https://doi.org/10.1038/s41592-023-01910-2

Towards effective adoption of novel image analysis methods

Talley Lambert & Jennifer Waters

The bridging of domains such as deep learning-driven image analysis and biology brings exciting promises of previously impossible discoveries as well as perils of misinterpretation and misapplication. We encourage continual communication between method developers and application scientists that emphases likely pitfalls and provides validation tools in conjunction with new techniques.

We are microscopists running a busy core facility. Our work sits at the interface between two domains: that of methodology, in which we have watched eagerly as technological advances have enabled faster and more precise measurements, and that of biological inquiry, in which we attempt to deliver the most promising techniques as robust workflows to biologists. To achieve the goal of enabling the adoption of new techniques, our experience has taught us the importance of providing tools that enable biologists to assess whether a method is the right tool for the job and encouraging them to temper excitement for a new method with a healthy level of skepticism. As the domain of deep learning and modern computer vision increasingly proffers groundbreaking solutions to the domain of biology, we would share a reminder of the importance of communication, validation and availability of information.

Imaging is an undeniably powerful technique with an unrivaled ability to extract spatial relationships. At the same time, it is a particularly fraught one: humans routinely misinterpret and misestimate image features, and the aphorism 'seeing is believing' belies the more complicated relationship between visual input and our error-prone brains. For the biologist evaluating a newly published technique, images and figures are often so visually compelling and failure modes are sufficiently subtle that all but the most skeptical adopters could be forgiven for missing critical details. Too often, we see this lead researchers toward techniques that are inappropriate or suboptimal for answering their biological question. As a cautionary example from the world of microscopy, consider the case of the super-resolution technique single-molecule localization microscopy (SMLM)[1].

Presenting SMLM results poses an interpretation challenge: the output 'images' are mere scatter plots or histograms of fluorophore localization events. They give the appearance of an image with infinite contrast: a sea of arbitrarily 'bright' spots on a black background. When compared with conventional images, the contrast is so pleasingly salient that one must actively remind oneself that resolution (the image quality we are ostensibly improving) is much harder to visually assess and is determined by less obvious parameters such as labeling density and localization precision and efficiency[2–4]. In other words, what we believe we are evaluating in the image (that is, resolution) can be easily conflated with ancillary features (that is, contrast), leading to suboptimal tool selection, wasted time and possibly erroneous conclusions when more-precise but lower-resolution tools are eschewed in favor of visually striking images.

The point is not to single out SMLM (a similar argument could be made for the visual preference for and overapplication of confocal microscopy, even when out-of-focus light is not an issue) nor is it to suggest that properly applied SMLM has been less than transformative. The point is to remind us that when 'intuitive' inference systems (such as the human visual system) are used in science, extra care must be taken to ensure that the correct tool is being selected for the question. Developers of new image analysis methods must effectively demand that their adopters do not fall prey to reasonable misinterpretations — particularly when those misinterpretations are visually appealing. Unfortunately, the task of pointing out failure modes and misapplications is often in direct conflict with the incentives that scientists may feel to promote their work: the more generalizable and less error-prone a particular method appears, the more likely it may be to pass peer review and be accepted into a journal. It would be particularly interesting to explore whether journals could establish 'risk–reward' metrics to evaluate the ease of validation and risk of misapplication for a given technique, in addition to the potential effects on scientific discovery. Some techniques that are generally just 'better' than traditional alternatives might offer clear rewards in terms of speed or accuracy, with lower risks for accidental misapplication (modern machine learning-based segmentation techniques come to mind), whereas the advances offered by more experimental or generative methodologies may be accompanied by increased risks or validation challenges for adopters.

Algorithm developers should not overestimate the amount of guidance that their biologist colleagues may need when adopting and validating their techniques. In communications and conferences within our own fields, failure modes and pitfalls of a given approach are often considered with such unconscious ease — pointing them out can even appear banal — that it can be easy for developers to neglect to explicitly state them when communicating across domains. It is often not enough to raise a few caveats in the discussion section of a paper and leave it to the reader to extrapolate them. In our experience teaching imaging courses and workshops, we repeatedly encounter excellent students and postdoctoral researchers who are surprised to learn the fundamental limitations of a popular technique, which suggests that these caveats are still not pervading the consciousness of the research scientists who developers rely upon to ultimately establish the impact of their technique.

