

ScienceDirect

Biotechnology

Defining the boundaries: challenges and advances in identifying cells in microscopy images

Nodar Gogoberidze and Beth A Cimini



Segmentation, or the outlining of objects within images, is a critical step in the measurement and analysis of cells within microscopy images. While improvements continue to be made in tools that rely on classical methods for segmentation, deep learning-based tools increasingly dominate advances in the technology. Specialist models such as Cellpose continue to improve in accuracy and user-friendliness, and segmentation challenges such as the Multi-Modality Cell Segmentation Challenge continue to push innovation in accuracy across widely varying test data as well as efficiency and usability. Increased attention on documentation, sharing, and evaluation standards is leading to increased user-friendliness and acceleration toward the goal of a truly universal method.

Address

Imaging Platform, Broad Institute, Cambridge, MA 02142, USA

Corresponding author: Cimini, Beth A (bcimini@broadinstitute.org)

Current Opinion in Biotechnology 2023, 85:103055

This review comes from a themed issue on Analytical Biotechnology

Edited by Andreas E. Vasdekis and Scott E. Baker

Available online xxxx

https://doi.org/10.1016/j.copbio.2023.103055

0958-1669/© 2023 Elsevier Ltd. All rights reserved.

Introduction

Segmentation plays a pivotal role in microscopy analysis and refers to the automatic delineation of individual objects (often cells or cellular components) within complex scientific images. It is an important step before measuring properties of those biological entities. Approaches for cell segmentation have benefitted from advancements in more general segmentation problems in traditional computer vision (CV), machine learning (ML), and in recent years deep learning (DL) [1,2]. Accurate segmentation allows the quantification and analysis of cellular features, such as morphology, staining intensity, and spatial relationships, which capture valuable cellular phenotypes. While computational methods now achieve better-than-human accuracy on a number of specific tasks, in general, given the wide range of cell types, imaging modalities, and experimental conditions, the problem remains an ongoing challenge.

As state-of-the-art (SOTA) methodologies for segmentation have progressed, the community has also tried to provide access to these methods to less-computational users in the form of user-friendly software interfaces and intuitive tools that improve reproducibility. Widespread adoption will require methods with few-or-no tunable parameters, models that are efficient in terms of computational runtime and memory requirements, and an ecosystem of tools for their use. The past two years reviewed here have seen a proliferation of new local and cloud-oriented software and workflows, the adoption of several user-oriented models, and the development of next-generation model architectures. We herein review progress in approaches utilizing classical CV techniques and specialist DL networks, as well as progress toward and current needs related to making high-qualityaccessible generalist networks that will reduce the 'time to science' for the broader community.

Progress in classical approaches

While advancements in segmentation accuracy are largely driven via DL approaches, they are not always a suitable solution, as some require large annotated datasets and interpretability (though often unnecessary for segmentation tasks) can be a challenge [3,4]. We therefore begin with advancements in non-deep ML and classic image processing. Kartezio [5] is a recent exemplar of non-deep ML, using Cartesian Genetic Programming to combine classic CV algorithms into a fully interpretable image pipeline for segmentation. It performs comparably to DL approaches such as Cellpose [6], StarDist [7], and Mask R-CNN [8], and importantly requires relatively few images to train a computationally efficient and explainable workflow.

Existing CV tools such as Fiji/ImageJ [9], CellProfiler [10], and Napari [11] are receiving continuous extensions via the growing ecosystem of plugins and integrations [1,12–15], allowing these tools to adapt to a broader range of tasks. One example, General Image Analysis of Nuclei-based Images [16], is a Fiji plugin for segmentation of cells in 3D microscopy images. Similarly, LABKIT [17] is a Fiji plugin specifically oriented toward efficient segmentation in large, multiterabyte, images. Other tools such as Tonga [18] prioritize ease of

installation and customization to a specific task to appeal to nontechnical and nonexpert users.

Lowering time-to-science in deep learning

As new DL models rapidly emerge, users who want to employ them must surmount computational hurdles. First, installing the models and their corresponding software dependencies can often be technically complex and time-consuming, often involving usage of the Command Line Interface (CLI), compiling source code, navigating environment-related conflicts, and other tasks that may confound nontechnical users. Second, pretrained models are most effective on the type of data they have been trained on, which may differ from a user's data in a number of ways such as the phenotype of interest, how the biological samples were prepared, the imaging modality (fluorescence, histological stains, phase contrast, etc.), and the experimental conditions. The model's generality is how well it is able to perform across these differences, without loss in segmentation accuracy. If the model's generality is not sufficient to perform well on the user's data, it must be fine-tuned by feeding in new training data (Figure 1). The complexity of this task can vary even more dramatically than installation. Depending on the tools made available by the model's developers, it may involve writing custom code, which can be a timeintensive task even for expert programmers, or may simply require the use of a purpose-built tool.

The less time spent installing, tuning, and configuring models and software, the more bandwidth is available to concentrate on addressing scientific questions. Unfortunately, usability varies wildly across tools and documentation for trained DL models, from only the model parameters (weights) without documentation or source code, to models that come with extensive documentation and entire libraries for utilizing the model, including data loading and processing, fine-tuning, and configuration, to models that come with several interfaces, including CLI or Graphical User Interfaces (GUI), easy-to-use installers, and guides or tutorials on how to fine-tune and configure the model. It is no accident that some of the most commonly used networks, discussed below, are those that in addition to high performance have emphasized usability.

