Trends in Cell Biology

CellPress

Review

Nucleus segmentation: towards automated solutions

Reka Hollandi,[1,8] Nikita Moshkov,[1,2,3,8] Lassi Paavolainen,[4] Ervin Tasnadi,[1,5] Filippo Piccinini,[6] and Peter Horvath[1,4,7,*]

Single nucleus segmentation is a frequent challenge of microscopy image processing, since it is the first step of many quantitative data analysis pipelines. The quality of tracking single cells, extracting features or classifying cellular phenotypes strongly depends on segmentation accuracy. Worldwide competitions have been held, aiming to improve segmentation, and recent years have definitely brought significant improvements: large annotated datasets are now freely available, several 2D segmentation strategies have been extended to 3D, and deep learning approaches have increased accuracy. However, even today, no generally accepted solution and benchmarking platform exist. We review the most recent single-cell segmentation tools, and provide an interactive method browser to select the most appropriate solution.

Highlights

Nucleus segmentation is one of the first steps of many microscopy image analysis pipelines.

Several large-scale competitions have yielded annotated datasets that are available for training and testing specific methods.

The 2D segmentation strategies cover a diverse range of image modalities; some of these are also available for 3D datasets.

Simple cases of segmentation, especially for 2D, are straightforward, while in more challenging cases improved accuracy has been achieved recently.

Deep learning has established a new level of image analysis, but the lack of uniform evaluation strategies makes quantitative comparison and relative performance determination highly challenging.

Towards robust and automated methods for nucleus segmentation

The history [1] of detecting and segmenting single cells goes along with the first digitized microscopy images. Many research fields utilizing microscopy, such as developmental biology [2], drug discovery [3], functional genomics [4] and pathology [5] are dependent on accurate cell and nucleus segmentation as a vital part of image analysis workflows. Since image analysis has moved from a methodological research area towards data science as a result of the recent machine learning revolution, annotated datasets have become essential regarding the performance of nuclear segmentation methods. Especially, modality-independent, generalizable, and robust machine learning-based nucleus segmentation models need heterogeneous and large collections of expert-annotated images [6,7].

The level of difficulty of single-cell detection in an image, let alone precise outlining, widely varies (see Figure 1). In most simple cases nuclei have high contrast and are separated by proper experimental conditions (referred to as 'easy' cases), hence their segmentation is not difficult, e.g., large **short-interfering RNA (siRNA)** (see Glossary) [8]. In other cases, segmentation is highly 'challenging', for instance in **3D**, label-free or thick tissue sections where cells touch, overlap, or have non-conventional morphology, intensity, or patterns. International competitions [6,9] have promoted the potential to overcome these issues, yet a genuinely general solution is still awaited. However, due to major advancements in this field in recent years, our community has reached an unprecedented improvement in detecting single nuclei [10]. Easy cases of segmentation, especially in **2D** are not problematic anymore [11,12], while accuracy has also improved in challenging cases [6]. In addition, 3D data analysis methods have progressed with extended 2D segmentation solutions [13] or with native 3D ones [14,15]. The community has accumulated large amounts of annotated data either by experts [6] or crowdsourcing [16] for training machine learning segmentation models, and to evaluate the methods in public benchmarking platforms [6,17]. This review describes the specific techniques biologists can exploit for single-cell analysis. However, we emphasize that no

[1]Synthetic and Systems Biology Unit, Biological Research Centre (BRC), H-6726, Szeged, Hungary
[2]Doctoral School of Interdisciplinary Medicine, University of Szeged, Szeged, Hungary
[3]Laboratory on AI for Computational Biology, Faculty of Computer Science, HSE University, Moscow, Russia
[4]Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, FI-00014 Helsinki, Finland
[5]Doctoral School of Computer Science, University of Szeged, Szeged, Hungary
[6]IRCCS Istituto Romagnolo per lo Studio dei Tumori (IRST) "Dino Amadori", I-47014 Meldola (FC), Italy

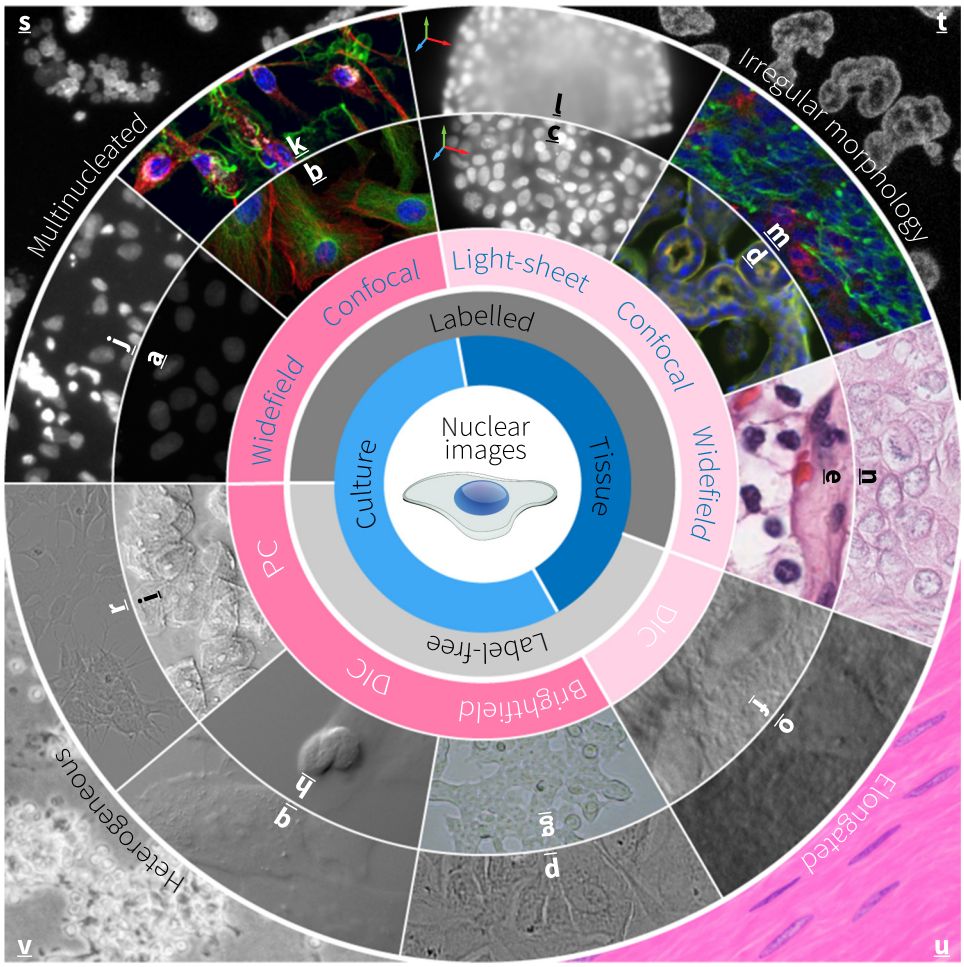

Figure 1. Diversity of optical microscopy images representing nuclei. The inner circle shows standard examples of each type (e.g., widefield, confocal, light-sheet, differential interference contrast (DIC), phase contrast (PC) images), while the outer circle presents more difficult cases. Finally, common challenging cases (e.g., multinucleated cells, irregular morphology, elongated shape, heterogeneous samples) regarding nucleus segmentation are reported in the corners. (C, D, L, F, G, O, T) Images from our laboratories/collaborators; (A, S, J) from the Broad Bioimage Benchmark Collection (BBBC); (E, N) from The Cancer Genome Atlas (TGCA) collection; (R) from the LIVECell dataset; the remaining images are from the internet (see Supplementary Table 1 for the sources).

standardized approach has been developed to date to properly compare different solutions before deciding which tool to use for a specific application.

First, the variety and extent of datasets currently available to test and train methods are presented. Next, a selection of annotation tools available for creation of training datasets for machine learning methods is introduced. Then the issues related to pre- and post-processing of images to reduce challenges inherent to complex data are briefly discussed (see Supplementary Material 1 for details on different techniques), followed by insights into 2D nuclear segmentation methods. Classical approaches that provide task-specific and general solutions for a wide variety of acquisition techniques are presented. However, most recent methods usually rely on **deep neural networks (DNNs)**, and since the target objective is related to image processing, **convolutional**

[7]Single-Cell Technologies Ltd., H-6726, Szeged, Hungary
[8]These authors contributed equally

*Correspondence:
horvath.peter@brc.hu (P. Horvath).

neural networks (CNNs) are most commonly applied to segment nuclei. As processing 3D data is one of the major challenges in single nucleus segmentation, a set of promising and successful methods appropriate to solve specific 3D segmentation tasks is discussed.

