2019

absO2luteU-Net: Tissue Oxygenation Calculation Using Photoacoustic Imaging and Convolutional Neural Networks

Kevin Hoffer-Hawlik
Kevin.A.Hoffer-Hawlik.19@Dartmouth.edu

Geoffrey P. Luke
Dartmouth College

absO$_2$luteU-Net: Tissue Oxygenation Calculation Using Photoacoustic Imaging and Convolutional Neural Networks

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of

Bachelor of Arts

in

Biomedical Engineering Sciences

by Kevin Hoffer-Hawlik

Thayer School of Engineering
Dartmouth College
Hanover, New Hampshire

Date:_____

Approved:_____
Advisor's Signature

_____
Signature of Author

THIS PAGE IS INTENTIONALLY LEFT BLANK,
UNCOUNTED AND UNNUMBERED

Abstract

Photoacoustic (PA) imaging is a novel biomedical imaging modality that uses incident light to generate ultrasound signals of photoabsorbers within tissues based on their absorption of optical energy. The images of optical absorption can then be used to calculate the concentration of specific photoabsorbers, which allow for measuring functional tissue parameters like blood oxygenation (sO_2). sO_2 is a valuable biomedical metric which aids in disease detection, prognosis, and treatment, such as metastatic cancer. However, calculating the optical absorption of a given tissue requires prior knowledge of the tissue to accurately estimate tissue fluence, or direct and costly computational methods to solve the non-linear fluence estimation problem. Recent work has shown that machine learning algorithms can estimate sO_2 with high accuracy and speed. absO_2luteU-Net, a convolutional neural network with a U-Net architecture designed to estimate tissue sO_2 from multispectral PA images, was trained, validated, and tested on Monte Carlo (MC) simulated PA data from randomized breast tissue. absO_2luteU-Net performed with higher accuracy compared to a baseline of simple linear unmixing and predicted sO_2 with remarkable speed (2 milliseconds per image), suggesting that machine learning algorithms can solve fluence estimation problems in PA imaging and can bring PA imaging from theoretical to clinical relevancy.

Preface

My research experiences at Dartmouth in biomedical engineering and medical imaging have challenged me through theoretical and technical difficulty, but have equivalently rewarded me through the opportunities for me to make my mark at Dartmouth in engineering. I originally studied fluorescent-guided surgery in Dr. Pogue's laboratory at Thayer, which gave me a background not only in medical imaging but in proper laboratory practices. After studying multiple other biomedical engineering fields, I decided to return to medical imaging to study photoacoustic imaging within Dr. Luke's Functional and Molecular Imaging Research Laboratory for my thesis project. This thesis project has been the culmination of my work in medical optics, machine learning, and biomedical science, and I could not have asked for a more interesting topic with powerful applications in cancer detection and treatment.

I would like to thank several people for their support of my senior honors thesis work. Firstly, I would like to thank Dr. Geoffrey Luke, Austin Van Namen, and the other PhD and undergraduate researchers in the FMI lab for their guidance and feedback. I also want to thank my professors and mentors for inspiring to take on this thesis project through their effort, teaching, direction, especially Dr. Pogue and Dr. Samkoe who initially oversaw my research in the Presidential Scholars Program. I am honored to give back to Dartmouth after gaining considerably from the college.

I would like to acknowledge the funding for my project generously donated by Undergraduate Advising & Research at Dartmouth through the Kaminsky Family Fund, as well as the Neukom Scholarship donated by the Neukom Institute for Computational Science.

Finally, I would also like to thank my friends and family that have supported me during my academic journey at Dartmouth. As I look back fondly on these past four years, I could not imagine a better supporting cast to cheer me on during the hard times, nor a group with which I would rather celebrate my successes. I look to the future with excitement, happy with what I am leaving behind and thrilled about what I have yet to experience.

Table of Contents

List of Tables

List of Figures

List of Acronyms

Photoacoustic	PA
Ultrasound	US
Machine Learning	ML
Monte Carlo	MC
Blood Oxygenation	sO_2
Mean Squared Error	MSE
Root Mean Squared Error	RMSE
Signal-Only Root Mean Squared Error	SO-RMSE
Simple Linear Unmixing	SLU
Signal to Noise Ratio	SNR

Introduction

Photoacoustic Imaging

Photoacoustic (PA) imaging is a novel biomedical imaging modality that uses incident light to generate ultrasound (US) signals in optical absorbers within tissues, which can be processed and imaged in real-time with high resolution. PA imaging takes advantage of the specificity of optical signal generators and the resolution and sensitivity of acoustic measurement instruments. Measuring the concentration of photoabsorbers allows for quantification of functional parameters within the tissue, e.g. blood oxygenation (sO_2). Quantification of these metrics would improve diagnostic and interventional medicine, such as tumor hypoxia as an adjunct therapy to chemotherapy and other radiation therapies (1, 2, 3).

Quantitative PA Imaging

A tissue's photoacoustic signal resulting from a single laser pulse, S_{PA}, is proportional to the rise in local pressure, p_0. If the laser pulse has a sufficiently short duration, p_0 is described by the following equation:

$$S_{PA}(x) \propto p_0(x) = \Gamma \cdot \mu_a(x) \cdot \phi(x, \mu_a, \mu_s, g) \quad (1)$$

where x is a given location within the tissue, Γ is the Grüneisen parameter (a thermodynamic property of the tissue describing how the tissue expands when exposed to heat), μ_a is the optical absorption coefficient (the total absorption due to chromophores), and ϕ is light fluence (laser energy per unit area). The optical absorption coefficient is typically proportional to the concentration of photoabsorbers of interest in a given tissue, and light fluence is a function of tissue location and the location's absorption, scattering, and anisotropy constants (which vary by the wavelength of the laser pulse) (4). However,

in order to quantitatively measure photoabsorber concentration more than a few millimeters deep within tissue, one must account for both the Grüneisen and fluence parameters. Specifically, fluence calculation poses a problem as fluence itself is a function of the tissue optical properties, and estimation models like Beer's Law and the Diffusion Approximation fail at increasing depths, in highly scattering media, and with decreased photoabsorber size (1).