Current progress toward useful specialist networks

As shown in Figure 1, models typically underperform on new, unseen data. Before fine-tuning, these data are considered out of distribution (OOD). Intuitively, OOD data can be thought of as being drawn from a different distribution than that of the original training set (such as the purple versus green cells in Figure 1). A shift between the in-distribution training set and OOD data can be referred to as a difference in *style* between the two sets of data, for instance, differences in acquisition

parameters, staining methods, or imaging modalities [19,20]. A process known as style transfer can be utilized to address changes in the data distribution by training a model that is able to pixel-wise map an image of one style to that of another, ideally with minimal loss in semantic content [21]. For instance, a style transfer model can be trained to transform an image of one modality, such as brightfield, to that of another, such as fluorescence; or a model can be trained to transform an image's annotation mask into the image itself. The nucleAlzer [22] model utilizes the latter approach as a means of achieving greater generalization capabilities, allowing the model to more easily be adapted to OOD data. Although developed over three years ago, it is still one of the top-performing models at cell segmentation [23]. In order to improve usability, a plugin was developed for CellProfiler 3 [24] allowing users of the tool to perform inference through the use of a GUI. While the plugin allows ease of use via the CellProfiler interface, manual installation of nucleAIzer and its dependencies are still necessary, which is a challenge for noncomputational users. A simple web interface is also available, but requires upload of the data to a central server, limiting use for large batches of images.

StarDist predates nucleAIzer, however, iterations of it are still being introduced [25,26], and general usage remains quite high due to a great deal of time being spent on documenting its usage, as well as making it available across environments and in a variety of graphical tools [12] such as Fiji/ImageJ, Napari, QuPath [27,28], Icy [29], CellProfiler, and KNIME [30]. While a variety of pretrained variants are available, it is still limited to specific modalities (such as fluorescence or histology stains) and even with fine-tuning is oriented to segmenting objects that are *star-convex* — shapes where line segments can be drawn from any point along the border to some single interior point — that makes it a poor choice for very irregular cell shapes or neurons, since shapes with very large bends or curves may not be star-convex.

A primary goal of Cellpose was to develop a generalist model by training on a large dataset of manually segmented images from a variety of modalities. Its preprocessing method, focusing on transforming input data to spatial gradients, allows it to generalize to a larger variety of shapes. While the architecture was developed for generalist purposes, fine-tuning is still often necessary. Cellpose 2.0 [31] was introduced as a package that included several pretrained models, a human-in-the-loop pipeline for fine-tuning custom models with small datasets, and an improved set of graphical software tools to aid in its usage. Omnipose [32] extends Cellpose to work better on elongated cells common in bacteria by adding distance field prediction similar to StarDist. It is similarly well-documented, packaged, and provides a number of interfaces in the form of a library, CLI, and



Figure 1

Segmentation and fine-tuning process. (a) The inference dataset contains a variety of samples: some are similar to the original training data the model was pretrained on, while others are slightly or very different, leading to low segmentation accuracy. (b) New training data matching the characteristics of the full distribution are annotated either manually, through software such as CellProfiler, or a human-in-the-loop model such as Cellpose, and used to fine-tune the model. (c) The fine-tuned model produces more accurate segmentations on the inference dataset, which can then be used for downstream tasks Figures created with BioRender.com.

GUI. While it allows for training, it is not available in Cellpose 2.0, and thus is not focused on human-in-theloop fine-tuning. Cellpose and Cellpose 2.0 include a custom GUI, and extensive documentation; broad usage is further supported by their availability via plugins from many of the same tools as StarDist.

Mesmer is a DL pipeline trained on the largest public tissue dataset of annotated nuclei and whole cells, TissueNet [33]. It provides access to a remotely hosted instance model through several interfaces, including a

web portal and plugins for Fiji/ImageJ and QuPath. It also provides a Docker container (described below, in Making Models Findable, Accessible, Interoperable, and Reusable (FAIR)) to run the model in a self-hosted manner, with access provided through a Jupyter Notebook or CLI.

A vision for the state-of-the-art

Many of those working on DL for image segmentation are aiming to create a truly generalist model, often referred to as a foundation model, capable of greater-than-human accuracy, across the wide variety of imaging modalities, in a parameter-free manner [34–36]. For biologists, this would mean a model that has no need for manual annotation or fine-tuning, preferably in a form factor that is easily accessible, configurable, and invocable. Although there is still much progress to be made on this front, a number of architectural developments can be highlighted as mile markers along the path.

Dataset availability

The effectiveness of DL models relies on sufficient similarity between the training data and the user's own data. The specific criteria for what constitutes 'sufficient similarity' will naturally differ based on the methods and architectural choices employed by the model, but it therefore follows that creating foundation models (capable of segmenting a wide variety of biological images) will require a diverse and comprehensive corpus of training data to ensure that the segmentation model can generalize effectively. The dataset should encompass a broad spectrum of microscope modalities, across a variety of imaging conditions, and include many distinct cell types.

The volume of publicly available datasets is ever-increasing, particularly driven from the development of specialist models and challenges wherein teams compete on segmentation-oriented tasks. TissueNet is the largest collection of annotated tissue images, while LIVECell [37] is the largest collection of high-quality, manually annotated, and expert-validated phase-contrast images. The Cell Tracking Challenge (CTC) [38] is an ongoing benchmark and reference in cell segmentation and tracking algorithms, which in recent years has extended the available benchmarks with the Cell Segmentation Challenge (MMCSC) [36] consolidated a modestly sized labeled dataset with a particular emphasis on diversity in modalities.

Challenges stand as an excellent pointer for future progress in a given area; trends among the top-ranking team's architectures and techniques often form the basis of future implementations available to the wider field. For instance, data augmentation — where existing training data are perturbed and transformed in various ways, such as rotations, scaling, intensity adjustments, or the infusion of random noise — was highlighted in the MMCSC as a particularly important feature in pretraining top-performing models, aiding them in their generalizability. Entirely synthetic datasets are also often useful, as demonstrated by their inclusion in a subset of the CTC datasets, and in the development of frameworks for their generation [39].

Next-generation models

Many architectures are being researched and explored in the quest for evermore general and robust models, as demonstrated in the MMCSC. The KIT-GE model [40] is among the top-3 performing models of CTC CSB, and was therefore used as one of the baselines in the MMCSC alongside Cellpose, Cellpose 2.0, and OmniPose [6,31,32]. Analysis of the top-ranking solutions in the MMCSC shows that choices in backbone networks are particularly important for next-generation models. While U-Net inspired architectures form the basis of the widely used contemporary models such as StarDist, Cellpose, and KIT-GE, the winning solutions in MMCSC employ backbones such as SegFormer [41], ConvNeXt [42], and ResNeXt [43]. CTC reports that segmentation performance increases with techniques such as self-configured neural networks (e.g. nnU-Net) [44], neural architecture search, and multibranch prediction.

