Supplementary Material

1 PARAMETER OPTIMIZATION FOR PHOTOACOUSTIC SPECTRAL ANALYSIS

To maximize the sensitivity, specificity, and accuracy of our approach (see Section 2.6 for metric definitions), the parameters for in-phase quadrature demodulation, PCA, and NN were first optimized through an iterative search. These parameters were the modulation frequency, filtered bandwidth, axial kernel size, number of principal components to use, and the k nearest neighbors used to determine the most common class in k-NN clustering. Each parameter was changed one at a time, and this search was repeated each time the parameter was changed. During each analysis, the optimal parameter found from the previous step was saved and used to find the new output until the optimal set of parameters were found. For the IQ-modulation, different combinations of bandwidth of 80-240% in intervals of 20% were tested with the modulation frequency which ranged from 2 MHz to 12 MHz in intervals of 2 MHz. Similarly, the axial kernel size was explored from 11 to 51 axial samples in increments of 2. For the PCA, the principal components were changed from 10 to 200 in steps of 10 and then from 1-10 in steps of 1. Finally, the NN classifier was analyzed by changing the classifier to 2-NN through 10-NN in increments of 1. The initial parameters were obtained from a previous publication in our group (1). The optimization process was conducted on datasets obtained with laser wavelengths ranging from 690 nm to 950 nm in 10 nm increments.

Fig. S1 shows two known criteria to determine the optimal number of principal components used in the dual-wavelength atlas method. Each scatter point and error bar represents the mean and one standard deviation, respectively, of eigenvalues (left plot) and explained variance (right plot) from the 351 wavelength pairs (see Section 1.2 of the manuscript) at a specific principal component. For the Kaiser rule (2), only the first principal component shows a distribution with values greater than 1.00, while the second principal component shows a mean eigenvalue less than 1.00. Similarly, the commonly accepted 80%-explained-variance threshold filter the first and second principal component when considering the mean value. However, when considering the standard deviation, the second principal component surpasses the variance threshold. Therefore, only the first principal component was used for the feature extraction in the dual-wavelength atlas method.

Fig. S2 shows the mean accuracy of classification results for the dual-wavelength atlas method before (left) and after (right) parameter optimization. The top right half of each image represents a triangle

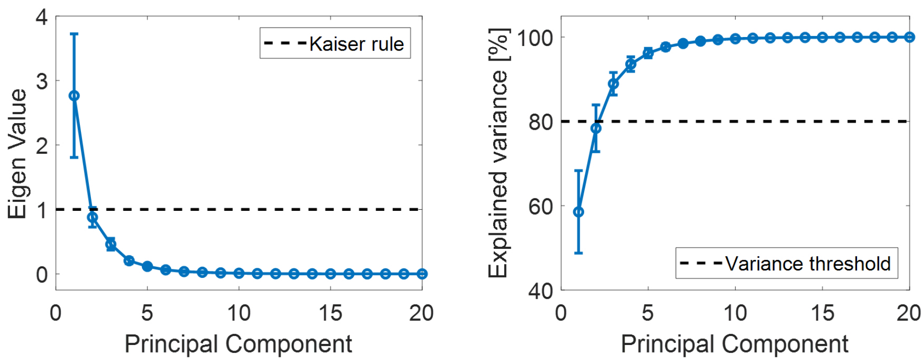

Figure S1. Criteria for determining the optimal number of principal components for the dual-wavelength atlas method using the (left) Kaiser rule (2) and (right) a 80% variance threshold.

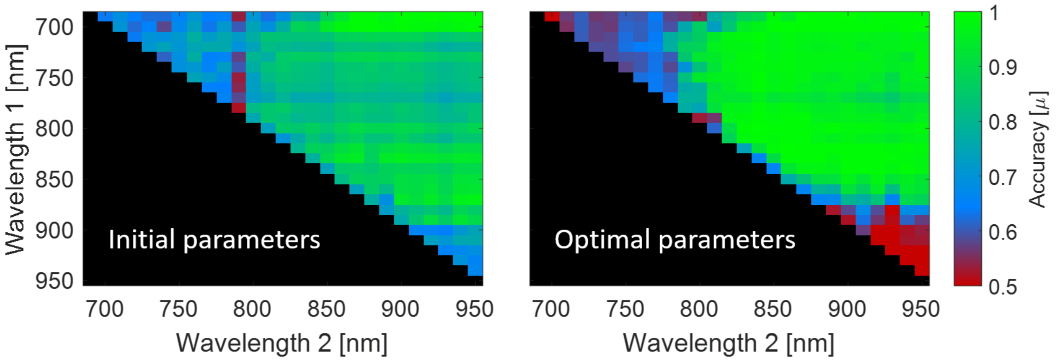

Figure S2. Mean accuracy classification results of the dual-wavelength atlas method per wavelength pair with a dataset of 10 frames per wavelength using the initial parameters of Gonzalez *et al.* (1) and the updated parameters.