Figure 2 supports a better understanding of the definitions of detection and segmentation tasks. When identifying single cells (objects) in microscopy images automatically, that is, using computer algorithms, the results may be either (i) 'detections' corresponding to the localization of the objects or (ii) 'segmentations' which separate independent image regions. The former is typically represented as bounding boxes, whereas the latter may be realized by either assigning a binary label to each pixel while dividing the image into not necessarily connected regions (semantic

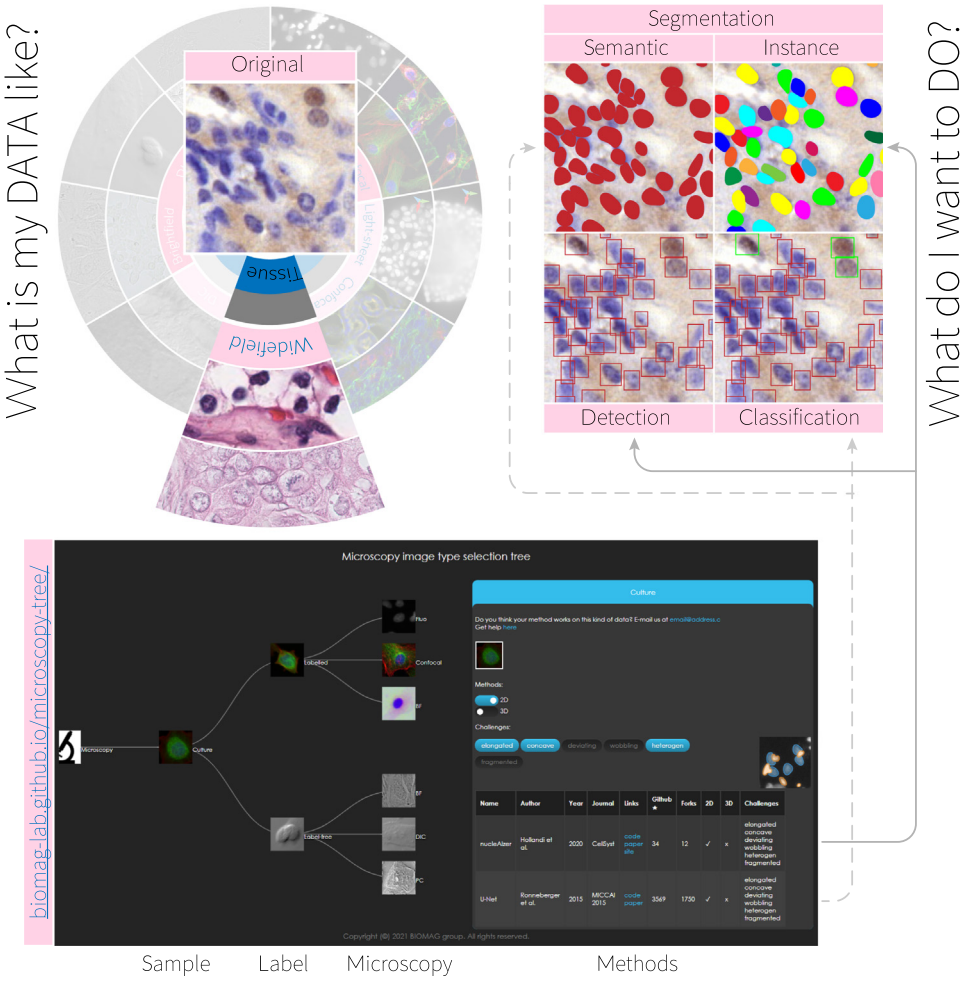

Figure 2. A sample-driven guide to select an appropriate method for single nucleus segmentation. Firstly, based on the images of the given experiment the user can determine the category [e.g., widefield, confocal, light-sheet, differential interference contrast (DIC), phase contrast (PC) images] and select the corresponding node in the interactive online tool *unbias*[i] according to the sample, label, and microscopy type. Then, a list of segmentation methods is shown in the table on the right, including the method description and implementation if available alongside pre-trained models. The list may be filtered with the buttons above the table by dimension (2D/3D) and challenging segmentation issues (e.g., elongated nucleus in smooth muscle tissue). Finally, the goal of the experiment (e.g., object-aware segmentation or additional phenotyping, i.e., classification) guides the user to select the appropriate segmentation method.

Glossary

2D: the term typically used to indicate the standard format in which images are acquired by a standard camera.
3D: the term typically used to indicate a z-stack of 2D images referring to different optical sections.
Application programming interface: a set of functions and procedures allowing the development of applications that access the features or data of an operating system, application, or other service.
Broad Bioimage Benchmark Collection (BBBC): an open microscopy image collection for scientific purposes.
Convolutional neural network (CNN): a class of deep neural networks including convolutional layers based on blocks responsible for appropriate image feature retrieval (via convolutions) and scaling (with pooling blocks).
DAPI: a widely used fluorescent stain that binds to adenine–thymine-rich regions of the DNA, thus labels the nucleus.
Data Science Bowl 2018 (DSB2018): the data science competition held in 2018 with a task to segment nuclei in microscopy images. The official, open dataset of the competition is also referred to as such, and is often used to benchmark nucleus segmentation methods.
Deep neural network (DNN): is an artificial neural network machine learning architecture that includes several hidden layers, and can be trained to solve more complex tasks on more complex data compared to shallow neural networks.
Differential interference contrast: a microscopy technique that introduces contrast to images of specimen with little or no contrast upon brightfield microscopy.
General public license (GPL): a series of widely used open-source licenses that guarantee end users the freedom to run, study, share, and modify the software.
Graphics processing unit (GPU): a specialized electronic circuit designed to rapidly manipulate memory to accelerate computations related primarily to graphics.
Hematoxylin and eosin (H&E): a combination of two histological stains: hematoxylin and eosin. Hematoxylin stains cell nuclei to purplish blue, and eosin stains the extracellular matrix and cytoplasm to pink.

segmentation) or by separating individual objects (instance segmentation). One may choose from a plethora of methods and software tools to perform nuclear segmentation (see section Segmentation methods and toolkits). Major challenges are discussed (see Outstanding questions). As many of the segmentation methods considered in this review utilize 'deep learning' approaches, this subset of machine learning is also introduced. Deep learning involves the training of DNNs for complex yet arbitrary tasks, such as detection, segmentation, and classification (not strictly in our domain of cellular image analysis, but also in natural image-, video- or audio-processing). Deep learning-based approaches are proven to perform excellently on the trained domain, with the potential of extension to unseen domains [18]. Their translation is limited by the lack of publicly available datasets related to less common modalities (see section Annotated nucleus datasets).

A portal[i] was developed to offer a graphical aid to select the most appropriate method for non-image analysis experts (see Figure 2). This portal has several advantages: (i) the imaging community can select the methods applicable/appropriate for their images of interest, and (ii) developers can submit the description and best practices for their methods. Currently, the most common optical microscopy categories are considered. Notably, the web portal is declared to be maintained by the authors, yet the community is encouraged to actively contribute and eventually propose extensions to it. For benchmarking the segmentation methods, users can exploit BIAFLOWS[ii] [17], a freely available web-based platform developed by the **Network of European BioImage Analysts (NEUBIAS)**. The assessment of new proposals has been commonly performed with limited datasets and arbitrary metrics (see Supplementary Materials 2); in contrast, NEUBIAS is a step forward to prevent such a biased evaluation. However, an unbiased quantitative method comparison is still impossible due to the lack of a comprehensive annotated dataset for training and testing the methods using a globally accepted benchmark platform and unified metrics (see section Segmentation methods and toolkits).

Annotated nucleus datasets

Annotated datasets are used in computer science to validate the accuracy of developed algorithms. In addition, nowadays annotated datasets are also used for training machine learning models for various tasks. One of the key factors influencing the performance of segmentation models is the composition of annotated data. Ideally, a trainable model yields optimal results on a test set sampled from the same domain as training data are collected from, hence domain-specific annotated datasets serve as a valuable asset, especially when they are expert-curated. Highly specific domain datasets are usually complemented with proper metadata [19,20], such as the experimental set-up, sample preparation or microscope device, and are expert-curated when it comes to annotations. However, they typically cover a narrow diversity, and are small in size. Open datasets may contain a varying number of images (Table 1). An important aspect for the user to consider when either training a new model or evaluating segmentation performance on publicly available datasets is that the corresponding annotations occasionally contain such segmentations yielded by automatic methods [19] that might not be refined by an expert, thus the results might be biased. Annotated nucleus datasets displayed in Table 1 show diversity in size (not only regarding the number of images, but also that of the objects) and content, focusing on those widely used as benchmarks or training data. Annotations may be realized as objects (instance aware) or binary masks (semantic), are primarily 2D, and the two most common imaging modalities cover fluorescence stained cell cultures and **hematoxylin and eosin (H&E)-**stained tissue sections.

