The Journal of Open Source Software

PATATO: a Python photoacoustic tomography analysis toolkit

Thomas R. Else [1,2], **Janek Gröhl** [1,2], **Lina Hacker** [1,2,3], and **Sarah E. Bohndiek** [1,2¶]

1 CRUK Cambridge Institute, University of Cambridge, United Kingdom **2** Department of Physics, University of Cambridge, United Kingdom **3** Department of Oncology, University of Oxford, United Kingdom ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Photoacoustic imaging (PAI) is an emerging scalable imaging technology that combines the high contrast of optical imaging with the spatiotemporal resolution of ultrasound (Beard, 2011). Using light absorption by endogenous molecules, such as haemoglobin in red blood cells, PAI can reveal the emergence of diseases ranging from inflammation to cancer in both preclinical animal models and in patients (Brown et al., 2019; Regensburger et al., 2021; Steinberg et al., 2019; Wang & Hu, 2012). Extracting accurate photoacoustic imaging biomarkers, such as blood oxygen saturation, from raw data requires a robust image reconstruction and analysis process, which is challenging due to the high dimensionality of the data across spatial, spectral and temporal domains. Here we introduce PATATO, a Python toolkit that offers fast implementations of commonly-used data analysis methods, including pre-processing, reconstruction and temporal data analysis, via a user-friendly command-line interface and Python API. The toolkit uses JAX, a modern machine learning tool, for GPU-accelerated pre-processing and image reconstruction, and NumPy for easy integration with other commonly used Python libraries. PATATO is open-source, hosted on GitHub and PyPi, and distributed under an MIT licence. We have designed PATATO to be modular and extendable to accommodate different data types, reconstruction methods, and custom analyses for specific scientific questions. We welcome contributions, bug reports, and feedback. Detailed examples, documentation, and an API reference are available at https://patato.readthedocs.io/en/latest/.

Statement of Need

Photoacoustic imaging (PAI) contrast arises from the absorption of light pulses by tissue chromophores, such as haemoglobin, melanin, lipids and water (Beard, 2011). The acoustic waves that arise from the photoacoustic effect are then captured by a detector array, giving raw acoustic time series data associated with each light pulse. These raw data are typically subject to i) pre-processing e.g. by filtering; ii) reconstruction into 2D or 3D visualisations; iii) spectral unmixing processes, to decompose the range of molecules that contributed to the absorption process; and iv) data visualisation and quantification, including drawing of regions of interest (ROIs) to extract both static and dynamic biomarker values (Figure 1).

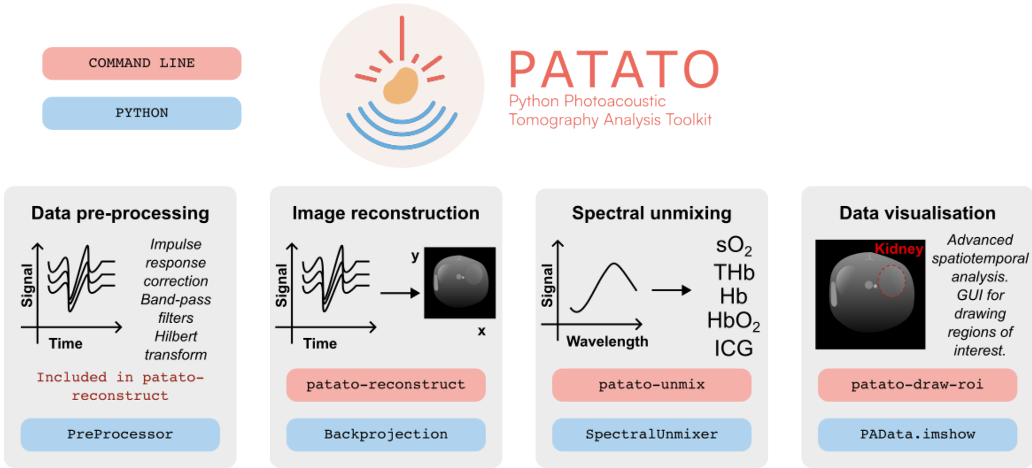

Figure 1: An overview of the key features of PATATO. Command line interfaces for the given feature are shown in red. Example Python interfaces for each feature are shown in blue.

