

Challenges and opportunities in bioimage analysis

Xinyang Li, Yuanlong Zhang, Jiamin Wu & Qionghai Dai



Advanced imaging techniques provide holistic observations of complicated biological phenomena across multiple scales while posing great challenges to data analysis. We summarize recent advances and trends in bioimage analysis, discuss current challenges toward better applicability, and envisage new possibilities.

“More is different.”¹ As indicated by Philip W. Anderson about 50 years ago, the interactions of large-scale elementary units may cause new properties to emerge that cannot be explained by the basic laws governing the elementary unit. Biology is a typical discipline exhibiting such a nature of hierarchical structures with heterogeneous properties across multiple spatial and temporal scales, ranging from genes, proteins, organelles and cells to tissues, organs and whole bodies. Most of these scales have been covered by advanced instruments, such as sequencing technology, cryo-electron microscopy, super-resolution fluorescence microscopy and magnetic resonance imaging, leading to many discoveries in the life sciences. However, there is a long-standing gap in mesoscale imaging, which involves linking information at the cell, tissue and organ scales. To fill this niche, the data throughput of fluorescence microscopy has increased by orders of magnitude in the past decade, opening a new horizon for investigating large-scale intracellular and intercellular interactions in various pathological or physiological states^{2–6}. Indeed, scientists from different areas have invested collaborative efforts to push forward this frontier for diverse applications, such as whole-brain vascular topology⁷, molecular heterogeneity of synaptic morphology⁸ and cell lineages during embryogenesis⁹. Therefore, learning how to process, analyze and understand large-scale imaging data efficiently to catalyze biological discoveries has become increasingly important.

Advances and trends

In the 2010s, deep convolutional networks demonstrated their dominance in computer vision for the first time¹⁰. Since then, deep learning has become a mainstay in contemporary image analysis as a result of its key advantages in both efficiency and performance. This revolution rapidly spread to the field of microscopy, and various intelligent methods have been proposed for different biological applications and imaging modalities to solve the problems of image enhancement, classification, segmentation, cell tracking and others with orders of magnitude improvement in the processing speed, which is vital for large-scale imaging data^{11,12}. Below we discuss important trends and advances for deep learning-based image analysis.

Self-supervised and unsupervised learning. Supervised learning has long been and is still the primary paradigm of deep-learning-based image analysis. With enough training images paired with ground truth at hand, it is not difficult to construct a model with good performance on a specific image analysis task. However, ground-truth data can be difficult or sometimes even impossible to obtain in microscopy, which has become an inevitable limitation of supervised learning. In recent years, the most impressive trend in image analysis is the transition from conventional supervised learning to self-supervised and unsupervised learning (Fig. 1). Using these new training mechanisms, networks can learn to perform specific tasks without requiring any paired ground truth for training^{13–15}. Moreover, a typical feature of optical microscopy is the great data variability between different model organisms, experimental conditions and laboratories. It is a common problem that pretrained models cannot be applied to new data and may generate artifacts. Self-supervised and unsupervised learning provide a better solution to train a customized model for a specific group of data. For large-scale image analysis, self-supervised and unsupervised methods have an inherent advantage because the raw data themselves are a large-scale training set, paving the way for training large models with better generalization.

Vision transformers. Network architecture also continues to undergo remarkable updates to go beyond classical convolutional networks. Vision transformers, a new type of architecture mainly using the self-attention mechanism to extract intrinsic features, have achieved state-of-the-art performance on a variety of computer vision tasks^{16,17}. Their capability to capture long-range dependencies can overcome the local receptive field of convolutional kernels, making them better at integrating global information. Just as in the success that transformers have attained in predicting genetic variants from long DNA sequences¹⁸, there is great potential to discover new phenomena by exploiting long-range spatiotemporal correlations in large-scale imaging data. For example, in neural functional imaging, transformers can help to reveal the causality between two distant events in long-term recording and characterize the relationship between two faraway neurons in mesoscale imaging.

