 https://doi.org/10.3311/PPee.44574

Periodica Polytechnica Electrical Engineering and Computer Science, 70(1), pp. 60–67, 2026

Standardization, Benchmarking, and Fine-Tuning in Deep Learning-Based Cell Segmentation and Tracking

Ágoston Richárd Kiss[1,2*]

[1] Department of Electronics Technology, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[2] HUN-REN Centre for Energy Research, Institute of Technical Physics and Materials Science, "Lendület" Nanobiosensorics Research Group, Konkoly-Thege Miklós út 29-33, H-1121 Budapest, Hungary
[*] Corresponding author, e-mail: kissagostonrichard@edu.bme.hu

Received: 05 May 2026, Accepted: 09 June 2026, Published online: 06 July 2026

Abstract

Bioimage analysis workflows for living cells typically involve a sequence of steps, including image acquisition, preprocessing, cell segmentation, object classification, and temporal tracking. Cell segmentation aims to identify and separate individual cells in microscopy images, whereas cell tracking links these segmented objects across time-lapse image sequences to quantify cell movement, morphology, and dynamic behavior. These tasks are challenging because microscopy datasets often vary in image quality and may contain imaging noise, variable staining, overlapping cells, heterogeneous cell morphologies, and changing acquisition conditions. Over the past decade, the field has shifted from heuristic and rule-based image processing toward deep learning-based approaches, supported by advances in convolutional neural networks, foundation models, and increasingly standardized datasets. This transition has been closely connected to the development of interoperable data formats, shared benchmarks, and open-source bioimage analysis ecosystems. The present review discusses this evolution with a focus on standardization, benchmarking, and data-centric model adaptation strategies. In this context, data-centric strategies refer to approaches that improve model performance primarily through better data selection, annotation, fine-tuning, and expert feedback rather than through architecture design alone. Particular attention is given to small-sample fine-tuning and human-in-the-loop workflows, which aim to adapt pretrained segmentation and tracking models to laboratory-specific microscopy data while reducing annotation effort and improving the reliability of downstream biological conclusions.

Keywords

cell segmentation, bioimage informatics, fine-tuning, human-in-the-loop

1 Introduction

In the past decade, bioimaging has become one of the emerging scientific fields, fundamentally transforming the quantitative and qualitative analysis of microorganisms and cellular activities. Early image analysis methods relied primarily on manual observations or heuristics based on simple, predefined features. For instance, global thresholding segments images by applying a single intensity cutoff across the entire image, while background correction compensates for uneven illumination or acquisition artifacts by subtracting an estimated background intensity. However, these methods frequently struggle with the inherent complexity of biological samples, such as the presence of outlier subpopulations at atypical proportions, intricate and transient cell-cell interactions, colliding cell pairs, and dimly stained or irregularly shaped cells. These limitations have contributed to the shift from heuristic image analysis toward trainable and deep learning-based approaches. At the same time, benchmark studies using both synthetic and experimentally acquired datasets have shown that robust model performance depends not only on algorithm design, but also on the availability of representative and standardized training data [1]. As a result, modern bioimage analysis increasingly relies on shared datasets, reproducible evaluation frameworks, and trainable models that can process larger and more heterogeneous image collections than earlier heuristic methods.

A typical bioimage analysis workflow starts with image acquisition and preprocessing, followed by cell segmentation, classification or phenotypic characterization, and, in time-lapse experiments, cell tracking. Because each step

Cite this article as: Kiss, Á. R. "Standardization, Benchmarking, and Fine-Tuning in Deep Learning-Based Cell Segmentation and Tracking", Periodica Polytechnica Electrical Engineering and Computer Science, 70(1), pp. 60–67, 2026. https://doi.org/10.3311/PPee.44574

depends on reliable data handling and task-specific evaluation, workflow-level challenges have made standardization, benchmarking, and iterative model adaptation central issues in modern bioimage analysis.