The generalist capabilities provided by Cellpose 2.0 rely on fine-tuning, which may cause the model to suffer from severe loss of performance on tasks outside of the fine-tuning dataset [36]. The top-performing model [45] of the MMCSC was able to outperform the pretrained generalist Cellpose and Omnipose models, as well as a Cellpose 2.0 model fine-tuned on the challenge's training data. The testing set included images that were distinct from the training data, and sourced from new biological experiments, meaning successful models needed to show a strong ability to generalize across data without additional fine-tuning.

While the winning solution of MMCSC and its associated code is available on GitHub, it remains to be seen whether it or any of the top-ranking models will be available with documentation and interface tools in a way comparable to StarDist, Cellpose 2.0, or Omnipose. If so, we may be one step closer to a truly general, easyto-set up, easy-to-use, one-click segmentation model, with no additional tuning. Short of that, alternative interfaces for model configuration may come to prominence in the form of dialog-driven large language models (LLMs) [46,47]. It is also as yet unclear whether future enhancements will be driven primarily by transformer architectures [34,48], whether advancements in convolutional networks [42] will keep pace, as demonstrated by the second- and third-place solutions in the MMCSC, or whether hybrid approaches will dominate [49,50].

It is also of great research and commercial interest to develop foundation models. Meta AI Research recently released a family of foundation models for segmentation, referred to as the Segment Anything Model [51] (SAM), the largest of which was trained on 1.1 billion high-quality segmentation labels, across 11 million high-resolution images. While the images included in the dataset were mostly photographs of natural scenes, it did include a small number of microscopy images taken from the 2018 Data Science Bowl [52]. In short order, Segment Anything for Microscopy [53] was developed, in which SAM was extended to generalize across many imaging modalities by fine-tuning the original model using a variety of datasets. Important to SAM's architecture is its interactive segmentation capabilities, where a subset of the user's data is first annotated with a small amount of either point annotations or rectangular bounding boxes. Annotations of this type are significantly less time-consuming than pixel-level mask annotations, and provide SAM with enough guidance to output full segmentation masks on the user's dataset. Segment Anything for Microscopy therefore adopts this capability and includes a Napari plugin for interactive and automatic segmentation. There is a mechanism for automatic segmentation, however, in order to get generalist accuracy above that of Cellpose, some manual annotations must be made.

Making models Findable, Accessible, Interoperable, and Reusable

Alongside progress in model development, there has been a greater push toward the dissemination of models such that they are FAIR [54-56]. Not only should models be available on publicly accessible platforms, but the associated code for asset loading, data preprocessing, data postprocessing, and model training and model inference, should also be made available in a well-packaged and documented form. Container platforms such as Docker [57] can alleviate many installation and setup complexities, providing an isolated and controlled environment in which software is installed, and a preconfigured installation process for the target software. This 'containerization' of software and its dependencies dramatically decreases the barriers for reuse. In addition, making models readily accessible, configurable, and tunable in a low- or no-code manner via interactive code notebooks or GUI encourages broader adoption of SOTA models.

The CTC proposed guidelines for algorithm developers to make their workflows both available and reproducible, while currently optional, they will be mandatory in the future. At minimum, the source code should be available on a public repository, and contain clear instructions for installing dependencies, initializing the model, loading weights, and training with new data. They also pushed for source code to be available via notebooks such as Jupyter and Google Colab. In the same vein, the MMCSC required all participants to place their solutions in Docker containers; the winning teams have made these available on public image registries and also made their algorithms publicly available on GitHub alongside processing source code. In addition, the top-three solutions were encouraged to develop Napari [11] plugins.

A missing component in full adherence to FAIR principles is the interoperability of models. While research and development in model architectures are healthy and vital, there is no agreed-upon specification on the inputs and outputs of models. The difference in the model outputs between, for example, StarDist and Cellpose, is stark, and the postprocessing that is needed is correspondingly distinct. While the outputs are necessary byproducts of the model architectures, the lack of a standard makes interoperability with existing tools difficult as custom code needs to be written to mirror the postprocessing steps.

Making models efficient

Models vary widely in terms of algorithmic efficiency, which will affect their adoption, especially in low-resource settings. While some challenges such as the CTC emphasize segmentation accuracy alone, others such as MMCSC evaluate efficiency as an explicit criterion, and the top-performing models had good trade-offs between accuracy and efficiency in runtime and memory usage. There are additional efforts in bringing model optimization tools to the bioimage community [58,59], as well as reducing runtime of bioimage analysis pipelines in general [59,60]. Efficient models, and model optimization tools will become increasingly important for training and inference tasks in local desktop and web-based tools [2,61], especially in contexts where moving data to the cloud is not viable or allowed.

Improving tool access and availability

Though local-first software and algorithms remain important to grow and maintain, cloud-oriented tools and resources for bioimage analysis are becoming more prevalent and easier to use as demand for large-scale cloudbased workflows increases. In addition to providing large storage capacity and high-performance hardware, cloudbased tools increase the availability and accessibility of models, and often move the technical complexity away from the end user.

Notebooks allow code, explanatory text, and interactive elements to live together in a single package, providing alternatives or complements to libraries, documentation, and GUI interfaces. ZeroCostDL4Mic [62] provides rigorously documented and annotated code notebooks with prewritten code that can be customized for specific workflows through the exposed settings. A major benefit to these notebooks is that they can be deployed either locally or on the Google Colab platform, which eases hardware requirements and allows running moderately sized workloads for free. Behind the scenes, a container is initialized in the cloud and the installation occurs via preconfigured installation scripts contained within the notebooks.