Simple Linear Unmixing

The current standard for multispectral quantitative PA imaging is simple linear unmixing (SLU), in which the relative concentrations of the photoabsorbers of interest, oxygenated (HbO_2) and deoxygenated hemoglobin (HbR or Hb) in this case, are calculated from PA signals at multiple wavelengths in the near infrared range.

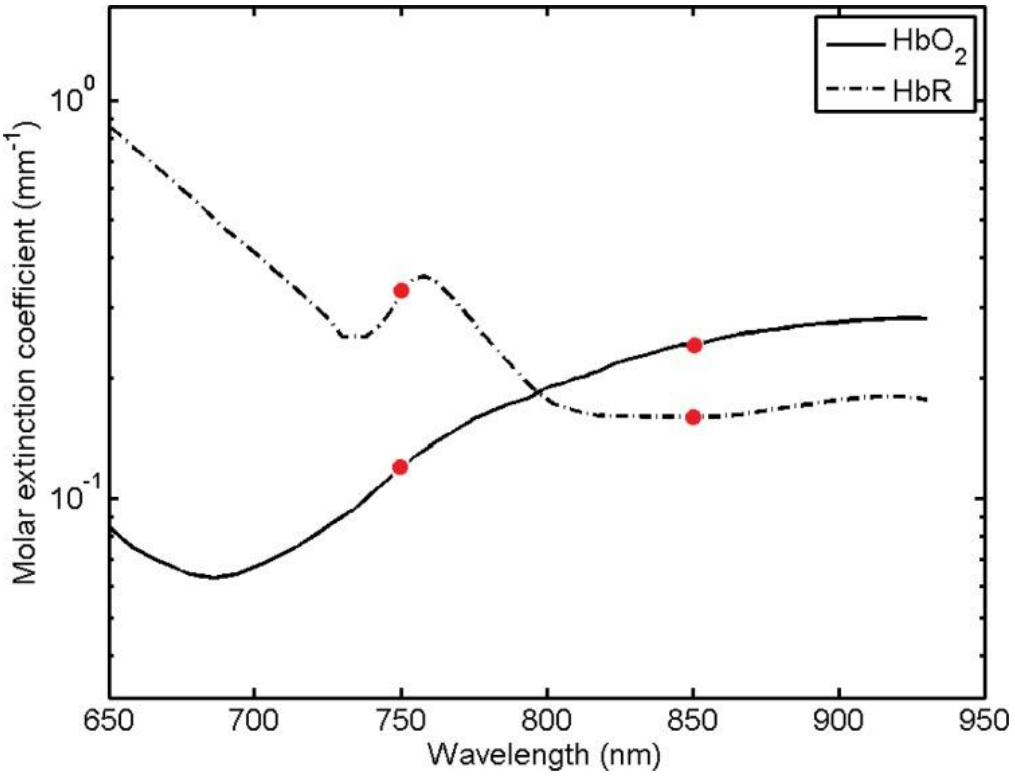

Figure 1: Absorption spectra for oxyhemoglobin (HbO_2) and deoxyhemoglobin (HbR) in the near-infrared spectrum (15).

Assuming only two optical absorbers are being estimated, two wavelengths can be used to estimate oxygen saturation with relatively high accuracy. The model does not compensate for local fluence, instead setting the quantity constant throughout the tissue, thus simplifying equation (1) to:

$$p_0(\lambda, x) = \phi \cdot [\varepsilon_{Hb}(\lambda) \cdot C_{Hb}(x) + \varepsilon_{HbO2}(\lambda) \cdot C_{HbO2}(x)] \quad (2)$$

where ε is the known molar extinction coefficient of HbO_2 and Hb at a given wavelength and C is the molar concentration of HbO_2 and Hb at each position x. After normalizing for the optical fluence and measuring PA signals from two wavelengths, equation (2) becomes a linear system of two equations with two unknowns. One can then calculate sO_2 from the estimated relative concentrations of HbO_2 and Hb using:

$$sO_2(x) = \frac{C_{HbO2}(x)}{C_{Hb}(x)+C_{HbO2}(x)}. \quad (3)$$

While it is straightforward and requires little computation, this basic model does not account for the variance of fluence at different wavelengths and position within the tissue, and thus SLU produces inaccurate results at increasing depths and with multiple present photoabsorbers (e.g. melanin in addition to HbO_2 and Hb) (2, 5).

Machine Learning Methods

Machine learning (ML) methods refer to training computer programs to complete digital tasks using loss-functions and iteratively changing calculational weights. Neural networks and deep learning refer to artificial networks of neurons acting as equations with changing weights, and neural networks with at least one hidden layer between input and output layers, respectively (6). More recent work has suggested that machine learning and neural network approaches can quickly and accurately solve the fluence estimation problem or bypass it entirely, directly calculating photoabsorber concentration from PA

signals (7, 8, 9). While the upfront time and the necessary amount of PA data needed to adequately train such ML models are high, the promise of quick functional tissue property predictions as well as high accuracy justifies this upfront cost.

Methods

Previous Approaches to ML and PA Imaging

While absO$_2$luteU-Net is not the first ML-based solution to fluence correction and oxygenation calculation based on multispectral PA images, the model builds upon prior work in the space. In 2017, Kirchner et al. used random forests to compute tissue fluence, estimate optical absorption, and subsequently calculate functional tissue parameters from multispectral data. While their work was the first machine learning based approach to quantitative PA imaging, the multiple steps between fluence estimation and oxygenation calculation could propagate errors, which can be prevented by directly calculating sO$_2$ rather than fluence. Additionally, convolutional neural networks (CNN) have since emerged as the leading approach to image recognition and prediction, and in theory can train with less data to perform with higher accuracy compared to random forests (7).

The same research group later used a simple fully-connected neural network to directly calculate sO$_2$ from a range of multispectral PA data. The group's results were of comparable accuracy compared to their previous approach. However, the use of 26 multispectral input images expanded the dataset, which likely lengthened the model's training time and could potentially limit its clinical applications given the need to image the same tissue 26 times to compute the sO$_2$. A simplified approach with two multispectral images should be sufficient to calculate oxygenation metrics. In addition, the fully connected neural network was potentially easier and quicker to train due to a simple

architecture, yet a CNN would improve accuracy with a minimal trade-off of marginally increased training time (8).