2 Standardization of datasets

The need for unified handling and sharing of large-scale, multidimensional microscopy image data, as well as for ensuring interoperability among different analysis tools, emerged early and led to efforts to standardize file formats and data models [2]. One of the key outcomes of these initiatives is the standardization effort initiated by the Open Microscopy Environment (OME), alongside which hierarchical data formats such as HDF5 and the OME's next-generation file format (OME-NGFF) have also gained importance. These formats differ not only in structure but also in data access strategies: while monolithic file formats (e.g., HDF5) rely on fast random access within a single file, next-generation, chunk-based approaches (e.g., OME-Zarr within OME-NGFF) enable scalable and efficient access in distributed and cloud-based environments. Within this framework the OME-TIFF format was developed, followed by the introduction of the OME-Zarr cloud-optimized format in response to the demands of cloud-based processing. The latter enables efficient, chunked storage of large-scale, multidimensional image data and allows fast, partial access [3]. Standardization is also closely linked to the advancement of image processing algorithms; in particular, deep learning-based approaches (such as those relying on convolutional neural networks or Transformer architectures) require increasingly large and diverse datasets, further reinforcing the need for unified data structures [4–7]. Typical bioimage analysis workflows can be broadly divided into three main tasks: segmentation, classification, and tracking, each of which introduces specific computational and data-related challenges (see Fig. 1).

In addition, the increasing scale and complexity of datasets have also driven the adoption of more informative visualization techniques, such as raincloud plots, which simultaneously display raw data points, distribution density, and summary statistics, thereby providing a more transparent representation of underlying data distributions compared to traditional summary plots.

Beyond file-format standardization, the increasing integration of automated image analysis into microscopy workflows also requires consistent data handling across acquisition, storage, preprocessing, and downstream analysis. This need for unified data structures is particularly

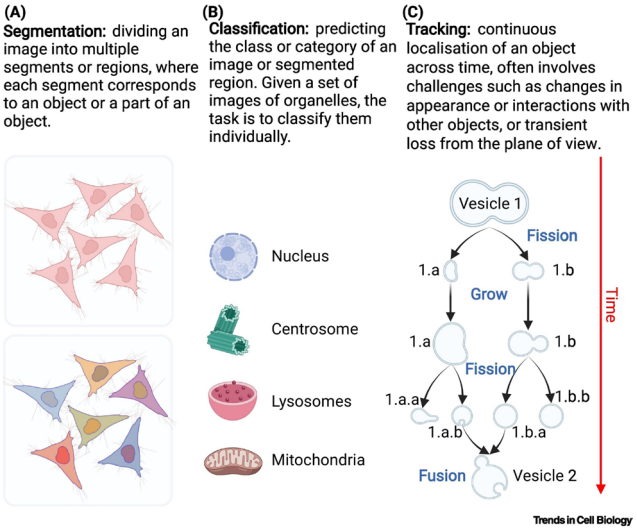

Fig. 1 Overview of core bioimage analysis workflows, including preprocessing, segmentation, classification, and tracking. The figure illustrates the main analytical steps and representative challenges; detailed task definitions and benchmarking aspects are discussed in the text. Adapted from [4].

important as deep learning-based methods become more data-intensive and are increasingly applied in biological and clinical imaging contexts. The growing combination of advanced microscopy, automated image analysis, and single-cell technologies further highlights the need for standardized and reproducible data handling in modern bioimage analysis workflows [8].

3 Ecosystems and algorithm benchmarking

In bioimage analysis, benchmarking refers to the systematic comparison of algorithms on shared datasets using predefined evaluation metrics and standardized ground truth annotations. For segmentation, commonly used metrics include precision, recall, Dice coefficient, intersection-over-union-based scores, and F1-score. For tracking, evaluation must additionally account for temporally consistent object identities, including linking accuracy, track fragmentation, missed detections, false detections, and errors around cell division events.