Beyond notebooks, other tools provide a larger degree of customization and control, albeit at the expense of additional complexity. The BioContainers project is an open-source and community-driven framework that provides cloud resources for defining, building, and distributing containers for biological tools [63]. The BioContainers Registry was developed with FAIR principles in mind, and provides both web and RESTful API interfaces to search for bioinformatics tools [64]. BIAFLOWS [65] is a community-driven, open-source web platform that allows deployment of and access to a wide variety of reproducible image analysis workflows. The platform provides a framework to import data, encapsulate workflows in container images, batch-process data, visualize data, and assess performance using widely accepted benchmark metrics on a large collection of public datasets. BioImageIT [66] is a more recent, plugin-oriented, workflow tool for data management and analysis. It has a unique emphasis on reconciling existing data management and data analysis tools, and although run locally has the ability to tap into remote data stores and job runners.

The BioImage Model Zoo [67] provides a communitydriven repository for pretrained DL models and promotes a standard model description format for describing metadata. Community partners can work with the Bio-Image Model Zoo to support execution of the models and include many common bioimage tools. In addition, model execution can be performed via the BioEngine application framework, on top of the ImJoy plugin framework [68], allowing inference both on the BioImage Model Zoo web application and other web applications using the client ImJoy software. Behind the scenes, multiple containers are being run and managed with a container orchestration tool called Kubernetes.

For moderately more technical users, who are comfortable with using tools for deploying their own container orchestration workloads, there are some additional options. DeepCell Kiosk [69] is a cloud-native tool for dynamic scaling of image analysis workflows, utilizing Kubernetes orchestration similarly to the BioEngine inference engine. The tool is managed from several interfaces, including a web portal and Fiji plugin. Distributed-Something [70] takes a script-based approach to scale and distribute arbitrary containerized jobs on AWS, automatically configuring the AWS infrastructure for container orchestration, monitoring, and data handling. It runs the work in a costeffective manner, and cleans up the infrastructure when the work has been completed.

Conclusion

The landscape of segmentation algorithms, enabling tools, workflow management systems, repositories, benchmarks, and challenges, is constantly shifting. This very active landscape makes it all the more important to create community standards for reporting on methods and robust segmentation quality metrics, on which there has been recent guidance [35,71–73]. While there is still much work to do, the past two years have seen essential strides made in democratizing the use of advanced segmentation methods through user-friendly interfaces and improved documentation. Integrating tools and scaling up reproducible workflows fosters a more collaborative and robust ecosystem; these continuing trends will empower researchers from diverse backgrounds to collectively explore the intricate universe of single-cell biology, ultimately accelerating the pace of discovery and innovation in this vital field of study.

Funding

This publication has been made possible in part by Chan Zuckerberge Initiative Grants 2020-225720 (DOI: 10.37921/977328pjvbca) and 2021-238657 (DOI: 10.37921/365498zdfyyk) from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation (funder DOI 10.13039/ 100014989). This work was also supported by the Center for Open Bioimage Analysis (COBA) funded by the National Institute of General Medical Sciences P41 GM135019. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the paper.

CRediT authorship contribution statement

Nodar Gogoberidze: Conceptualization, Visualization, Writing – original draft, Writing – review & editing. **Beth A Cimini:** Conceptualization, Visualization, Writing – review & editing, Funding acquisition.

Data Availability

No data were used for the research described in the article.

Declaration of Competing Interest

The authors declare that there are no competing interests associated with the paper.

Acknowledgements

The authors thank the members of the Imaging Platform for helpful discussions and contributions.

Figures created with BioRender.com

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- •• of outstanding interest
- Hollandi R, Moshkov N, Paavolainen L, Tasnadi E, Piccinini F,
 Horvath P: Nucleus segmentation: towards automated solutions. *Trends Cell Biol* 2022, 32:295-310, https://doi.org/10. 1016/j.tcb.2021.12.004.

A comprehensive review, up to 2021, on the state of nucleus segmentation.

- Lucas AM, Ryder PV, Li B, Cimini BA, Eliceiri KW, Carpenter AE: Open-source deep-learning software for bioimage segmentation. *Mol Biol Cell* 2021, 32:823-829, https://doi.org/10. 1091/mbc.E20-10-0660
- Rudin C: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 2019, 1:206-215, https://doi.org/10.1038/s42256-019-0048-x
- Karim MR, Islam T, Shajalal Md, Beyan O, Lange C, Cochez M, Rebholz-Schuhmann D, Decker S: Explainable AI for bioinformatics: methods, tools and applications. *Brief Bioinform* 2023, 24:bbad236, https://doi.org/10.1093/bib/bbad236
- Cortacero K, McKenzie B, Müller S, Khazen R, Lafouresse F,
 Corsaut G, Van Acker N, Frenois F-X, Lamant L, Meyer N, Vergier B, Wilson DG, Luga H, Staufer O, Dustin ML, Valitutti S, Cussat-Blanc S, Kartezio: evolutionary design of explainable pipelines for biomedical image analysis, *arXiv* [cs.CV]; 2023. http://arxiv.org/ abs/2302.14762

Introduction of a tool that uses machine learning techniques to develop pipelines composed of combinations of classical, interpretable image processing methods for segmentations.

