruction with a nonlinear optical model and morphological information permit to discriminate malignant lesions from benign ones.

© 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.27.3.036003](https://doi.org/10.1117/1.JBO.27.3.036003)]

Keywords: diffuse optical tomography; ultrasound; lesion classification; breast cancer; breast digital phantom.

Paper 210385GRR received Dec. 6, 2021; accepted for publication Feb. 28, 2022; published online Mar. 24, 2022.

1 Introduction

Diffuse optical tomography (DOT) is a technique of medical imaging based on the injection and collection of near-infrared and visible red photons undergoing scattering and absorption events in tissues.^{1,2} The evolution of the related technology has recently seen the miniaturization of the electronic components involved, thus leading to the design of handheld probes.^{3–5} However, the mathematical reconstruction problem stemming from DOT is severely ill-posed and reconstructed images are low resolution and low contrast, and present artifacts, which hinders the wide use of the technique in clinical scenarios.

*Address all correspondence to Giuseppe Di Sciacca, giuseppe.sciacca.17@ucl.ac.uk

Nonetheless, DOT presents qualities that make it an active and growing research topic, e.g., the spectral nature of DOT, which allows functional analysis of tissues and is crucial for clinical diagnosis. Other qualities are its noninvasiveness and cost-effectiveness.^{6–10} Among its applications, DOT is currently utilized for breast cancer diagnosis.^{11,12} In the field, x-ray mammography is often referred to as the state of the art for breast cancer screening, but it presents decreased sensitivity with dense breasts. In such cases, ultrasound (US) imaging is often used as an adjunct technique. The diagnosis via US B-mode images is done at a morphological level and allows the clinician to reach a sensitivity close to 100%.¹³ With DOT, US shares the non-invasiveness and cost-effectiveness, but it is also characterized by higher resolution. For these reasons, the combination of DOT with a well-posed, high-resolution techniques such as US imaging was proposed early in the field^{14,15} together with the question on how to best achieve an improvement of DOT out of this combination in the wider topic of multimodal imaging.^{16–19}

In general, the inverse problem in DOT is solved via the minimization of an objective function in the following form:

$$f^{\text{recon}} = \operatorname{argmin} \mathcal{D}(\mathcal{A}f, g) + \alpha\Psi(f), \quad (1)$$

where \mathcal{D} is a—usually quadratic—data-fitting term, \mathcal{A} is the forward operator, g is the data, Ψ is a regularizer, and α is the hyperparameter weighting the regularization with respect to the data-fitting term. The forward model of propagation \mathcal{A} is chosen based on, e.g., linearity, time dependency, domain discretization. Many approaches have been proposed in literature for the combination of DOT with other techniques: US information can be incorporated either in the regularizer^{18,20,21} or directly in the model.^{22–25} Often, the technical specifications of a probe and the main objective of a study define the best reconstruction strategy to adopt, but newly designed physical devices require a time span of months or even years before being assessed on clinically relevant cases and extract statistical figures.^{26,27} Having realistic simulations would, in some degree, help assessing proposed reconstruction approaches before clinical evaluation also from a statistical point of view. Recently, VICTRE was proposed as a tool for tomosynthesis of realistic breast digital phantoms for applications in z-ray mammography²⁸ and was subsequently adopted as a test database for other optical imaging applications.^{29–31} Hereby, we propose a simulation pipeline based on VICTRE for the generation of functional phantoms for a combined US+DOT hand held probe, adopting—as an example—the integrated probe developed in the SOLUS Project.³² Based on this, we identify a reconstruction method and a classification procedure with the aim of best discriminating benign lesions from malignant ones. To our knowledge, it is the first time that US and DOT standalone simulations are combined for the characterization of such a multimodal probe. A strategy for the combination of acoustic and optical properties of tissues has instead already been presented in Refs. 30 and 33. In Secs. 2.1 and 2.2, we show the procedure used to generate respectively US simulations and optical data from a VICTRE digital phantom. In Secs. 2.2.1 and 2.3, we present the reconstruction method that has been adopted to retrieve the optical properties of the simulated lesions and the classification method to separate them into benign and malignant.

2 Methods

VICTRE is a simulation pipeline aiming at giving a realistic tomosynthesis of breast digital phantoms.^{28,31} A digital phantom is a voxelized image, where each voxel is assigned to tissue class: glandular, adipose, artery, vein, terminal duct lobular unit, duct, nipple, skin, muscle, and ligament.^{30,34} For simplicity, in the following discussion, we refer to a VICTRE phantom as an operator $\mathcal{V}: \Omega \rightarrow \mathbb{N}$ assigning an index i representing a tissue to a region in 3D space so $i = \mathcal{V}\mathbf{r}$. The effect is a decomposition of the domain Ω in subsets Ω_i composed of a single tissue so $\Omega = \cup_i \Omega_i$. The software also allows for the compression that the breast undergoes in a standard mammography exam and for the generation of tumor-like shapes.³⁵ We simulated a total of $N_{\text{ph}} \sim 700$ left breasts where the fat-fraction was the main parameter that was changed to simulate the most frequent types of breasts. The resolution was set to 0.5 mm as a compromise between the resolution of US and computational effort.

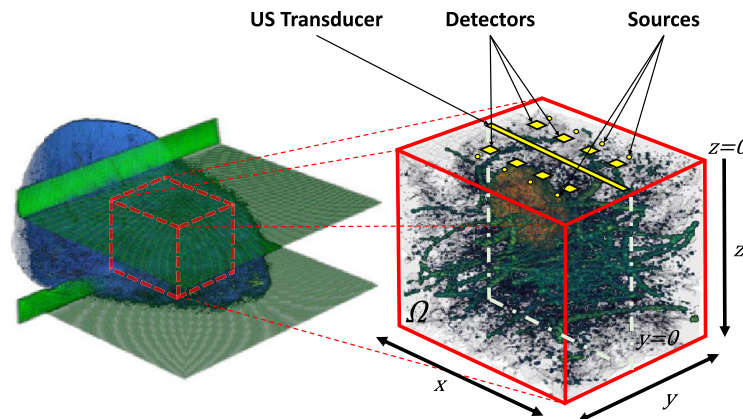


Fig. 1 Geometry definition. A VICTRE phantom is compressed between two paddles to 45 mm. A cuboid region Ω is extracted to serve as functional ground truth in the discussion. A SOLUS-like probe is simulated on the top surface of the considered domain. The plane $y = 0$ is set to be the imaging plane of the US transducer.

Here, we choose to compress the breast to 45 mm along the axis going from head to feet to simulate the mild compression given by a handheld probe with flat surface. We note that this arrangement does not take into account some peculiarities of handheld devices such as the effects of the chest wall. For each digital phantom, a lesion of average radius varying from 6 to 13.5 mm was generated and inserted in the tissues after compression.

Building on VICTRE, we present here a simulation pipeline for a dual-mode probe inspired by the multimodal device developed in the EC funded SOLUS Project.⁵ A whole acquisition is composed by a B-mode image set along the plane $y = 0$ and by a set of time diffuse optical data obtained in reflectance geometry by eight sources and eight detectors placed on the plane $z = 0$, identifying the top surface of the breast. The center of the US transducer and of the matrix of detectors is set to be at the origin of the reference system. We show a schematic of the simulated acquisition system in Fig. 1. In detail, the sources are disposed on two rows of four elements each at a distance of 17.5 mm from and parallel to the US imaging plane. The distance between each source on the same line is 12 mm. An analog geometry is selected for the detectors, but at a distance 10.8 mm from the plane $y = 0$.

For simplicity, each lesion was placed in the phantom so as to have its center of mass at a maximum distance of $0.3l$ from the US imaging plane $y = 0$, with l the maximum elongation of the lesion along y . In the following, we present a simulation approach to generate US and optical data from the VICTRE phantoms. In Sec. 2.2.1, we present an optical reconstruction method based on a two-region nonlinear model and, in Sec. 2.3, we apply some commonly used classification methods to assess the separability of the reconstructed lesions into benign and malignant classes. The whole procedure of generation and compression of the digital phantoms took circa 3 h on a virtual machine of 20 GB of RAM assembled on a computational cluster.

2.1 US Modeling

Computational US is an ever-growing branch of research since field-II made its first appearance in 1996.³⁶ Subsequently, many US simulators have been proposed and each of them differs from its competitors in a variety of factors such as the model of propagation for acoustic waves, the numerical method used or the physical model for tissues. For this work, we considered as simulators Field-II,³⁶ SIMUS,³⁷ CREANUIS,³⁸ and k-wave.³⁹ A thorough comparison of these software is out of scope in this paper, thus we summarize here the main differences on methodology and aims. Apart from k-Wave, all the considered simulators have a mesh-free numerical solver: acoustic scatterers with defined properties and positions are placed in the computational domain and an acoustic field is then propagated. The signal recorded at the transducer is the result of the interaction of the acoustic field with the scatterers. This approach allows a certain flexibility in simulating the position and translation of scatterers in tissues. Speckles in the simulated B-mode

images arise from the random position of scatterers, while the general acoustic properties of the medium are usually kept constant.