One might suggest that the responsibility for the proper use of an image analysis method lies solely with the biologist who wishes to use the technique. But, realistically, given the fast pace of research,

biologists rarely have the time to become experts in the many different techniques that are often required to make important scientific advances. Instead, method developers should strive to provide researchers with an answer to the question "How do I know this is working for my dataset?" Method developers are themselves often intimately familiar with such validation techniques by the time they are ready to publish: conveying these validation 'meta-methods' should be as prioritized as conveying the technique itself. Journals that cater to a cross-domain audience should look for explicit validation guidelines in the review process of a new method.

The most successful workflows almost always involve direct collaboration between developers of image analysis methods and biologists. Although it is becoming more common to hear promotions of 'no-code' solutions for biologists, with simple click-button access to cutting-edge technology that cuts out the need for the developer, the truth is that assembly-line efficiency and independence is not always precise or nuanced enough to work on the messy frontiers of biological research. This is not to say that we cannot eventually make some methods turn-key; it is to say that we should be very cautious about doing so if it comes at the loss of communication between domain experts. The increasing number of courses that are designed to train biologists in the methods and vocabularies of modern machine learning and deep learning are an important step in the right direction, but the goal should not be independence — it should be literacy and more efficient collaboration.

More often than not, biologists who adopt new methods do not have access to local experts who are broadly versed in modern image analysis. This is a problem in itself, one that calls for funding and cultural changes to incentivize career research-support positions and retain academic bioimage analysts. But while we can dream of the 'perfect' — a collaborative utopia in which every biologist has a local image analyst — we must also strategize and aim for the 'good': a community framework in which promising new methods can proliferate, but with an acute awareness of the remaining uncertainties, and, most importantly, a scientific record that provides a mechanism for future reevaluation of conclusions. It should be non-negotiable that raw data and associated code be made available with publications. That availability allows future researchers to fairly evaluate whether they should build upon past results. This is an almost cliché recommendation at this point, and yet the process of making data available is far from simple and ubiquitous. Many journals do require some sort of statement on data availability but could go further to require availability and facilitate the process of selecting and uploading to data repositories.

Lest it appear otherwise, let us emphasize that we are inspired and enthused by the direction and potential of modern bioimage analysis methods. Images have a tremendous information density — much of which is probably lost to human visual inference — and recent developments in imaging (such as light-sheet microscopy) have only increased the deluge of data sitting on hard drives, waiting to be used. As we look ahead, it may not only be that 'seeing' is not 'believing': it may be that seeing is simply not possible, owing both to the sheer volume of data and to the subtleties of biological phenomenon that are hidden beyond reach of our limited visual inference — although they are undoubtedly present in the data. Only through the careful adoption of modern image processing and analysis techniques will the real promises of recent advances in acquisition be realized. We maintain that true progress will require continued close collaboration with the image analysis community as we move toward a future in which the products of our acquisition labor are more frequently converted into truly actionable biological insight.

Talley Lambert [1,2] ✉ **& Jennifer Waters** [1,2] ✉
[1]Department of Cell Biology, Harvard Medical School, Boston, MA, USA. [2]Department of Systems Biology, Harvard Medical School, Boston, MA, USA.
✉e-mail: talley@hms.harvard.edu; jennifer@hms.harvard.edu

Published online: 11 July 2023

References

1. Huang, B., Bates, M. & Zhuang, X. *Annu. Rev. Biochem.* **78**, 993–1016 (2009).
2. Nieuwenhuizen, R. P. J. et al. *Nat. Methods* **10**, 557–562 (2013).
3. Legant, W. R. et al. *Nat. Methods* **13**, 359–365 (2016).
4. Lambert, T. J. & Waters, J. C. *J. Cell Biol.* **216**, 53–63 (2017).

Acknowledgements

J.W. is funded by the Chan Zuckerberg Initiative.

Competing interests

The authors declare no competing interests.