A separate group, Cai et al., presented an end-to-end model for sO_2 calculation from PA data named ResU-net in 2018. The group implemented a deeper neural network than previous works and adopted a residual learning mechanism to prevent a vanishing loss gradient preventing model convergence during training. While the model performed with mean errors below 1% and far below those of a SLU baseline, the model's performance deteriorated as the noise level within the PA image increased; that is to say, ResU-net did not optimally handle PA image noise below a signal-to-noise ratio (SNR) of 20 dB when reconstructing sO_2 for a given tissue. Using a shallower neural network could decrease the training time from hours to minutes, and training on noisier data (SNR below 40 dB) could increase the performance of a ML model on noisy data (9)

This thesis project set out to build a ML-based model to directly calculate sO_2 from multispectral PA signals. To that end, the project utilized similar approaches to data generation as the aforementioned works, namely Monte Carlo (MC) simulated PA signal images with added white Gaussian noise as an *in-silico* data set for training, validating, and testing. However, in comparison to previous ML models, absO_2luteU-Net was designed to train faster and with less data, and, more importantly, to predict sO_2 in real-time—on the scale of milliseconds rather than seconds (7) or hundredths of seconds (9)—with comparable if not better accuracy and noise suppression ability.

Convolutional Neural Network Model

We present absO_2luteU-Net, a CNN for calculating tissue sO_2 at each point from a PA image. CNNs, usually used to analyze images, are composed of multiple convolutional

layers that use filters to learn important image feature information. More specifically, absO$_2$luteU-Net is a modified implementation of a recent CNN architecture for biomedical image segmentation, the U-Net, depicted in Figure 2 (10).

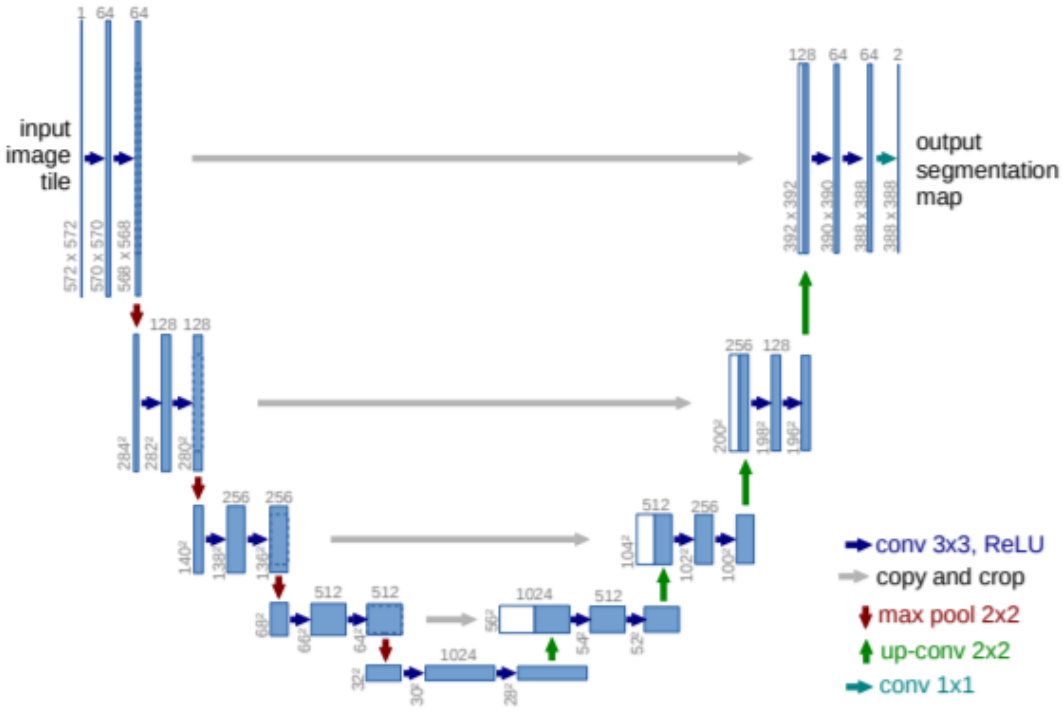

Figure 2: The original U-Net architecture, from Ronneberger et al. (10) Each blue box represents a multi-channel feature map with the channel number above and the dimension size on the bottom left. Each white box represents a copied feature map from the symmetric layer in the contracting path. The arrows represent different operations.

Similar to Ronneberger's implementation of a U-Net, absO$_2$luteU-Net is a CNN that features contracting and expanding paths. In contrast to the original U-net, absO$_2$luteU-Net features exponential linear units (ELU) instead of rectified linear units (ReLU) between convolutional blocks (two consecutive convolutional layers), and the input and output image size for absO$_2$luteU-Net is 128x128 and 128x128 rather than 572x572 and 388x388, respectively (11).

The contracting path consists of convolutional blocks of two consecutive 3x3 convolutions, followed by an exponential linear unit (ELU) and a 2x2 max pooling

operator. The 3x3 convolutional layers use filters to capture spatial dependencies and reduce the output image to only include those features. The max-pooling layers decrease the computational power needed for each successive layer in the contracting path by reducing image dimensions in half through selecting max values in every 2x2 window.

The expansive path is a symmetric inverse of the contracting path, consisting of a 2x2 up-convolution and two 3x3 convolutions followed by an ELU. The 2x2 up-convolution layers serve to perform the converse operation of the max-pooling layers, scaling the image back up. Between each layer is a dropout layer with dropout rates varying between 0.1 and 0.3, depending on the location within the U-Net, to prevent overfitting during training. Each block of the expanding arm is concatenated with the same block from the contracting arm to transfer contextual information. The number of feature channels doubles after each down-sampling step and is halved after each up-sampling step. Finally, a 1x1 convolution layer produces the output image (10).

