

HistoJS: Web-Based Analytical Tool for Advancing Multiplexed Images

Mohamed Masoud^{1¶}, David Gutman³, and Sergey Plis^{1,2}

¹ Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State University, Georgia Institute of Technology, Emory University, Atlanta, United States of America ² Department of Computer Science, Georgia State University, Atlanta, United States of America ³ Department of Pathology, Emory University School of Medicine, Atlanta, United States of America ¶ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Open Journals](#)

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Advances in multiplexed imaging technologies enable us to capture spatial single-cell proteomics and transcriptomics data in unprecedented detail and with high spatial resolution. This large volume of image data presents a challenge in accurately isolating and quantifying distinct cell types to understand disease complexity, neurological disorders, potential biomarkers, and targets for drug development. Therefore, there is an incremental demand and challenges to developing and validating cutting-edge quantitative image analysis tools for diagnosis, prognosis, and therapy response prediction and assessment in neurological and oncological diseases. **HistoJS** is a newly developed web-based tool that aims to overcome the challenges of utilizing highly-multiplexed immunofluorescence (mIF) images [lin2015highly] for spatial biology research. It provides open-source and extensible tools for analyzing spatial-molecular patterns, enabling a deeper view of the single-cell spatial relationship, along with machine learning algorithms in an easy-to-use interactive interface for the biomedical community.

Statement of need

Single-cell data have the potential to enhance our understanding of biological systems, shedding light on disease mechanisms and reactions. The data from single-cell sequencing and multiplexed imaging technologies have distinctive capabilities for cell phenotyping, clustering, and landscape analysis. However, handling large-scale multiplexed image datasets generated from advanced imaging technologies and narrowing the lag between those technologies and the existing analytical tools remains challenging. There is a need to effectively utilize subcellular imaging data by storing, retrieving, visualizing, and performing quantitative analysis of those big data. Addressing the challenges posed by highly multiplexed image analysis, HistoJS emerges as an open-source and adaptable tool for visualizing and analyzing intricate biological processes with spatial subcellular resolution. It offers a graphical user interface for effortless navigation through stored multiplexed images, allowing dynamic selection of image channels for composite views. HistoJS encompasses a diverse set of image processing and machine learning algorithms to support essential analysis, including real-time cell segmentation, phenotyping, classification, correlations, spatial analysis, and quantification of cell types to unveil interactions within the tissue samples. These functionalities are provided in an easy-to-use interactive interface to help biomedical users and related groups understand the progression of neurological and oncological diseases and find clinical outcomes. Other commercial tools like the VisioPharm suite are expensive and complicated to be customized by the informatics community to meet the specific needs of cancer researchers. QuPath (Bankhead et al., 2017), another popular open-source tool, has some useful features for the analysis that are not web-based and need more key

43 features for whole slide image analysis. Positioned as a web-based solution, HistoJS focuses
 44 on enhancing usability, accessibility, sustainability, scalability, and collaboration. Boosting user-
 45 friendly interfaces and cross-platform compatibility, HistoJS ensures a seamless and accessible
 46 experience. The sustainability of HistoJS is realized through centralized updates and reduced
 47 hardware dependencies, streamlining management and minimizing hardware requirements.