International challenges, such as the annual **International Symposium on Biomedical Imaging (ISBI)** or competitions hosted by for example, Kaggle with industry partners, inspire those in the field of research and development to propose new technologies and methods or combine existing

Immunofluorescence: a staining which utilizes fluorescent-labeled antibodies to detect specific target antigens
International Symposium on Biomedical Imaging (ISBI): a scientific conference series dedicated to mathematical, algorithmic, and computational aspects of biological and biomedical imaging.
Iterative thresholding: an algorithm used to define the background and foreground in an image.
Line-of-sight (LoS): the straight line between the object and the target.
Mean average precision: a popular metric related to measuring the accuracy of object detectors.
Medical Imaging Interaction Toolkit (MITK): a software suite designed for medical image analysis.
Network of European BioImage Analysts (NEUBIAS): a network of experts in life sciences for image data analysis.
Phase contrast: an optical microscopy technique that converts phase shifts in light passing through a transparent specimen to brightness changes in the image.
Region proposal network (RPN): a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position.
Short-interfering RNA (siRNA): a class of double-stranded, non-coding RNA molecules, similar to miRNA, operating within the RNA interference (RNAi) pathway.
Test-time augmentation: the aggregation of predictions across transformed versions of a test input.
The Cancer Genome Atlas (TCGA): a huge cancer genomic program which covers many cancer types with a patient-based, open dataset including genomic, proteomic, imaging, etc. data.
Whole slide image (WSI): scanned image of an entire histopathology tissue section, usually of gigapixel size, resulting in file size of gigabytes, which is difficult to handle by an image processing software.

Table 1. Open datasets of annotated nucleus for single-cell analysis purposes

Name	Individual objects (O) or binary masks (BM)	2D/3D	Microscopy	Staining	Sample	No. of images	No. of objects	Refs
BBBC032	O	3D	Confocal	Fluo	Mouse embryo blastocyst cells	1 (172[a])	1220	[21]
BBBC033	O	3D	Confocal	Fluo	Mouse trophoblast stem cells	1 (32[a])	585	[21]
BBBC034	O	3D	Brightfield/fluorescent	x/3 fluo[b]	hiPSC 3D	1 (52[a])	790	[19]
Scaffold-A549 dataset	O	3D	Fluorescent	Hoechst+DiL	Lung cancer tissue	21	800 (+10 000 w/o labels)	[22]
BBBC039	O	2D	Fluorescent	Hoechst	U2OS cells	200	23 615	[10]
CoNSeP	O	2D	Brightfield	H&E	Colon tissue	41	24 319	[23]
CryoNuSeg (TCGA[iii])	O	2D	Brightfield	H&E	Various tissues	30	7596	[24]
DSB2018	O	2D	Various	Various	Various tissues and cells	841	37 530	[6]
Janowczyk *et al.*[vi]	BM	2D	Brightfield	H&E	Breast tissue	141	~12 000	[25]
LIVECell	O	2D	Phase contrast	Label-free	Cell cultures	5239	1 686 352	[26]
Lizard	O	2D	Brightfield	H&E	Colon tissue	291	495 179	[27]
MoNuSeg2018	O	2D	Brightfield	H&E	Various tissues	44	28 846	[28]
NuCLS	O[c]	2D	Brightfield	H&E	Breast tissue	N/A	222 396	[16]
NucMM	O	3D	Electron microscopy/micro-CT	Label-free	Brain tissue	2	~170 000 + ~7000	[29]
PanNuke	O	2D	Brightfield	H&E	Various tissues	481	205 343	[30]
S-BSST265	O	2D	Fluorescent/ confocal	**Immunofluorescence/DAPI**	Various tissues and cells	79	7813	[20]
TCGA[iii] images processed by Irshad *et al.*	N/A	2D	Brightfield	H&E	Kidney clear cell renal carcinoma tissue	63	N/A	[31]
TCGA[iii] images processed by Kumar *et al.*	O	2D	Brightfield	H&E	Various tissues	30	21 623	[28]
TNBC	O	2D	Brightfield	H&E	Breast tissue	50	4022	[32]
Wienert *et al.*	O	2D	Brightfield	H&E	Various tissues	36	7931	[33]
TissueNet Version v1.0	O	2D	Fluorescent	Various stainings	Various tissues	6990	~1200 000	[7]

[a]The number of slices is presented; one image is available.
[b]CellMask Deep Red plasma membrane, EGFP beta-actin, Hoechst DNA.
[c]Object contours or bounding boxes with class label.

ones for a new purpose. One of the most successful and widely used segmentation methods, *U-Net* [34] (see Supplementary Material 4) arose from the 2015 ISBI Cell Tracking Challenge [35], and has been the basis for several novel CNN architectures ever since (see section Segmentation methods and toolkits). Similar competitions contribute to the development of this field with invaluable collections of microscopy images, on which developers may

benchmark their novel approaches according to standard evaluation metrics (typically **mean average precision)** in a fairly comparable way. In recent years the **Data Science Bowl 2018 (DSB2018)** [6] dataset has been applied as such, since its image set comprises various types of microscopy modalities, magnifications, labels, sources, etc. This might also provide insight into the expected model performance. Generally, datasets originating from challenges are carefully validated by field-expert annotators [6,15] (usually biologists and pathologists), promoting their further applicability to train new models. Notably, annotations of the training set are usually released instantly, while test set annotations may remain private even after the challenge is concluded [35]. Dataset size strongly depends on the task, for example, a competition in 2D instance segmentation (like DSB2018) generally has a larger number of annotated images than a tracking [35] or a 3D segmentation task [19,21].

Conclusively, a key contribution of the bioimage analysis community to this field is the release of open datasets of annotated images, in as many varying imaging modalities as possible. Data sharing is highly encouraged, especially in case of intrinsically challenging microscopy types, such as label-free imaging (notably, LIVECell [26] is a promising step in this direction) or generally in 3D. Provided in an open way, these annotated datasets could inspire method developers to increase their focus on less frequent modalities, and release pre-trained models for those as well. Also, they enable users to benchmark (evaluate the performance of) available methods on this data. Additionally, experiment-specific unlabeled image sets [e.g., **The Cancer Genome Atlas (TCGA)**[iii]] may also promote progress in case an annotated subset is shared later independently [24,28,31]. Finally, as annotated datasets require an appropriate software tool that the experts (or generally, annotators) can use to create the labels, various annotation software solutions are collected in the following section.

Tools for annotation

Countless software tools are available to create annotations for single-cell segmentation training or validation, with a widely varying spectrum of functionality. These tools are designed either for specialists, such as biologists and pathologists, or for method developers. Options for annotation typically include freehand drawing, point, ellipse or polygon labeling, all of which may be exported to formats suitable for different applications. The finer the representation (annotation) of the object is, the more information it provides for a model when used as training data. While object location marked simply by a center point or bounding box coordinates is sufficient for detection or even classification training, contours (boundaries) labeled either semantically (binary) or in an instance-aware way are usually used to train segmentation. Equivalently, the same types of annotated data may be utilized to assess the accuracy of different methods.

Even though labeling several images tends to be time-consuming for a single expert, even students [16] can learn how to create accurate annotations when curated by experts, yielding large annotated datasets via joint and shared efforts. Semi-automatic annotation achieved by initial segmentation methods offers a convenient solution to speed up the annotation process for experts, and is often preferred by the community. Such annotation methods also help to increase [36,37] the agreement between experts, which is a common problem source in annotation. Alternatively, a consensus of multiple annotators may be used [6,31] at the object- or pixel level; crowdsourced annotations [16] are easier to combine this way. Commercial solutions and free-to-use software, including but not limited to those applied in cell biology, are described in detail in Supplementary Materials 3 and Supplementary Table 2.

Plugins or extensions to existing open-source software, such as *ImageJ/Fiji* [38,39] or **Medical Imaging Interaction Toolkit (*MITK*)** [40] are popular choices preferred by bioimage analysts

already experienced with the given software. The *Fiji* plugins *Trainable Weka Segmentation* [41] and *LabKit*[iv] use machine learning to train pixel classification similarly to *ilastik* [42] (see Supplmentary Materials 3 and 4), while *AnnotatorJ* [36] applies a *U-Net* to assist contour annotation. Assistance in the *MITK* plugin *3D-Cell-Annotator* [43] exploits active surfaces with shape descriptors in 3D, while *NuClick* [18] uses its own CNN for histopathology images.