Reinforcement learning. As a computational model rooted in the decision-making process of humans and other animals, reinforcement learning is widely used to build intelligent agents with the ability to interact with the environment by rewarding desired behaviors and punishing incorrect ones. Combined with image analysis, reinforcement learning promises to uncover hidden patterns in large-scale imaging data. For example, by treating the migrating cell as an agent and the other cells as the environment, deep reinforcement learning can infer the mechanism of cell migration during embryo development¹⁹. Reinforcement learning is also suitable for deciphering the neural mechanisms of animal behaviors. If we have a large amount of neural imaging data and synchronize

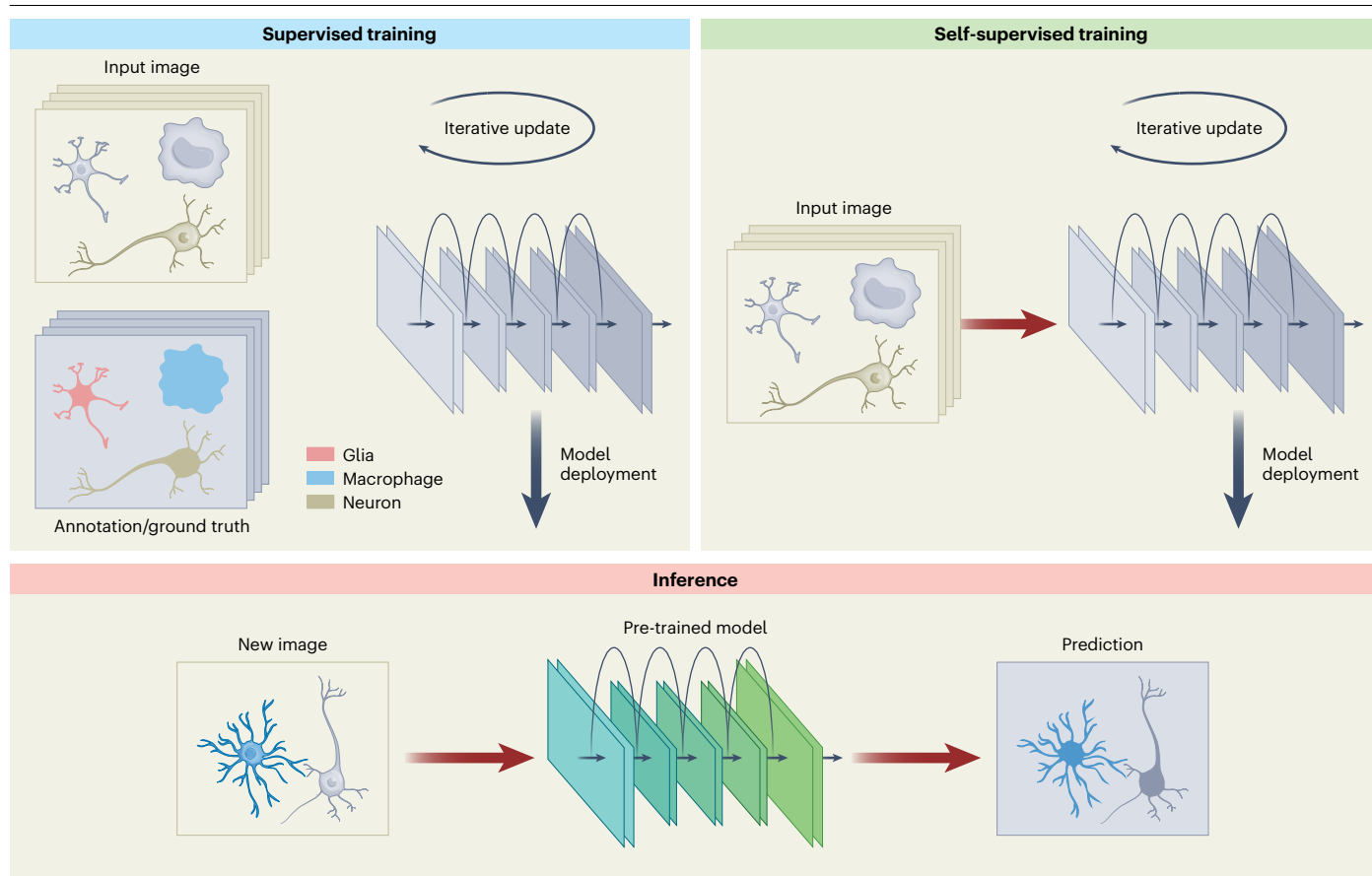


Fig. 1 | Supervised and self-supervised or unsupervised learning for image analysis. In supervised training (top left), each input image must have corresponding annotations (ground truth) to guide the iterative update of network parameters. Self-supervised and unsupervised training do not need any paired ground truth; they can learn to update network parameters by

mining the endogenous features and correlations inside the input images. After convergence, pretrained models can be used to predict the result of unseen images. This is an illustrative example based on cell segmentation that also works for other tasks such as cell tracking and pose estimation.

visual input and sensory stimuli, it is possible to build cognitive models in a data-driven manner using reinforcement learning. Such a paradigm can be extended to diverse fields to extract huge amounts of biological discovery through the screening of large-scale imaging data during different pathological or physiological states in different organs. Another success of deep reinforcement learning is to train intelligent agents to complete various challenging tasks, such as playing video games²⁰. This inspires us because by designing the reward properly and learning from a large amount of experiment data iteratively, a microscope can learn to interact with the sample, complete a given imaging task with optimal parameters, and discover interesting phenomena automatically.