Field-II is based on the concept of a transducer's impulse response function, the main assumptions being: a linear wave equation, scatterers acting as monopole sources and weak scattering phenomena. By the Huygens' principle, in a homogeneous medium the spatial impulse response function $h(\mathbf{r}_1, t)$ at \mathbf{r}_1 from an aperture surface S is:⁴⁰

$$h(\mathbf{r}_1, t) = \int_S \frac{\delta\left(t - \frac{|\mathbf{r}_1 - \mathbf{r}_2|}{v_s}\right)}{2\pi|\mathbf{r}_1 - \mathbf{r}_2|} dS, \tag{2}$$

where $\delta(t)$ is a point-like source, \mathbf{r}_2 is the position of the transducer, and v_s is the speed of sound in the medium. Thus, the emitted pressure field can be retrieved by the convolution: $p(\mathbf{r}_1, t) = \rho_0 \frac{\partial}{\partial t} |\mathbf{u}| *_r h(\mathbf{r}_1, t)$ with $*_r$ spatial convolution operator, thus getting the Rayleigh integral,³⁶ with ρ_0 density of the medium and \mathbf{u} is the velocity of the front face of the transducer. In this scenario, a point-like change in pressure and sound speed in the medium can be introduced by a term $f_m(\mathbf{r}_1)$ so that:

$$p(\mathbf{r}_1, t) = v_{pe}(t) *_t f_m(\mathbf{r}_1) *_r h_{pe}(\mathbf{r}_1, t), \tag{3}$$

where v_{pe} is the electronic impulse response function, h_{pe} is a two-ways impulse response function taking into account the transmitting and receiving impulses, and $*_r$ denotes a convolution in time. Equation (3) is solved for a number of scatterers placed in the medium, and the results are summed to generate one B-mode scan line. We refer to Refs. 41 and 42 for details. The simulation assumptions used in SIMUS are the same as Field-II.³⁷ The main computation is performed in Fourier space and it is fully open source. This brings a set of numerical advantages and ease of use, especially for what regards frequency-dependent characteristics of ultrasound imaging. However, as Field-II, SIMUS is limited to linear propagation of acoustic waves and to single scattering phenomena that do not allow to reproduce, artifacts such as reverberation, shadowing, or mirror image artifacts. The addition of a heterogeneous nonlinear parameter and a GPU-based solver are the main features brought by CREANUIS.^{38,43,44} The software is still particle-based, however, under the approximation of weak nonlinear phenomena, two pressure fields are propagated, one that is linear $p^{(1)}$ and a second nonlinear one $p^{(2)}$:

$$\left(\nabla^2 - \frac{1}{v_s} \frac{\partial}{\partial t}\right) p^{(1)} = 0, \tag{4}$$

$$\left(\nabla^2 - \frac{1}{v_s} \frac{\partial}{\partial t}\right) p^{(2)} = -\left(1 + \frac{B}{2A}\right) \frac{1}{\rho_0 v_s^4} \frac{\partial^2}{\partial t^2} p^{(1)2}, \tag{5}$$

where v_s is the speed of sound in the medium, p_0 is the background pressure, and $\frac{B}{A}$ is the non-linearity parameter. The propagation of the two fields is made in Fourier domain and solved via the GASM method.⁴⁴ The second harmonic field is dependent on the value of the fundamental one. As with Field-II and SIMUS, the scan lines are obtained by placing scattering particles in the medium and backpropagating both fields to the transducer. The main limitations of the particle-based simulators described above is the small flexibility in simulating a heterogeneous speed of sound, density, and other spatial properties in tissues, that are only handled by the position and amplitude of scatterers. In turns, this limits the number of real-like artifacts that can be obtained from a simulation.

k-Wave is a mesh-based nonlinear simulator based on the set of acoustic equations.^{39,45,46}

$$\begin{cases} \frac{\partial}{\partial t} \mathbf{u} = -\frac{1}{\rho_0} \nabla p \\ \frac{\partial}{\partial t} \rho = (2\rho - \rho_0) \nabla \cdot \mathbf{u} - \mathbf{u} \cdot \nabla \rho_0 \\ p = v_s^2 \left(\rho + \frac{B}{2A} \frac{\rho^2}{\rho_0} + \mathbf{d} \cdot \rho_0 - L\rho \right) \end{cases}, \tag{6}$$

where ρ_0 is the background mass density of the medium, ρ , p , and \mathbf{u} are the acoustic mass density, pressure, and particle speed, respectively. L is a fractional Laplacian term accounting

Table 1 Summary of the properties of the considered simulators.

	Field-II	SIMUS	CREANUIS	k-Wave
Numerical method	Meshfree	Meshfree	Meshfree	Mesh-based
Governing acoustic equations	Linear	Linear	Non-linear	Non-linear
Time domain	Temporal	Harmonic	Temporal	Both ⁴⁸
Space domain	x, y, z	x, y, z	k_x, k_y, k_z	$x, y, z + k$
Medium	Homo.	Homo.	Hete.	Hete.
Scattering	Weak	Weak	Weak	Multiple
Artifacts	No	No	No	Yes

for dispersion and acoustic absorption. Nonlinear acoustic effects are governed by the terms 2ρ and $\frac{B}{2A}\rho^2$ with $\frac{B}{A}$ nonlinearity parameter.⁴⁷ Acoustic sources are simulated by adding a source term to the first and/or second equation. The set of equations are solved with a pseudospectral method with the spatial components of the acoustic quantities that are calculated in k-space. One B-mode scan-line is obtained by defining a grid, a source, a detector, a medium and solving the set of Eq. (6) in the system. In principle, the approach followed in k-Wave allows one to simulate heterogeneous media and artifacts such as speckles, shadowing, and reverberations that come naturally from the wave equation and allow for more realistic simulations.⁴⁸

A schematic comparison of the considered US simulators, modeled after the one in Ref. 48 can be found in Table 1. Given its characteristics, k-Wave has been chosen as the main tool for US simulations in this paper.

2.1.1 Definition of the medium

The system in Eq. (6) describes the contrast in a US image as the result of spatially changing acoustic properties—speed and density—in tissues.⁴⁹ Two different scales of contrast are observed in a B-mode breast image: a macroscopic one, visible at the interfaces of tissues and a microscopic one, that is due to changes in characteristic impedance at the cellular level and results in regions of different brightness in the images depending on the tissues.⁵⁰ We present a VICTRE-based acoustic digital phantom that aims at having the same dual scale of heterogeneities. To each tissue identified by the index $i = \mathcal{V}\mathbf{r}$ is assigned an average sound speed and standard deviation such that for each point \mathbf{r} in Ω_i the speed of sound $v_s(\mathbf{r})$ for $\mathbf{r} \in \Omega_i$ is defined to be a stationary Gaussian process \mathcal{GP} such that

$$\begin{aligned} \bar{v}_{s,i} &\sim \mathcal{N}(\mu_i, \sigma_{i,\text{macro}}) \\ \{v_s\}_{\mathbf{r} \in \Omega_i} &\sim \mathcal{GP}(\bar{v}_{s,i}, \sigma_{i,\text{micro}}), \end{aligned} \quad (7)$$

where $\mathcal{N}(\mu, \sigma)$ is a normal distribution of mean μ and standard deviation σ , $\bar{v}_{s,i}$ defines the average speed of sound in the tissue i , and it is extracted from a normal distribution of mean μ_i and standard deviation of $\sigma_{i,\text{macro}}$. The standard deviation $\sigma_{i,\text{micro}}$ acts as a source of microscopic contrast depending on the tissues. The mean speed of sound μ and macroscopic standard deviation σ_{macro} of adipose tissues, glandular tissues, and lesions have been assigned following Ref. 51. In Figs. 4(a) and 4(b), we give an example of simulated maps for \bar{v}_s and the piecewise function σ_{micro} over the imaging plane $y = 0$ of Ω . The resulting speed of sound can be found in Fig. 4(c). The acoustic properties of the remaining tissues were drawn from the same distribution as the glandular ones. The microscopic contrast in the lesions was selected to be lower than the level of noise of all the surrounding tissues, as often happens in US B-mode images where lesions are less bright. The microscopic standard deviation for each of the other tissues was randomly selected to have a standard deviation between 1% and 5% of the average speed of sound of the tissues. To reduce the computational cost of the simulations, only the active

Table 2 Main simulation parameters for ultrasound B-mode images.

f_c	7 MHz
Focus	15 mm
Elevation plane focal point	16 mm
Fractional bandwidth	80%
N elements	256
Height	4 mm
Pitch	0.2 mm
Kerf	0.001 mm
Width	0.198 mm
Scan lines	200
Grid size point	$516 \times 172 \times 86$
Grid size	$45 \times 15 \times 7.5$ mm
$\Delta x, \Delta y, \Delta z$	0.0872 mm
Perfectly matched layer	$30 \times 10 \times 5$
Δt	8.9 ns
T	67 μ s
$v_{s,0}$	1465 m/s

elements (42 out of 256 in total) of the whole transducer are simulated at the same time; for each scan line the smaller transducer is moved along the scanning direction. Thus, it is possible to simulate a portion of the simulated medium for each scan line, so it is possible to have a finer grid allowing to simulate higher US frequencies.³⁹ With these settings, a B-mode image took ~ 7 h to run on a virtual machine with a 10 GB GPU assembled on a computational cluster. In Table 2, we give the main simulation parameters used. Where possible the simulations parameter have been set to be equal to those of the final SOLUS probe. Even though some speckles are present in the images, we note that to model fully developed speckles, a high number of scatterers per wavelength would be required. We choose here to limit the number of scatterers to reduce the computational effort.