Larger image analysis projects not primarily intended for annotation, but for a rather more comprehensive evaluation of the sample images (*Cytomine* [44,45], *ilastik* [42], *DeepCell* [46], *QuPath* [47]), including, for example, the segmentation or classification of cells, may also provide convenient solutions for annotation. Still, each has its target application: for example, *QuPath* is a desktop tool suitable for **whole slide image (WSI)** analysis, while *Cytomine* processes WSIs online in a collaborative way, and DeepCell improves its segmentation DNN with annotation collaboration.

Standalone software packages (*Diffgram*, *LabelImg, Segmentor* [37]) offer a lightweight, specific solution for annotation: *Segmentor* [37] is intended for 3D annotation, *Make Sense* and *Diffgram* have additional online interfaces, and the latter also supports deep learning. Online tools (*VGG Image Annotator*, *Kaibu*, *supervise.ly, Piximi annotator*) require no installation and have no specific hardware requirements. However, it is worth noting that online service-based platforms (*Lionbridge.AI* or *Hive*) require that raw data are sent out of the laboratory, which might be undesirable in case of sensitive (e.g., patient-related) images.

Nonetheless, genuinely general-purpose image editing applications, such as *GIMP* [**general public licence (GPL)**, free] or *Photoshop* (Adobe, commercial) may also be used to create annotations at the expense of more cumbersome export, for example, in the case of instance annotation labels.

Conclusively, several options are available, depending on the specific requirements of a project or experiment. Tools that provide multiple implementations (e.g., both local and online) might be ideal for more users.

Segmentation methods and toolkits

Single nucleus segmentation methods may work with raw images, but in more challenging cases (e.g., Figure 1 J–V) the quality of the analysis (and specifically that of single nucleus segmentation) benefits from additional pre- and post-processing steps (e.g. illumination correction [48,49] or denoising [50] prior to the analysis, mask refinement or **test time augmentation** [51] applied as post-processing). Application of these methods depends on the task and the desired quality of the result; some of the most commonly used processing steps are described in Supplementary Materials 1.

Nucleus segmentation is traditionally performed using a data-specific workflow that contains various filtering and thresholding methods, followed by morphological operations and processing steps (*ImageJ/Fiji* [38,39], *QuPath* [47], *CellProfiler* [52]). Segmentation using pixel classification, based on classical machine learning methods has been used for challenging data for a decade, with early versions of tools including, for example, *DeepMIB* [53] and *ilastik* [42]. The fundamental difference between classical image processing-based nucleus segmentation and that with classical machine learning is the input required from the user: in the former case, manual parameter setting and fine-tuning is expected in different processing modules in the pipeline, which is still capable of yielding very high accuracy at the expense of time-consuming re-parameterization for each new experiment. The latter enables users to rely on automated feature extraction and learning by still providing examples manually, which most likely also need to be repeated in

experiments. Notably, appropriate pre-processing of input images (e.g., intensity scaling) can help to unify the range of optimal parameters in both cases. The nuclear segmentation task has moved towards robust and automated approaches with *U-Net* [34] (see in Supplementary Materials 4), which was a breakthrough for deep learning-based nucleus segmentation (and in the field of deep learning-based segmentation in general). In contrast to image processing and classical machine learning, deep learning-based methods require fewer input parameters from the user, and are generally more straightforward to apply between experiments than in the case of classical approaches. Nonetheless, pre-processing also increases the accuracy of CNNs in most cases. *U-Net* still serves as a baseline for semantic segmentation tasks, and is (i) used as part of recent general nucleus/cell segmentation pipelines, such as *Cellpose* [12] and *StarDist* [54], and (ii) utilized or further developed in *nnU-Net* [55] and *UNet++* [56]. Even though *U-Net* is a semantic segmentation framework, it can be extended to instance segmentation with post-processing. One typical solution is to classify pixels into three classes where one class represents nuclear edges, and as such, it can aid instance segmentation [10]. Computationally *U-Net* is relatively simple, thus it is possible to train a basic *U-Net* on workstations or even laptops with a **graphics processing unit (GPU)**.

Another breakthrough in deep learning-based instance segmentation was *Mask R-CNN* [57]. This network was designed for the segmentation of natural images; however, it has been adapted for nucleus segmentation in methods such as *nucleAIzer* [11]. *Mask R-CNN* is built over a *CNN* feature extraction backbone and **regional proposal network (RPN)** [58] to suggest possible object regions. These proposals are classified and used for binary mask prediction. *Mask R-CNN* outputs a list of masks allowing overlaps, whereas the output of *U-Net* is an image with no overlaps. However, two recent extensions to *U-Net*-based *StarDist, MultiStar* [59] and *SplineDist* [60], enable segmentation of overlapping objects. *NuSeT* [61] combines RPN, *U-Net* and watershed post-processing to optimize segmentation of crowded cells. *Mask R-CNN* requires more computational resources than *U-Net*; still it can be trained on a modern workstation or laptop.

Even though many segmentation methods are not deep learning-based (*MINS* [62,63], *XPIWIT* [64], etc.), the field has recently tended to shift towards approaches based on deep learning (e.g., *ilastik* [42] now offers DNNs). This includes bundles of specific deep learning methods for segmentation and pre-processing which could be used on Google Colab [*ZeroCostDL4Mic* [65], *Segmentation of stochastic optical reconstruction microscopy (STORM) images* [66]], or other client-server architecture (*ImJoy* [67], *DeepCell Kiosk* [46,68], *HistomicsML2* [69]) with provided separate pre-trained models (*CDeep3M* [70], *nucleAIzer* [11], *Cellpose* [12]). *ImageJ* users can also utilize deep learning-based segmentation with plugins and pre-trained models (*DeepImageJ* [71]). The majority of the methods discussed here are deep learning-based (see Table 2), which require hardware resources due to the parallelizable and heavy computational costs of DNNs, hence GPU acceleration is advised, especially for training. Cloud-based solutions often meet this requirement.

Several methods mentioned in the preceding text could be used for 3D datasets (see Table 2). Segmentation of 3D nuclear images with deep learning is not straightforward. The major limitation is that the annotated data in the field are less abundant compared with the planar case. There are several deep learning-based methods developed by the medical image analysis community facing a similar challenge. However, in the case of medical images, usually only one or a few objects need to be segmented. This task is different from and less difficult than nucleus segmentation, where hundreds of instances should be segmented even when they touch. For example, segmenting a medical image by combining the segmentations of 2D images may provide

Table 2. Relevant tools for nucleus segmentation[a]

2D/3D/Both	Tool name	Pipeline/algorithm/platform	Code availability	Year	Reference	GUI/Tutorial/Biaflows/GPU/Cloud
2D	*U-Net*	Algorithm	Yes	2015	Ronneberger *et al.* [34]	N/N/Y/Y/N
2D	*SegNet*	Algorithm	Yes	2015	Badrinarayanan *et al.* [87]	N/N/?/Y/N
2D	*Mask R-CNN*	Algorithm	Yes	2017	He *et al.* [57]	N/Y/Y/Y/N
2D	*QuPath*	Platform	Yes	2017	Bankhead *et al.* [47]	Y/Y/N/Y/N
2D	*UNet++*	Algorithm	Yes	2018	Zhou *et al.* [56]	N/N/N/Y/N
2D	Segmentation of nuclei in histopathology images by deep regression of the distance map	Algorithm	Yes	2018	Naylor *et al.* [88]	N/Y/N/Y/N
2D	Multi-scale cell instance segmentation with keypoint graph-based bounding boxes	Algorithm	Yes	2019	Yi et al. [89]	N/N/?/Y/N
2D	*HoVer-Net*	Algorithm	Yes	2019	Graham *et al.* [23]	N/Y/N/Y/N
2D	*CIA-Net*	Algorithm	No	2019	Zhou *et al.* [90]	N/N/N/Y/N
2D	*Bend-Net*	Algorithm	No	2020	Wang *et al.* [91]	N/N/N/Y/N
2D	*nucleAIzer*	Algorithm, Pipeline	Yes	2020	Hollandi *et al.* [11]	Y/Y/N/Y/Y
2D	*MultiStar*	Algorithm	Yes	2020	Walter *et al.* [59]	N/N/N/Y/N
2D	Instance-aware self-supervised learning for nuclei segmentation	Algorithm	No	2020	Xie *et al.* [85]	N/N/N/Y/N
2D	Self-supervised nuclei segmentation in histopathological images using attention	Algorithm	Yes	2020	Sahasrabudhe *et al.* [84,85]	N/N/N/Y/N
2D	*Triple U-Net*	Algorithm	Yes	2020	Zhao *et al.* [92]	N/N/N/Y/N
2D	High-resolution deep transferred ASPPU-Net for nuclei segmentation of histopathology images	Algorithm	No	2021	Chanchal *et al.* [93]	N/N/N/Y/N
2D	*NucleiSegNet*	Algorithm	Yes	2021	Lal *et al.* [94]	N/Y/N/Y/N
2D	*SplineDist*	Algorithm	Yes	2021	Mandal *et al.* [60]	N/N/N/Y/N
2D	Contour proposal network	Algorithm	Yes	2021	Upschulte *et al.* [95]	N/N/N/Y/N
2D	*HistomicsML2*	Pipeline, Platform	Yes	2021	Lee *et al.* [69]	Y/Y/N/Y/Y
2D	*STORM*	Pipeline	Yes	2021	Mela *et al.* [66]	N/N/N/Y/Y
2D	*MSRF-Net*	Algorithm	Yes	2021	Srivastava *et al.* [96]	N/N/N/Y/N
3D	3D cell nuclei segmentation based on gradient flow tracking	Algorithm	No	2007	Li *et al.* [97]	N/N/N/N/N
3D	*Vaa3D*	Platform	Yes	2010	Peng *et al.* [79]	Y/Y/Y/Y/N
3D	*IT3DImageJSuite*	Platform	Yes	2013	Ollion *et al.* [98]	Y/Y/N/N/N
3D	*LoS*	Algorithm	Yes	2013	Asafi *et al.* [74]	N/Y/N/N/N
3D	Automated cell segmentation with 3D fluorescence microscopy images	Algorithm	No	2015	Kong *et al.* [99]	N/N/N/N/N
3D	*OpenSegSPIM*	Platform	Yes	2016	Gole *et al.* [75]	Y/Y/N/N/N
3D	*RACE*	Platform	Yes	2016	Stegmaier *et al.* [76]	Y/Y/N/Y/N
3D	*U-Net (3D)*	Algorithm	Yes	2016	Cicek *et al.* [100]	N/N/N/Y/N
3D	Segmentation of fluorescence microscopy images using 3D active contours with inhomogeneity correction	Algorithm	No	2017	Lee *et al.* [14]	N/N/N/N/N
3D	*DeepSynth*	Algorithm	No	2019	Dunn *et al.* [101]	N/N/N/Y/N