Within the bioimaging software ecosystem, open-source and commercial solutions coexist in a complementary manner. Community-driven open-source tools and libraries play a central role in facilitating interoperability between workflows and analytical platforms, while increasingly adopting standardized data formats. Recent work has introduced specialized models designed to address specific computational challenges; for example, Scl-Yolo provides a lightweight architecture for blood cell detection, with strength in resolving overlapping cells in

complex imaging conditions. At the same time, generalist algorithms such as Cellpose3 have evolved to integrate image restoration with improved cellular segmentation performance. More recently, the emergence of foundation models, such as Segment Anything for Microscopy and CellSAM has marked a shift toward large-scale pretrained approaches capable of delivering robust segmentation across diverse biological datasets [9–12]. In contrast, commercial bioimaging platforms, particularly those deployed in clinical and high-throughput industrial settings are designed as tightly integrated systems that combine image acquisition, hardware control, and analysis within a single environment. The practical benchmarking of these ecosystems is particularly vital in high-throughput environments, such as microfluidic single-cell analysis platforms. Studies focusing on continuous cell monitoring have demonstrated that standardized algorithms must maintain high precision even under varying flow conditions and temporal discontinuities [13, 14]. These solutions typically provide validated performance, streamlined workflows, and guaranteed compatibility with specific imaging hardware, which is critical in regulated applications. However, this tight integration often relies on proprietary data formats and limited extensibility, which can restrict data reuse, complicate interoperability with external tools, and hinder transparent algorithm comparison across platforms. As a result, user flexibility and reproducibility may be constrained in practice [15–17]. To address these limitations, the bioimaging community has increasingly shifted toward open and standardized data structures and interoperable frameworks. This transition has not only facilitated data sharing and reuse but has also enabled the systematic benchmarking of algorithms using objective and comparable evaluation metrics, thereby supporting more rigorous methodological development [18].

In response, a landmark study introduced a framework that enabled the comparison of different algorithms using unified datasets and evaluation metrics. One of the most important innovations of this approach was the creation of a publicly available database consisting of both real and simulated microscopy images, covering various cell types, imaging modalities, and differing levels of noise and complexity [18]. A key finding of the study was that the performance of different algorithms strongly depends on data characteristics, such as noise level, cell density, and three-dimensional structure, and that no universally optimal method exists for all scenarios [18]. This clearly highlights that progress in the field requires not only new algorithms but also standardized evaluation environments and shared reference datasets. These developments coincided with the emergence of deep learning methods in bioimage processing.

4 Deep learning

Uniformly designed evaluation frameworks have not only enabled the objective comparison of algorithms but have also provided standardized, annotated datasets and evaluation protocols that are essential for training and reliably assessing data-intensive deep neural networks. Cell tracking methods clearly reflect these developments: beyond their initial role as reference frameworks, such initiatives have continuously expanded with new, more diverse datasets and evaluation possibilities. Developments after 2017, particularly the rise of comprehensive segmentation-focused benchmarks, have significantly increased the volume and diversity of annotated data available to the research community. This expansion has directly enabled the application of data-intensive deep learning approaches by providing the scale needed for robust model training. At the same time, the growing dominance of deep neural networks has introduced new challenges, particularly regarding stringent data requirements and the complexity of training strategies [19].

Clinical hematology provides a practical example of how deep learning-based image analysis can be integrated into standardized diagnostic workflows. The convolutional neural network (CNN) based Morphogo system, developed for bone marrow cell analysis, combines automated whole-slide imaging with expert-validated morphological classification and is capable of identifying more than 25 types of nucleated cells. Its reported performance, including an overall accuracy of 99.01%, a sensitivity of 80.95%, and a specificity of 99.48%, illustrates how deep learning can support high-throughput, standardized cell classification in clinically relevant settings [20].

The efficacy of deep learning architectures is not limited to bone marrow analysis; similar robust performances have been observed in the morphological characterization of peripheral blood cells. Neural network-based automated pipelines effectively address challenges such as high variance in cell shapes and staining artifacts, thereby successfully bridging the gap between raw experimental observations and quantitative biological insights [21].

Compared with segmentation, cell tracking introduces additional temporal challenges. Errors in a single frame may propagate across the entire trajectory, and tracking algorithms must handle cell migration, mitosis, temporary

occlusion, changes in morphology, and missed or false detections. Therefore, tracking performance cannot be assessed solely by segmentation accuracy; it also requires metrics that evaluate identity preservation and the correctness of temporal links. This distinction is particularly important in biological studies where downstream conclusions depend on dynamic parameters such as migration speed, lineage relationships, proliferation, or time-dependent morphological changes.

