The Journal of Open Source Software

Image Marker

Ryan Walker[1*], **Andi Kisare**[1*], and **Lindsey Bleem**[1]

1 Argonne National Laboratory, United States * These authors contributed equally.

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: David Stansby

Submitted: 24 April 2025
Published: unpublished

Summary

A wide range of scientific imaging datasets benefit from human inspection for purposes ranging from prosaic—such as fault identification and quality inspection—to profound, enabling the discovery of new phenomena. As such, these datasets come in a wide variety of forms, with diverse inspection needs. In this paper we present a software package, `Image Marker`, designed to help facilitate human categorization of images. The software allows for quick seeking through images and enables flexible marking and logging of up to 9 different classes of features and their locations in files of FITS, PNG, TIFF, and JPEG format. Additional tools are provided to add text-based comments to the marking logs and for displaying external *mark* datasets on images during the classification process. As our primary use case will be the identification of features in astronomical survey data, `Image Marker` will also utilize standard World Coordinate Systems (WCS) headers embedded in FITS headers and TIFF metadata when available. The lightweight software, based on the Qt Framework to build the GUI application, enables efficient marking of thousands of images on personal-scale computers. We provide `Image Marker` as a Python package, and as Mac and Windows 11 executables. It is available on GitHub or via pip installation.

Statement of need

The rapid advancement in detector technology across all fields of science has led to larger and larger datasets without an equal increase in the number of scientists available to analyze the data. This imbalance of available work to available workers has led to a need for developing more efficient methods of parsing data. In response to large datasets in astronomy, projects like DES Exposure Checker (Melchior et al., 2016), and Space Warps (Marshall et al., 2016) and Galaxy Zoo (Fortson et al., 2012) using the Zooniverse framework (Simpson et al., 2014) emerged to crowdsource classification and identification tasks in large datasets. Zooniverse offers the ability to easily outsource image identification and advanced classification statistics through the power of citizen science. This level of sophistication is not required, however, for projects which may involve fewer collaborators or for low-level data or algorithmic phases that are not suitable for a broader audience. Zooniverse also requires an internet connection. FitsMap (Hausen & Robertson, 2022) takes a different approach with a focus on large images and their associated catalogs by hosting a web client on the user's local machine and displaying a reduced-scale image with catalog objects overlaid. While FitsMap has broad functionality, it does not contain a method for scanning many images quickly, saving feature coordinates, or methods for crowdsourcing efforts. Other software for viewing and analyzing data, like the widely-used SAO-DS9 (Joye, 2017), handle smaller datasets best.

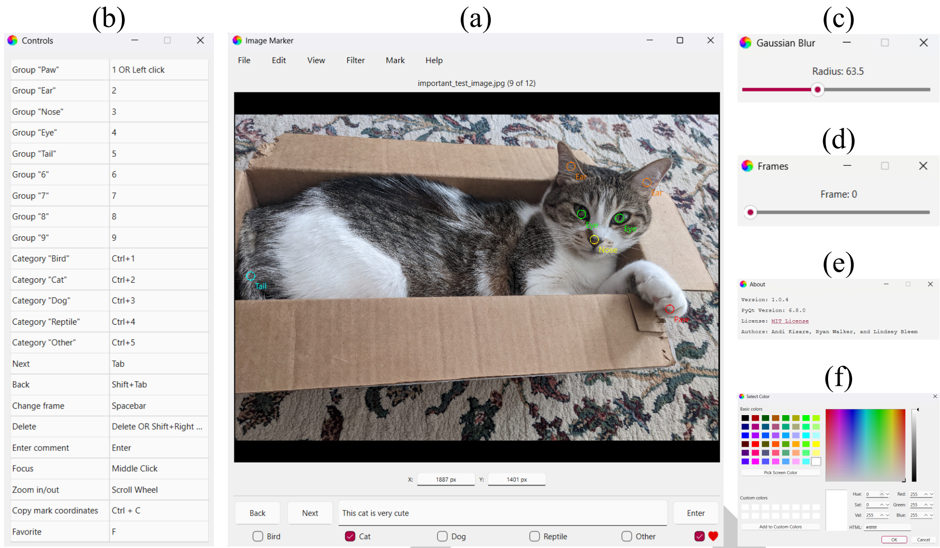

Figure 1: Diagram of `Image Marker` windows highlighting main features of the application: (a) main window; nine different group marks are available for tagging features. Below the image users can read the pixel coordinates (and, if available, WCS coordinates) of the cursor. Note that a comment has been written in the main comment box in the center. (b) controls window; updates group and category names when they are customized, helping keep track of what buttons are for which group or category. Other shortcuts are shown as well. (c) Gaussian blur window; note that blur has not been applied to the example image in (a). (d) frames window; for selecting frames in multi-extension images/FITS files. (e) about window; displays basic information about the user's installation of `Image Marker`. (f) color picker window; used to select the color of an imported *catalog* of *marks*. Window themes are dependent on the user's operating system.