2.2 Optical Modeling

Photon propagation of monochromatic light in tissues is well described by the radiative transfer equation:⁵²

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \hat{\mathbf{s}} \cdot \nabla + \mu_a + \mu_s \right) L_\nu(\mathbf{r}, \hat{\mathbf{s}}, t) = \mu_s \int_{\hat{\Omega}'} L_\nu(\mathbf{r}, \hat{\mathbf{s}}', t) \Theta(\hat{\mathbf{s}} \cdot \hat{\mathbf{s}}') d\hat{\Omega}' + Q(\mathbf{r}, \hat{\mathbf{s}}, t), \quad (8)$$

where L_ν is the spectral radiance, Q is a source term, $\hat{\mathbf{s}}$ is the direction of propagation, Θ is a phase function, and μ_s and μ_a are the scattering and absorption coefficients. Even though some analytical solutions for simple systems are known, the RTE can be effectively solved only with Monte Carlo methods. A valid alternative to the RTE in human tissues is given by the diffusion equation (DE) with Robin boundary conditions:⁵³

$$\begin{cases} \frac{1}{c} \frac{\partial}{\partial t} \Phi(\mathbf{r}, t) - \nabla \cdot \kappa \nabla \Phi(\mathbf{r}, t) + \mu_a \Phi(\mathbf{r}, t) = Q(\mathbf{r}, t) & \text{in } \Omega \times [0, T], \\ \Phi(\mathbf{r}, t) + 2A\kappa \frac{\partial}{\partial n} \Phi(\mathbf{r}, t) = 0 & \text{on } \partial\Omega, \end{cases} \quad (9)$$

where photons migration is described via a diffusive process of the fluence Φ in a time T and is dependent on the properties of absorption μ_a and diffusivity $\kappa = \frac{1}{3\mu_s}$ of the medium,^{52,54-57} where μ_s' is the reduced scattering coefficient. Here, $A = \frac{1+R}{1-R}$ with R reflection coefficient. The DE can be analytically solved for simple geometries and homogeneous optical coefficients. The solution of the DE for complex systems and heterogeneous coefficients requires the use of numerical solvers such as the finite element method (FEM).⁵⁸ In general, the forward model to calculate the signal observed by a detector in position \mathbf{r}_j due to a source i can be expressed as a forward operator:

$$y_{i,j} := \Phi_i(\mathbf{r}_j, t) = \mathcal{A}_{i,j} \begin{bmatrix} \mu_a(\mathbf{r}) \\ \mu_s'(\mathbf{r}) \end{bmatrix}. \tag{10}$$

Experimental conditions are incorporated in the model by convolving the exitance with the experimental impulse response function measured with a time-resolved diffuse spectroscopy laboratory research prototype.⁵⁹ Among the many numerical tools available for the implementation of FEM, we choose the software TOAST++.⁶⁰

The optical coefficients μ_a and μ_s' can be linked to the functional properties of the tissues via a spectral model:⁶¹⁻⁶³

$$\mu_a(\mathbf{r}, \lambda) = \sum_i \epsilon_i(\lambda) C_i(\mathbf{r}), \tag{11}$$

$$\mu_s'(s, \lambda) = a(\mathbf{r}) \left(\frac{\lambda}{\lambda_0} \right)^{-b(\mathbf{r})}. \tag{12}$$

These expressions define a spectral model that relates the functional properties of tissues to a set of $2 \times N_\lambda$ optical coefficients, where N_λ is the number of the probing wavelengths. We use the functional information of each VICTRE phantom to assign an optical ground truth to each tissue. The wavelengths chosen for the simulations of the system are 635, 670, 685, 785, 905, 930, 975, 1060 nm as explored in Ref. 64. This set of wavelengths is suitable for the retrieval of a number of chromophores: hemoglobin, oxyhemoglobin, lipids, water, and collagen. For each tissue present in the digital phantoms, a chromophore concentration was drawn and then inserted in Eq. (11) to retrieve its absorption spectrum. In Fig. 2, we show the simulated absorption spectra for each tissue.

The coefficients a were drawn from a normal distribution such that $a \sim \mathcal{N}(1.5 \text{ mm}^{-1}, 0.25 \text{ mm}^{-1})$ for benign lesions and $a \sim \mathcal{N}(1.4 \text{ mm}^{-1}, 0.25 \text{ mm}^{-1})$ for malignant ones. In addition, for each λ a term $\eta \sim \mathcal{N}(0, 0.1\mu_s'(\lambda))$ was added to $\mu_s'(\lambda)$ as a source of noise in the model in Eq. (11).

A set of optical data was simulated without any lesion. A nonlinear global fit of the homogeneous optical coefficients μ_a and μ_s' was then applied.⁶⁵ The average reconstructed optical coefficients are shown in Fig. 3.

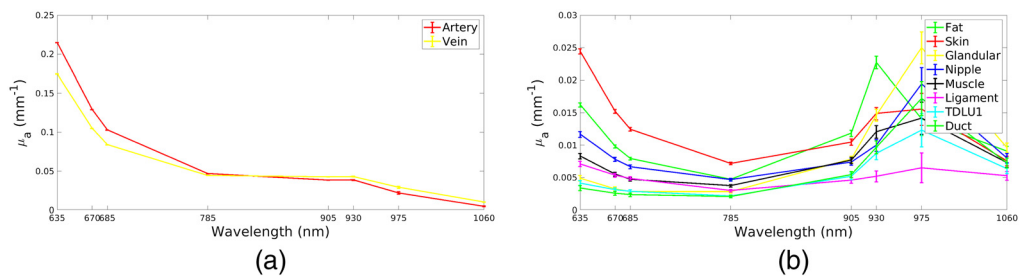


Fig. 2 Spectra used in simulations for the components generated by VICTRE. Blood vessels and other tissues are presented in different images for visualization purposes due to the scale. (a) Absorption spectra of blood vessels. A smaller saturation has been chosen for veins. (b) Absorption spectra simulated for the other tissues.

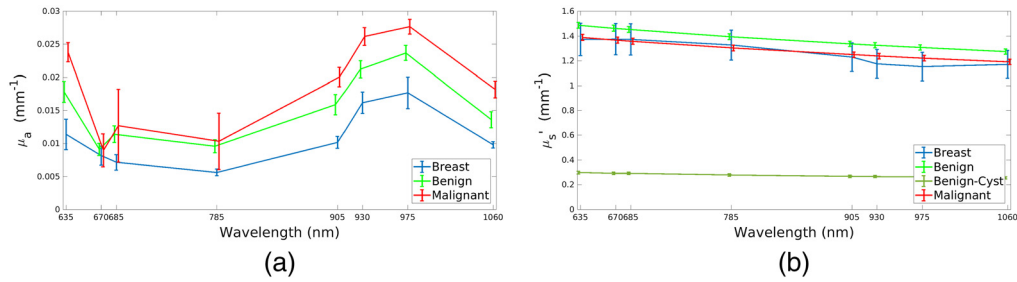


Fig. 3 Spectral plot of optical coefficients for breast bulk (blue), benign (green), and malignant (red) inclusions. The effective optical coefficient values for the breast have been obtained by fitting a homogeneous analytical model to the reference optical data obtained by simulating a breast with no inclusion. (a) Absorption spectra of malignant lesions, benign lesions, and healthy breast tissues. (b) Scattering spectra of malignant lesions, benign lesions (excluding cysts), cysts, and healthy breast tissues.

For the optical properties of breast lesions and their connection to benign and malignant nature, we refer to Ref. 9 to assign their optical properties. To take into account also cysts, in the simulations 25% of benign lesions were set to have $a \sim \mathcal{N}(0.3 \text{ mm}^{-1}, 0.01 \text{ mm}^{-1})$. The number of benign and malignant lesions was 349 and 379, respectively. No spatial variation of the optical coefficients inside each tissue was simulated at this stage. Unless specified otherwise, we refer to the simulated optical properties of lesions as μ_a^{truth} and μ_s^{truth} (i.e., ground truth). In Fig. 3, we display the set of the absorption and scattering properties simulated for benign lesions, malignant lesions and the effective optical properties for breast healthy tissues. The latter have been obtained generating optical data from the breast without inclusion and fitting an homogeneous analytical optical model to the data.

Optical data were simulated with a cuboid domain of computation of side $64 \times 58 \times 30$ mm with cubic voxels of side 2 mm. Time stepping was performed by an implicit Eulerian scheme with a time step of 25 ps for a total of 400 steps. Sources were defined to be of the Neumann type, with Gaussian profile of width 0.5 mm at $t = 0$ and to be zero for $t > 0$. Detectors were defined to have a Gaussian profile of width 1 mm. Noise was modeled by sampling a Poisson variable with mean and variance equal to an expected number of detected photons N^{expect} . This number was estimated as the number of photons detectable at the largest source–detector distance for a 1-min clinical examination with large area SIPM detectors^{66,67} and is equal to the area under the curve, i.e., $N^{\text{expect}} = \int_{-\infty}^{\infty} y_{i,j} dt$. The photon counts for other source–detector distances were rescaled linearly by their individual integrals over time. We note that, as a result of the compression, the dimensions of the phantoms are on average 6.3% smaller than the domain of computation for the optical data. For ease of computation, the part occupied by air was assigned to adipose or glandular tissue depending on which tissue was prevalent in the remainder of the breast.