5 Iterative training workflows and the human-in-the-loop ecosystem

In bioimaging, the training of neural networks typically follows an iterative, progressively refined workflow in which data preparation, model fitting, evaluation, and method optimization occur in repeated cycles [22]. This cyclic perspective is also reflected in practical protocols: training neural networks rarely involves a single model run, but rather follows a process involving data, model, evaluation, and deployment, where dataset and evaluation protocols are at least as critical as the chosen architecture.

The typical steps in image processing tasks are as follows. First, data collection and preprocessing are performed, including image acquisition (2D/3D formats), intensity normalization, noise reduction, contrast enhancement, and tiling in the case of images with large field-of-view [22, 23]. This is followed by labeling or annotation and by dividing the data into training, validation, and test subsets. In segmentation, cell detection, or cell tracking tasks, annotation is typically the most resource-intensive step, requiring not only labels but also consistent definitions (e.g., what constitutes an object, how overlaps are handled, and how specific cell-cell interactions are treated). Frameworks aimed at small-sample adaptation also emphasize that manual annotation of large datasets can realistically take weeks; consequently, fine-tuning approaches requiring minimal labeling are gaining increasing importance in practice [24–26].

The third step is model selection and training. In the field of segmentation, one of the key methodological breakthroughs was the introduction of the U-Net architecture, which combines an encoder-decoder structure with skip connections, enabling the capture of both global context and fine structural details. This approach, complemented by strong data augmentation strategies, allows effective training even on relatively small datasets, which is particularly important in biological applications [27]. Subsequent developments have focused not only on architecture but also on complete training and processing frameworks. An example is nnU-Net, which implements the entire processing pipeline, from preprocessing through architecture and training parameters to postprocessing, within a unified framework, significantly reducing the need for expert manual tuning [28]. As a result, pre-trained and directly applicable models have emerged, such as the open-source Cellpose, which are designed to work across various types of microscopy images with minimal training cycles and limited fine-tuning. These systems do not introduce entirely new architectures but instead build upon existing ones to provide generalizable models that operate robustly under diverse experimental conditions. This approach is particularly advantageous in everyday laboratory applications, where fast and reliable processing is essential [10, 29].

Finally, user-friendly tools requiring fewer or no programming have significantly improved accessibility. For example, deepImageJ, as a Fiji/ImageJ plugin, enables the direct execution of pre-trained neural networks while maintaining compatibility with multiple deep learning frameworks. Such solutions bridge the gap between traditional image processing environments and modern Python-based deep learning tools, enabling the integrated use of different technologies. This substantially reduces development effort and facilitates broader accessibility of advanced methods to the research community [30].

This progress is largely driven by the open-source ecosystem, which ensures free access to software and promotes community-driven integration. Within this framework, the "human-in-the-loop" paradigm emerges at stages where full automation is either unreliable or not cost-effective. Such tasks include the definition of quality criteria, annotation, labeling, and validation, error analysis, and iterative refinement, as well as expert approval in clinical applications [31]. In cell segmentation and tracking, it is particularly evident that the challenges of deep learning-based methods often stem not from model inaccuracy, but from the difficulty of correcting errors and feeding these corrections back into the learning process. In addition to the high annotation demands of supervised learning, many approaches lack mechanisms to efficiently correct erroneous associations and incorporate these corrections through fine-tuning. Furthermore, interactive, human-in-the-loop tools that would support continuous adjustment of model behavior are often missing. As a result, manual cell tracking often remains the preferred solution in practice, since post hoc correction of automated results frequently does not yield meaningful time savings [25].

To improve time efficiency, the human-in-the-loop approach has evolved toward annotating not entire datasets but only the most informative data points. Based on these targeted annotations, the model is fine-tuned, and performance is gradually improved through iterative feedback cycles. This integrated approach, combining automated selection and expert intervention, creates a synergistic environment for model optimization (see Fig. 2) [32].