48 Background

49 Advanced High-plex imaging technologies, such as Multiplexed Immunofluorescence (MxIF)
 50 ([Gerdes et al., 2013](#)), tissue-based cyclic immunofluorescence (t-CyCIF) ([Lin et al., 2018](#)),
 51 CO-Detection by indEXing (CODEX) ([Goltsev et al., 2018](#)), and Multiplexed Ion Beam Imaging
 52 (MIBI) ([Angelo et al., 2014](#)), facilitate the study of intricate biological processes with precise
 53 spatial subcellular resolution. While MIBI utilizes a mass spectrometry-based approach, the
 54 other technologies utilize conventional fluorescence microscopes, and all enable the simultaneous
 55 detection of 50+ antigens within a single tissue section. These technologies offer invaluable
 56 insights into the complexities of biological systems. The process typically commences with
 57 sample preparation of high-quality formaldehyde-fixed and paraffin-embedded (FFPE) tissue
 58 sections (e.g., biopsy, surgical specimen) that can pass through either an iterative, multicycle
 59 image acquisition pipeline as CODEX or a non-iterative single imaging cycle, such as MIBI.
 60 Either approach produces a stack of raw microscopy image tiles that need stitching with
 61 drift correction and then registration across channels to generate large-scale mosaic images in
 62 OME-TIFF (open microscopy environment-tagged image file format). Each channel represents
 63 the spectral signal from a specific marker or antigen in the tissue. Other preprocessing steps
 64 include image deconvolution to minimize image blurring, noise reduction to reduce background
 65 and autofluorescence noise, and artifact removal that addresses image artifacts (e.g., axial and
 66 lateral tile drift). Preprocessing steps are essential to enhance the quality of the raw data and
 67 prepare it for downstream analysis. The goal is to identify, localize, and count different cell
 68 types to study their population and interactions.

69 Pipeline

70 For hosting and organizing user data, the Digital Slide Archive (DSA), an open-source platform,
 71 is used as a data hosting environment ([Gutman et al., 2017](#)). DSA is a reliable containerized
 72 web-based platform that can store and manage large image datasets such as immunofluorescence
 73 image data. DSA components include a MongoDB database, a job execution/scheduler (girder
 74 worker), and a Secure data management system (Girder) that provides RESTful APIs to allow
 75 programmatic control over image data and metadata while providing user access controls.
 76 DSA is easy to install locally on the user side or remotely on the cloud. HistoJS uses DSA as
 77 a backend to manage user images locally or remotely and store analysis results according to
 78 end-user preferences. To visualize mIF images in HistoJS, OpenSeadragon ([OpenSeadragon
 79 dev. team, 2022](#)), an open-source web-based viewer, is used. HistoJS has rich JavaScript
 80 functions to support different types of canvas rendering global Composite operations, giving the
 81 best insight into data with panning, zooming, and overlay options. To that end, we were able
 82 to make HistoJS suitable for building a customized stack of protein channels and composite
 83 them in a high resolution and full scale as illustrated in [Figure 1](#). The analytical components
 84 of HistoJS include cell segmentation, phenotyping, classification, correlations, spatial analysis,
 85 and quantification of cell types to discover cell interactions. We tackled the challenges in
 86 real time; fast and reliable techniques are used to expedite the extraction of cell boundaries
 87 ([Schmidt et al., 2018](#)) and morphological features (e.g., solidity, eccentricity, orientation, etc.).
 88 Cell neighbor detections are tackled with spatial plotting and analysis by computing Delaunay
 89 neighborhood graphs ([Gabriel & Sokal, 1969](#)).



Figure 1: HistoJS graphical interface overview. Biological statistical tasks such as biological cell biomarkers histogram, sample statistics quartiles, cell classification, correlations, spatial analysis, and quantification of specific marker expression are available for cell analysis and discovering the cell interactions. (Dataset (Rashid & others, 2019)).

Although the tool is suitable for cluster-based deployment, it can also be deployed on the cloud or locally on the client side. With an average GPU GeForce GTX 1050 Ti of 768 cores/ 4GB buffer, 7Gbps memory speed, Intel® Core™ i7-8700 CPU @ 3.20GHz × 12, and a system memory of 16 GB, HistoJS shows in general fast response while rendering and processing immunofluorescence images, with the potential for better performance thanks to the TensorFlow.js (Smilkov et al., 2019) Web Graphics Library (WebGL-2) backbone.

Code availability

HistoJS source code is publicly accessible in the GitHub repository (<https://github.com/Mmasoud1/HistoJS>). Multiple DSA online servers can be accessed from HistoJS such as <https://styx.neurology.emory.edu/girder/> to load mIF data samples and test the tool performance. Researchers could visualize and analyze the expression patterns of key biomarkers associated with diseases and disorders. The platform's interactive features facilitated the identification of disease-specific signatures, providing valuable insights into the molecular basis of diseases.