(continued on next page)

Table 2. (continued)

2D/3D/Both	Tool name	Pipeline/algorithm/platform	Code availability	Year	Reference	GUI/Tutorial/Biaflows/GPU/Cloud
3D	3D segmentation and reconstruction of neuronal nuclei in confocal microscopic images	Algorithm	Yes	2019	Ruszczycki *et al.* [15]	N/N/N/N/N
3D	Semi supervised segmentation and graph-based tracking of 3D nuclei in time-lapse microscopy	Algorithm	Yes	2020	Shailja *et al.* [102]	N/N/N/Y/N
3D	A deep learning pipeline for nucleus segmentation	Pipeline	No	2020	Zaki *et al.* [103]	N/N/N/Y/N
3D	Combined detection and segmentation of cell nuclei in microscopy images using deep learning	Algorithm	No	2020	Ram *et al.* [104]	N/N/N/Y/N
3D	*QCANet*	Algorithm	Yes	2020	Tokuoka *et al.* [81]	N/Y/N/Y/N
3D	*Allen cell and structure segmenter*	Platform	Yes	2020	Chen *et al.* [105]	Y/Y/N/Y/N
3D	*3D-Cell-Annotator*	Platform	Yes	2020	Tasnadi *et al.* [43]	Y/Y/N/Y/N
3D	Nuclei detection for 3D microscopy with a fully convolutional regression network"	Algorithm	No	2021	Lapierre-Landry *et al.* [83]	N/N/N/Y/N
3D	*3DeeCellTracker*	Platform	Yes	2021	Wen *et al.* [82]	N/Y/N/Y/N
Both	*MINS*	Platform	Yes	2014	Lou *et al.* [62,63]	Y/Y/N/N/N
Both	*XPIWIT*	Algorithm	Yes	2016	Bartschat *et al.* [64]	Y/Y/N/Y/N
Both	*ilastik*	Platform	Yes	2018	Berg *et al.* [42]	Y/Y/Y/Y/N
Both	*DeepImageJ*	Platform	Yes	2019	Gómez-de-Mariscal *et al.* [71]	Y/Y/N/Y/N
Both	*ImJoy*	Platform	Yes	2019	Ouyang *et al.* [67]	Y/Y/N/Y/Y
Both	A coarse-to-fine data generation method for 2D and 3D cell nucleus segmentation	Algorithm	No	2020	Zhao *et al.* [106]	N/N/N/Y/N
Both	*Cellpose*	Algorithm	Yes	2020	Stringer *et al.* [12]	Y/Y/Y/Y/Y
Both	*CDeep3M*	Platform	Yes	2020	Haberl *et al.* [70]	Y/Y/N/Y/Y
Both	*StarDist*	Algorithm	Yes	2020	Shmidt *et al.* [13]; Weigert *et al.* [54]	N/Y/Y/Y/N
Both	*NuSeT*	Platform	Yes	2020	Yang *et al.* [61]	Y/Y/N/Y/N
Both	*nnU-Net*	Platform	Yes	2021	Isensee *et al.* [55]	N/Y/N/Y/N
Both	*DeepMIB*	Platform	Yes	2021	Belevich *et al.* [53]	Y/Y/N/Y/N
Both	*InstantDL*	Pipeline, Platform	Yes	2021	Waibel *et al.* [107]	N/Y/N/Y/N
Both	*ZeroCostDL4Mic*	Pipeline, Platform	Yes	2021	von Chamier *et al.* [65]	Y/Y/N/Y/Y
Both	*DeepCell Kiosk*	Pipeline, Platform	Yes	2021; 2016	Bannon *et al.* [46,68]; Van Valen *et al.* [46,68]	Y/Y/Y/Y/Y
Both	*AD-GAN*	Algorithm	No	2021	Yao *et al.* [86]	N/N/N/Y/N
Both	Embedding-based instance segmentation in microscopy	Algorithm	Yes	2021	Lalit *et al.* [108]	N/Y/?/Y/N

[a]Algorithm: a complete method to segment nuclei. An algorithm can be shared as a source code for developers in e.g., a GitHub repository or can be implemented as a user-accessible method in a platform. Pipeline: a workflow of image processing algorithms to segment nuclei, allowing the user to set parameters for each step of the workflow or even change the included algorithms to optimize segmentation tailored to the specific data. Platform: a software package that includes multiple algorithms or pipelines for nucleus segmentation, and often has a defined **application programming interface** to include additional methods as well.

acceptable accuracy. In contrast, nucleus segmentation is an instance segmentation task where this approach alone is less likely to work in crowded parts of the image, but the connected components of the stacked 2D segmentations can be used as a seed image for the watershed

transform to compute the final 3D instance segmentation [72]. Besides, 3D segmentation is more demanding in terms of computational resources (especially GPU memory and file sizes) when a dense 2D method is extended directly to process 3D images. Introduction of a further dimension may lead to substantially growing complexity (e.g., in case of differential geometry-based approaches) and more complex spatial dependencies in case of CNNs, however, this phenomenon termed 'the curse of dimensionality' is especially problematic, thus more training data and more computational resources are required. Still, several tools are specifically developed for the 3D segmentation task [73], and some deep learning-based methods developed for 2D segmentation are also extended to 3D. The *IT3DImageJSuite* is an *ImageJ* (*Fiji*) [38,39] plugin that involves several algorithms (including **iterative thresholding** and watershed). **Line of sight (*LoS*)** [74] approximates the convex decomposition of the objects with spectral clustering. *OpenSegSPIM* [75] is a *MATLAB* application which performs instance segmentation by applying a pipeline of filters in a semi-automatic manner. *RACE* [76] and Ruszczycki *et al*. [15] first compute the 2D segmentation on the z-slices, and then combine them to 3D objects. Similarly to *BioImageXD* [77], *Fiji* [38,39] and *Icy* [78], *Vaa3D* [79] uses a pipeline consisting of Gaussian filtering, adaptive thresholding, distance transformation and 3D watershed [80], while the MITK plugin *3D-Cell-Annotator* [43] uses active contours for semi-automatic 3D segmentation. In contrast, most recent methods apply deep learning techniques to segment nuclei. These include *QCANet* [81], developed to analyze mouse embryos in 3D, *3DeeCellTracker* [82], intended for tracking after the segmentation of nucleus instances, and the algorithm proposed in Lapierre-Landry *et al*. [83] which performs watershed segmentation on the probability map, and supervoxel clustering to achieve the final instance segmentation.