2.2.1 Reconstruction

For the optical reconstruction, we choose a two-region nonlinear diffusion forward model in the form:

$$y_{i,j} = [\mathcal{A}_{i,j}[\mathbf{h}]](t + t_{\text{shift}}) \quad \text{with } \mathbf{h} := \begin{bmatrix} \mu_a^{\text{in}} \\ \mu_a^{\text{bulk}} \\ \mu_s^{\text{in}} \\ \mu_s^{\text{bulk}} \end{bmatrix}, \quad (13)$$

where μ_a^{in} and μ_s^{in} are the coefficients inside the inclusion, μ_a^{bulk} , μ_s^{bulk} are the optical coefficients in the bulk, and t_{shift} is a fictitious time shift parameter that is included to improve convergence in reconstruction. The operator \mathcal{A} is the same as Eq. (10) and the operator χ applied to \mathbf{h}

returns the optical coefficients over the whole domain. χ supposes prior knowledge on the morphology of the domain, e.g., by means of ultrasound information.

As with the forward model used in simulations, we include the instrumental impulse response function to the forward model as a convolution. To limit nonidealities present in any experimental system such as noise and amplitude mismatches, we test here an approach based on automatic selection of a region of interest (ROI), binning and self-normalization of data. For each source–detector couple, the preprocessed data results in a curve of N_{TW} temporal steps and area 1. Here, we automatically select a temporal ROI ranging from T_1 where the curve first reaches the 0.1 of its peak to $T_{N_{\text{TW}}}$ where it last reaches values of 0.01 of the peak. A set of $N_{\text{TW}} = 80$ equally spaced bins (T_k, T_{k+1}) for $k = 1, \dots, N_{\text{TW}}$ along the ROI were selected. We can define a pre-processing operator

$$\mathcal{P}_k[y(t)] := \frac{\int_{T_k}^{T_{k+1}} y(t) dt}{\sum_{k=1}^{N_{\text{TW}}} \int_{T_k}^{T_{k+1}} y(t) dt}, \quad (14)$$

which extracts the k 'th bin from the self-normalized measurements.

Reconstructions for each wavelength are performed by minimization of an objective function of the form:

$$\mathcal{L}(\mathbf{h}) = \frac{1}{2} \sum_{i=1}^{N_Q} \sum_{j=1}^{M_D} \sum_{k=1}^{N_{\text{TW}}} \frac{(\mathcal{P}_k[y_{i,j}(t)] - \mathcal{P}_k[\mathcal{A}_{i,j}[\chi\mathbf{h}](t + t_{\text{shift}})])^2}{\sigma_{i,j,k}^2(t) dt}, \quad (15)$$

where $\sigma_{i,j,k}^2 = \frac{\mathcal{P}_k[y_{i,j}(t)]}{\sum_{k=1}^{N_{\text{TW}}} \int_{T_k}^{T_{k+1}} y_{i,j}(t) dt}$ under the assumption of Poisson noise. We minimize the objective function by means of the MATLAB function `lsqcurvefit` with constraints on the values of \mathbf{h} so $0 < \mu_a < 0.06 \text{ mm}^{-1}$ and $0 < \mu_s' < 2.2 \text{ mm}^{-1}$.⁶⁸ The absolute values of the fictitious time parameter were limited to 25 ps. Only source–detector distances bigger than 27 mm have been considered. Thus, a subset of all the available curves was selected to allow for convergence of the absorption parameters over scattering⁵² as well as to mitigate the effect of local perturbations, e.g., because of highly absorbing blood vessels that hinder the validity of a two-region model. The fitting procedure took circa 60 min per wavelength for each phantom on a virtual machine with 300 MB of GPU. Overall, the procedure of simulation and reconstruction took circa 18 h per breast, with the most important hardware requirements needed for the simulation of B-scan mode images.

2.3 Classification

We investigate the feasibility of machine learning classification methods to the reconstructed optical properties μ_a^{in} and μ_s^{in} of the inclusions for a total of 16 features per lesion. To ease the visualization of 16-dimensional data, a dimensionality reduction with principal components analysis (PCA) was employed. As a result of this preliminary analysis, log-normalization of data was applied to guarantee a higher degree of separability. Among the many techniques that are available in literature, three main methods have been used for classification: logistic regression, support vector machines (SVM), and a fully connected network.⁶⁹ A nonlinear SVM was implemented with the Python library `scikit-learn`.⁷⁰ The SVM was optimized by choosing which kernel amongst Gaussian, polynomial, and sigmoidal and what regularization parameter C ranging from 0.5 and 5 guaranteed the minimum accuracy error on the training dataset. Separation of the dataset for logistic regression and SVM was done with a split ratio of 60:40 between training and test datasets. A fully connected neural network (FCN) was implemented with the Python library `Tensorflow`⁷¹ and is composed of four layers with ReLU activation function for the hidden layers and a sigmoid for the final layer. The width of the fully connected layers were 16, 32, 16, and 1, respectively. Hinge-loss was selected as loss function.⁷² Separation of the dataset was done with a split ratio of 60:20:20 between training validation and test datasets. Quantification of the performances of the mentioned techniques was performed by means of accuracy (acc.), precision (prec.), recall (rec.), and $F1$ -score ($F1$) defined as

$$\begin{aligned}
 \text{acc.} &= \frac{\text{TM} + \text{TB}}{\text{TB} + \text{TM} + \text{FB} + \text{FM}} \\
 \text{prec.} &= \frac{\text{TM}}{\text{TM} + \text{FM}} \\
 \text{rec.} &= \frac{\text{TM}}{\text{TM} + \text{FB}} \\
 F1 &= \frac{2\text{prec.} \times \text{rec.}}{\text{acc.} + \text{prec.}}, \tag{16}
 \end{aligned}$$

where TM is the number of correctly identified malignant lesions, TB is the number of correctly benign ones, while FM and FB are the numbers of incorrectly identified malignant and benign lesions, respectively.

3 Results

We assess our results from three points of view. First, we qualitatively show the simulated US B-mode images. After that, we apply the prior extraction procedure as described in Ref. 73. As a second assessment, we analyze the performances of the optical reconstructions. As a last step, we show how the reconstruction impacts on the separability of the lesions into benign and malignant classes.

3.1 Assessment of US Simulations and Distance Transform

In Figs. 4 and 5, we show some of the B-mode images simulated via the method presented in Sec. 2 and the main acoustic properties of the media. The images present characteristics of real B-mode images, such as shadowing and speckles. The method is able to simulate breast structures as they are present in the corresponding VICTRE digital phantom. The simulated

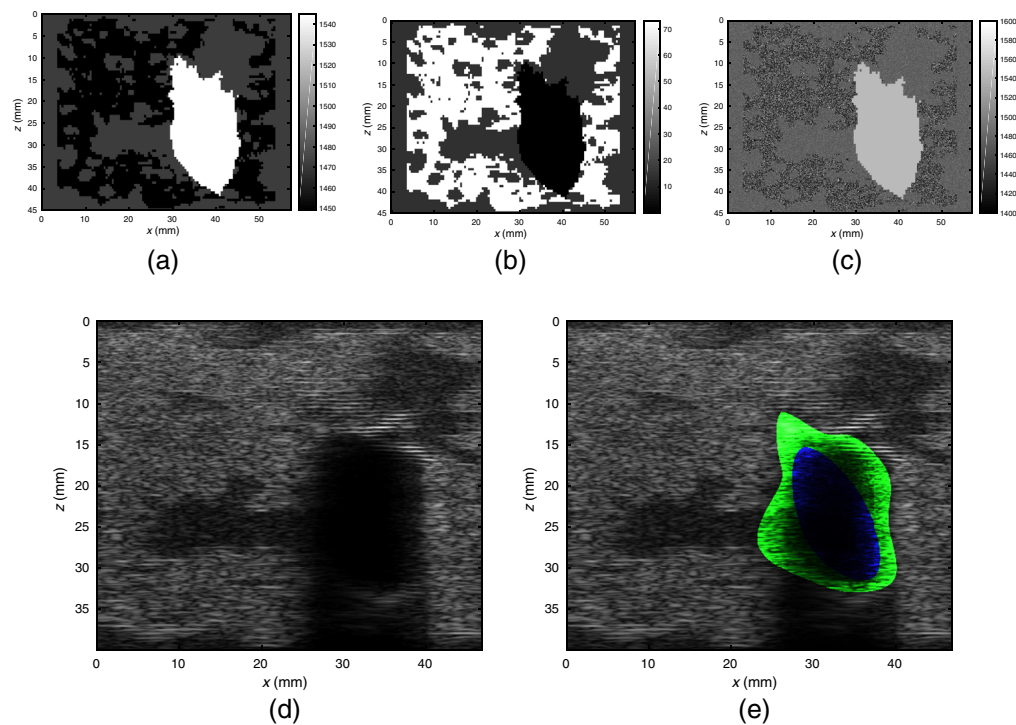


Fig. 4 B-mode image generation: Example 1. (a) Map of \bar{v}_s over $\Omega|_{y=0}$. (b) Map of σ_{micro} over $\Omega|_{y=0}$. (c) $v_s(\mathbf{r})$ for $\mathbf{r} \in \Omega|_{y=0}$. (d) B-mode image. (e) Segmentation. Blue is the user-defined segmentation, green is the final one. $\text{SDI} = 0.80$, $dA = -0.26$.