PA Data Generation Pipeline

Tissue Creation

Using a MATLAB script (Appendix A), 125 128x128x128 digital phantom tissue cubes were randomly created to contain and simulate the optical properties of 3.84 cm-length cube tissues containing homogenous layers of breast tissue (dermis, epidermis, and breast tissue). 16 tissue types were constructed to represent air, water, epidermis, dermis, breast, and incrementally oxygenated blood vessels. The absorption coefficient, scattering coefficient, and anisotropy was calculated from the significant optical absorber components (volume fraction of fat, water, blood, and oxygen saturation of hemoglobin), for each voxel in the simulated tissue cube. One to three cylindrical blood vessels of radius

between 0.5 and 4 mm were randomly added and randomly oriented within the tissue, with sO2 values ranging from 0 to 100% in increments of 10%. A default tissue type schematic from mcxyz is depicted in Figure 3(a).

mcxyz Simulations

Mcxyz, a MC light transport simulation program written by S. Jacques, was modified to generate *in-silico* PA signals for each pixel in a digital phantom tissue. A PA image from a 10-minute, 825,000-photon simulation for the default tissue is shown in Figure 3(b) (12). 700 nm and 900 nm wavelengths were chosen for spectral unmixing, and each simulation was run twice to generate PA signal maps for the same tissue for both wavelengths (5). 10^6 photons per simulation were ultimately used to balance between accuracy and simulation time (13). Each simulation was run for 30 minutes to reach at least 10^6 photons. Simulations were run on the MATLAB generated tissues on a 0.3 mm simulation grid, for a final tissue dimension size of $(38.4\text{mm})^3$. The simulated laser source was configured to match a converging pair of 0.22 numeric aperture, 36 mm by 1.5 mm rectangular fibers held 1 cm above the tissue surface at a 30-degree angle. As two laser sources could not be simulated at once within mcxyz, the beam profile of the laser pair was calculated, and a simulated rectangular beam was constructed to have the same profile.

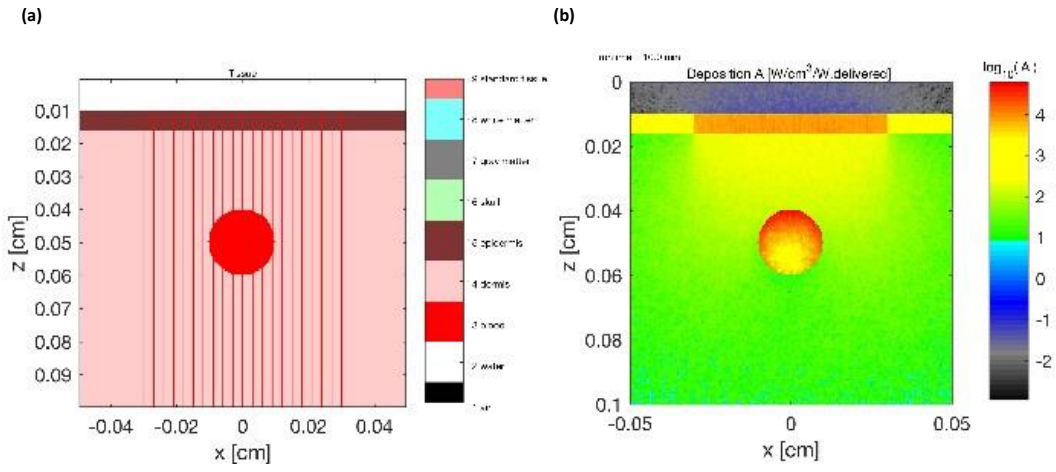

Figure 3. Sample input phantom tissue and PA output image from mcxyz by S. Jacques. Both images are cross-sections of the digital phantom tissue cube at y = 64 (halfway in the y-direction). (a) is a schematic of the different tissue types, and (b) is a map of the PA signal at each pixel.

Data Processing

A 3x3x3 median filter was used to process each PA signal map to suppress noise arising from simulating a finite number of photons. The 3D 128x128x128 images were cropped into 2D 128x128 images at halfway cross-sections to match the lab's clinical imaging system based on a linear array ultrasound transducer. Each 700 nm and 900 nm image pair was normalized by the maximum absorbance value found in both images. Finally, to mimic natural randomness throughout the PA image, white Gaussian noise was added with variable SNRs:

$$SNR_{dB} = 10 \log_{10} \frac{S_{PA}^2}{\sigma^2} \quad (4)$$

where σ represents the standard deviation of the Gaussian distribution characterizing the noise found within the PA image. The SNR for the training images was randomly chosen between 0dB ($S_{PA} = \sigma$) and 20dB ($S_{PA} = 10 \times \sigma$), and the SNR for testing images was chosen between 0dB, 4dB, 8dB, 12dB, 16dB, and 20dB, with two images per noise level.

Training Data Shaping

In order for absO$_2$luteU-Net to be trained to reconstruct sO$_2$ at all points within the tissue from ultrasound-measured PA signals, the data was organized into an input set of *x* of tuples containing the 700 nm and 900 nm simulated PA images, and an output set *y* of ground truth sO$_2$ pixel-wise maps, shown in Figure 4.