- Stringer C, Wang T, Michaelos M, Pachitariu M: Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 2021, 18:100-106, https://doi.org/10.1038/s41592-020-01018-x
- Schmidt U, Weigert M, Broaddus C, Myers G: Cell detection with Star-Convex polygons. Springer International Publishing; 2018:265-273, https://doi.org/10.1007/978-3-030-00934-2_30
- He K, Gkioxari G, Dollar P, Girshick R: Mask R-CNN. IEEE Trans Pattern Anal Mach Intell 2020, 42:386-397, https://doi.org/10.1109/ TPAMI.2018.2844175
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A: Fiji: an open-source platform for biological-image analysis. Nat Methods 2012, 9:676-682, https://doi.org/10.1038/nmeth.2019
- Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A: CellProfiler 4: improvements in speed, utility and usability. BMC Bioinform 2021, 22:433, https://doi.org/10. 1186/s12859-021-04344-9
- 11. Ahlers J, Althviz Moré D, Amsalem O, Anderson A, Bokota G, Boone P, Bragantini J, Buckley G, Burt A, Bussonnier M, Can Solak A, Caporal C, Doncila Pop D, Evans K, Freeman J, Gaifas L, Gohlke C, Gunalan K, Har-Gil H, Harfouche M, Harrington KIS, Hilsenstein V, Hutchings K, Lambert T, Lauer J, Lichtner G, Liu Z, Liu L, Lowe A, Marconato L, Martin S, McGovern A, Migas L, Miller N, Muñoz H, Müller J-H, Nauroth-Kreß C, Nunez-Iglesias J, Pape C, Pevey K, Peña-Castellanos G, Pierré A, Rodríguez-Guerra J, Ross D, Royer L, Russell CT, Selzer G, Smith P, Sobolewski P, Sofiiuk K, Sofroniew N, Stansby D, Sweet A, Vierdag W-M, Wadhwa P, Weber Mendonça M, Windhager J, Winston P, Yamauchi K: napari: a multi-dimensional image viewer for Python. Zenodo 2023,1-2, https://doi.org/10.5281/ZENODO.3555620
- Haase R, Fazeli E, Legland D, Doube M, Culley S, Belevich I,
 Jokitalo E, Schorb M, Klemm A, Tischer C: A Hitchhiker's guide through the bio-image analysis software universe. *FEBS Lett* 2022, 596:2472-2485, https://doi.org/10.1002/1873-3468.14451.
 A review of the state of bio-image analysis tools, for segmentation and beyond, up to 2022.
- Selzer GJ, Rueden CT, Hiner MC, Evans EL 3rd, Harrington KIS, Eliceiri KW: napari-imagej: ImageJ ecosystem access from napari. Nat Methods 2023, 20:1443-1444, https://doi.org/10.1038/ s41592-023-01990-0
- Weisbart E, Tromans-Coia C, Diaz-Rohrer B, Stirling DR, Garcia-Fossa F, Senft RA, Hiner MC, de Jesus MB, Eliceiri KW, Cimini BA: CellProfiler plugins – an easy image analysis platform integration for containers and Python tools. *J Microsc* 2023, 1:1-8, https://doi.org/10.1111/jmi.13223
- 15. Rueden CT, Hiner MC, Evans EL 3rd, Pinkert MA, Lucas AM, Carpenter AE, Cimini BA, Eliceiri KW: **PyImageJ: a library for**

integrating ImageJ and Python. Nat Methods 2022, 19:1326-1327, https://doi.org/10.1038/s41592-022-01655-4

- Barry DJ, Gerri C, Bell DM, D'Antuono R, Niakan KK: GIANI open-source software for automated analysis of 3D microscopy images. J Cell Sci 2022, 135:jcs259511(https:// journals.biologists.com/jcs/article-abstract/135/10/jcs259511/ 275435).
- 17. Arzt M, Deschamps J, Schmied C, Pietzsch T, Schmidt D, Tomancak P, Haase R, Jug F: LABKIT: labeling and segmentation Toolkit for Big Image Data. Front Comput Sci 2022, 4:1-12, https://doi.org/10.3389/fcomp.2022
- Ritchie A, Laitinen S, Katajisto P, Englund JI: "Tonga": a novel toolbox for straightforward bioimage analysis. Front Comput Sci 2022, 4:1-9, https://doi.org/10.3389/fcomp.2022
- Uhlmann V, Donati L, Sage D: A practical guide to supervised deep learning for bioimage analysis: challenges and good practices. *IEEE Signal Process Mag* 2022, 39:73-86, https://doi. org/10.1109/msp.2021.3123589
- Tian J, Hsu Y-C, Shen Y, Jin H, Kira Z: Exploring covariate and concept shift for detection and calibration of out-ofdistribution data. arXiv [cs.LG]; 2021. (http://arxiv.org/abs/2110. 15231).
- Isola P, Zhu J-Y, Zhou T, Efros AA: Image-to-image translation with conditional adversarial networks. arXiv [cs.CV]; 2016:1125–1134. (http://openaccess.thecvf.com/content_cvpr_ 2017/html/sola_Image-To-Image_Translation_With_CVPR_2017_ paper.html) [Accessed 27 November 2023].
- Hollandi R, Szkalisity A, Toth T, Tasnadi E, Molnar C, Mathe B, Grexa I, Molnar J, Balind A, Gorbe M, Kovacs M, Migh E, Goodman A, Balassa T, Koos K, Wang W, Caicedo JC, Bara N, Kovacs F, Paavolainen L, Danka T, Kriston A, Carpenter AE, Smith K, Horvath P: nucleAlzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer. *Cell Syst* 2020, 10:453-458.e6, https://doi.org/10.1016/j.cels.2020.04.003
- Lee MY, Bedia JS, Bhate SS, Barlow GL, Phillips D, Fantl WJ, Nolan GP, Schürch CM: CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images. *BMC Bioinform* 2022, 23:46, https://doi.org/10.1186/s12859-022-04570-9
- McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, Doan M, Ding L, Rafelski SM, Thirstrup D, Wiegraebe W, Singh S, Becker T, Caicedo JC, Carpenter AE: CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol* 2018, 16:e2005970, https://doi.org/10.1371/journal.pbio.2005970
- Mandal S, Uhlmann V: SplineDist: automated cell segmentation with spline curves, *bioRxiv*; 2021. 2020.10.27.357640. (https:// doi.org/10.1101/2020.10.27.357640).
- Walter FC, Damrich S, Hamprecht FA: MultiStar: instance segmentation of overlapping objects with star-convex polygons. arXiv [cs.CV]; 2020. http://arxiv.org/abs/2011.13228.
- Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, James JA, Salto-Tellez M, Hamilton PW: QuPath: open source software for digital pathology image analysis. *Sci Rep* 2017, 7:16878, https://doi.org/10.1038/s41598-017-17204-5
- Humphries MP, Maxwell P, Salto-Tellez M: QuPath: the global impact of an open source digital pathology system. Comput Struct Biotechnol J 2021, 19:852-859, https://doi.org/10.1016/j. csbj.2021.01.022
- de Chaumont F, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-Yedid V, Pankajakshan P, Lecomte T, Le Montagner Y, Lagache T, Dufour A, Olivo-Marin J-C: Icy: an open bioimage informatics platform for extended reproducible research. Nat Methods 2012, 9:690-696, https://doi.org/10.1038/ nmeth.2075
- Fillbrunn A, Dietz C, Pfeuffer J, Rahn R, Landrum GA, Berthold MR: KNIME for reproducible cross-domain analysis of life science data. J Biotechnol 2017, 261:149-156, https://doi.org/10.1016/j. jbiotec.2017.07.028