The PAI data processing pipeline is computationally intensive and the output values are highly susceptible to parameter changes (Hochuli et al., 2019; Shen et al., 2020). Unfortunately, existing PAI data analysis typically relies on commercial software packages, or custom in-house unpublished codebases. Commercial software is generally optimised for image reconstruction from a specific instrument marketed by the vendor, enabling only a limited subset of predefined analyses and making analysis incompatible with open-access research mandates. Similarly, closed-source code reduces the transparency and reproducibility of research and impedes the widespread adoption of new algorithms.

Some open access code for PAI backprojection and model-based reconstruction is available (pyoat, https://github.com/berkanlafci/pyoat and RAFT (O'Kelly et al., 2021)). Still, implementation is restricted to a few use cases, documentation is limited and GPU acceleration lacking. Similar limitations exist for open access spectral processing code (Gröhl et al., 2021; Kirchner & Frenz, 2021).

We developed PATATO to enhance the transparency, reproducibility and consistency of PAI data analysis. PATATO is designed to be fast, extendable, and compatible with different data formats and systems, enabling users to easily go beyond the limited capabilities of commercial software packages and have complete control of their data processing pipeline. By providing an extendable platform for reproducible analysis, we hope to improve the uptake and dissemination of new analysis algorithms across both application-focused users and computational researchers, a goal that has garnered significant community support through the International Photoacoustic Standardization Consortium (Gröhl et al., 2022).

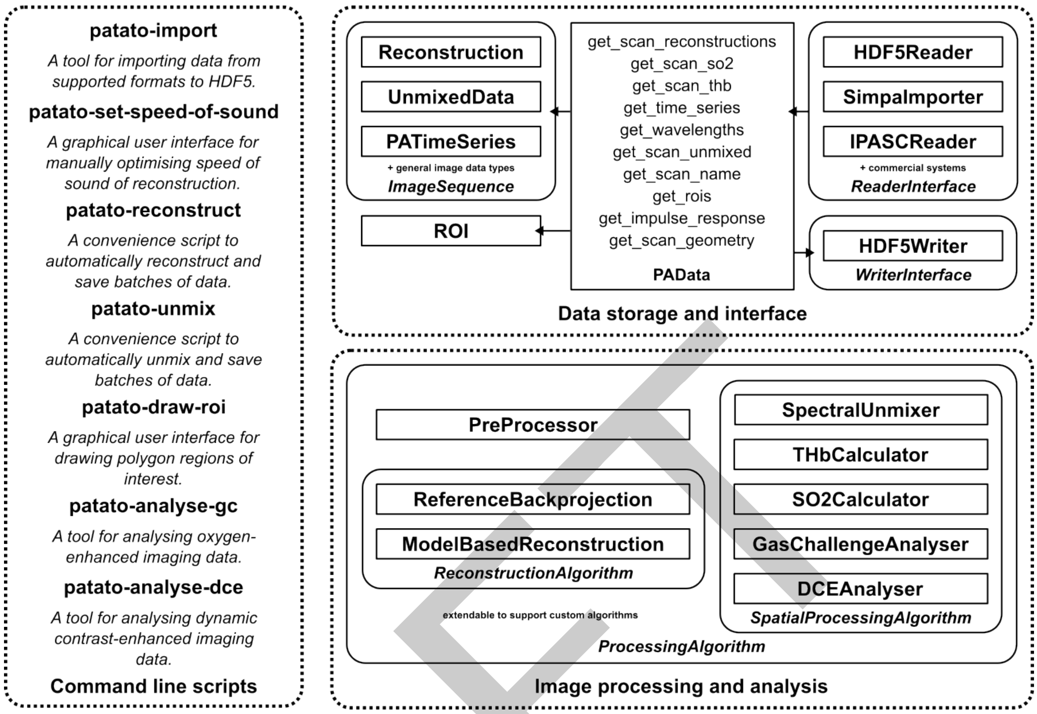

Figure 2: An outline of the software architecture of PATATO showing the key features. Abstract base classes are shown in bold italics. Classes are shown in boxes in bold. Functions are shown with a normal font. Command line functions are shown in bold without a box.