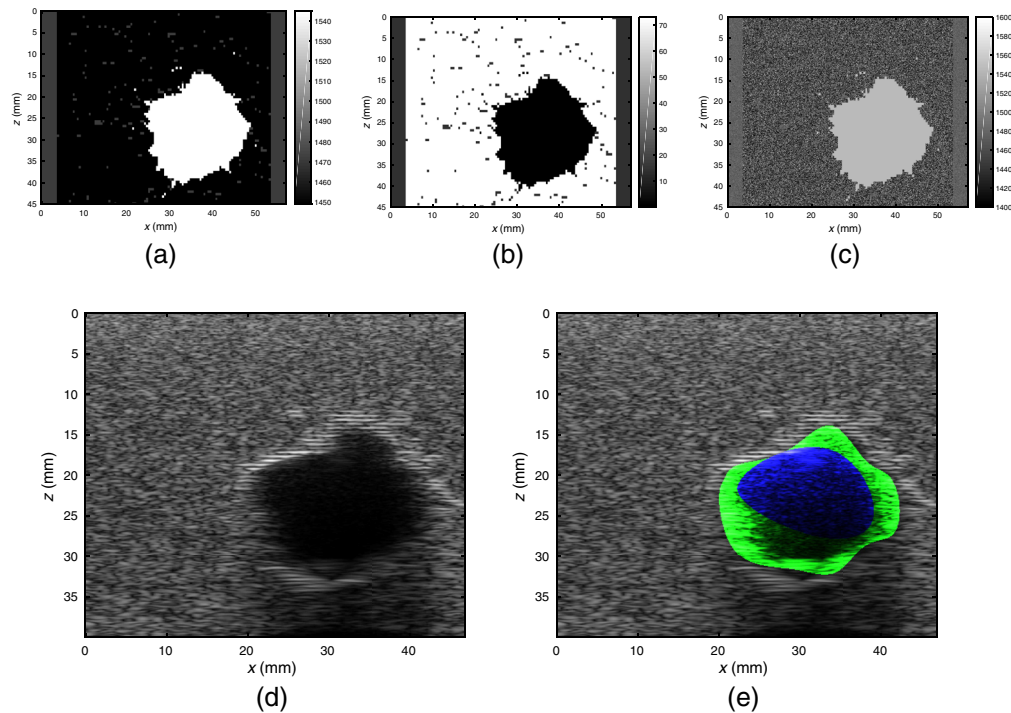


Fig. 5 B-mode image generation: Example 2. (a) Map of \bar{v}_s over $\Omega|_{y=0}$. (b) Map of σ_{micro} over $\Omega|_{y=0}$. (c) $v_s(\mathbf{r})$ for $\mathbf{r} \in \Omega|_{y=0}$. (d) B-mode image. (e) Segmentation. Blue is the user-defined segmentation, green is the final one. SDI = 0.76, dA = -0.32.

inclusions result to be less scattering than the rest of the tissues in the background and their borders are highlighted by means of the simulated contrast in characteristic impedance with respect to the surrounding area. For each inclusion, a set of three to four control points was selected close to the internal border of the inclusion itself. An algorithm based on snake fitting^{73,74} retrieved a segmentation of the image as shown in Figs. 6(b) and 6(d). In Fig. 6, we show the results of the extrapolation routine described in Ref. 73. An assessment of the extrapolation procedure was performed by means of the Sørensen-Dice index (SDI):⁷⁵⁻⁷⁷

$$\text{SDI}(X, Y) = \frac{2\sum_i |X_i \times Y_i|}{\sum_i |X_i| + \sum_i |Y_i|}, \quad (17)$$

where X_i and Y_i are the i 'th elements of two binary images X and Y ranging from 0 and 1, of the volume mismatch and of the displacement between the center of mass of ground truth and of the retrieved images. An analogous quantification has been performed for the segmentation procedure alone. Results are shown in Table 3. The SDI shows that the 3D extrapolation procedure introduces a source of error with respect to the previous step of segmentation. Other considered metrics were the relative mismatch in area dA and volume dV and the displacement of the center of the lesions. While the displacement of the inclusion results to be negligible with respect to the computational grid chosen for the optical reconstruction, there is a general underestimation of the size of the inclusions by about 50% of the ground truth value. This is mainly due to the different speed of sound of the inclusion with respect to the selected $v_{s,0}$ used in the image formation.

3.2 Assessment of Optical Reconstruction and Classification

A first assessment aimed at retrieving μ_a and μ_s' inside the inclusion with the two regions defined by the ground truth lesion. In Fig. 7, we plot all the retrieved values of absorption and scattering coefficients obtained in reconstruction versus the ground truth value. The black dashed line in Fig. 7 shows the ideal results for which the reconstructed properties are identical to the ground

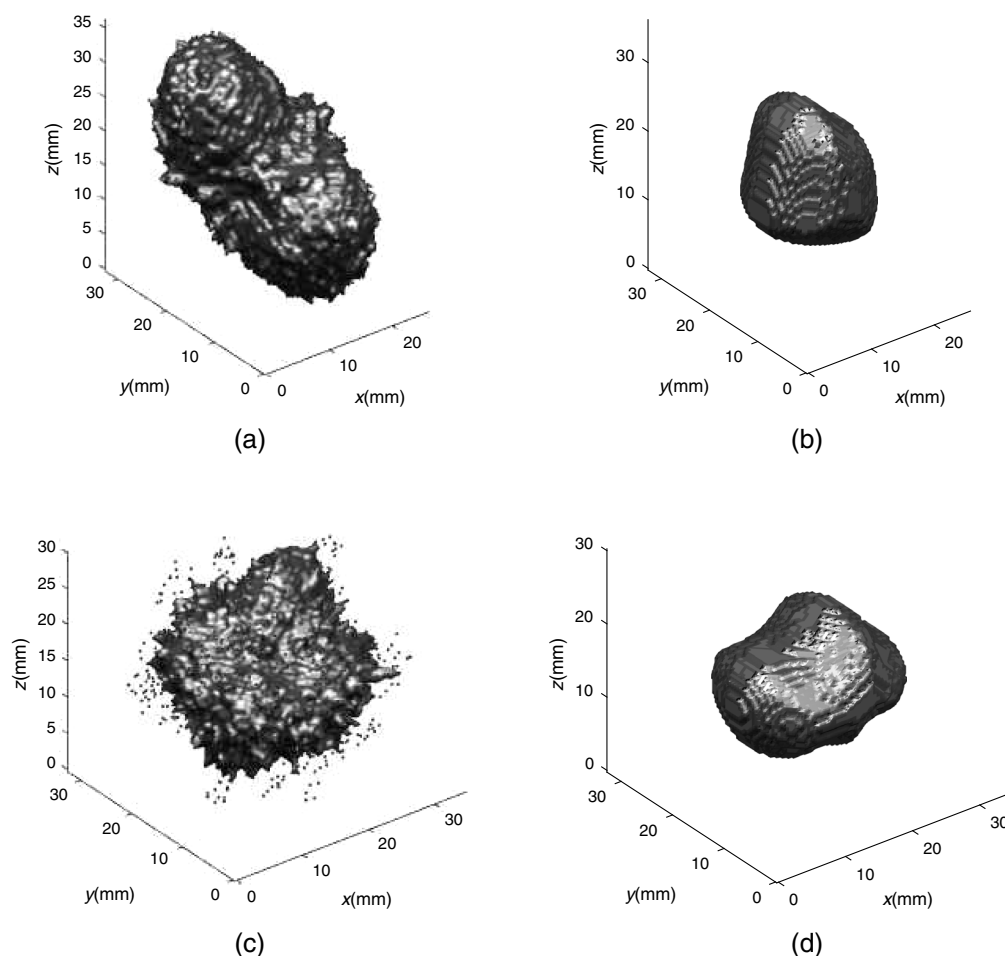


Fig. 6 Ground truth and final extrapolation for example 1 from Fig. 4 and example 2 from Fig. 5. (a) Example 1: ground truth lesion. (b) Example 1: distance transform extrapolated shape, $SDI = 0.46$, $dV = -0.68$. (c) Example 2: ground truth lesion. (d) Example 2: distance transform extrapolated shape, $SDI = 0.68$, $dV = -0.39$.

Table 3 Table displaying the performances of the segmentation and extrapolation procedure first presented in Ref. 73. In general, both the segmentation and the extrapolation procedure underestimate the dimensions of the ground truth shape. As expected, the SDI of the extrapolation procedure show a worse agreement between retrieved shape and ground truth with respect to the sole segmentation. Displacements in the lesion position negligible with respect to the grid resolution used here for DOT reconstruction.

	SDI	dA dV	Displacement (mm)		
Segmentation	0.71(0.13)	-0.33(0.16)	$\Delta x = 0.03(0.13)$	—	$\Delta z = 0.15(0.07)$
Extrapolation	0.55(0.14)	-0.5(0.17)	$\Delta x = 0.04(0.13)$	$\Delta y = 0.08(0.05)$	$\Delta z = 0.14(0.08)$

truth. Especially with respect to absorption, a clustering of points along the bisector can be observed. This effect is still present but less prevalent for scattering. In Table 4, we show the average quantification errors for the lesions in the cases of two regions defined by the ground truth lesion shapes and defined by the US extrapolation procedure. As expected, the latter exhibits larger errors. Twenty-five percent of the retrieved values have a relative difference with the ground truth higher than 50%. The percentage decreases to 18% when the ground truth lesion shape is available. As a second step, reconstructions were performed by defining the two regions with the information retrieved from the US.