Self-supervised and unsupervised learning approaches decrease or even eliminate the need of annotated training data. A few of such methods for nuclear segmentation have appeared recently [84–86]. These methods show competitive results, although their accuracy does not exceed that of the supervised state-of-the-art methods. Self-supervised segmentation for histopathology images [85] uses *ResUnet-101* and requires a minimum of annotated data for fine-tuning. Another approach [84] uses an attention mechanism, and does not require annotated data. *AD-GAN* [86] uses a sophisticated training approach based on GAN, does not require annotated data, and also works for both 2D and 3D.

Table 2 and Supplementary Table 3 report the list of tools mentioned previously, whilst Supplementary Materials 4 includes their short descriptions.

Most of the listed tools require some effort from the user to install, prepare the environment, do the pre-processing of the input if needed, and finally to run it. The amount of time and effort primarily depends on the computational background of the user, and on the tool itself. Cloud-based tools [usually supplied with web graphical user interface (GUI)] could be the primary starter choices for life scientists. However, there is a trade-off: cloud-based versions of tools have limited customizability, while local versions are more flexible, and the user does not need to share the data with third-party services. In the latter case the quality of the documentation also matters to assure proper set-up. In Table 2 we provide information on whether the tool is documented properly (only official documentation was taken into account). The algorithms quite often lack detailed official documentation, though provide the most flexibility (usually are parts of the complex pipelines), and for the most popular ones unofficial documentation or tutorials and third-party implementations exist too. The potential performance of a tool is obviously an important concern for the user, and it might be challenging to decide on choosing the appropriate tool. The user may decide based on the community's preferences. Alternatively, a reliable comparison of the performance of the different tools can support decision-making. However, apart from BIAFLOWS[ii] and

automatic challenge submission systems (e.g., Kaggle or ISBI[v]), the microscopy image analyst community lacks (i) an evaluation platform for the objective comparison of nucleus segmentation methods, using (ii) a standardized evaluation metric in a transparent way. Thus, a consensus on utilizing a single, standardized platform is eagerly awaited. Since challenge portals only provide this functionality for the datasets of given challenges, a more inclusive platform, such as BIAFLOWS is suggested. Even though the relevance of newly published methods is usually supported by some quantitative segmentation results, it has several shortcomings from the user's point of view as follows. (i) The test dataset might not suggest relevant performance when the dataset size is too small, or covers a single imaging modality only. However, approaches developed for specific microscopy images (such as H&E or fluorescence confocal images) or segmentation scenarios (e.g., crowded cell culture) are intended to work in their given domain of images, and should not be expected to perform just as well on more extensive or general datasets. (ii) When comparison to prior methods is performed and reported, the number of tested methods is usually low, and (iii) additional model- or data-specific modifications might have been applied to the compared methods (or the test images as pre-processing), thus merely literature-based comparisons of accuracy scores may confuse the user (see Supplementary Materials 2).

Outstanding questions

Can existing tools reliably segment touching/overlapping/cluttered or morphologically challenging nuclei? Do training datasets include such examples of nuclei to enable learning of the visual representation?

How to make general-purpose models that work accurately for many different imaging modalities?

How to define a standardized approach to quantitatively compare the existing solutions, and properly assess new methods to overcome current comparison challenges?

What are the future steps developers should take towards a general solution for nucleus segmentation?

Concluding remarks

Recent years have brought significant improvements in nucleus segmentation, including large annotated datasets, new high-accuracy 2D/3D strategies, deep learning approaches, and segmentation benchmarking platforms, however, establishing a genuinely general solution for nucleus segmentation is still an unmet need. In this review and the accompanying web-based portal[i] we aimed to cover the missing link between recent advancements and users' needs by providing a detailed overview on the available means for nucleus segmentation. The concluding remarks are focused on crucial limitations and future goals.

The first crucial point is to cover more modalities of microscopy data for both 2D, and especially 3D, with open datasets of annotated images. Current methods are expected to work when trained on additional microscopy data modalities [18,109]. Most datasets include H&E-stained tissues or fluorescently labeled cell cultures (see Table 1) which are two of the most widely used modalities in practice. However, further microscopy types (e.g., **differential interference contrast**, light-sheet or **phase contrast**) lack such publicly available annotations, except a recently published, large, label-free dataset [26]. Even though researchers can train existing deep learning methods on their own nowadays, these models remain private (unless released on e.g., GitHub, zenodo, Kaggle or in a Napari [110] plugin; on the first three platforms datasets may also be deposited [32]) and the initial datasets are small, resulting in suboptimal model generalization. For a given modality of interest, generalization is also a crucial point for medical applications: the data should be as diverse as possible to promote robust models. Diversity from a computer vision point of view would include various regions of tissue with the distinct visual appearance of both the target objects and the surroundings, as well as covering several phenotypes of cells, different batches or slightly different experimental set-ups. An extensive annotated dataset including most (if not all) modalities occurring in single-cell analysis experiments with respect to the type of microscopy, sample, and label could definitely improve existing trainable methods. Besides, it would offer the possibility of releasing genuinely general pre-trained models, and would also serve as a standard dataset, similarly to the widely used COCO dataset [111] in computer vision.

The second crucial point relates to solving common microscopy challenges for both 2D and 3D data, such as: touching, overlapping, and irregularly shaped nuclei [54,59–61]. Either dataset design or model architecture can be beneficial for a solution. Current methods achieve various levels of success in overcoming these issues, thus further developments are needed.

The third crucial point is the lack of (i) a globally accepted benchmark platform for comparison, and (ii) a unified metric for tool evaluation. BIAFLOWS and Kaggle are available solutions to overcome these issues. However, still most publications presenting novel methods or tools typically provide limited comparisons (either in terms of the data used in evaluation or the number of methods compared) and use non-standardized metrics. Accordingly, the results published by different authors are often difficult to compare.

The ultimate goal is to develop an algorithm, and train it so that the resulting single model would be able to accurately segment nuclei in a variety of microscopy modalities. Some of the available algorithms and models are aimed to meet this requirement [11,12], and the field is moving towards a generally applicable solution. While a quantitative comparison of the methods available for each modality is beyond our intention, it is worth mentioning that deep learning tends to provide fine accuracy in segmenting nuclei in images obtained with different microscopy techniques, as shown at the DSB2018 challenge [6,112].

Acknowledgments

We would like to thank Krisztián Koós and Ákos Diósdi (BIOMAG, Szeged, Hungary) for sharing images used in the figures and their important opinions on the segmentation tools; Michael Reil (Kutay lab, Institute of Biochemistry, Department of Biology, ETH Zurich, Zurich, Switzerland), Marie Follo (Lighthouse Core Facility, Medical Center, University of Freiburg, Freiburg, Germany), microscope-microscope.org, Frits Thorsen (Molecular Imaging Center, Department of Biomedicine, University of Bergen, Bergen, Norway), Dr Markus Kerr (Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, Australia) for sharing images; Csaba Molnár (Broad Institute, Cambridge, MA, USA) for helping with the implementation of *unbias*. E.T. and N.M. acknowledge support from the Cooperative Doctoral Programme of the Ministry for Innovation and Technology (KDP-2020). F.P. acknowledges support from the Union for International Cancer Control (UICC) for a UICC Yamagiwa-Yoshida (YY) Memorial International Cancer Study Grant (ref: UICC-YY/678329). P.H., E.T., N.M., and R.H. acknowledge support from the LENDULET-BIOMAG Grant (2018-342), European Regional Development Funds (GINOP-2.3.2-15-2016-00006, GINOP-2.3.2-15-2016-00026, GINOP-2.3.2-15-2016-00037), COMPASS-ERAPerMed H2020, CZI Deep Visual Proteomics, H2020-DiscovAir, H2020-FAIR CHARM-101016457, Thematic Excellence Programme Grant (TKP2021-EGA-09), Eotvos Lorand Research Network (ELKH SZBK-SZTE ELI), the Fast track programme Grant (2020-1.1.5-GYORSÍTÓSÁV), the National Research, Development and Innovation Fund (OTKA SNN-21 - 139455), and from the ELKH-Excellence grant. L.P. acknowledges the FIMM High Content Imaging and Analysis unit (HiLIFE, University of Helsinki and Biocenter Finland) and the grants from Academy of Finland (340273, 337036). The authors thank Dora Bokor, PharmD (Szeged, Hungary), for proofreading the manuscript.

Declaration of interests

No interests are declared.

Supplemental information

Supplemental information associated with this article can be found online https://doi.org/10.1016/j.tcb.2021.12.004.