`Image Marker` is a tool specifically designed for quickly scanning images and tagging locations in the images or the images themselves. It is run locally, has a user-friendly interface, a fail-safe saving mechanism, and also includes a suite of features and options for customization, image manipulation, and testing user consistency (see Figure 1 for application interface). By sharing an `Image Marker` configuration file and data with other users, `Image Marker` also allows joint analyses of datasets at the expense of requiring manual sharing of files after marking is completed. This, however, enables trained observers to quickly scan through images with loading times not limited by internet connections and thus faster identifications or classifications.

Our use case

The SPT-3G camera has surveyed ~10,000 square degrees of the Southern sky (Prabhu et al., 2024; Sobrin et al., 2022) at millimeter-wavelengths. Two objectives of these observations are to identify a sample of galaxy clusters through the Sunyaev-Zel'dovich (SZ) effect (Sunyaev & Zeldovich, 1972) and to use this sample to constrain cosmology (Chaubal et al., 2022; Raghunathan, 2022). As part of this process, one must select the centers of the galaxy clusters in order to enable connection of cluster observables to theoretical models (using e.g., weak gravitational lensing, see reviews in Allen et al. (2011); Umetsu (2020)). The most commonly adopted choice for such centers are cluster galaxies known as "brightest cluster galaxies" (BCGs). Automatic BCG selection algorithms typically fail 10-20% of the time, however, and human inspection plays an important role in both validating these algorithms and improving centering choices when they fail (Ding et al., 2025; Kelly et al., 2024; Rozo & Rykoff, 2014). Our first use case for `Image Marker` is the identification of BCGs in the SPT-3G cluster sample (see Figure 2) using optical image data from DeCALS (Dey et al., 2019). This human-generated BCG dataset will be analyzed and compared to results from algorithms such as redMaPPer (Rykoff et al., 2014) and MCMF (Klein et al., 2019) that will also be run on the sample. The validated BCG dataset from thousands of clusters ($>5,000$ galaxy clusters expected in the SPT-3G main survey alone, Benson et al. (2014)) also offers the opportunity to develop and test machine learning BCG identification tools that can be applied not only to the SPT-3G sample but also cluster samples from other surveys (e.g., Hilton et al. (2021); Bulbul et al. (2024); LSST Science Collaboration et al. (2009)). We are currently testing the use of convolutional neural networks (CNN; Krizhevsky et al. (2012)) and vision transformers (ViT; Dosovitskiy et al. (2020)) for this task.

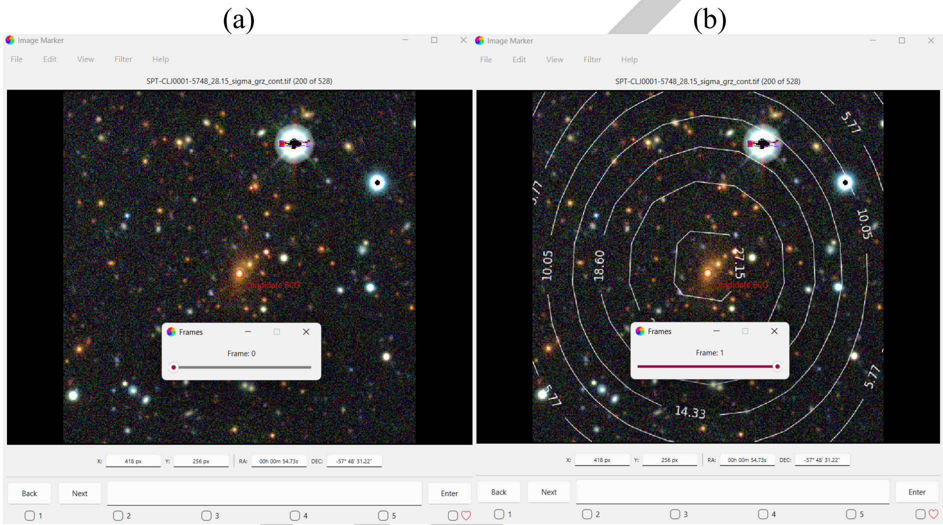

Figure 2: Optical *grz* band image of a galaxy cluster from the SPT-3G survey (Optical images from DeCALS (Dey et al., 2019)). (a, Left) We display the first frame of the image file with just optical image data. (b, Right) The second frame of the image file, which contains the optical image data with contours overlaid indicating the SZ detection signal-to-noise from SPT-3G. The human-selected candidate BCG is denoted by the red mark in both images.