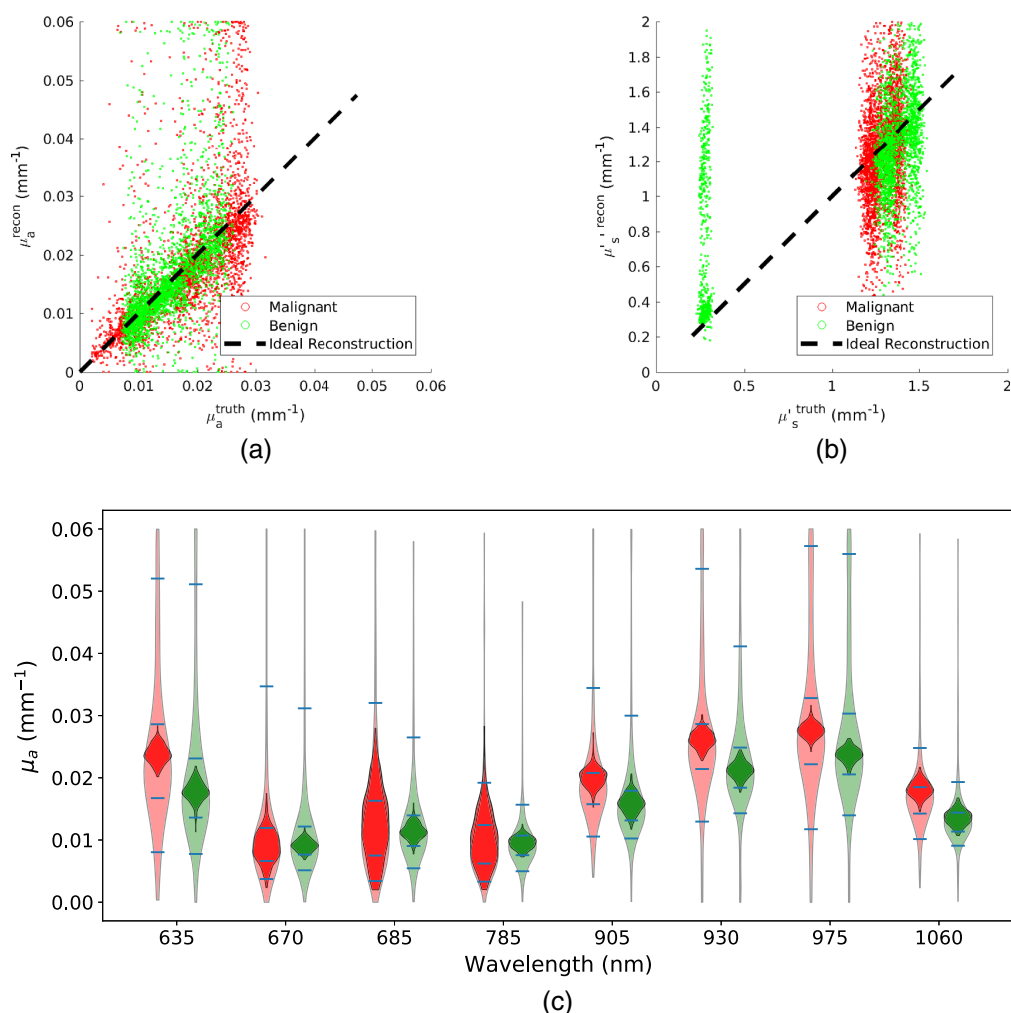


Fig. 7 Comparison of ground truth optical coefficients and reconstructed ones using the ground truth shape as two-region delimiter. In (a) and (b), a scatter plot of reconstructed inclusions versus ground truth values are shown. Each scatter plot consists of 724 lesions \times 8 wavelengths scatter points. As can be seen, points tend to cluster around the optimal behavior highlighted by the black dashed line. As expected, this behavior is more accentuated for absorption. (a) Scatter plot of ground truth lesion absorption versus reconstructed lesion absorption. (b) Scatter plot of ground truth lesion scattering versus reconstructed lesion scattering. (c) Violin plot of ground truth (dark) and retrieved (light) absorptions by wavelength. Green represents benign lesions and red malignant ones. Blue ticks represent the 5th, 25th, 75th and 95th percentiles of the distribution of the retrieved absorptions. (c) Display of the statistics of the ground truth and retrieved values by wavelength.

Table 4 Display of reconstruction error for the optical coefficients of the inclusion for a total of $N_{ph} \times N_\lambda \sim 5600$ reconstructions. While the average error for both absorption is around limited to 13%, variability in the performances is higher as highlighted by the standard deviation in parenthesis. We also show for both coefficients the fraction of reconstructions with absolute error higher than 50%. The value of this is around 18% of reconstructions when using the ground truth lesions shape in the fit and increases to 25% when using US.

Two-region	$\epsilon_{\mu_a} = \frac{\mu_a^{recon} - \mu_a^{truth}}{\mu_a^{truth}}$	$\frac{\#(\epsilon_{\mu_a} > 0.5)}{N_\lambda \times N_{ph}}$	$\epsilon_{\mu_s} = \frac{\mu_s^{recon} - \mu_s^{truth}}{\mu_s^{truth}}$	$\frac{\#(\epsilon_{\mu_s} > 0.5)}{N_\lambda \times N_{ph}}$
Ground	0.08 (0.64)	0.18	0.35 (0.79)	0.1
US	0.13 (0.76)	0.25	0.16 (0.89)	0.13

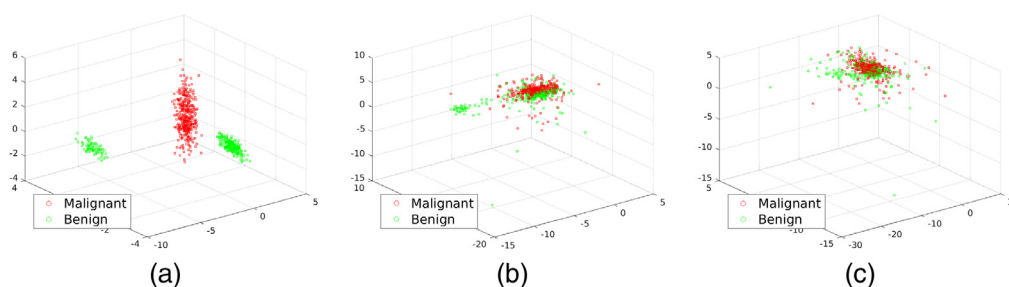


Fig. 8 PCA of ground truth and reconstructed values of the inclusion with the three axes representing the first three principal components of the dataset. The separation is neat on the ground truth. A certain degree of separation can also be observed after reconstruction with the ground truth shape defining the two regions. Results are apparently worse when defining the two-region extrapolating a shape from the US B-mode simulations. However, the dataset can have better separability in higher dimensions or applying a nonlinear transformation. (a) PCA of ground truth—log normalization. (b) PCA of nonlinear model fit—log normalization. (c) PCA of nonlinear model fit with US prior—log normalization.

Table 5 Overall results for classification. All considered methods can correctly classify all the lesions by giving their ground truth as input. Upon reconstruction with the ideal two-region definition, accuracy decreases to 83.5% in the best-case scenario, given by FCN. When the two-region definition is retrieved by the US image performances generally decrease. The best-case scenario is given by SVMs that reach an accuracy of 78%.

Model	Two-region	FCN (%)				SVM (%)				Log. Regr. (%)			
		Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1	Acc.	Prec.	Rec.	F1
Ground	—	100	100	100	100	100	100	100	100	100	100	100	100
Sep.Wav.	Ground	83.5	85.7	83.0	84.4	81.4	79.6	85.5	82.4	76.9	73.6	85.5	79.1
Sep.Wav.	US	75.4	76.1	79.7	77.9	78.0	79.3	76.7	78.0	69.1	68.5	72.5	70.4

In Fig. 8, we show the scatter plot of the first three components for (i) ground truth, (ii) optical coefficients in the case of reconstruction with ideal two-region model, and (iii) reconstruction with US-extrapolated two-region. The separation of the phantoms into benign and malignant is clear for the ground truth. A noticeable decrease in separation is observed upon optical reconstruction from optical data, especially in the case of US-extrapolated two-region model.

Results of the application of logistic regression, SVM, and the FCN described in Sec. 2.3 are shown in Table 5. In general, the FCN performs better than other proposed methods, at a cost of larger training time (5 min) and fine tuning of the hyperparameters. All the applied methods show an optimal prediction for what concerns the ground truth. When the coefficients are retrieved with our reconstruction method, all the figures appreciably decrease. The accuracy, when the two regions are identified by the ground truth, is over 80%. In general, recall is higher than the precision, so malignant lesions are more easily identified as such with respect to benign ones.

4 Conclusion

We proposed a simulation pipeline based on VICTRE for the generation of realistic US B-mode images and DOT data. The method allows to generate a dataset for a newly designed probe on a shorter time scale and with larger variability than that which is obtainable experimentally in clinics. We generated a total of 349 benign and 379 malignant samples complete of B-mode images and optical data. We assessed a two-region fitting reconstruction method on the

simulated data where the two-region system was extracted from the B-mode images. The procedure of segmentation and 3D prior extraction was evaluated. On average, the retrieved volumes resulted to underestimate the physical dimensions of the ground truth lesions of about 50% in volume. The displacement with respect to the center of mass of the lesions resulted to be less than the resolution of the computational grid of the optical reconstructions of 2 mm used here. The utilized optical reconstruction method does not suffer from model approximation error as the optical contrast between bulk and inclusion increases. In general, the reconstructed values of the inclusions approximate the actual values of the ground truth with somewhat frequent exceptions, absorption relative errors higher than 50% are present in 18% of reconstructions with an ideal two-region definition and 25% when the two regions are extracted from a US B-mode image. The number of reconstruction with relative errors higher than 50% generally increase when using also measurements coming from shorter source–detector distances in reconstruction, suggesting a limit of the two-region model. Nonetheless, a certain degree of separation between benign and malignant inclusions is still observable when doing principal component analysis on the log-normalized data. Three classification methods have been tested on the results of the reconstructions: an SVM, an FCN, and logistic regression. SVM and logistic regression show comparable results, while FCN has a better classification performances at the cost of a higher training time. The classification displays an accuracy close to 84% when using an ideal two-region reconstruction. When the two-region system is inferred by the US, a worsening in the performances of classification is observable, bringing the accuracy to 78%. Future steps will explore other methods for the extrapolation of the two regions from the US B-mode images. The availability of large datasets of lesion shapes and of corresponding US B-mode images allows to speculate the refinement of the extrapolation procedure with 3D shape reconstruction techniques utilizing supervised machine learning approaches.⁷⁸ Other approaches for a possible improvement for the quantification of lesions optical properties may include the integration of a spectral model also in reconstruction as in Ref. 79. Future steps in the simulations may include more realistic ways to simulate the compression of breasts to make it better resemble the one given by handheld probes.