with each pixel displaying the average accuracy of classifying methylene blue and blood for a specific wavelength pair over 10 frames. The new parameter set consisted of a modulation frequency of 2 MHz, bandwidth of 140%, 30 axial samples, 1 principal component, and 1-NN. For each triangle, we define the "low-wavelength region" as wavelength pairs <800 nm, the "high-wavelength region" as wavelength pairs >850 nm, and the "mid-wavelength region" as wavelengths that do not belong to either the low- or high-wavelength regions. For wavelength pairs located in the mid-wavelength region, the mean accuracy increased from 85.45% to 95.67%. In contrast, for wavelength pairs that located in either the low- or high-wavelength region, the mean accuracy decreased from 74.67% to 67.40%. In clinical practice, we envision the use of a reduced set of wavelength pairs, choosing those that maximize the sensitivity, specificity and accuracy (i.e., mostly occurring in the mid-wavelength region) and otherwise omitting wavelength pairs that result in poor classification performance.

Fig. S3 shows a summary of sensitivity, specificity, and accuracy among the 351 wavelength pairs before and after parameter optimization. The mean ± one standard deviation of sensitivity showed no significant change from 91.6 ± 16.8% to 90.9 ± 14.8%. However, an increase in specificity was observed from 70.7 ± 16.9% to 81.7 ± 20.9%, which in turn resulted in an overall increase of specificity from 81.5 ± 10.4% to 85.9 ± 17.1%. Because blood was considered as negative samples for the specificity metric, the results suggest that the optimized version of the dual-wavelength atlas method can identify blood with more accuracy than with the initial parameter set used in Gonzalez *et al.* (1).

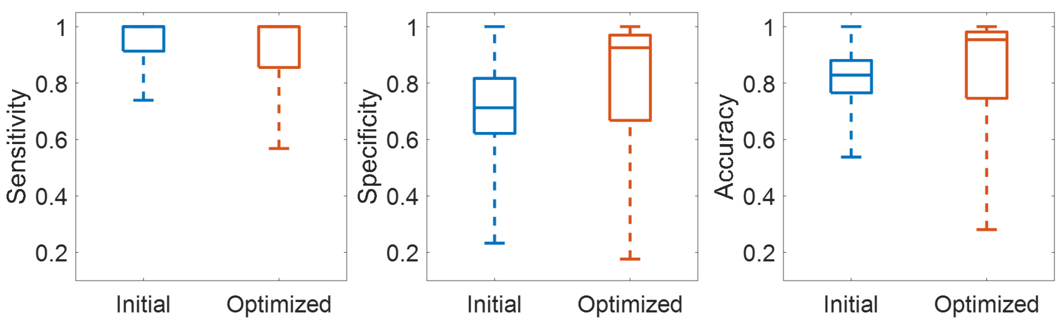

Figure S3. Summary of sensitivity, specificity, and accuracy results of the dual-wavelength atlas method before and after parameter optimization.

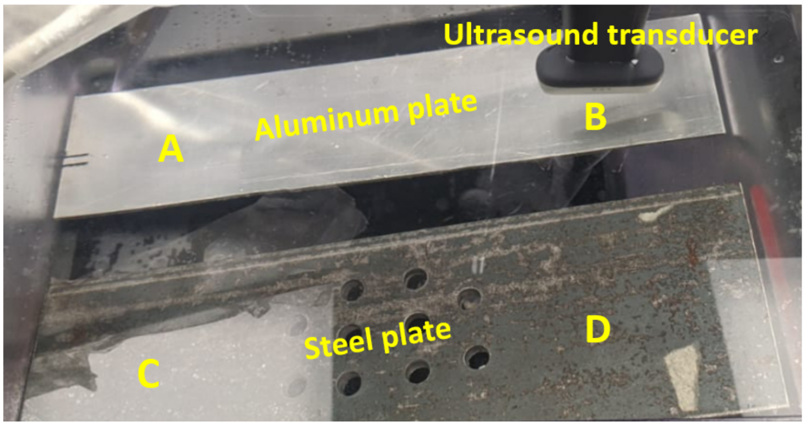

Figure S4. Setup for reference spectra acquisition. A and B represent the imaging depths of 5 and 9 cm for aluminum, respectively. C and D represent the imaging depths of 5 and 9 cm for steel, respectively.