31. Pachitariu M, Stringer C: Cellpose 2.0: how to train your own

 model. Nat Methods 2022, 19:1634-1641, https://doi.org/10.1038/ s41592-022-01663-4.

User-friendly software for human-in-the-loop fine-tuning of DL models applied to cell segmentation.

- Cutler KJ, Stringer C, Lo TW, Rappez L, Stroustrup N, Brook Peterson S, Wiggins PA, Mougous JD: Omnipose: a highprecision morphology-independent solution for bacterial cell segmentation. Nat Methods 2022, 19:1438-1448, https://doi.org/ 10.1038/s41592-022-01639-4
- Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, Fullaway CC, McIntosh BJ, Leow KX, Schwartz MS, Pavelchek C, Cui S, Camplisson I, Bar-Tal O, Singh J, Fong M, Chaudhry G, Abraham Z, Moseley J, Warshawsky S, Soon E, Greenbaum S, Risom T, Hollmann T, Bendall SC, Keren L, Graf W, Angelo M, Van Valen D: Whole-cell segmentation of tissue images with humanlevel performance using large-scale data annotation and deep learning. Nat Biotechnol 2022, 40:555-565, https://doi.org/10. 1038/s41587-021-01094-0
- Ma J, Wang B: Towards foundation models of biological image segmentation. Nat Methods 2023, 20:953-955, https://doi.org/10. 1038/s41592-023-01885-0
- 35. Laine RF, Arganda-Carreras I, Henriques R, Jacquemet G:
 Avoiding a replication crisis in deep-learning-based bioimage analysis. Nat Methods 2021, 18:1136-1144, https://doi.org/10. 1038/s41592-021-01284-3.

Opinion on best practices for implementing and reporting on the use of development of DL image analysis tools.

36. Ma J, Xie R, Ayyadhury S, Ge C, Gupta A, Gupta R, Gu S, Zhang Y,
Lee G, Kim J, Lou W, Li H, Upschulte E, Dickscheid T, de Almeida JG, Wang Y, Han L, Yang X, Labagnara M, Rahi SJ, Kempster C, Pollitt A, Espinosa L, Mignot T, Middeke JM, Eckardt J-N, Li W, Li Z, Cai X, Bai B, Greenwald NF, Van Valen D, Weisbart E, Cimini BA, Li Z, Zuo C, Brück O, Bader GD, Wang B: The multi-modality cell segmentation challenge: towards universal solutions. arXiv [eess.IV]; 2023. (http://arxiv.org/abs/2308.05864).

Report on the results of the Multi-modality Cell Segmentation Challenger, wherein teams competed to develop state-of-the-art generalist models for cell segmentation, utilizing the latest advancements in developing DL architectures, preprocessing techniques, and data augmentation.

- Edlund C, Jackson TR, Khalid N, Bevan N, Dale T, Dengel A, Ahmed S, Trygg J, Sjögren R: LIVECell – a large-scale dataset for label-free live cell segmentation. *Nat Methods* 2021, 18:1038-1045, https://doi.org/10.1038/s41592-021-01249-6
- Maška M, Ulman V, Delgado-Rodriguez P, Gómez-de-Mariscal E, Nečasová T, Guerrero Peña FA, Ren TI, Meyerowitz EM, Scherr T, Löffler K, Mikut R, Guo T, Wang Y, Allebach JP, Bao R, Al-Shakarji NM, Rahmon G, Toubal IE, Palaniappan K, Lux F, Matula P, Sugawara K, Magnusson KEG, Aho L, Cohen AR, Arbelle A, Ben-Haim T, Raviv TR, Isensee F, Jäger PF, Maier-Hein KH, Zhu Y, Ederra C, Urbiola A, Meijering E, Cunha A, Muñoz-Barrutia A, Kozubek M, Ortiz-de-Solórzano C: The cell tracking challenge: 10 years of objective benchmarking. Nat Methods 2023, 20:1010-1020, https://doi.org/10.1038/s41592-023-01879-y
- Dey N, Mazdak Abulnaga S, Billot B, Turk EA, Ellen Grant P, Dalca AV, Golland P: AnyStar: Domain randomized universal starconvex 3D instance segmentation. arXiv [cs.CV]; 2023. (http:// arxiv.org/abs/2307.07044).
- Scherr T, Löffler K, Böhland M, Mikut R: Cell segmentation and tracking using CNN-based distance predictions and a graphbased matching strategy. *PLoS One* 2020, 15:e0243219, https:// doi.org/10.1371/journal.pone.0243219
- Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P: SegFormer: simple and efficient design for semantic segmentation with transformers. Adv Neural Inf Process Syst 2021, 34:12077-12090 (https://proceedings.neurips.cc/paper/ 2021/hash/64f1f27bf1b4ec22924fd0acb550c235-Abstract.html).
- Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S: A ConvNet for the 2020s. arXiv [cs.CV]; 2022:11976–11986. (http:// openaccess.thecvf.com/content/CVPR2022/html/Liu_A_ConvNet_ for_the_2020s_CVPR_2022_paper.html) [Accessed 28 September 2023].