Software Pipeline

PATATO is written in Python (currently supporting versions 3.9, 3.10 and 3.11) incorporating the strengths of standard numerical programming libraries, including NumPy, SciPy and matplotlib to access fast, well-tested numerical algorithms. PATATO can run without advanced hardware, as GPU dependencies are optional and memory requirements minimal, promoting accessibility and flexibility for the maximum number of users.

Photoacoustic data can be large in size ($\gg$ 1 GB), impeding data import and data sharing. To enable fast handling of large datasets, PATATO implements batch processing and stores output in an HDF5 format, which allows seamless transfer of large data sets between fixed storage and memory. With HDF5, users can transfer data from PATATO to other tools and programming languages. PATATO includes dedicated wrappers for a number of data sources, e.g., for the IPASC data format (Gröhl et al., 2022), while also enabling user-defined wrappers.

PATATO features a modular design (Figure 2) reflecting the four main steps of PAI data processing (Figure 1). Raw time-series pre-processing and backprojection (Xu & Wang, 2005) are implemented using JAX, a high-performance numerical computing library that enables GPU acceleration. JAX uses the same code for CPUs and GPUs, removing potential inconsistencies between different platforms. PATATO enables linear spectral unmixing based on the NumPy matrix pseudo-inverse (Hochuli et al., 2019), including reference optical absorption spectra for common chromophores such as deoxyhaemoglobin, oxyhaemoglobin and melanin, and the contrast agent indocyanine green (ICG) (Prahl, 2018). Users can adapt existing algorithms for any part of the processing pipeline or implement their own algorithms by extending the appropriate class.

Strengths

PATATO is designed to be user-friendly. Users are not required to write code to use PATATO for typical image processing tasks. We also include command line tools for data importation, speed of sound optimisation, reconstruction, unmixing, temporal analysis, and region of interest analysis. Graphical user interfaces based on matplotlib are provided. Custom data processing algorithms can easily be added and examples are presented in the documentation.

Limitations and caveats

Spectral analysis in PATATO is currently limited to linear spectral unmixing. This is an approximate method that does not account for changes in light fluence.

Current data examples are restricted to 2D PAI, however, the reconstruction and analysis algorithms do support 3D data.

Example results

To demonstrate the main features and enable benchmarking of different algorithms, we have included a selection of data sets (Else et al., 2023) with PATATO that were collected using a cylindrical-array pre-clinical (small animal) PAI system and a handheld clinical PAI system (MSOT inVision 256 and MSOT Acuity Echo respectively; both iThera Medical GmbH, Munich, Germany). Animal procedures were conducted under project licence PE12C2B96 and personal licence I33984279, issued under the United Kingdom Animals (Scientific Procedures) Act, 1986, and were approved locally under compliance form number CFSB2022. Detailed methods for this procedure have been published previously (Tomaszewski et al., 2018).

The typical photoacoustic analysis procedure in PATATO can be illustrated using the mouse dataset described above. In this study, mice were implanted with tumours and photoacoustic images were acquired to interrogate the blood perfusion of the tumours. We perturbed the distribution of absorbing molecules in the mouse by changing the breathing gas to oxygen, thereby causing a change in the blood oxygenation, and by injecting the contrast agent indocyanine green (ICG). PATATO allows the streamlined analysis of such datasets (Figure 3). Firstly, we reconstructed the photoacoustic images by backprojection. We then drew polygon regions of interest around three regions of the mouse: the two implanted tumours and the spine (Figure 3 A). To obtain maps of the blood oxygenation (sO_2), total haemoglobin (THb), and ICG content we applied linear spectral unmixing (Figure 3 B). Plots of the sO_2 and ICG levels in the three regions over time were made and the enhancement level was calculated (Figure 3 C). Maps of the signal enhancement (ΔsO_2 or ΔICG) were then made, revealing useful biomarkers related to hypoxia and blood perfusion respectively (Figure 3) (Tomaszewski et al., 2018).