6 Fine-tuning and annotation-efficient model adaptation

In this context, fine-tuning refers to adapting a pre-trained model (a general-purpose or task-specific segmentation neural network) to a new microscopy environment using a small amount of data from critical regions. This allows the method to be applied under diverse conditions, including different cell types, staining protocols, noise levels, and focus settings. The importance of fine-tuning lies in the fact that, although "generalist" segmenters (e.g., Cellpose) exist, the heterogeneity of real laboratory data often leads to performance degradation in specific tissues or imaging conditions, and domain-specific adaptation can yield measurable improvements [10, 25, 29, 33].

Fine-tuning strategies can follow different approaches, primarily differing in how many model parameters are updated during training. A common approach to fine-tuning is to update only a subset of the model parameters, such as the output layers, which enables more stable training and faster convergence even with limited annotations. At the other extreme is full fine-tuning, where most or all of the network parameters are updated; while this provides high flexibility, it is computationally demanding and may lead to overfitting when only small datasets are available. In recent years, so-called parameter-efficient adaptation methods have gained increasing attention. These approaches aim to achieve significant performance improvements with only a minimal number of trainable parameters. The core idea behind LoRA (Low-Rank Adaptation) is that retraining the entire model is unnecessary; instead, it is sufficient to learn low-rank modifications relative to the fixed weights, significantly reducing computational and memory requirements [24, 25, 34]. Although this approach was originally introduced for language models, the underlying principle is particularly promising for bioimaging tasks, where annotated data are often limited.

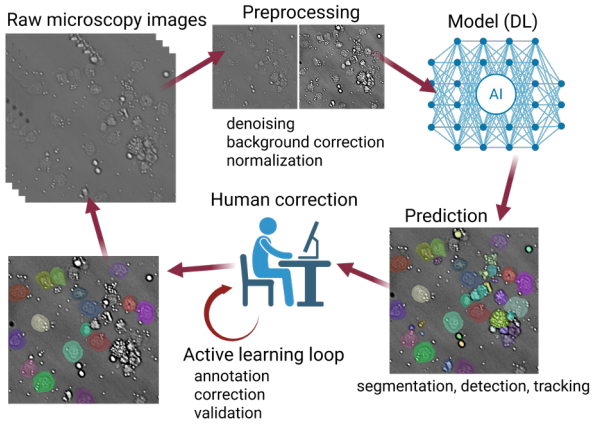

Fig. 2 Schematic representation of the iterative human-in-the-loop (HITL) fine-tuning workflow. The pipeline utilizes a pre-trained model to extract feature vectors from the raw microscopy dataset, followed by the Centrality-Guided Maximum Diversity (CGMD) algorithm to identify the most diverse and representative samples. These targeted images are then expert-annotated and incorporated into the model through fine-tuning, ensuring high-accuracy domain adaptation with minimal manual effort. Concept adapted from [32].

The efficiency of fine-tuning can be further improved through targeted sample selection. For example, the CGMD (Centrality-Guided Maximum Diversity) pipeline proposes an image selection strategy that enables effective adaptation even under extremely low annotation budgets: the selected images are both representative and diverse in the model's representation space, thereby providing maximal new information during fine-tuning [32]. In this approach, centrality favors samples that are representative of the dataset distribution, whereas maximum diversity reduces redundancy among the selected images. As a result, a small annotation set can cover a broader range of relevant morphologies, image qualities, and experimental conditions, making fine-tuning more efficient under limited annotation budgets. Among recent small-sample adaptation strategies, CSRefiner is discussed separately because it provides a concrete example of how limited manual annotation can be translated into measurable segmentation improvement in whole-tissue images.