- S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He: Aggregated residual transformations for deep neural networks. arXiv [cs.CV]; 2016:1492–1500. (http://openaccess.thecvf.com/content_cvpr_ 2017/html/Xie_Aggregated_Residual_Transformations_CVPR_ 2017_paper.html) [Accessed September 28, 2023].
- 44. Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH: nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021, 18:203-211, https://doi.org/10.1038/s41592-020-01008-z
- G. Lee, S. Kim, J. Kim, S.-Y. Yun: MEDIAR: harmony of datacentric and model-centric for multi-modality microscopy. arXiv [cs.CV]; 2022. (http://arxiv.org/abs/2212.03465).
- Royer LA: The future of bioimage analysis: a dialog between mind and machine. Nat Methods 2023, 20:951-952, https://doi. org/10.1038/s41592-023-01930-y
- C. Wu, S. Yin, W. Qi, X. Wang, Z. Tang, N. Duan: Visual ChatGPT: talking, drawing and editing with visual foundation models. arXiv [cs.CV]; 2023. (http://arxiv.org/abs/2303.04671).
- Li X, Zhang Y, Wu J, Dai Q: Challenges and opportunities in bioimage analysis. Nat Methods 2023, 20:958-961, https://doi. org/10.1038/s41592-023-01900-4
- Gao Y, Zhou M, Metaxas DN: UTNet: a hybrid transformer architecture for medical image segmentation. Springer International Publishing; 2021:61-71, https://doi.org/10.1007/978-3-030-87199-4_6
- Wang R, Lei T, Cui R, Zhang B, Meng H, Nandi AK: Medical image segmentation using deep learning: a survey. *IET Image Proc* 2022, 16:1243-1267, https://doi.org/10.1049/ipr2.12419
- A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A.C. Berg, W.-Y. Lo, P. Dollár, R. Girshick: Segment anything. arXiv [cs.CV]; 2023. (http://arxiv.org/abs/2304.02643).
- Caicedo JC, Goodman A, Karhohs KW, Cimini BA, Ackerman J, Haghighi M, Heng C, Becker T, Doan M, McQuin C, Rohban M, Singh S, Carpenter AE: Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. Nat Methods 2019, 16:1247-1253, https://doi.org/10.1038/s41592-019-0612-7
- A. Archit, S. Nair, N. Khalid, P. Hilt, V. Rajashekar, M. Freitag, S. Gupta, A. Dengel, S. Ahmed, C. Pape: Segment anything for microscopy. *bioRxiv*; 2023: 2023.08.21.554208. (https://doi.org/10.1101/2023.08.21.554208).
- Paul-Gilloteaux P: Bioimage informatics: Investing in software usability is essential. *PLoS Biol* 2023, 21:e3002213, https://doi. org/10.1371/journal.pbio.3002213
- Kemmer I, Keppler A, Serrano-Solano B, Rybina A, Özdemir B, Bischof J, El Ghadraoui A, Eriksson JE, Mathur A: Building a FAIR image data ecosystem for microscopy communities. *Histochem Cell Biol* 2023, 160:199-209, https://doi.org/10.1007/s00418-023-02203-7
- 56. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016, **3**:160018, https://doi.org/10.1038/sdata.2016.18
- Merkel D: Docker: lightweight Linux containers for consistent development and deployment. Linux J 2014;:2, https://doi.org/ 10.5555/2600239.2600241
- Y. Zhou, J. Sonneck, S. Banerjee, S. Dörr, A. Grüneboom, K. Lorenz, J. Chen: EfficientBioAl: making bioimaging Al models efficient in energy, latency and representation. *arXiv* [cs.LG]; 2023. (http://arxiv.org/abs/2306.06152).
- 59. B.M. Saraiva, I.M. Cunha, A.D. Brito, G. Follain, R. Portela, R. Haase, P.M. Pereira, G. Jacquemet, R. Henriques: **NanoPyx:**

super-fast bioimage analysis powered by adaptive machine learning. *bioRxiv*; 2023: 2023.08.13.553080. (https://doi.org/10.1101/2023.08.13.553080).

- R. Haase, A. Jain, S. Rigaud, D. Vorkel, P. Rajasekhar, T. Suckert, T.J. Lambert, J. Nunez-Iglesias, D.P. Poole, P. Tomancak, E.W. Myers: Interactive design of GPU-accelerated Image Data Flow Graphs and cross-platform deployment using multi-lingual code generation. *bioRxiv*; 2020: 2020.11.19.3865655. (https://doi. org/10.1101/2020.11.19.386565).
- 61. Ouyang W, Eliceiri KW, Cimini BA: Moving beyond the desktop: prospects for practical bioimage analysis via the web. Front Bioinform 2023, 3:1233748, https://doi.org/10.3389/fbinf.2023. 1233748
- von Chamier L, Laine RF, Jukkala J, Spahn C, Krentzel D, Nehme E, Lerche M, Hernández-Pérez S, Mattila PK, Karinou E, Holden S, Solak AC, Krull A, Buchholz T-O, Jones ML, Royer LA, Leterrier C, Shechtman Y, Jug F, Heilemann M, Jacquemet G, Henriques R: Democratising deep learning for microscopy with ZeroCostDL4Mic. Nat Commun 2021, 12:2276, https://doi.org/10. 1038/s41467-021-22518-0
- 63. da Veiga Leprevost F, Grüning BA, Alves Aflitos S, Röst HL, Uszkoreit J, Barsnes H, Vaudel M, Moreno P, Gatto L, Weber J, Bai M, Jimenez RC, Sachsenberg T, Pfeuffer J, Vera Alvarez R, Griss J, Nesvizhskii Al, Perez-Riverol Y: BioContainers: an open-source and community-driven framework for software standardization. *Bioinformatics* 2017, 33:2580-2582, https://doi. org/10.1093/bioinformatics/btx192
- Bai J, Bandla C, Guo J, Vera Alvarez R, Bai M, Vizcaíno JA, Moreno P, Grüning B, Sallou O, Perez-Riverol Y: BioContainers registry: searching bioinformatics and proteomics tools, packages, and containers. *J Proteome Res* 2021, 20:2056-2061, https://doi.org/ 10.1021/acs.jproteome.0c00904
- 65. Rubens U, Mormont R, Paavolainen L, Bäcker V, Pavie B, Scholz LA, Michiels G, Maška M, Ünay D, Ball G, Hoyoux R, Vandaele R, Golani O, Stanciu SG, Sladoje N, Paul-Gilloteaux P, Marée R, Tosi S: BIAFLOWS: a collaborative framework to reproducibly deploy and benchmark bioimage analysis workflows. *Patterns* 2020, 1:100040, https://doi.org/10.1016/j.patter.2020. 100040
- Prigent S, Valades-Cruz CA, Leconte L, Maury L, Salamero J, Kervrann C: BiolmageIT: open-source framework for integration of image data management with analysis. *Nat Methods* 2022, 19:1328-1330, https://doi.org/10.1038/s41592-022-01642-9
- 67. W. Ouyang, F. Beuttenmueller, E. Gómez-de-Mariscal, C. Pape, T.
 Burke, C. Garcia-López-de-Haro, C. Russell, L. Moya-Sans, C. dela-Torre-Gutiérrez, D. Schmidt, D. Kutra, M. Novikov, M. Weigert, U. Schmidt, P. Bankhead, G. Jacquemet, D. Sage, R. Henriques, A. Muñoz-Barrutia, E. Lundberg, F. Jug, A. Kreshuk: Biolmage Model Zoo: a community-driven resource for accessible deep learning in bioimage analysis. *bioRxiv*; 2022: 2022.06.07.495102. (https://doi.org/10.1101/2022.06.07.495102).