Broader use cases

While `Image Marker` was initially designed with the above use cases in mind, we have found it valuable as a general tool for inspecting data products and validating algorithmic development. As a second example usage, the ability to rapidly scan hundreds of small thumbnail cutouts in a matter of minutes, mark problematic locations, and easily read in lists of these locations, helped us to improve data cleaning for an upcoming analysis of SPT-3G data in the Euclid Deep Field South region. This broad applicability motivated us to publicly release the software.

Acknowledgements

We thank Keren Sharon, Mike Gladders, and Giulia Campitiello for helpful suggestions on features to include in `Image Marker`, Will Hicks for help compiling `Image Marker` on an Intel based Mac, and Florian Keruzore for providing helpful comments. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships (SULI) program. Work at Argonne National Lab is supported by UChicago Argonne LLC, Operator of Argonne National Laboratory (Argonne). Argonne, a U.S. Department of Energy Office of Science Laboratory, is operated under contract no. DE-AC02-06CH11357. We thank

the SPT-3G collaboration for the use of the Sunyaev Zel'dovich detection contours displayed on the image in Figure 2. This work made use of Astropy[1]: a community-developed core Python package and an ecosystem of tools and resources for astronomy (Astropy Collaboration et al., 2013, 2018, 2022); Pillow (Clark, 2015); SciPy (Virtanen et al., 2020); NumPy (Harris et al., 2020); and PyQt (Limited, 2025).

References

Allen, S. W., Evrard, A. E., & Mantz, A. B. (2011). Cosmological Parameters from Observations of Galaxy Clusters. *Annual Review of Astronomy and Astrophysics*, *49*(1), 409–470. https://doi.org/10.1146/annurev-astro-081710-102514

Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C. E., Nöthe, M., Donath, A., Tollerud, E., Morris, B. M., Ginsburg, A., Vaher, E., Weaver, B. A., Tocknell, J., Jamieson, W., … Astropy Project Contributors. (2022). The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. *The Astrophysical Journal*, *935*(2), 167. https://doi.org/10.3847/1538-4357/ac7c74

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., VanderPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., … Astropy Contributors. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *The Astronomical Journal*, *156*(3), 123. https://doi.org/10.3847/1538-3881/aabc4f

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., … Streicher, O. (2013). Astropy: A community Python package for astronomy. *Astronomy and Astrophysics*, *558*, A33. https://doi.org/10.1051/0004-6361/201322068

Benson, B. A., Ade, P. A. R., Ahmed, Z., Allen, S. W., Arnold, K., Austermann, J. E., Bender, A. N., Bleem, L. E., Carlstrom, J. E., Chang, C. L., Cho, H. M., Cliche, J. F., Crawford, T. M., Cukierman, A., de Haan, T., Dobbs, M. A., Dutcher, D., Everett, W., Gilbert, A., … Yoon, K. W. (2014). SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope. In W. S. Holland & J. Zmuidzinas (Eds.), *Millimeter, submillimeter, and far-infrared detectors and instrumentation for astronomy VII* (Vol. 9153, p. 91531P). https://doi.org/10.1117/12.2057305

Bulbul, E., Liu, A., Kluge, M., Zhang, X., Sanders, J. S., Bahar, Y. E., Ghirardini, V., Artis, E., Seppi, R., Garrel, C., Ramos-Ceja, M. E., Comparat, J., Balzer, F., Böckmann, K., Brüggen, M., Clerc, N., Dennerl, K., Dolag, K., Freyberg, M., … Zelmer, S. (2024). The SRG/eROSITA All-Sky Survey. The first catalog of galaxy clusters and groups in the Western Galactic Hemisphere. *Astronomy and Astrophysics*, *685*, A106. https://doi.org/10.1051/0004-6361/202348264

Chaubal, P. S., Reichardt, C. L., Gupta, N., Ansarinejad, B., Aylor, K., Balkenhol, L., Baxter, E. J., Bianchini, F., Benson, B. A., Bleem, L. E., Bocquet, S., Carlstrom, J. E., Chang, C. L., Crawford, T. M., Crites, A. T., de Haan, T., Dobbs, M. A., Everett, W. B., Floyd, B., … Williamson, R. (2022). Improving Cosmological Constraints from Galaxy Cluster Number Counts with CMB-cluster-lensing Data: Results from the SPT-SZ Survey and Forecasts for the Future. *The Astrophysical Journal*, *931*(2), 139. https://doi.org/10.3847/1538-4357/ac6a55

Clark, A. (2015). *Pillow (PIL fork) documentation*. readthedocs. https://buildmedia.