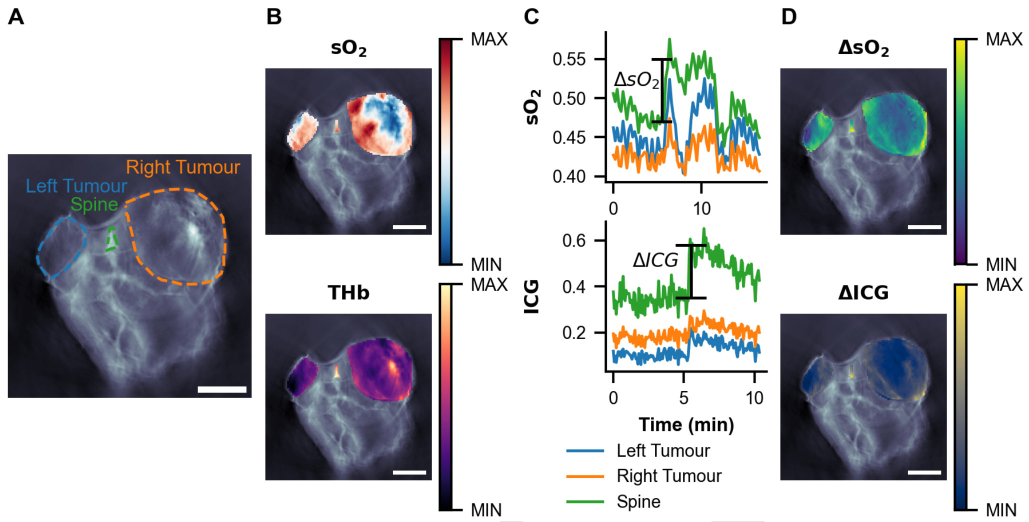

Figure 3: An example of the analysis process in PATATO using a two-dimensional cross-sectional image of a tumour-bearing mouse. Regions of interest were drawn around key features of the image, the spine and both tumours (A). We applied linear spectral unmixing, giving maps of the blood oxygenation (sO_2) and total haemoglobin (THb) (B). We acquired dynamic imaging data in response to changing breathing gas or contrast agent administration, showing the effects of oxygen enhancement or contrast enhancement respectively (C). We calculated two enhancement metrics, ΔsO_2 and $\Delta \mathrm{ICG}$ (D). Scale bar = 5 mm.

Future developments

PATATO will be developed on a continuous basis, and we welcome collaborative contributions from the PAI community, particularly to implement wrappers for different data formats and in adding image reconstruction and analysis tools. Contributions are particularly encouraged for model-based reconstructions and advanced spectral unmixing tools.

Acknowledgements

We would like to thank Mariam-Eleni Oraiopoulou, Ellie Bunce and Thierry Lefebvre for their helpful feedback as early users of the toolkit. This work was funded by: Cancer Research UK (SB, TRE; C9545/A29580); the MedAccel program of the National Physical Laboratory financed by the Department for Business, Energy and Industrial Strategy's Industrial Strategy Challenge Fund (LH); and the Walter Benjamin Stipendium of the Deutsche Forschungsgemeinschaft (JG). We would like to thank Elly Pugh for design of the PATATO logo.