Disclosures

The authors have no conflicts of interest to declare.

Acknowledgments

This project has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No. 731877 and EPSRC grant EP/T014369/1. The authors wish to thank Professor Ben Cox and Professor Brad Treeby for helpful advice on US modeling systems.

Code, Data, and Materials availability

The code for the generation of digital phantoms and for optical and acoustic ground can be found at: https://github.com/giudisciacca/victre_codes. The code builds on the software suite VICTRE that can be found at: <https://github.com/DIDSR/VICTRE>. The code for the generation of US B-mode images requires the installation of K-wave and can be found at: https://github.com/giudisciacca/CLUSTER_COMPUTATION/tree/master/k-wave/myscripts/USfromVitre. The code for the segmentation of the images is at https://github.com/giudisciacca/CLUSTER_COMPUTATION/tree/master/k-wave/myscripts/USfromVitre, the generation of Diffuse Optical data and the can be found at: https://github.com/giudisciacca/CLUSTER_COMPUTATION/tree/master/SOLUS/example/VICTRE_SIMULATION_AND_RECON. It requires the installation of TOAST++ <http://web4.cs.ucl.ac.uk/research/vis/toast/> and builds upon the software suite SOLUS: <https://github.com/andrefarina/SOLUS>. Codes for classification can be found in: https://github.com/giudisciacca/victre_codes/tree/main/pyClassify.

References

1. A. Yodh and B. Chance, "Spectroscopy and imaging with diffusing light," *Phys. Today*, **48**, 34–41 (1995).
2. D. A. Boas et al., "Imaging the body with diffuse optical tomography," *IEEE Signal Process. Mag.* **18**(6), 57–75 (2001).
3. A. Pifferi et al., "New frontiers in time-domain diffuse optics, a review," *J. Biomed. Opt.* **21**(9), 091310 (2016).
4. A. D. Mora et al., "Towards next-generation time-domain diffuse optics for extreme depth penetration and sensitivity," *Biomed. Opt. Express* **6**(5), 1749–1760 (2015).
5. A. Pifferi et al., "SOLUS: an innovative multimodal imaging system to improve breast cancer diagnosis through diffuse optics and ultrasounds," *Proc. SPIE* **11639**, 116390C (2021).
6. S. K. V. Sekar et al., "Broadband (600–1350 nm) time-resolved diffuse optical spectrometer for clinical use," *IEEE J. Sel. Top. Quantum Electron.* **22**, 7100609 (2016).
7. V. Quaresima and M. Ferrari, "Functional near-infrared spectroscopy (fNIRS) for assessing cerebral cortex function during human behavior in natural/social situations: a concise review," *Organ. Res. Methods* **22**, 46–68 (2019).
8. T. Durduran et al., "Diffuse optics for tissue monitoring and tomography," *Rep. Prog. Phys.* **73**(7), 076701 (2010).
9. G. Quarto et al., "Estimate of tissue composition in malignant and benign breast lesions by time-domain optical mammography," *Biomed. Opt. Express* **5**(10), 3684–3698 (2014).
10. S. Fantini, B. Frederick, and A. Sassaroli, "Perspective: prospects of non-invasive sensing of the human brain with diffuse optical imaging," *APL Photonics* **3**, 110901 (2018).
11. H. Vavadi et al., "Compact ultrasound-guided diffuse optical tomography system for breast cancer imaging," *J. Biomed. Opt.* **24**(2), 021203 (2018).
12. M. L. Flexman et al., "Monitoring early tumor response to drug therapy with diffuse optical tomography," *J. Biomed. Opt.* **17**(1), 016014 (2012).
13. W.-H. Yuan et al., "Supplemental breast cancer-screening ultrasonography in women with dense breasts: a systematic review and meta-analysis," *Br. J. Cancer* **123**(4), 673–688 (2020).
14. S. R. Arridge and M. Schweiger, "Reconstruction in optical tomography using MRI based prior knowledge," in *Information Processing in Medical Imaging '95*, Y. Bizais, C. Barillot, and R. diPaola, Eds., pp. 77–88, Kluwer Academic Publishers, Dordrecht (1995).
15. Q. Zhu et al., "Imager that combines near-infrared diffusive light and ultrasound," *Opt. Lett.* **24**(15), 1050–1052 (1999).
16. V. Ntziachristos, X. H. Ma, and B. Chance, "Time-correlated single photon counting imager for simultaneous magnetic resonance and near-infrared mammography," *Rev. Sci. Instrum.* **69**, 4221 (1998).
17. V. Ntziachristos et al., "MRI-guided diffuse optical spectroscopy of malignant and benign breast lesions," *Neoplasia* **4**, 347–354 (2002).
18. L. Zhang et al., "Direct regularization from co-registered anatomical images for MRI-guided near-infrared spectral tomographic image reconstruction," *Biomed. Opt. Express* **6**, 3618–3630 (2015).
19. S. R. Cherry, "Multimodality in vivo imaging systems: twice the power or double the trouble?" *Annu. Rev. Biomed. Eng.* **8**, 35–62 (2006).
20. M. Althobaiti, H. Vavadi, and Q. Zhu, "Diffuse optical tomography reconstruction method using ultrasound images as prior for regularization matrix," *J. Biomed. Opt.* **22**(2), 026002 (2017).
21. S. R. Arridge, M. J. Ehrhardt, and K. Thielemans, "(An overview of) synergistic reconstruction for multimodality/multichannel imaging methods," *Philos. Trans. R. Soc. A* **379**, 20200205 (2020).
22. M. Huang and Q. Zhu, "Dual-mesh optical tomography reconstruction method with a depth correction that uses a priori ultrasound information," *Appl. Opt.* **43**(8), 1654–1662 (2004).
23. N. G. Chen et al., "Simultaneous near-infrared diffusive light and ultrasound imaging," *Appl. Opt.* **40**(34), 6367–6380 (2001).

24. Q. Zhu, N. G. Chen, and S. H. Kurtzman, "Imaging tumor angiogenesis by use of combined near-infrared diffusive light and ultrasound," *Opt. Lett.* **28**(5), 337–339 (2003).
25. S. Li, M. Zhang, and Q. Zhu, "Ultrasound segmentation-guided edge artifact reduction in diffuse optical tomography using connected component analysis," *Biomed. Opt. Express* **12**(8), 5320–5336 (2021).
26. A. Badano et al., "In silico imaging clinical trials for regulatory evaluation: initial considerations for VICTRE, a demonstration study," *Proc. SPIE* **10132**, 1013220 (2017).
27. B. A. Lau et al., "A statistically defined anthropomorphic software breast phantom," *Med. Phys.* **39**(6Part1), 3375–3385 (2012).
28. A. Badano et al., "Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial," *JAMA Network Open* **1**(7), e185474 (2018).
29. T. Han et al., "A three-dimensional modeling method for quantitative photoacoustic breast imaging with handheld probe," *Photoacoustics* **21**, 100222 (2021).
30. Y. Bao et al., "Development of a digital breast phantom for photoacoustic computed tomography," *Biomed. Opt. Express* **12**(3), 1391–1406 (2021).
31. S. Li et al., "Edge artifact removal in diffuse optical tomography using the heterogeneous VICTRE digital breast model," *Proc. SPIE* **11639**, 1163905 (2021).
32. "Smart optical and ultrasound diagnostics of breast cancer," <http://www.solus-project.eu> (accessed 6 December 2021).
33. S. Park et al., "Realistic three-dimensional optoacoustic tomography imaging trials using the VICTRE breast phantom of FDA (conference presentation)," *Proc. SPIE* **11240**, 112401H (2020).
34. D. Sharma et al., "In silico imaging tools from the victre clinical trial," *Med. Phys.* **46**(9), 3924–3928 (2019).
35. L. de Sisternes et al., "A computational model to generate simulated three-dimensional breast masses," *Med. Phys.* **42**(2), 1098–1118 (2015).
36. J. A. Jensen, "Simulation of advanced ultrasound systems using field II," in *2nd IEEE Int. Symp. Biomed. Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, IEEE, pp. 636–639 (2004).
37. D. Garcia, "SIMUS: an open-source simulator for ultrasound imaging. Part I: Theory & examples," in *Computer Methods and Programs in Biomedicine*, p. 106726, Elsevier (2022).
38. F. Varray et al., "Creanuis: a non-linear radiofrequency ultrasound image simulator," *Ultrasound Med. Biol.* **39**(10), 1915–1924 (2013).
39. B. E. Treeby and B. T. Cox, "k-wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields," *J. Biomed. Opt.* **15**(2), 021314 (2010).
40. P. R. Stepanishen, "Transient radiation from pistons in an infinite planar baffle," *J. Acoust. Soc. Am.* **49**(5B), 1629–1638 (1971).
41. J. A. Jensen, *Linear Description of Ultrasound Imaging Systems: Notes for the International Summer School on Advanced Ultrasound Imaging at the Technical University of Denmark*, Technical University of Denmark, Department of Electrical Engineering (1999).
42. J. A. Jensen, "A model for the propagation and scattering of ultrasound in tissue," *J. Acoust. Soc. Am.* **89**(1), 182–190 (1991).
43. F. Varray et al., "Fundamental and second-harmonic ultrasound field computation of inhomogeneous nonlinear medium with a generalized angular spectrum method," *IEEE Trans. Ultrasonics, Ferroelectr. Freq. Control* **58**(7), 1366–1376 (2011).
44. F. Varray et al., "Simulation of ultrasound nonlinear propagation on GPU using a generalized angular spectrum method," *EURASIP J. Image Video Process.* **2011**(1), 1–6 (2011).
45. B. E. Treeby et al., "Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method," *J. Acoust. Soc. Am.* **131**(6), 4324–4336 (2012).
46. B. E. Treeby and B. T. Cox, "Modeling power law absorption and dispersion for acoustic propagation using the fractional laplacian," *J. Acoust. Soc. Am.* **127**(5), 2741–2748 (2010).
47. R. T. Beyer, "Parameter B/A," *Nonlinear Acoustics*, Vol. 25, p. 40, Academic Press (1997).