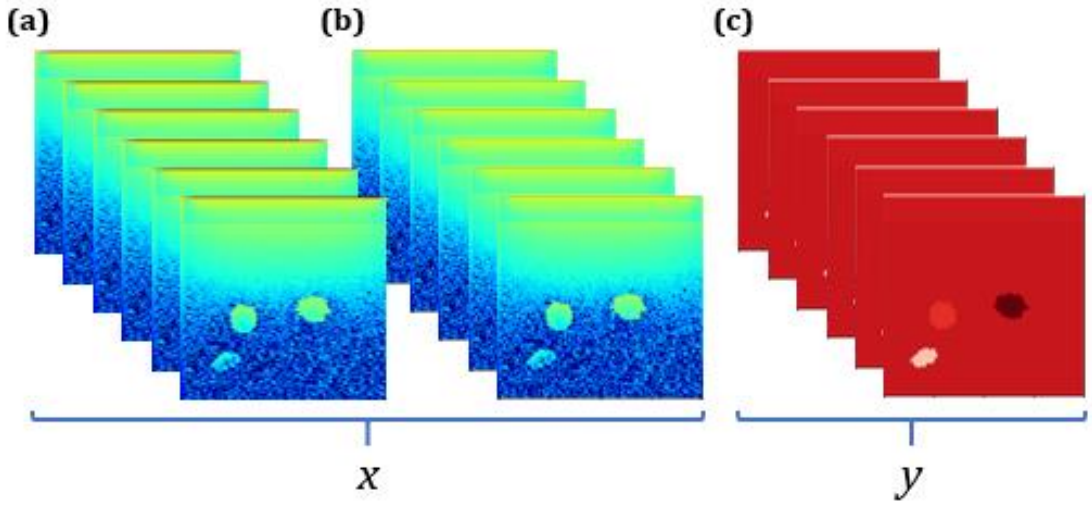

Figure 4. The 125 data set separated into *x* input PA images and *y* ground truth sO_2 output maps. (a) and (b) are the mcxyz-generated PA pressure maps for 700 nm and 900 nm light, respectively. (c) are the ground truth sO_2 maps, represented as cross-sections of variously oxygenated blood vessels within a constantly oxygenated background of breast tissue.

Baseline Model

Simple linear unmixing (SLU) was used as the baseline model. A MATLAB script was created to normalize each image by the maximum absorption value (thereby accounting for optical fluence) and solve the linear spectral fitting equations for HbO_2 and Hb for each image. Oxygenation was then calculated according to equation (3) (2, 5).

Training and Validation

absO_2luteU-Net was created, trained, validated, and tested in Keras using a TensorFlow backend. An 80/10/10 split of the 125 PA signal data was implemented, for 100 training data, 13 validating data, and 12 testing data. Mean squared error (MSE) was used for the loss function, and an Adam optimizer was used with parameters set to the suggested values in the original paper (14). absO_2luteU-Net was trained for 100 epochs with a batch size of 1 image (for stochastic gradient descent training), for a total of 7.83 minutes of training, on a 4GB NVIDIA Quadro M4000 GPU.

Accuracy and Computational Speed Metrics

Both pixel-wise root mean squared error (RMSE) and a signal-only adaptation of pixel-wise root mean squared error (SO-RMSE) were used to quantify the accuracy of absO$_2$luteU-Net, and prediction time per image was used to quantify the speed of absO$_2$luteU-Net's predictions. RMSE gives the absolute error in predicted sO$_2$ throughout the entire image, in terms of percentage sO$_2$. This metric is useful for evaluating the overall effectiveness of models in detecting background tissue oxygenation. Additionally, as the background breast tissue for each simulation featured constant tissue types, a SO-RMSE calculation was defined to limit the RMSE calculation to regions with pixels containing only blood vessels and no constant background tissue. This metric is useful for evaluating the ability of each model to detect hypoxic cores and sO$_2$ in specific regions of interest (i.e., blood vessels). Both RMSE and SO-RMSE were calculated for the 12 testing images using both absO$_2$luteU-Net's predicting methods and SLU, and averages for both metrics were taken for the 12 testing images.

For each method, a program was implemented to measure how long each prediction would take. absO$_2$luteU-Net's predicting method and the SLU MATLAB program were run on the 12 previously unseen images in Python and MATLAB, respectively, with a timer set to end after sO$_2$ predictions were made for all 12 test images. The resulting time was divided by 12 to get an average prediction speed per image metric.

Results

Computational Speed and Accuracy Comparison

After training and validation, absO$_2$luteU-Net predicted pixel-wise quantitative sO$_2$ maps for the 12 previously unseen test PA image tuples in 2 milliseconds per image on average,

with a per-image average RMSE of 4.49% and SO-RMSE of 18.4%. SLU predicted sO_2 in 20 milliseconds per image on average, with 75.5% RMSE per-image and 64.8% SO-RMSE per image. A tabular summary of the accuracy and speed metrics for each method is shown in Table 1. A visual comparison for a representative test datum is shown in Figure 5. An error map for the representative datum is shown in Figure 6.

Table 1. Summary of Performance Metrics for absO2luteU-Net and SLU (baseline)

Performance Metric	**Average absO2luteU-Net Performance**	**Average SLU Performance**
Training Time	7.83 minutes	N/A
Average RMSE (n = 12)	4.49 %	75.5 %
Average SO-RMSE (n = 12)	18.4 %	64.8 %
Average Reconstruction Time (n = 12)	2 ms	20 ms