2 SELECTION OF REFERENCE SPECTRA FOR K-MEANS CLUSTERING

Fig. S4 illustrates the ultrasound acquisition setup used to evaluate the performance variability of the method proposed by Cao *et al.* (3). In particular, eight ultrasound reference spectra were acquired from two materials: (1) an aluminium plate and (2) a carbon-steel plate. These two materials contain relatively flat, reflective surfaces, which make them ideal for use as reference spectra (4). These two materials also have acoustic properties that produce different pulse-echo responses (5, 6), which affects the normalization process applied to the acquired spectra and results in different classification accuracies.

Acquisitions were obtained after independently placing each material at axial depths of 5 cm or 9 cm from the L3-8 ultrasound transducer, tilted 45° or 0° relative to the elevation-lateral plane of the transducer. Each reference spectra was calculated by averaging the fast Fourier transform of scan lines across the lateral dimension. Then, a comparison of the accuracy results was conducted only for the subsets obtained from the optimal laser wavelength λ, defined by Eq. 7 in the manuscript. The reference spectra that produced the highest median classification accuracy was selected for comparison of the method by Cao *et al.* (3) with other methods reported throughout the manuscript.

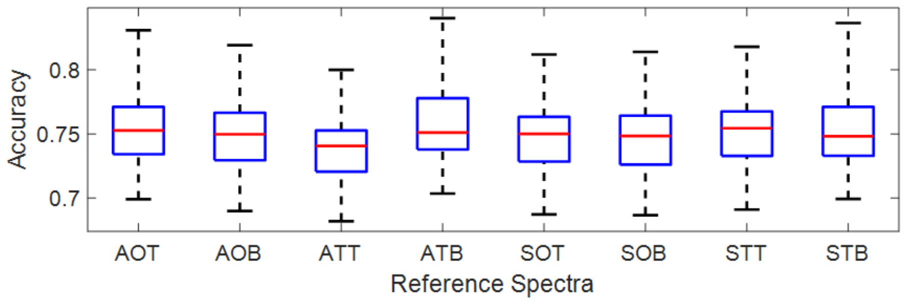

Figure S5. Mean classification accuracy of k-means clustering with different reference spectra. The box plots represent the group of frames evaluated with laser wavelength that achieved the overall best accuracy among reference spectra (i.e., 840 nm). The three letter codes on the abscissa represent the material of the reference (i.e., A = Aluminum or S = Steel), followed by the orientation of the material relative to the ultrasound transducer (i.e., O = Orthogonal or T= Tilted), followed by the imaging depth of 5 cm or 9 cm (i.e., T = Top or B = Bottom, respectively).

Fig. S5 shows the classification accuracy achieved after performing the k-means clustering step of the method of proposed by Cao *et al.* (3), obtained with a laser wavelength of 890 nm (which was the optimal parameter λ defined by Eq. 7 in the manuscript) when using the reference spectra described above. The reference spectra reported for the k-means clustering method in Fig. 6 of the manuscript was obtained from the tilted aluminum reference with an imaging depth of 9 cm (i.e, ATB in Fig. S5). This reference was chosen because it produced the greatest classification accuracy (i.e., 84.0%).

REFERENCES

[1]Gonzalez EA, Bell MAL. Acoustic frequency-based differentiation of photoacoustic signals from surgical biomarkers. *2020 IEEE International Ultrasonics Symposium (IUS)* (IEEE) (2020), 1–4.

[2]Kaiser HF. The application of electronic computers to factor analysis. *Educational and psychological measurement* **20** (1960) 141–151.

[3]Cao Y, Kole A, Lan L, Wang P, Hui J, Sturek M, et al. Spectral analysis assisted photoacoustic imaging for lipid composition differentiation. *Photoacoustics* **7** (2017) 12–19.

[4]Baddour RE, Sherar MD, Hunt J, Czarnota G, Kolios MC. High-frequency ultrasound scattering from microspheres and single cells. *The Journal of the Acoustical Society of America* **117** (2005) 934–943.

[5]De Morais N, Brett C. Influence of power ultrasound on the corrosion of aluminium and high speed steel. *Journal of Applied Electrochemistry* **32** (2002) 653–660.

[6]Netshidavhini N, Raymond BM. Effects of various couplants on carbon steel and aluminium materials using ultrasonic testing. *18th World Conference on Nondestructive Testing* (2012), 16–20.