6.1 Small-sample fine-tuning for tissue segmentation (CSRefiner)

CSRefiner represents a significant methodological focal point, as it provides a small-sample fine-tuning workflow for scenarios requiring robust segmentation across entire tissue sections (whole tissue). This approach is particularly critical where methodological errors could directly impact biological conclusions, such as the identification of rare cell populations or the detection of subpopulations with proportions strongly deviating from physiological norms [5, 34–36]. A key strength of this framework is that algorithms like

StarDist are highly effective at segmenting samples with high morphological variance, offering a clear parallel to the diverse morphological characteristics of leukocyte populations. The workflow is exceptionally practical for laboratory applications due to its minimal data requirements; approximately 300 annotated cells are sufficient for effective model adaptation. The performance improvements enabled by CSRefiner are substantial: in the case of the StarDist algorithm, the F1-score, defined as the harmonic mean of precision and recall, increased from 0.42 to 0.87, while the overall processing speed was enhanced by a factor of 36 [34].

In another approach, data augmentation can be achieved by generating synthetic samples (e.g., through mixup-type and graph-based techniques). This is particularly useful in cases such as cancer cells, where certain subpopulations may appear in atypical proportions within the tissue [35]. While such model-based approaches aim to improve generalization, CSRefiner primarily operates on the data side, adapting to the specific tissue and image quality through targeted, small-scale fine-tuning. Together, these approaches illustrate that performance gains often arise not from a single new architecture but from the combination of architectural innovations and fine-tuning strategies.

One of the most important practical advantages of CSRefiner is the drastic reduction in annotation cost. Compared to full manual annotation of entire samples, annotating targeted regions is orders of magnitude faster while still being sufficient for effective model adaptation. This makes fine-tuning truly practical in laboratory settings, where model performance is ultimately evaluated based on the quality of biological conclusions.

7 Summary

In the field of bioimage processing, particularly in cell segmentation and cell tracking, a significant methodological transformation has taken place over the past decade. This shift has been driven by the rise of deep learning methods, the emergence of standardized data formats, and the development of community-based evaluation frameworks. Compared to earlier heuristic-based approaches, trainable models provide greater robustness and generalizability; however, they are strongly data-dependent, which has brought increased attention to the importance of standardized data handling and annotation.

Initiatives such as the cell tracking challenge have enabled objective comparison of algorithms through unified datasets and evaluation metrics and have laid the foundation for the creation of large-scale annotated databases.

In current trends, both the development of model architectures (e.g., data augmentation strategies) and data-centric approaches (e.g., fine-tuning and targeted annotation) play a decisive role. From a practical perspective, the emergence of user-friendly tools and human-in-the-loop approaches is particularly important, as they enable iterative model improvement. Together, these elements form an integrated ecosystem in which deep learning can be effectively applied in biological research and clinical practice. Overall, the field is moving toward integrated workflows in which standardized data formats, shared benchmarks, pretrained models, targeted annotation, and human-in-the-loop fine-tuning jointly support more reproducible and biologically reliable bioimage analysis.

Acknowledgement

This work was supported by the National Research, Development and Innovation Office of Hungary (NKFIH) under grants no. PD 134195 (for Z.Sz.), the "Élvonal" KKP_19 funding scheme, the TKP2022-EGA-04 project, and the HUN-REN (formerly ELKH) thematic fund.

The research was further supported by the Doctoral Excellence Scholarship Program (DKÖP) of the Ministry of Culture and Innovation, implemented with the support provided from the National Research, Development and Innovation Fund of Hungary and the Budapest University of Technology and Economics, under the support agreement with the National Research, Development and Innovation Office.

References

[1] Maška, M., Ulman, V., Svoboda, D., Matula, P., Matula, P., ..., Ortiz-de-Solorzano, C. "A benchmark for comparison of cell tracking algorithms", Bioinformatics, 30(11), pp. 1609–1617, 2014. https://doi.org/10.1093/bioinformatics/btu080

[2] Moore, J., Allan, C., Besson, S., Burel, J.-M., Diel, E., ..., Swedlow, J. R. "OME-NGFF: a next-generation file format for expanding bioimaging data-access strategies", Nature Methods, 18(12), pp. 1496–1498, 2021. https://doi.org/10.1038/s41592-021-01326-w

[3] Moore, J., Basurto-Lozada, S., Besson, S., Bogovic, J., Bragantini, J., ..., Swedlow, J. R. "OME-Zarr: a cloud-optimized bioimaging file format with international community support", Histochemistry and Cell Biology, 160(3), pp. 223–251, 2023. https://doi.org/10.1007/s00418-023-02209-1