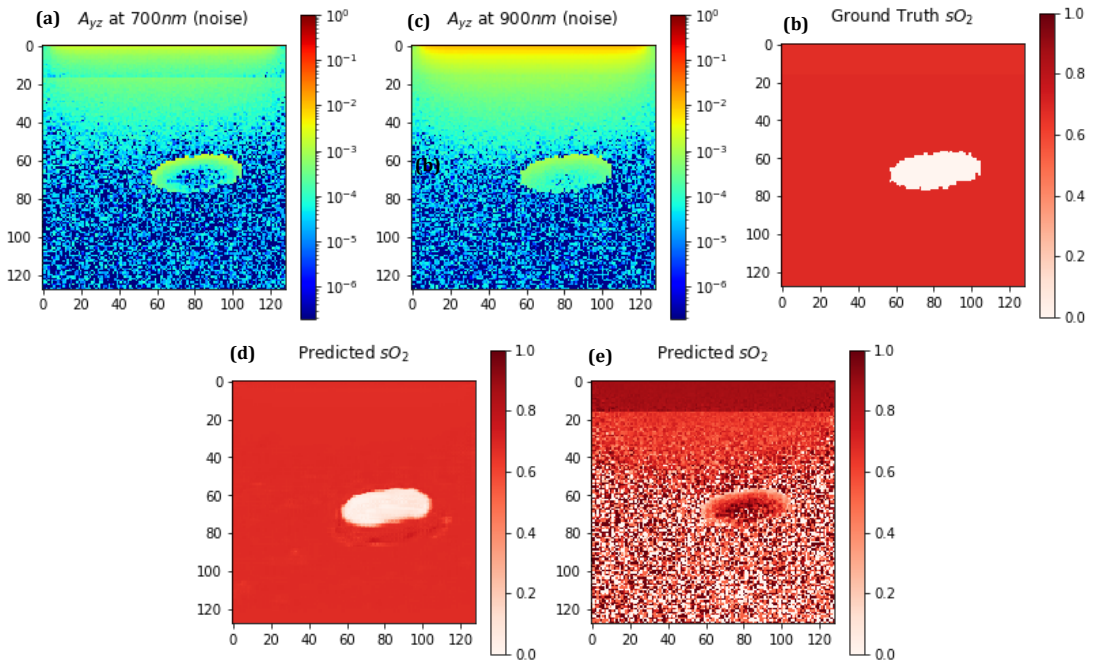

Figure 5. Representative sO_2 prediction output from absO2luteU-Net compared to SLU baseline. (a) and (b) are from the input tuple of 700 nm and 900 nm PA images, respectively. (c) is the ground truth used to evaluate prediction accuracy for the two models. (d) is the predicted sO_2 map from absO2luteU-Net, and (e) is the predicted sO_2 map from SLU. The image had an SNR of 8 dB. RMSE was 3.92% for absO2luteU-Net and 79.8% for SLU. SO-RMSE was 16.7% for absO2luteU-Net and 63.6% for SLU.

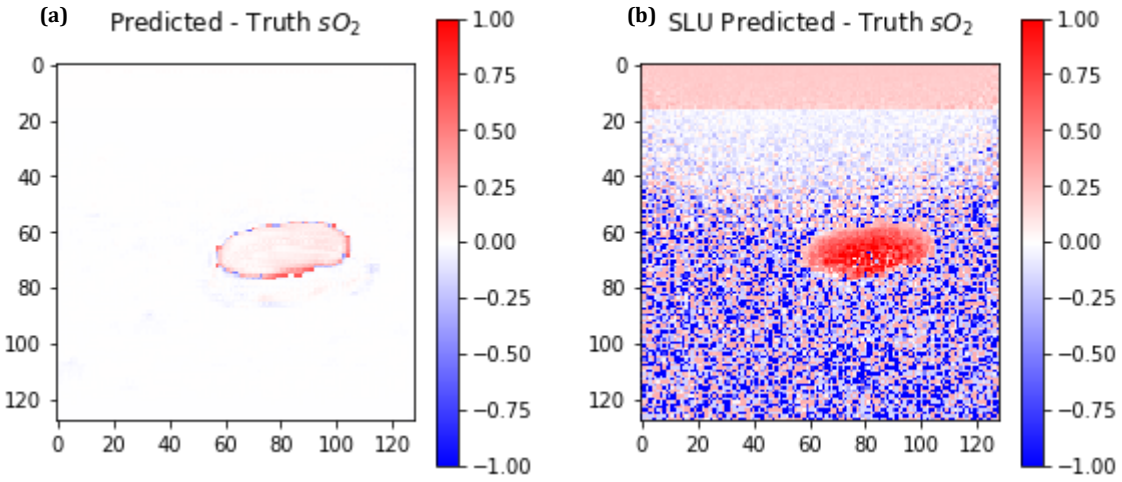

Figure 6: sO_2 Prediction Error Maps for absO2luteU-Net and SLU. The error maps for absO2luteU-Net (a) and SLU (b) plot the difference between the predicted and ground truth sO_2 values, thus displaying how far off the prediction was from the truth in terms of sO_2.

Noise Suppression

For the SLU baseline, pixel-wise RMSE decreased as image SNR increased from 0dB to 20dB, as expected, but there was no clear relationship between SO-RMSE and SNR. For absO2luteU-Net, pixel-wise RMSE remained consistently at or below 6%, and there was no clear relationship between SO-RMSE and SNR. Graphs displaying trends between RMSE, SO-RMSE, and SNR are shown in Figures 7 and 8.

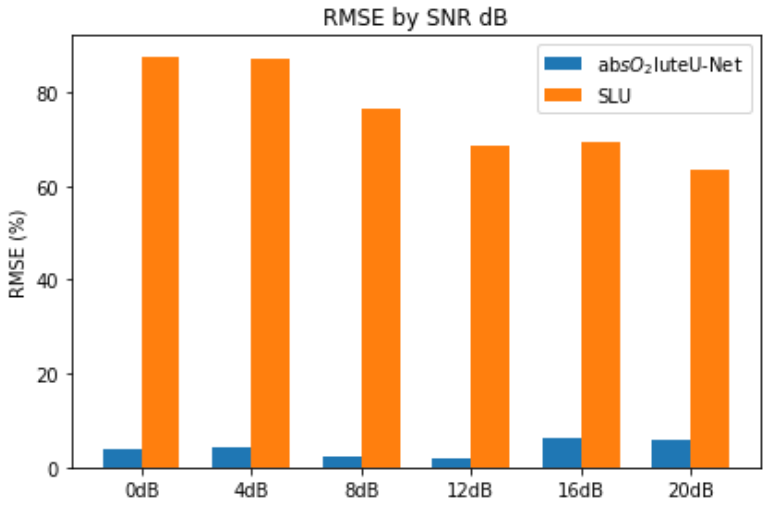

Figure 7: RMSE as a function of PA image SNR. For each SNR level, n = 2.

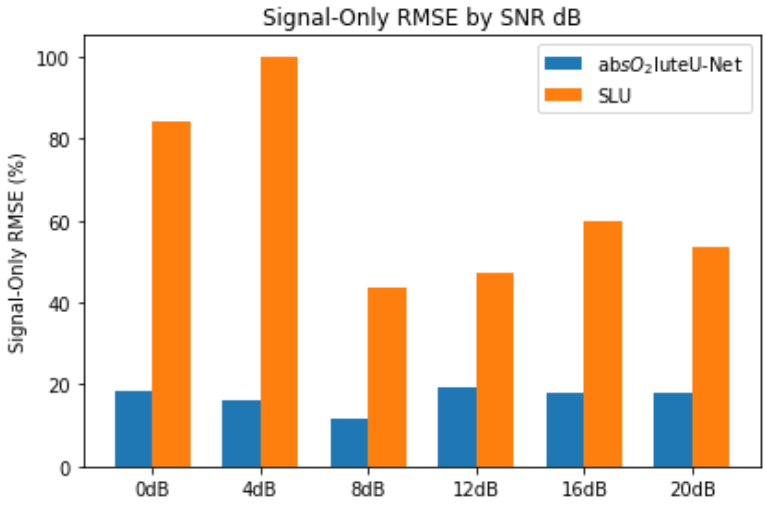

Figure 8: SO-RMSE as a function of PA image SNR. For each SNR level, n = 2.