Resources

[i]https://biomag-lab.github.io/microscopy-tree/

[ii]https://biaflows.neubias.org

[iii]http://cancergenome.nih.gov/

[iv]https://imagej.net/plugins/labkit

[v]https://biomedicalimaging.org/

[vi]www.andrewjanowczyk.com/deep-learning/

References

1. Meijering, E. (2012) Cell segmentation: 50 years down the road [life sciences]. *IEEE Signal Process. Mag.* 29, 140–145
2. Hallou, A. *et al.* (2021) Deep learning for bioimage analysis in developmental biology. *Development* 148, dev199616
3. Chandrasekaran, S.N. *et al.* (2021) Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nat. Rev. Drug Discov.* 20, 145–159
4. Dhar, R. *et al.* (2019) Single cell functional genomics reveals the importance of mitochondria in cell-to-cell phenotypic variation. *eLife* 8, e38904

5. Hayakawa, T. *et al.* (2021) Computational nuclei segmentation methods in digital pathology: a survey. *Arch. Comput. Methods Eng.* 28, 1–13
6. Caicedo, J.C. *et al.* (2019) Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl. *Nat. Methods* 16, 1247–1253
7. Greenwald, N.F. *et al.* (2021) Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol.* Published online November 18, 2021. https://doi.org/10.1038/s41587-021-01094-0
8. Pelkmans, L. *et al.* (2005) Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. *Nature* 436, 78–86
9. Kumar, N. *et al.* (2020) A multi-organ nucleus segmentation challenge. *IEEE Trans. Med. Imaging* 39, 1380–1391
10. Caicedo, J.C. *et al.* (2019) Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytometry A* 95, 952–965
11. Hollandi, R. *et al.* (2020) NucleAIzer: a parameter-free deep learning framework for nucleus segmentation using image style transfer. *Cell Syst.* 10, 453–458.e6
12. Stringer, C. *et al.* (2021) Cellpose: a generalist algorithm for cellular segmentation. *Nat. Methods* 18, 100–106
13. Weigert, M. *et al.* (2020) Star-convex polyhedra for 3D object detection and segmentation in microscopy. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, IEEE
14. Lee, S. *et al.* (2017) Segmentation of fluorescence microscopy images using three dimensional active contours with inhomogeneity correction. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 709–713
15. Ruszczycki, B. *et al.* (2019) Three-dimensional segmentation and reconstruction of neuronal nuclei in confocal microscopic images. *Front. Neuroanat.* 13, 81
16. Amgad, M. *et al.* NuCLS: a scalable crowdsourcing, deep learning approach and dataset for nucleus classification, localization and segmentation. *arXiv [cs.CV]* Published online February 18, 2021, arxiv.org/abs/2102.09099
17. Rubens, U. *et al.* (2020) BIAFLOWS: A collaborative framework to reproducibly deploy and benchmark bioimage analysis workflows. *Patterns (N Y)* 1, 100040
18. Alemi Koohbanani, N. *et al.* (2020) NuClick: a deep learning framework for interactive segmentation of microscopic images. *Med. Image Anal.* 65, 101771
19. Ljosa, V. *et al.* (2012) Annotated high-throughput microscopy image sets for validation. *Nat. Methods* 9, 637
20. Kromp, F. *et al.* (2020) An annotated fluorescence image dataset for training nuclear segmentation methods. *Sci. Data* 7, 262
21. Rivron, N.C. *et al.* (2018) Blastocyst-like structures generated solely from stem cells. *Nature* 557, 106–111
22. Yao, K. *et al.* (2021) Scaffold-A549: a benchmark 3D fluorescence image dataset for unsupervised nuclei segmentation. *Cognit. Comput.* 13, 1603–1608
23. Graham, S. *et al.* (2019) Hover-Net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Med. Image Anal.* 58, 101563
24. Mahbod, A. *et al.* (2021) CryoNuSeg: a dataset for nuclei instance segmentation of cryosectioned H&E-stained histological images. *Comput. Biol. Med.* 132, 104349
25. Janowczyk, A. and Madabhushi, A. (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J. Pathol. Inform.* 7, 29
26. Edlund, C. *et al.* (2021) LIVECell – a large-scale dataset for label-free live cell segmentation. *Nat. Methods* 18, 1038–1045
27. Graham, S. *et al.* Lizard: a large-scale dataset for colonic nuclear instance segmentation and classification. *arXiv [cs.CV]* Published online August 25, 2021, arxiv.org/abs/2108.11195
28. Kumar, N. *et al.* (2017) A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans. Med. Imaging* 36, 1550–1560
29. Lin, Z. *et al.* (2021) NucMM dataset: 3D neuronal nuclei instance segmentation at sub-cubic millimeter scale. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, pp. 164–174
30. Gamper, J. *et al.* (2019) PanNuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In *Digital Pathology. ECDP 2019. Lecture Notes in Computer Science* (vol 11435) (Reyes-Aldasoro, C. *et al.*, eds), pp. 11–19, Springer
31. Irshad, H. *et al.* (2015) Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd. *Pac. Symp. Biocomput.* 2015, 294–305
32. Naylor, P.J. *et al.* (2018) Segmentation of nuclei in histopathology images by deep regression of the distance map. *Zenodo* v1.1. https://zenodo.org/record/2579118#.YcB01dKES70
33. Wienert, S. *et al.* (2012) Detection and segmentation of cell nuclei in virtual microscopy images: a minimum-model approach. *Sci. Rep.* 2, 503
34. Ronneberger, O. *et al.* (2015) U-Net: convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science*, pp. 234–241, Springer International Publishing
35. Ulman, V. *et al.* (2017) An objective comparison of cell-tracking algorithms. *Nat. Methods* 14, 1141–1152
36. Hollandi, R. *et al.* (2020) AnnotatorJ: an ImageJ plugin to ease hand annotation of cellular compartments. *Mol. Biol. Cell* 31, 2179–2186
37. Borland, D. *et al.* (2021) Segmentor: a tool for manual refinement of 3D microscopy annotations. *BMC Bioinformatics* 22, 260
38. Schneider, C.A. *et al.* (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675
39. Schindelin, J. *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682
40. Nolden, M. *et al.* (2013) The Medical Imaging Interaction Toolkit: challenges and advances: 10 years of open-source development. *Int. J. Comput. Assist. Radiol. Surg.* 8, 607–620
41. Arganda-Carreras, I. *et al.* (2017) Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* 33, 2424–2426
42. Berg, S. *et al.* (2019) ilastik: interactive machine learning for (bio)image analysis. *Nat. Methods* 16, 1226–1232
43. Tasnadi, E.A. *et al.* (2020) 3D-Cell-Annotator: an open-source active surface tool for single-cell segmentation in 3D microscopy images. *Bioinformatics* 36, 2948–2949
44. Marée, R. *et al.* (2016) Collaborative analysis of multi-gigapixel imaging data using Cytomine. *Bioinformatics* 32, 1395–1401
45. Rubens, U. *et al.* (2019) Cytomine: toward an open and collaborative software platform for digital pathology bridged to molecular investigations. *Proteomics Clin. Appl.* 13, e1800057
46. Bannon, D. *et al.* (2021) DeepCell Kiosk: scaling deep learning-enabled cellular image analysis with Kubernetes. *Nat. Methods* 18, 43–45
47. Bankhead, P. *et al.* (2017) QuPath: open source software for digital pathology image analysis. *Sci. Rep.* 7, 16878
48. Peng, T. *et al.* (2017) A BaSiC tool for background and shading correction of optical microscopy images. *Nat. Commun.* 8, 14836
49. Smith, K. *et al.* (2015) CIDRE: an illumination-correction method for optical microscopy. *Nat. Methods* 12, 404–406
50. Goyal, B. *et al.* (2020) Image denoising review: from classical to state-of-the-art approaches. *Inf. Fusion* 55, 220–244
51. Moshkov, N. *et al.* (2020) Test-time augmentation for deep learning-based cell segmentation on microscopy images. *Sci. Rep.* 10, 5068
52. McQuin, C. *et al.* (2018) CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol.* 16, e2005970
53. Belevich, I. and Jokitalo, E. (2021) DeepMIB: user-friendly and open-source software for training of deep learning network for biological image segmentation. *PLoS Comput. Biol.* 17, e1008374
54. Schmidt, U. *et al.* (2018) Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pp. 265–273, Springer International Publishing
55. Isensee, F. *et al.* (2021) nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* 18, 203–211