Physics-informed methods. Imaging is a rigorous optical process that needs to be reliable enough to support scientific discoveries. In fact, optical imaging is a means of quantitative measurement. The intensity of each pixel has a specific biophysical and biochemical meaning, such as ion concentration in calcium imaging, membrane potential in voltage imaging, and gene expression in spatial transcriptomics. Preserving the quantitative property in processed images is therefore critical for researchers to decode underlying biological phenomena. The analysis of imaging data, especially image processing, should

maintain this reliability as much as possible. For low-level vision tasks aimed at improving image quality (for example, denoising, super-resolution reconstruction or deblurring), incorporating the physics of image formation into the processing framework can increase the confidence level of the results²¹. It has been broadly verified that better modeling of the imaging process, such as optical aberrations²² and scattering²³, leads to better performance, especially for computational imaging methods.

Challenges toward better applicability

Although impressive achievements have been made in bioimage analysis, there are still some issues that hinder the applicability of these tools in optical microscopy. Here we summarize several key points and hope that concerted efforts could be made by the research community to resolve these issues.

Standard and metrics. Publishing standard validation datasets and setting corresponding performance metrics can promote the rigorous development of bioimage analysis methods. Since images from different imaging modalities vary greatly, a specific dataset should be archived for each modality. These validation datasets must be