Paper introducing a repository of bioimage-specific DL models, a platform for their execution, and a framework for interoperability with bioimage analysis tools.

- Ouyang W, Mueller F, Hjelmare M, Lundberg E, Zimmer C: ImJoy: an open-source computational platform for the deep learning era. Nat Methods 2019, 16:1199-1200, https://doi.org/10.1038/ s41592-019-0627-0
- Bannon D, Moen E, Schwartz M, Borba E, Kudo T, Greenwald N, Vijayakumar V, Chang B, Pao E, Osterman E, Graf W, Van Valen D: DeepCell Kiosk: scaling deep learning-enabled cellular image analysis with Kubernetes. Nat Methods 2021, 18:43-45, https:// doi.org/10.1038/s41592-020-01023-0
- Weisbart E, Cimini BA: Distributed-Something: scripts to leverage AWS storage and computing for distributed workflows at scale. Nat Methods 2023, 20:1120-1121, https://doi. org/10.1038/s41592-023-01918-8
- 71. Schmied C, Nelson MS, Avilov S, Bakker G-J, Bertocchi C, Bischof J, Boehm U, Brocher J, Carvalho MT, Chiritescu C, Christopher J, Cimini BA, Conde-Sousa E, Ebner M, Ecker R, Eliceiri K, Fernandez-Rodriguez J, Gaudreault N, Gelman L, Grunwald D, Gu T, Halidi N, Hammer M, Hartley M, Held M, Jug F, Kapoor V, Koksoy AA, Lacoste J, Le Dévédec S, Le Guyader S, Liu P, Martins GG, Mathur A, Miura K, Montero Llopis P, Nitschke R, North A, Parslow AC, Payne-Dwyer A, Plantard L, Ali R, Schroth-Diez B, Schütz L, Scott RT, Seitz A, Selchow O, Sharma VP, Spitaler M, Srinivasan S, Strambio-De-Castillia C, Taatjes D, Tischer C, Jambor HK: Community-developed checklists for publishing images and image analyses. Nat Methods 2023,1-9, https://doi.org/10. 1038/s41592-023-01987-9
- Hirling D, Tasnadi E, Caicedo J, Caroprese MV, Sjögren R, Aubreville M, Koos K, Horvath P: Segmentation metric misinterpretations in bioimage analysis. Nat Methods 2023, https://doi.org/10.1038/s41592-023-01942-8
- L. Maier-Hein, A. Reinke, P. Godau, M.D. Tizabi, F. Buettner, E.
 Christodoulou, B. Glocker, F. Isensee, J. Kleesiek, M. Kozubek, M. Reyes, M.A. Riegler, M. Wiesenfarth, A. Emre Kavur, C.H. Sudre, M. Baumgartner, M. Eisenmann, D. Heckmann-Nötzel, A. Tim Rädsch, L. Acion, M. Antonelli, T. Arbel, S. Bakas, A. Benis, M. Blaschko, M. Jorge Cardoso, V. Cheplygina, B.A. Cimini, G.S. Collins, K. Farahani, L. Ferrer, A. Galdran, B. van Ginneken, R. Haase, D.A. Hashimoto, M.M. Hoffman, M. Huisman, P. Jannin, C.E. Kahn, D. Kainmueller, B. Kainz, A. Karargyris, A. Karthikesalingam, H. Kenngott, F. Kofler, A. Kopp-Schneider, A. Kreshuk, T. Kurc, B.A. Landman, G. Litjens, A. Madani, K. Maier-Hein, A.L. Martel, P. Mattson, E. Meijering, B. Menze, K.G.M. Moons, H. Müller, B. Nichyporuk, F. Nickel, J. Petersen, N. Rajpoot, N. Rieke, J. Saez-Rodriguez, C.I. Sánchez, S. Shetty, M. van Smeden, R.M. Summers, A.A. Taha, A. Tiulpin, S.A. Tsaftaris, B. Van Calster, G. Varoquaux, P.F. Jäger: Metrics reloaded: recommendations for image analysis validation. arXiv [cs.CV]; 2022. (http://arxiv.org/ abs/2206.01653)

A wide-sweeping and comprehensive guide on the use of metrics for biomedical analysis tasks in Machine Learning contexts.