48. A. Cigier, F. Varray, and D. Garcia, "SIMUS: an open-source simulator for ultrasound imaging. Part II: comparison with three popular simulators," arXiv:2103.05521 (2021).
49. C. M. Sehgal et al., "A review of breast ultrasound," *J. Mammary Gland Biol. Neoplasia* **11**(2), 113–123 (2006).
50. J. Mamou and M. L. Oelze, *Quantitative Ultrasound in Soft Tissues*, Springer (2013).
51. T. Hopp, N. V. Ruiten, and N. Duric, "Breast tissue characterization by sound speed: correlation with mammograms using a 2D/3D image registration," in *IEEE Int. Ultrasonics Symp.*, IEEE, pp. 1–4 (2012).
52. F. Martelli et al., *Light Propagation Through Biological Tissue and Other Diffusive Media: Theory, Solutions and Software*, SPIE Press, Bellingham, Washington (2010). DOI: [10.1117/3.824746](https://doi.org/10.1117/3.824746)
53. S. R. Arridge, "Methods in diffuse optical imaging," *Philos. Trans. R. Soc. A* **369**, 4558–4576 (2011).
54. R. Graaff and J. J. Ten Bosch, "Diffusion coefficient in photon diffusion theory," *Opt. Lett.* **25**(1), 43–45 (2000).
55. R. Aronson and N. Corngold, "Photon diffusion coefficient in an absorbing medium," *J. Opt. Soc. Am. A* **16**(5), 1066–1071 (1999).
56. M. Bassani et al., "Independence of the diffusion coefficient from absorption: experimental and numerical evidence," *Opt. Lett.* **22**(12), 853–855 (1997).
57. K. Furutsu and Y. Yamada, "Diffusion approximation for a dissipative random medium and the applications," *Phys. Rev. E* **50**(5), 3634 (1994).
58. S. R. Arridge et al., "A finite element approach for modeling photon transport in tissue," *Med. Phys.* **20**(2), 299–309 (1993).
59. E. Martinenghi et al., "Spectrally resolved single-photon timing of silicon photomultipliers for time-domain diffuse spectroscopy," *IEEE Photonics J.* **7**(4), 1–12 (2015).
60. M. Schweiger and S. R. Arridge, "The Toast++ software suite for forward and inverse modeling in optical tomography," *J. Biomed. Opt.* **19**(4), 040801 (2014).
61. A. Corlu et al., "Diffuse optical tomography with spectral constraints and wavelength optimization," *Appl. Opt.* **44**(11), 2082–2093 (2005).
62. C. D'Andrea et al., "Time-resolved spectrally constrained method for the quantification of chromophore concentrations and scattering parameters in diffusing media," *Opt. Express* **14**(5), 1888–1898 (2006).
63. A. Torricelli et al., "In vivo optical characterization of human tissues from 610 to 1010 nm by time-resolved reflectance spectroscopy," *Phys. Med. Biol.* **46**(8), 2227 (2001).
64. P. Taroni, "Deliverable 2.1: definition of paradigms representing exemplary breast lesions cases," 2017, <http://solus-project.eu/storage/app/media/D2.1%20Definition%20of%20paradigms%20representing%20exemplary%20breast%20lesions%20cases.pdf>.
65. M. Schweiger, S. R. Arridge, and I. Nissilä, "Gauss–Newton method for image reconstruction in diffuse optical tomography," *Phys. Med. Biol.* **50**(10), 2365 (2005).
66. L. Spinelli et al., "Bulk optical properties and tissue components in the female breast from multi-wavelength time-resolved optical mammography," *J. Biomed. Opt.* **9**(6), 1137–1142 (2004).
67. A. D. Mora et al., "The SiPM revolution in time-domain diffuse optics," *Nucl. Instrum. Methods Phys. Res. Sect. A: Accel., Spectrom., Detect. Assoc. Equip.* **978**, 164411 (2020).
68. The MathWorks, *Matlab Optimization Toolbox*, The MathWorks, Natick, Massachusetts (2017).
69. K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, MIT Press (2012).
70. F. Pedregosa et al., "Scikit-learn: machine learning in Python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
71. M. Abadi et al., "TensorFlow: large-scale machine learning on heterogeneous systems," 2015, [tensorflow.org](https://www.tensorflow.org).
72. K. Janocha and W. M. Czarnecki, "On loss functions for deep neural networks in classification," <http://arxiv.org/abs/1702.05659> (2017).
73. G. Di Sciacca et al., "Enhanced diffuse optical tomographic reconstruction using concurrent ultrasound information," *Philos. Trans. R. Soc. A* **379**(2204), 20200195 (2021).
74. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vision* **1**(4), 321–331 (1988).

75. L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).
76. A. P. Zijdenbos et al., "Morphometric analysis of white matter lesions in mr images: method and validation," *IEEE Trans. Med. Imaging* **13**(4), 716–724 (1994).
77. A. Carass et al., "Evaluating white matter lesion segmentations with refined Sørensen-Dice analysis," *Sci. Rep.* **10**(1), 8242 (2020).
78. X. Han, H. Laga, and M. Bennamoun, "Image-based 3D object reconstruction: state-of-the-art and trends in the deep learning era," *IEEE Trans. Pattern Anal. Mach. Intell.* **43**, 1–27 (2021).
79. G. D. Sciacca et al., "Fitting a spectral model for component analysis in diffuse optical tomography," *Proc. SPIE* **11074**, 1107426 (2019).

Giuseppe Di Sciacca is a PhD student at the Computer Science Department of University College London. He received his BS and MS degrees in engineering physics from Politecnico di Milano in 2014 and 2016, respectively. His research focuses in inverse problems arising in diffuse optical tomography and in multimodal imaging. Other research interests include the application of machine learning for imaging through diffusive media.

Giulia Maffeis is currently pursuing a PhD in physics at Politecnico di Milano (Italy), from which she previously received also her BS and MS degrees in engineering physics in 2016 and 2019, respectively. Her research interests include photon migration in turbid media, near-infrared non-invasive diffuse spectroscopy and imaging with time-resolved systems. She focuses on instruments and analytical models for optical mammography, being involved in clinical trials for breast cancer therapy monitoring and lesion discrimination.

Andrea Farina received his master's degree (cum laude) in electronics engineering from Politecnico di Milano, Italy, in 2005, and his PhD in physics from the same institution in 2009. Since December 2011, he has a permanent position as researcher at the Italian National Research Council (CNR) in Milan. His current research includes the development and the application of time-resolved techniques for the imaging and spectral characterization of diffusive media.

Alberto Dalla Mora received his MS degree in electronics engineering and his PhD in information and communication technology from Politecnico di Milano, Italy, in 2006 and 2010, respectively. He is currently an associate professor at the Physics Department of Politecnico di Milano. He has co-authored about 180 publications between peer-reviewed journals and conference proceedings. His main research interests are related to the development of innovative time-resolved diffuse optics techniques and instrumentation for biomedical applications.

Antonio Pifferi is a full-time professor in the Department of Physics, Politecnico di Milano, Italy. He received his MS degree in nuclear engineering from Politecnico di Milano in 1991, and his PhD in physics from Politecnico di Torino, Italy, in 1995. He is the author of about 200 papers in international peer-reviewed journals. His current research interests are based on time-domain diffuse optics at large, including tissue and material spectroscopy, diffuse Raman, and diffuse correlation spectroscopy.

Paola Taroni is full professor of physics at Politecnico di Milano (Milan, Italy) and was head of the PhD program in physics in 2013–2018. He is a co-author of about 140 scientific papers on international peer-reviewed journals (H-index 45). In the last years, her research activity focused on time domain diffuse optical spectroscopy and imaging, both for instrument development and diagnostic applications, with special attention to breast cancer management (lesion characterization, cancer risk estimate, therapy monitoring).

Simon Arridge is a professor of image processing in the University College London. He received his BA (Hons) in physics from Cambridge University in 1982 and his PhD in medical physics from University College London in 1990. His research interests include inverse problems, and he has worked two on applications to different medical imaging areas including optical tomography, photoacoustics, diffusion tensor MRI, SPECT, PET, and digital tomosynthesis.