Beard, P. (2011). Biomedical photoacoustic imaging. *Interface Focus*, *1*(4), 602–631. https://doi.org/10.1098/rsfs.2011.0028

Brown, E., Brunker, J., & Bohndiek, S. E. (2019). Photoacoustic imaging as a tool to probe the tumour microenvironment. *Disease Models & Mechanisms*, *12*(7), dmm039636. https://doi.org/10.1242/dmm.039636

Else, T., Groehl, J., Hacker, L., & Bohndiek, S. (2023). *Dataset for: PATATO: A Python Photoacoustic Analysis Toolkit*. https://doi.org/10.17863/CAM.93181

Gröhl, J., Dreher, K. K., Schellenberg, M., Rix, T., Holzwarth, N., Vieten, P., Ayala, L., Bohndiek, S. E., Seitel, A., & Maier-Hein, L. (2022). SIMPA: An open-source toolkit for simulation and image processing for photonics and acoustics. *Journal of Biomedical Optics*, *27*(08), 083010. https://doi.org/10.1117/1.JBO.27.8.083010

Gröhl, J., Kirchner, T., Adler, T. J., Hacker, L., Holzwarth, N., Hernández-Aguilera, A., Herrera, M. A., Santos, E., Bohndiek, S. E., & Maier-Hein, L. (2021). Learned spectral decoloring enables photoacoustic oximetry. *Scientific Reports*, *11*(1), 6565. https://doi.org/10.1038/s41598-021-83405-8

Hochuli, R., An, L., Beard, P. C., & Cox, B. T. (2019). Estimating blood oxygenation from photoacoustic images: Can a simple linear spectroscopic inversion ever work? *Journal of Biomedical Optics*, *24*(12), 1. https://doi.org/10.1117/1.JBO.24.12.121914

Kirchner, T., & Frenz, M. (2021). Multiple illumination learned spectral decoloring for quantitative optoacoustic oximetry imaging. *Https://Doi.org/10.1117/1.JBO.26.8.085001*, *26*(8), 085001. https://doi.org/10.1117/1.JBO.26.8.085001

O'Kelly, D., Campbell, J., Gerberich, J. L., Karbasi, P., Malladi, V., Jamieson, A., Wang, L., & Mason, R. P. (2021). A scalable open-source MATLAB toolbox for reconstruction and analysis of multispectral optoacoustic tomography data. *Scientific Reports 2021 11:1*, *11*(1), 1–18. https://doi.org/10.1038/s41598-021-97726-1

Prahl, S. (2018). *Assorted Spectra*. https://omlc.org/spectra/index.html.

Regensburger, A. P., Brown, E., Krönke, G., Waldner, M. J., & Knieling, F. (2021). Optoacoustic Imaging in Inflammation. *Biomedicines*, *9*(5), 483. https://doi.org/10.3390/biomedicines9050483

Shen, K., Liu, S., Feng, T., Yuan, J., Zhu, B., & Tian, C. (2020). Negativity artifacts in back-projection based photoacoustic tomography. *Journal of Physics D: Applied Physics*, *54*(7), 074001. https://doi.org/10.1088/1361-6463/ABC37D

Steinberg, I., Huland, D. M., Vermesh, O., Frostig, H. E., Tummers, W. S., & Gambhir, S. S. (2019). Photoacoustic clinical imaging. *Photoacoustics*, *14*, 77–98. https://doi.org/10.1016/j.pacs.2019.05.001

Tomaszewski, M. R., Gehrung, M., Joseph, J., Quiros Gonzalez, I., Disselhorst, J. A., & Bohndiek, S. E. (2018). Oxygen-enhanced and dynamic contrast-enhanced optoacoustic tomography provide surrogate biomarkers of tumour vascular function, hypoxia and necrosis. *Cancer Research*, *78*(20), canres.1033.2018. https://doi.org/10.1158/0008-5472.CAN-18-1033

Wang, L. V., & Hu, S. (2012). Photoacoustic Tomography: In Vivo Imaging from Organelles to Organs. *Science*, *335*(6075), 1458–1462. https://doi.org/10.1126/science.1216210

Xu, M., & Wang, L. V. (2005). Universal back-projection algorithm for photoacoustic computed tomography. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, *71*(1). https://doi.org/10.1103/PhysRevE.71.016706