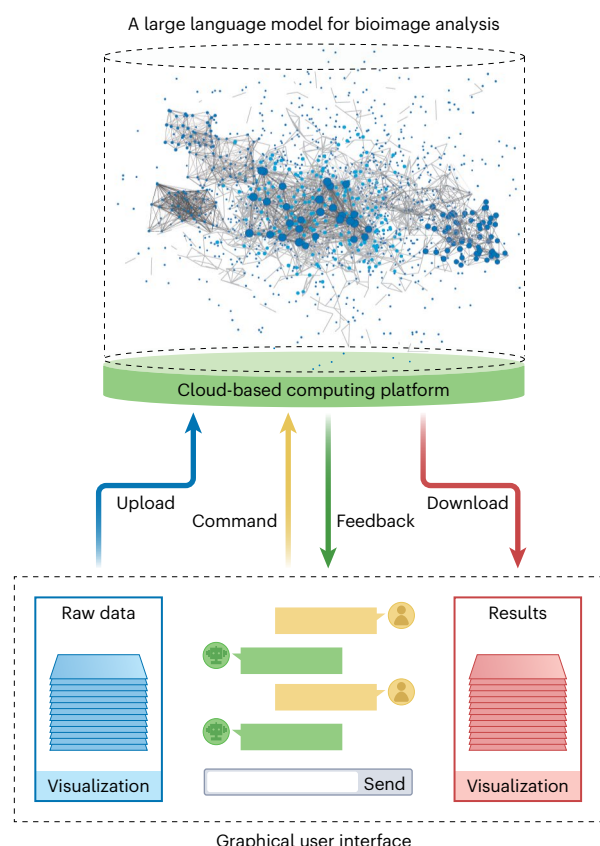


Fig. 2 | Using large language models for analyzing bioimaging data. A pretrained large language model is deployed on a cloud-based computing platform. Users can send commands and get results by communicating with the model through a dialog box.

representative enough and cover a wide enough range of samples and imaging conditions. For intuitive comparison and ranking, the performance metric of each dataset should preferably be a number that can reflect comprehensive capability of a method. Efficiency must be emphasized for large-scale analysis to facilitate practical applications.

Interpretability and reliability. Despite the superior performance, deep learning in microscopy suffers from a trust crisis because of its black-box nature. It is a long-standing challenge to endow deep neural networks with interpretability. Realizing interpretable deep learning requires innovations in the most fundamental concepts. But some technical treatments such as feature visualization and physics-based modeling can improve the reliability to some extent^{21,24}. The biggest concern of most researchers is potential artifacts in the results. It will help relieve the trust crisis if we have practical approaches to evaluate these errors quantitatively. For example, in addition to the results, trained models that provide the corresponding confidence levels at the same time would be advantageous.

General platforms. State-of-the-art image analysis methods are built on the latest advances in computer vision. Using these methods requires strong programming skills and professional background, which can be troublesome for biologists without computational

expertise. The most successful platform for biological image analysis is Fiji, based on ImageJ²⁵. There is an urgent demand for a new interactive platform or updated Fiji to implement deep-learning-based image analysis. A good platform should not only enable the deployment of pretrained models²⁶ but, more importantly, support the training of new models. These operations are all computationally demanding, so the platform must consider how to have easy access to local or cloud-based computing power²⁷. Since most current deep-learning methods are based on Python, the new platform should be compatible with Python to make full use of open-source resources²⁸.

Data sharing. The huge amount of data in biological imaging makes data sharing quite difficult. To facilitate sharing of large-scale datasets, first, web platforms supporting online preview and download of high-dimensional imaging data are needed. Second, high-efficiency, lossless compression should be applied to reduce the demands on transmission bandwidth and storage devices. Moreover, a new data format that divides original large-scale dataset into many units would be useful, as each unit would be independently available and meaningful. The metadata should include not only the basic description of datasets but also thumbnails. Users can then have a concrete understanding of the data without downloading the whole dataset.

New possibilities offered by emerging technologies

Recent research in artificial intelligence suggests that large language models (LLMs) can have human-level performance in language comprehension, reasoning and programming²⁹. Among them, ChatGPT and multimodal GPT-4 have aroused wide attention owing to their ability to understand users and respond fluently on various topics. The great potential of LLMs provides the possibility to build a professional model to help us analyze imaging data (Fig. 2). This would be an artificial intelligence engineer that can understand our needs and process the data automatically, such as writing custom scripts and training specific deep learning models. All we would need to do is to upload our data, enter our requirements in the dialog box and give some intermediate comments, which effectively relieves the heavy burden of large-scale image analysis. However, there is still a long way to go to accumulate enough training examples and train such an intelligent and professional model. Open access to code and data of published papers is critical to realizing this long-term goal. Additionally, a widespread concern is that current LLMs sometimes give fictitious answers. Strict criteria and validations must be in place to ensure they are properly used in image analysis.