[1] http://www.astropy.org

readthedocs.org/media/pdf/pillow/latest/pillow.pdf

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D., Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., … Zhou, Z. (2019). Overview of the DESI Legacy Imaging Surveys. *The Astronomical Journal*, *157*(5), 168. https://doi.org/10.3847/1538-3881/ab089d

Ding, J., Dalal, R., Sunayama, T., Strauss, M. A., Oguri, M., Okabe, N., Hilton, M., Monteiro-Oliveira, R., Sifón, C., & Staggs, S. T. (2025). Miscentring of optical galaxy clusters based on Sunyaev-Zeldovich counterparts. *Monthly Notices of the Royal Astronomical Society*, *536*(1), 572–591. https://doi.org/10.1093/mnras/stae2601

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *CoRR*, *abs/2010.11929*. https://arxiv.org/abs/2010.11929

Fortson, L., Masters, K., Nichol, R., Borne, K. D., Edmondson, E. M., Lintott, C., Raddick, J., Schawinski, K., & Wallin, J. (2012). Galaxy Zoo: Morphological Classification and Citizen Science. In M. J. Way, J. D. Scargle, K. M. Ali, & A. N. Srivastava (Eds.), *Advances in machine learning and data mining for astronomy* (pp. 213–236). https://doi.org/10.48550/arXiv.1104.5513

Harris, C. R., Millman, K. J., Walt, S. J. van der, Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. van, Brett, M., Haldane, A., Río, J. F. del, Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hausen, R., & Robertson, B. E. (2022). FitsMap: A simple, lightweight tool for displaying interactive astronomical image and catalog data. *Astronomy and Computing*, *39*, 100586. https://doi.org/https://doi.org/10.1016/j.ascom.2022.100586

Hilton, M., Sifón, C., Naess, S., Madhavacheril, M., Oguri, M., Rozo, E., Rykoff, E., Abbott, T. M. C., Adhikari, S., Aguena, M., Aiola, S., Allam, S., Amodeo, S., Amon, A., Annis, J., Ansarinejad, B., Aros-Bunster, C., Austermann, J. E., Avila, S., … Zhang, Y. (2021). The Atacama Cosmology Telescope: A Catalog of >4000 Sunyaev-Zel'dovich Galaxy Clusters. *The Astrophysical Journal Supplement Series*, *253*(1), 3. https://doi.org/10.3847/1538-4365/abd023

Joye, W. (2017). *SAOImageDS9/SAOImageDS9* (Version v7.6b8). Zenodo. https://doi.org/10.5281/zenodo.1041781

Kelly, P. M., Jobel, J., Eiger, O., Abd, A., Jeltema, T. E., Giles, P., Hollowood, D. L., Wilkinson, R. D., Turner, D. J., Bhargava, S., Everett, S., Farahi, A., Romer, A. K., Rykoff, E. S., Wang, F., Bocquet, S., Cross, D., Faridjoo, R., Franco, J., … Collaboration, D. (2024). Dark energy survey year 3 results: Miscentring calibration and x-ray-richness scaling relations in redMaPPer clusters. *Monthly Notices of the Royal Astronomical Society*, *533*(1), 572–588. https://doi.org/10.1093/mnras/stae1786

Klein, M., Grandis, S., Mohr, J. J., Paulus, M., Abbott, T. M. C., Annis, J., Avila, S., Bertin, E., Brooks, D., Buckley-Geer, E., Rosell, A. C., Kind, M. C., Carretero, J., Castander, F. J., Cunha, C. E., D'Andrea, C. B., da Costa, L. N., De Vicente, J., Desai, S., … DES Collaboration. (2019). A new RASS galaxy cluster catalogue with low contamination extending to z $\sim$ 1 in the DES overlap region. *Monthly Notices of the Royal Astronomical Society*, *488*(1), 739–769. https://doi.org/10.1093/mnras/stz1463

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In F. Pereira, C. J. Burges, L. Bottou, & K.

