

Where imaging and metrics meet



When it comes to bioimaging and image analysis, details matter. Papers in this issue offer guidance for improved robustness and reproducibility.

In July of 2023 we published a **Focus** issue on the future of bioimage analysis in which we asked experts to share their thoughts and visions for the near and distant future of the field. Common themes emerged, including the importance of computer vision to the future of bioimaging, the necessity of doing appropriate data analysis, and the need for improved data sharing. In this issue, we continue this momentum with papers describing cutting-edge applications of deep learning in microscopy, highlighting the importance of proper metrics for analyzing the performance of bioimage analysis algorithms, and offering practical guidance for reporting microscopy data.

Advances in microscopy have been driven by both improved hardware and computational tools. Underscoring this point are two research papers in this issue that combine advanced microscopes with deep learning for distinct applications. In one Article¹, Jan Huisken and colleagues use deep learning to combine the convenience of imaging GFP with the power of near-infrared imaging for improved imaging at depth in developing animals. In another, Giovanucci and Legant and colleagues introduce smart lattice-light sheet microscopy², an approach that uses artificial intelligence to automatically switch between epifluorescence imaging to survey cells and lattice light sheet microscopy to do fast, high resolution, volumetric imaging of cells undergoing processes of interest.

Given these trends, it seems assured that automated tools involving artificial intelligence will become commonplace in routine applications of microscopy, especially for image analysis tasks including object segmentation, detection and classification. As such, we are at a crucial juncture for understanding and comparing the performance of these tools. This is especially true when one envisages a ‘smart microscopy’ future in which computers are increasingly autonomous from human decision-making in deciding which algorithms are optimal for processing image data.

In 2018, Lena Maier-Hein and colleagues published a study examining 150 large biomedical image analysis competitions and showing, among other things, that competition results were highly sensitive to design choices, including datasets used, how the data were annotated, and the metrics chosen for ranking performance³. Given how visible and influential the results of such competitions are within the field and to the broader community, the concerns they raised and guidance they offered for improving competition robustness, reproducibility and long-term impact are vitally important.

In the years that have followed, Maier-Hein, Annike Reinke, Paul Jäger, Minu Tizabi and a group of nearly 70 collaborators working across multiple disciplines have continued this work. Their team noted that there are few resources available that offer practical guidance in choosing optimal metrics for common image analysis tasks, and they sought to fill this gap – work that has culminated in two Perspective articles in this issue. These Perspectives emphasize the importance of proper image analysis metrics for furthering scientific progress and translation of artificial intelligence advances into practice.

In the first, the researchers present a detailed and comprehensive assessment of metrics used to assess performance of computational algorithms in four common image analysis tasks: image-level classification, semantic segmentation, object detection and instance segmentation⁴. For each of dozens of metrics, they describe associated pitfalls, drawbacks and limitations regarding their use in specific applications. The team also created a common taxonomy to characterize these pitfalls. This piece serves as a unique and invaluable guide for potential users to learn whether using a certain metric for a specific task is associated with known problems before use.

The second piece describes Metrics Reloaded, a comprehensive guide to help users select task-appropriate metrics in bioimage analysis for the same tasks covered in the sister piece⁵. The researchers put forth the idea of a ‘problem fingerprint’ specific to a given image analysis task, which helps guide users to appropriate metrics on the basis of this fingerprint. This framework is also implemented as an online tool for ease of use.

Optimal metrics for image analysis are also covered in a Brief Communication in this issue from Peter Horvath and colleagues⁶. Here, the researchers focus on metrics used to assess the performance of algorithms in segmentation tasks. A highlight of the work is that they identify six and five different interpretations of the popular ‘average precision’ (AP) metric and mean AP (mAP), respectively. They further show that different interpretations can have profound impact on ranking algorithm performance on the same tasks, highlighting again how important attention to detail and proper metric choice are.

Finally, this issue contains a Perspective from Christian Schmied, Christian Tischer, Helena Jambor and colleagues describing a series of checklists for preparing images as publication-quality figures and sharing the supporting methodological details⁷. The piece presents checklists both for image preparation and for image analysis workflows and stands as a landmark piece for those seeking to meet the highest standards in microscopy reporting.

Analyzing, presenting and sharing bioimaging data are commonplace tasks, and many associated field standards are already in place. However, quantitative microscopy and bioimage analysis are becoming increasingly complex, and there is a growing consensus for revised best practices coming from the bioimaging community, as reflected in the papers featured here. Thus, we think there is no time like the present to shake up the status quo in order to advance the field.

We hope these pieces inform future studies and are taken as guides at the planning stages of experimental endeavors to ensure that quantitative bioimaging stands on solid ground as the basis for future biological discovery.

Published online: 12 February 2024

References

1. Gritti, N. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-023-02127-z> (2024).
2. Shi, Y. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-023-02126-0> (2024).
3. Maier-Hein, L. et al. *Nat. Commun.* **9**, 5217 (2018).
4. Reinke, A. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-023-02150-0> (2024).
5. Maier-Hein, L. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-023-02151-z> (2024).
6. Hirling, D. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-023-01942-8> (2023).
7. Schmied, C. et al. *Nat. Methods* <https://doi.org/10.1038/s41592-023-01987-9> (2023).