The growing demand for large-scale image analysis poses another great challenge to computing power. Conventional silicon processors can hardly satisfy the requirement for high-speed processing. Optical computing is an emerging technology that processes information at the speed of light by using photons instead of electrons for computation³⁰. Not only the inference of deep neural networks but also basic matrix operations could be implemented by integrated photonic circuits³¹. Combining image analysis with optical computing promises to improve the processing speed to a much higher level and enable data-heavy and high-throughput applications such as single-cell sequencing and image-based high-content screening.

Xinyang Li[✉], Yuanlong Zhang, Jiamin Wu[✉] & Qionghai Dai[✉]
 Department of Automation, Tsinghua University, Beijing, China.
 ✉e-mail: wujiamin@tsinghua.edu.cn; qhdai@tsinghua.edu.cn

Published online: 11 July 2023

References

1. Anderson, P. W. *Science* **177**, 393–396 (1972).
2. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. *eLife* **5**, e14472 (2016).
3. Fan, J. et al. *Nat. Photonics* **13**, 809–816 (2019).
4. McConnell, G. et al. *eLife* **5**, e18659 (2016).
5. Wu, J. et al. *Cell* **184**, 3318–3332.e17 (2021).
6. Kim, T. H. & Schnitzer, M. J. *Cell* **185**, 9–41 (2022).
7. Todorov, M. I. et al. *Nat. Methods* **17**, 442–449 (2020).
8. Helm, M. S. et al. *Nat. Neurosci.* **24**, 1151–1162 (2021).
9. Ma, X. et al. *Nat. Methods* **18**, 893–902 (2021).
10. Krizhevsky, A., Sutskever, I. & Hinton, G. E. *Commun. ACM* **60**, 84–90 (2017).
11. Moen, E. et al. *Nat. Methods* **16**, 1233–1246 (2019).
12. Belthangady, C. & Royer, L. A. *Nat. Methods* **16**, 1215–1225 (2019).
13. Li, X. et al. *Nat. Methods* **18**, 1395–1400 (2021).
14. Midtvedt, B. et al. *Nat. Commun.* **13**, 7492 (2022).
15. Li, X. et al. *Light Sci. Appl.* **10**, 44 (2021).
16. Han, K. et al. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**, 87–110 (2023).
17. Wang, Z., Xie, Y. & Ji, S. *Nat. Mach. Intell.* **3**, 161–171 (2021).
18. Avsec, Ž. et al. *Nat. Methods* **18**, 1196–1203 (2021).
19. Wang, Z., Xu, Y., Wang, D., Yang, J. & Bao, Z. *Nat. Mach. Intell.* **4**, 73–83 (2022).
20. Mnih, V. et al. *Nature* **518**, 529–533 (2015).
21. Li, Y. et al. *Nat. Methods* **19**, 1427–1437 (2022).
22. Wu, J. et al. *Nature* **612**, 62–71 (2022).
23. Zhang, Y. et al. *Nat. Commun.* **12**, 6391 (2021).
24. Zeiler, M. D. & Fergus, R. In *European Conf. Computer Vision* (eds Fleet, D., Pajdla, T., Schiele, B. & Tuytelaars, T.) 818–833 (Springer, 2014).
25. Schindelin, J. et al. *Nat. Methods* **9**, 676–682 (2012).
26. Gómez-de-Mariscal, E. et al. *Nat. Methods* **18**, 1192–1195 (2021).
27. Haase, R. et al. *Nat. Methods* **17**, 5–6 (2020).
28. Rueden, C. T. et al. *Nat. Methods* **19**, 1326–1327 (2022).
29. Li, Y. et al. *Science* **378**, 1092–1097 (2022).
30. Lin, X. et al. *Science* **361**, 1004–1008 (2018).
31. Bogaerts, W. et al. *Nature* **586**, 207–216 (2020).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (62088102, 62222508).

Author contributions

All authors contributed equally to this work.

Competing interests

The authors declare no competing interests.