Q. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 25). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Limited, R. C. (2025). *PyQt documentation*. https://www.riverbankcomputing.com/static/Docs/PyQt6/

LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S., Bailey, S., Ballantyne, D. R., Bankert, J. R., Barkhouse, W. A., Barr, J. D., Barrientos, L. F., Barth, A. J., Bartlett, J. G., Becker, A. C., … Schmidt, S. (2009). LSST Science Book, Version 2.0. *arXiv e-Prints*, arXiv:0912.0201. https://doi.org/10.48550/arXiv.0912.0201

Marshall, P. J., Verma, A., More, A., Davis, C. P., More, S., Kapadia, A., Parrish, M., Snyder, C., Wilcox, J., Baeten, E., Macmillan, C., Cornen, C., Baumer, M., Simpson, E., Lintott, C. J., Miller, D., Paget, E., Simpson, R., Smith, A. M., … Collett, T. E. (2016). SPACE WARPS - I. Crowdsourcing the discovery of gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, *455*(2), 1171–1190. https://doi.org/10.1093/mnras/stv2009

Melchior, P., Sheldon, E., Drlica-Wagner, A., Rykoff, E. S., Abbott, T. M. C., Abdalla, F. B., Allam, S., Benoit-Lévy, A., Brooks, D., Buckley-Geer, E., Carnero Rosell, A., Carrasco Kind, M., Carretero, J., Crocce, M., D'Andrea, C. B., da Costa, L. N., Desai, S., Doel, P., Evrard, A. E., … Zhang, Y. (2016). Crowdsourcing quality control for Dark Energy Survey images. *Astronomy and Computing*, *16*, 99–108. https://doi.org/10.1016/j.ascom.2016.04.003

Prabhu, K., Raghunathan, S., Millea, M., Lynch, G. P., Ade, P. A. R., Anderes, E., Anderson, A. J., Ansarinejad, B., Archipley, M., Balkenhol, L., Benabed, K., Bender, A. N., Benson, B. A., Bianchini, F., Bleem, L. E., Bouchet, F. R., Bryant, L., Camphuis, E., Carlstrom, J. E., … Zebrowski, J. A. (2024). Testing the ΛCDM Cosmological Model with Forthcoming Measurements of the Cosmic Microwave Background with SPT-3G. *The Astrophysical Journal*, *973*(1), 4. https://doi.org/10.3847/1538-4357/ad5ff1

Raghunathan, S. (2022). Assessing the Importance of Noise from Thermal Sunyaev-Zel'dovich Signals for CMB Cluster Surveys and Cluster Cosmology. *The Astrophysical Journal*, *928*(1), 16. https://doi.org/10.3847/1538-4357/ac510f

Rozo, E., & Rykoff, E. S. (2014). redMaPPer II: X-Ray and SZ Performance Benchmarks for the SDSS Catalog. *The Astrophysical Journal*, *783*(2), 80. https://doi.org/10.1088/0004-637X/783/2/80

Rykoff, E. S., Rozo, E., Busha, M. T., Cunha, C. E., Finoguenov, A., Evrard, A., Hao, J., Koester, B. P., Leauthaud, A., Nord, B., Pierre, M., Reddick, R., Sadibekova, T., Sheldon, E. S., & Wechsler, R. H. (2014). redMaPPer. I. Algorithm and SDSS DR8 Catalog. *The Astrophysical Journal*, *785*(2), 104. https://doi.org/10.1088/0004-637X/785/2/104

Simpson, R., Page, K. R., & De Roure, D. (2014). Zooniverse: Observing the world's largest citizen science platform. *Proceedings of the 23rd International Conference on World Wide Web*, 1049–1054. https://doi.org/10.1145/2567948.2579215

Sobrin, J. A., Anderson, A. J., Bender, A. N., Benson, B. A., Dutcher, D., Foster, A., Goeckner-Wald, N., Montgomery, J., Nadolski, A., Rahlin, A., Ade, P. A. R., Ahmed, Z., Anderes, E., Archipley, M., Austermann, J. E., Avva, J. S., Aylor, K., Balkenhol, L., Barry, P. S., … Young, M. R. (2022). The Design and Integrated Performance of SPT-3G. *The Astrophysical Journal Supplement Series*, *258*(2), 42. https://doi.org/10.3847/1538-4365/ac374f

Sunyaev, R. A., & Zeldovich, Ya. B. (1972). The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *Comments on Astrophysics and Space Physics*, *4*, 173.

Umetsu, K. (2020). Cluster-galaxy weak lensing. *Astronomy and Astrophysics Reviews*, *28*(1), 7. https://doi.org/10.1007/s00159-020-00129-w

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2