Discussion

absO$_2$luteU-Net outperformed the baseline of SLU, as its predictions had average RMSE and SO-RMSE values both less than those of SLU. SLU did not take into account the melanin in the epidermal layer of the simulated breast tissue, so the model detected much higher levels of optical absorbance and reported close to 100% levels of sO$_2$. absO$_2$luteU-Net, in contrast, could and did adapt to account for unknown photoabsorbers like melanin. absO$_2$luteU-Net predicted background (non-signal) sO$_2$ with very low error, as shown by the error plots in Figure 6. Moreover, the CNN shows preliminary noise suppression, as the model makes predictions with less than 10% RMSE across samples with SNRs as low as 0dB.

The accuracy of absO$_2$luteU-Net is matched by its computational efficiency. For a given laboratory PA imaging setup, a laser can pulse with a maximum repetition rate of 10 Hz, for a laser pulse every 0.1 seconds; absO$_2$luteU-Net can predict sO$_2$ maps from the resulting US image in 2 milliseconds on average, 50 times faster than the maximum rate of PA data generation for that given setup. In theory, absO$_2$luteU-Net can be trained

beforehand and imported into a PA imaging device in order to mitigate the upfront cost of training and maximize the efficiency of its predictions.

Conclusions

In this report we describe absO$_2$luteU-Net, a convolutional neural network taking advantage of recent advances in biomedical image segmentation to solve the fluence estimation problem in photoacoustic imaging. Previous models either sacrifice accuracy for speed by discounting variation in local fluence, or sacrifice speed for accuracy by using computational expensive methods, but absO$_2$luteU-Net can make sO$_2$ predictions with both high accuracy and low processing time. We believe absO$_2$luteU-Net has the potential to further bring quantitative PA imaging from the laboratory to clinical application (7, 8, 9).

While other recent deep neural network approaches to quantitative PA imaging have yielded positive results, absO$_2$luteU-Net further improves upon efficiency metrics for sO$_2$ calculation, with a training speed of less than 10 minutes and predicting speeds of 2 ms on average. Further research should compare the performance of absO$_2$luteU-Net against other machine learning and nonlinear spectral unmixing techniques (7, 8, 9).

Nonetheless, absO$_2$luteU-Net's application is limited to breast tissue, which has inherent differences in optical properties compared to other tissue types. Additionally, the background tissue had an assumed constant oxygenation and melanin concentration, which could have led to a quick convergent solution during the initial model training. And most notably, the purely *in-silico* data set renders absO$_2$luteU-Net unreliable when predicting sO$_2$ generally from *ex-vivo* phantom tissues or *in-vivo* animal tissue, due to inherent differences between simulations and empirically measured data. Further experiments can rectify these shortcomings by accounting for tissue types different than breast tissue, better

modeling the variation in background tissue optical properties, and including *ex-vivo* and *in-vivo* data in the data set.

Future research using this model should use additional *in-silico* data to continue training and validating the model, as well as *ex-vivo* and eventually *in-vivo* PA data to further develop the model's predictive ability and generalizability. The U-Net architecture has been shown to reduce the amount of data necessary to train an efficient and effective convolutional network by implementing data augmentation, meaning that even doubling the dataset with *ex-vivo* data may significantly improve performance and generalizability to real-life PA data (10). Great lengths were taken to make the mcxyz simulation parameters and dimensions match as closely as possible to an *ex-vivo* PA image, and transfer learning can allow for faster training and ultimately more accurate predictive ability compared to *ex-vivo* training alone. The same analysis applies to *in-vivo* data for eventual animal studies and clinical studies.

In sum, the *in-silico* experiments of this paper served as proof of concept for the application of U-Nets and other CNNs for functional tissue parameter estimation using multispectral PA imaging. We hope that future work builds upon the results of absO_2luteU-Net, both in inspiration and through additional training and transfer learning using the current model. We look forward to witnessing the continuously developing field of quantitative PA imaging as it grows from a nascent technology to a robust clinical imaging modality.