Detailed step-by-step [documentation](#) is provided along with HistoJS [live demo](#).

Author contributions

We describe contributions to this paper using the CRediT taxonomy (Brand et al., 2015). Writing – Original Draft: M.M.; Writing – Review & Editing: M.M., and S.P.; Conceptualization and methodology: M.M., and D.G.; Software and data curation: M.M.; Validation: M.M., and S.P.; Resources: D.G.; Visualization: M.M.; Project Administration: M.M.;

References

- Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., Levenson, R. M., Lowe, J. B., Liu, S. D., Zhao, S., & others. (2014). Multiplexed ion beam imaging of human breast tumors. *Nature Medicine*, 20(4), 436–442. <https://doi.org/10.1038/nm.3488>

- 114 Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P.
 115 D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., & others. (2017). QuPath:
 116 Open source software for digital pathology image analysis. *Scientific Reports*, 7(1), 1–7.
 117 <https://doi.org/10.1038/s41598-017-17204-5>
- 118 Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. (2015). Beyond authorship:
 119 Attribution, contribution, collaboration, and credit. *Learned Publishing*, 28(2), 151–155.
 120 <https://doi.org/10.1087/20150211>
- 121 Gabriel, K. R., & Sokal, R. R. (1969). A new statistical approach to geographic variation
 122 analysis. *Systematic Zoology*, 18(3), 259–278. <https://doi.org/10.2307/2412323>
- 123 Gerdes, M. J., Sevinsky, C. J., Sood, A., Adak, S., Bello, M. O., Bordwell, A., Can, A., Corwin,
 124 A., Dinn, S., Filkins, R. J., & others. (2013). Highly multiplexed single-cell analysis of
 125 formalin-fixed, paraffin-embedded cancer tissue. *Proceedings of the National Academy of
 126 Sciences*, 110(29), 11982–11987. <https://doi.org/10.1073/pnas.1300136110>
- 127 Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S., &
 128 Nolan, G. P. (2018). Deep profiling of mouse splenic architecture with CODEX multiplexed
 129 imaging. *Cell*, 174(4), 968–981. <https://doi.org/10.1016/j.cell.2018.07.010>
- 130 Gutman, D. A., Khalilia, M., Lee, S., Nalisnik, M., Mullen, Z., Beezley, J., Chittajallu, D. R.,
 131 Manthey, D., & Cooper, L. A. (2017). The digital slide archive: A software platform for
 132 management, integration, and analysis of histology for cancer research. *Cancer Research*,
 133 77(21), e75–e78. <https://doi.org/10.1158/0008-5472.CAN-17-0629>
- 134 Lin, J.-R., Izar, B., Wang, S., Yapp, C., Mei, S., Shah, P. M., Santagata, S., & Sorger, P.
 135 K. (2018). Highly multiplexed immunofluorescence imaging of human tissues and tumors
 136 using t-CyCIF and conventional optical microscopes. *Elife*, 7. <https://doi.org/10.7554/eLife.31657>
- 137
- 138 OpenSeadragon dev. team. (2022). *OpenSeadragon*. GitHub Pages. <http://openseadragon.github.io/>
 139
- 140 Rashid, R., & others. (2019). *Lung dataset-1*. Synapse repository. <https://doi.org/10.7303/syn17865732>
 141
- 142 Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell detection with star-convex
 143 polygons. *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018:
 144 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part
 145 II 11*, 265–273. https://doi.org/10.1007/978-3-030-00934-2_30
- 146 Smilkov, D., Thorat, N., Assogba, Y., Nicholson, C., Kreeger, N., Yu, P., Cai, S., Nielsen, E.,
 147 Soegel, D., & others. (2019). TensorFlow.js: Machine learning for the web and beyond.
 148 *arXiv*. <https://doi.org/10.48550/arXiv.1901.05350>