56. Zhou, Z. *et al.* (2018) UNet++: a nested U-Net architecture for medical image segmentation. *Deep Learn. Med. Image Anal. Multimodal. Learn. Clin. Decis. Support* 11045, 3–11
57. He, K. *et al.* (2017) Mask R-CNN. In *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE
58. Ren, S. *et al.* (2017) Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149
59. Walter, F.C. *et al.* (2021) Multistar: instance segmentation of overlapping objects with star-convex polygons. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, IEEE
60. Mandal, S. and Uhlmann, V. (2021) Splinedist: automated cell segmentation with spline curves. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 1082–1086, IEEE
61. Yang, L. *et al.* (2020) NuSeT: a deep learning tool for reliably separating and analyzing crowded cells. *PLoS Comput. Biol.* 16, e1008193
62. Lou, X. *et al.* (2014) A rapid and efficient 2D/3D nuclear segmentation method for analysis of early mouse embryo and stem cell image data. *Stem Cell Rep.* 2, 382–397
63. Saiz, N. *et al.* (2016) Quantitative analysis of protein expression to study lineage specification in mouse preimplantation embryos. *J. Vis. Exp.* 108, 53654
64. Bartschat, A. *et al.* (2016) XPIWIT – an XML pipeline wrapper for the Insight Toolkit. *Bioinformatics* 32, 315–317
65. von Chamier, L. *et al.* (2021) Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat. Commun.* 12, 2276
66. Mela, C.A. and Liu, Y. (2021) Application of convolutional neural networks towards nuclei segmentation in localization-based super-resolution fluorescence microscopy images. *BMC Bioinformatics* 22, 325
67. Ouyang, W. *et al.* (2019) ImJoy: an open-source computational platform for the deep learning era. *Nat. Methods* 16, 1199–1200
68. Van Valen, D.A. *et al.* (2016) Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS Comput. Biol.* 12, e1005177
69. Lee, S. *et al.* (2021) Interactive classification of whole-slide imaging data for cancer researchers. *Cancer Res.* 81, 1171–1177
70. Haberl, M.G. *et al.* (2018) CDeep3M-Plug-and-Play cloud-based deep learning for image segmentation. *Nat. Methods* 15, 677–680
71. Gómez-de-Mariscal, E. *et al.* DeepImageJ: a user-friendly environment to run deep learning models in ImageJ. *bioRxiv*, Published online October 16, 2019, https://doi.org/10.1101/799270
72. Kar, A. *et al.* Assessment of deep learning algorithms for 3D instance segmentation of confocal image datasets. *bioRxiv*, Published online June 10, 2021, https://doi.org/10.1101/2021.06.09.447748
73. Piccinini, F. *et al.* (2020) Software tools for 3D nuclei segmentation and quantitative analysis in multicellular aggregates. *Comput. Struct. Biotechnol. J.* 18, 1287–1300
74. Asafi, S. *et al.* (2013) Weak convex decomposition by lines-of-sight. *Comput. Graph. Forum* 32, 23–31
75. Gole, L. *et al.* (2016) OpenSegSPIM: a user-friendly segmentation tool for SPIM data. *Bioinformatics* 32, 2075–2077
76. Stegmaier, J. *et al.* (2016) Real-time three-dimensional cell segmentation in large-scale microscopy data of developing embryos. *Dev. Cell* 36, 225–240
77. Kankaanpää, P. *et al.* (2012) BioImageXD: an open, general-purpose and high-throughput image-processing platform. *Nat. Methods* 9, 683–689
78. de Chaumont, F. *et al.* (2012) Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* 9, 690–696
79. Peng, H. *et al.* (2010) V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nat. Biotechnol.* 28, 348–353
80. Long, F. *et al.* (2009) A 3D digital atlas of C. elegans and its application to single-cell analyses. *Nat. Methods* 6, 667–672
81. Tokuoka, Y. *et al.* (2020) 3D convolutional neural networks-based segmentation to acquire quantitative criteria of the nucleus during mouse embryogenesis. *NPJ Syst. Biol. Appl.* 6, 32
82. Wen, C. *et al.* (2021) 3DeeCellTracker, a deep learning-based pipeline for segmenting and tracking cells in 3D time lapse images. *eLife* 10, e59187
83. Lapierre-Landry, M. *et al.* (2021) Nuclei detection for 3D microscopy with a fully convolutional regression network. *IEEE Access* 9, 60396–60408
84. Sahasrabudhe, M. *et al.* (2020) Self-supervised nuclei segmentation in histopathological images using attention. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 393–402, Springer International Publishing
85. Xie, X. *et al.* (2020) Instance-aware self-supervised learning for nuclei segmentation. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, pp. 341–350, Springer International Publishing
86. Yao, K. *et al.* AD-GAN: end-to-end unsupervised nuclei segmentation with aligned disentangling training. *arXiv* Published online July 23, 2021, arxiv.org/abs/2107.11022
87. Badrinarayanan, V. *et al.* (2017) SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 2481–2495
88. Naylor, P. *et al.* (2019) Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Trans. Med. Imaging* 38, 448–459
89. Yi, J. *et al.* (2019) Multi-scale cell instance segmentation with keypoint graph based bounding boxes. In *Lecture Notes in Computer Science*, pp. 369–377, Springer International Publishing
90. Zhou, Y. *et al.* (2019) CIA-net: robust nuclei instance segmentation with contour-aware information aggregation. In *Lecture Notes in Computer Science*, pp. 682–693, Springer International Publishing
91. Wang, H. *et al.* (2020) Bending loss regularized network for nuclei segmentation in histopathology images. *Proc. IEEE Int. Symp. Biomed. Imaging* 2020, 258–262
92. Zhao, B. *et al.* (2020) Triple U-net: hematoxylin-aware nuclei segmentation with progressive dense feature aggregation. *Med. Image Anal.* 65, 101786
93. Chanchal, A.K. *et al.* (2021) High-resolution deep transferred ASPPU-Net for nuclei segmentation of histopathology images. *Int. J. Comput. Assist. Radiol. Surg.* 16, 2159–2175
94. Lal, S. *et al.* (2021) NucleiSegNet: robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images. *Comput. Biol. Med.* 128, 104075
95. Upschulte, E. *et al.* Contour proposal networks for biomedical instance segmentation. *arXiv [cs.CV]* Published online April 7, 2021, arxiv.org/abs/2104.03393
96. Srivastava, A. *et al.* MSRF-Net: a multi-scale residual fusion network for biomedical image segmentation. *arXiv* Published online May 16, 2021, arxiv.org/abs/2105.07451
97. Li, G. *et al.* (2007) 3D cell nuclei segmentation based on gradient flow tracking. *BMC Cell Biol.* 8, 40
98. Ollion, J. *et al.* (2013) TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics* 29, 1840–1841
99. Kong, J. *et al.* (2015) Automated cell segmentation with 3D fluorescence microscopy images. *Proc. IEEE Int. Symp. Biomed. Imaging* 2015, 1212–1215
100. Çiçek, Ö. *et al.* (2016) 3D U-net: learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, pp. 424–432, Springer International Publishing
101. Dunn, K.W. *et al.* (2019) DeepSynth: three-dimensional nuclear segmentation of biological images using neural networks trained with synthetic data. *Sci. Rep.* 9, 18295
102. Shailja, S. *et al.* (2021) Semi supervised segmentation and graph-based tracking of 3D nuclei in time-lapse microscopy. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pp. 385–389, IEEE
103. Zaki, G. *et al.* (2020) A deep learning pipeline for nucleus segmentation. *Cytometry A* 97, 1248–1264
104. Ram, S. *et al.* (2020) Combined detection and segmentation of cell nuclei in microscopy images using deep learning.

In *2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 26–29, IEEE

105. Chen, J. *et al.* The Allen Cell and Structure Segmenter: a new open source toolkit for segmenting 3D intracellular structures in fluorescence microscopy images. *bioRxiv*, Published online December 8, 2018, https://doi.org/10.1101/491035
106. Zhao, Z. *et al.* (2020) A coarse-to-fine data generation method for 2D and 3D cell nucleus segmentation. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 41–46, IEEE
107. Waibel, D.J.E. *et al.* (2021) InstantDL: an easy-to-use deep learning pipeline for image segmentation and classification. *BMC Bioinformatics* 22, 103
108. Lalit, M. *et al.* Embedding-based instance segmentation in microscopy. *arXiv* Published online January 25, 2021, arxiv.org/abs/2101.10033
109. Fishman, D. *et al.* (2021) Practical segmentation of nuclei in brightfield cell images with neural networks trained on fluorescently labelled samples. *J. Microsc.* 284, 12–24
110. Sofroniew, N. *et al.* (2021) *napari/napari: 0.4.11*, *Zenodo*. https://zenodo.org/record/4136688#.YcB-DNKES70
111. Lin, T.-Y. *et al.* Microsoft COCO: Common Objects in Context. *arXiv* Published online May 1, 2014, arxiv.org/abs/1405.0312
112. Verma, R. *et al.* (2021) MoNuSAC2020: a multi-organ nuclei segmentation and classification challenge. *IEEE Trans. Med. Imaging* 40, 3413–3423