Appendices

Appendix A- Tissue Creation MATLAB code

```
function T = KHHmaketissue(look)
% KHHmaketissue.m
%   Creates a 3D shape of optical property pointers, T(y,x,z)
%
%   Note: mcxyz.c can use optical properties in cm^-1 or mm^-1 or m^-1,
%       if the bin size (binsize) is specified in cm or mm or m,
%       respectively.
%
%  Steven L. Jacques. updated Aug 21, 2014.
%  Kevin Hoffer-Hawlik. updated Mar 6, 2019.

if nargin == 0, look = 0; end

format compact
clc
home

success      = 0;
Nbins       = 128;        % # of bins in each dimension of cube
binsize     = 0.03;       % size of each bin, eg. [cm] or [mm]; changed from
0.0003?
% .3 mm between pixels, with 128 lateral columns, so 38.4 mm

% 1 to 3 blood vessels, with blood from 0% to 100% oxygenation, and radius
% between .5mm and 4mm in radius
num_vessel = randi(3,1);
blood1 = 5 + randi(11,1);
radius1 = 0.05 + .35*rand;
blood2 = 5 + randi(11,1);
radius2 = 0.05 + .35*rand;
blood3 = 5 + randi(11,1);
radius3 = 0.05 + .35*rand;
angle2 = 30; % decrease if necessary

zsurf = 1.02; % how far to pad with water/air
epiderm_thick = 0.03; % [cm] only 1 pixel thick, which is super thick
derm_thick = 0.47; % [cm]
zoffset = zsurf + epiderm_thick + derm_thick;

% Specify Monte Carlo parameters
Nx = Nbins;
Ny = Nbins;
Nz = Nbins;
dx = binsize;
dy = binsize;
dz = binsize;
x  = ([1:2*Nx]'-Nx)*dx;
y  = ([1:2*Ny]'-Ny)*dy;
z  = [1:2*Nz]'*dz;

    function T_prime = bloodvessel(T_prime,radius,blood)
        % blood vessel @ xc, zc, radius, oriented along y axis
        xc = (-Nx/4+rand*Nx/2)*dx;
        zc = (3*Nz/4+rand*Nz/2)*dz;
        for iz=1:2*Nz % for every depth z(iz)
            for ix=1:2*Nx
```

```
                    xd = x(ix) - xc;        % vessel, x distance from vessel
center
                    zd = z(iz) - zc;        % vessel, z distance from vessel
center
                    r  = sqrt(xd^2 + zd^2);        % r from vessel center
                    if (r<=radius)
                        T_prime(:,ix,iz) = blood;
                    end
            end %ix
        end % iz
    end % bloodvessel function

    function T_prime = rotate(T_prime,angle2)
        T_prime = imrotate3(T_prime,rand*180,[0 0
1],'nearest','crop','FillValues',5);
        T_prime = imrotate3(T_prime,rand*angle2,[1 1
0],'nearest','crop','FillValues',5);
    end % rotate function

%%%%%%
% CREATE TISSUE STRUCTURE T(y,x,z)
%   Create T(y,x,z) by specifying a tissue type (an integer)
%   for each voxel in T. Generate until no intersection with top bounds.
counter = uint64(0);
while ~success
    if counter > 0
        fprintf('failure # %u\n',counter)
    end

    % T_prime so that when we rotate the blood vessel doesn't get cropped
    T_prime = double(zeros(2*Ny,2*Nx,2*Nz));
    T_prime = T_prime + 5; % fill background with breast tissue

    % create first blood vessel and rotate
    T_prime = bloodvessel(T_prime,radius1,blood1);
    T_prime = rotate(T_prime,angle2);
    % create second and third blood vessels and rotate
    if num_vessel > 1
        T_prime = bloodvessel(T_prime, radius2, blood2);
        T_prime = rotate(T_prime,angle2);
        if num_vessel > 2
            T_prime = bloodvessel(T_prime,radius3, blood3);
            T_prime = rotate(T_prime,angle2);
        end
    end

    % check if blood vessels don't intersect dermis
    success = 1;
    for iz=Nz/2:Nz/2+ceil(zoffset/dz)
       for iy=Ny/2:3*Ny/2
           for ix=Nx/2:3*Nx/2
               if T_prime(iy,ix,iz)==blood1, success=0; end
               if T_prime(iy,ix,iz)==blood2, success=0; end
               if T_prime(iy,ix,iz)==blood3, success=0; end
           end %ix
       end %iy
    end %iz

    counter = counter + 1;
end % while

% crop into T(y,x,z)
T = double(zeros(Ny,Nx,Nz+34));
```

```
T(1:Ny,1:Nx,35:162) = T_prime(Ny/2+1:3*Ny/2, Nx/2+1:3*Nx/2, Nz/2+1:3*Nz/2);

% water (assuming surface was accounted for) (=2)
if zsurf~=0
    for iz=1:round(zsurf/dz)
        T(:,:,iz) = 2;
    end
end

% epidermis (=3) only one pixel layer thick however
T(:,:,round(zsurf/dz)+1) = 3;
%for iz=(round(zsurf/dz)+1):ceil((zsurf+epiderm_thick)/dz)
%    T(:,:,iz) = 3;
%end

% dermis (=4)
for iz=round(zsurf/dz)+2:round(zoffset/dz)
    T(:,:,iz) = 4;
end

% do 3D visualization to double check
if look
    figure(2)
    p=patch(isosurface(T,5));
    isonormals(T,p)
    p.FaceColor = 'red';
    p.EdgeColor = 'none';
    daspect([1 1 1])
    view(3);
    axis([0 128 0 128 0 162])
    camlight
    lighting gouraud
end %look

end % parent function
```

Published with MATLAB® R2018b

References

1. B.T. Cox, J.G. Laufer, and P.C. Beard, S.R. Arridge, “Quantitative spectroscopic photoacoustic imaging, a review,” J. Biomed Opt. (1 June 2012).

2. M. Li and Y. Tang, “Photoacoustic tomography of blood oxygenation: A mini review,” Photoacoustics (31 May 2018).

3. C.L. Bayr, G.P. Luke, S.Y. Emelianov, “Photoacoustic imaging for medical diagnostics,” Acoustics Today (Oct 2012).

4. S. Jacques, "Optical properties of biological tissues: a review," Physics in Medicine & Biology (10 May 2013).

5. G.P. Luke et. al, “Optical wavelength selection for improved spectroscopic photoacoustic imaging,” Photoacoustics (May 2013).

6. I. Goodfellow et. al, “Deep Learning,” MIT Press (2016).

7. T. Kirchner, J. Gröhl, and L. Maier-Hein, “Context encoding enables machine learning based quantitative photoacoustics,” arXiv (29 Jan 2018).

8. J. Gröhl, T. Kirchner, T. Adler, and L. Maier-Hein, “Estimation of blood oxygenation with learned spectral decoloring for quantitative photoacoustic imaging (LSD-qPAI),” arXiv (15 Feb 2019).

9. C. Cai, K. Deng, C. Ma, and J. Luo, “End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging,” Opt. Letter (2018).

10. O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” MICCAI (18 May 2015).

11. D. Clevert, T. Unterthiner, and S. Hochreiter, “Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs),” ICLR (22 Feb 2016).

12. S. Jacques, “Coupling 3D Monte Carlo light transport in optically heterogenous tissues to photoacoustic signal generation,” Photoacoustics (2 September 2014).

13. L. Wang, S. Jacques, L. Zheng, “MCML—Monte Carlo modeling of light transport in multi-layered tissue,” Computer Methods and Programs in Biomedicine (July 1995).

14. D.P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” 3rd International Conference for Learning Representations (30 Jan 2017).

15. Y. Zhan et. al, “Singular value decomposition based regularization prior to spectral mixing improves crosstalk in dynamic imaging using spectral diffused optical

tomography," Biomedical Optics Express (September 2012).