[4] Chai, B., Efstathiou, C., Yue, H., Draviam, V. M. "Opportunities and challenges for deep learning in cell dynamics research", Trends in Cell Biology, 34(11), pp. 955–967, 2024. https://doi.org/10.1016/j.tcb.2023.10.010

[5] Salem, A., Teama, M., Kassem, H. A., Vakilzadehian, N., Ali, A. M. A., Venugopal, D., Khalifa, A. M. "Artificial Intelligence in Hematologic Malignancies: Opportunities, Challenges, and Clinical Integration", Cureus, 18(1), 100950, 2026. https://doi.org/10.7759/cureus.100950

[6] Greenwald, N. F., Miller, G., Moen, E., Kong, A., Kagel, T., ..., Van Valen, D. "Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning", Nature Biotechnology, 40(4), pp. 555–565, 2022. https://doi.org/10.1038/s41587-021-01094-0

[7] Baydargil, H. B., Bocklitz, T. "Unstained Blood Smear Analysis: A Review of Rule-Based, Machine Learning, and Deep Learning Techniques", [preprint] Preprints.org, 17 March 2025. https://doi.org/10.22541/au.174223532.21153944/v1

[8] Szittner, Z., Péter, B., Kurunczi, S., Székács, I., Horvath, R. "Functional blood cell analysis by label-free biosensors and single-cell technologies", Advances in Colloid and Interface Science, 308, 102727, 2022. https://doi.org/10.1016/j.cis.2022.102727

[9] Stringer, C., Pachitariu, M. "Cellpose3: one-click image restoration for improved cellular segmentation", Nature Methods, 22(3), pp. 592–599, 2025. https://doi.org/10.1038/s41592-025-02595-5

[10] Yang, F., Yonghong, W. "Scl-Yolo: A Lightweight Model Based on Improved Yolov11n and its Application in Blood Cell Object Detection", [preprint] SSRN, 05 March 2025. https://doi.org/10.2139/ssrn.5165561

[11] Archit, A., Freckmann, L., Nair, S., Khalid, N., Pilt, P., ..., Pape, C. "Segment Anything for Microscopy", Nature Methods, 22(3), pp. 579–591, 2025. https://doi.org/10.1038/s41592-024-02580-4

[12] Israel, U., Marks, M., Dilip, R., Li, Q., Yu, C., ..., Van Valen, D. "CellSAM: A Foundation Model for Cell Segmentation", [preprint] bioRxiv, 17 February 2025. https://doi.org/10.1101/2023.11.17.567630

[13] Beres, B., Kovacs, K. D., Kanyo, N., Peter, B., Szekacs, I., Horvath, R. "Label-Free Single-Cell Cancer Classification from the Spatial Distribution of Adhesion Contact Kinetics", ACS Sensors, 9(11), pp. 5815–5827, 2024. https://doi.org/10.1021/acssensors.4c01139

[14] Kovacs, K. D., Beres, B., Kanyo, N., Szabó, B., Peter, B., Bősze, S., Szekacs, I., Horvath, R. "Single-cell classification based on label-free high-resolution optical data of cell adhesion kinetics", Scientific Reports, 14(1), 11231, 2024. https://doi.org/10.1038/s41598-024-61257-2

[15] Katz, B., Feldman, M. D., Tessema, M., Benisty, D., Toles, G. S., ..., Pozdnyakova, O. "Evaluation of Scopio Labs X100 Full Field PBS: The first high-resolution full field viewing of peripheral blood specimens combined with artificial intelligence-based morphological analysis", International Journal of Laboratory Hematology, 43(6), pp. 1408–1416, 2021. https://doi.org/10.1111/ijlh.13681

[16] Lee, L. H., Mansoor, A., Wood, B., Nelson, H., Higa, D., Naugler, C. "Performance of CellaVision DM96 in leukocyte classification", Journal of Pathology Informatics, 4(1), 14, 2013. https://doi.org/10.4103/2153-3539.114205

[17] Gilda, J. E., Ko, J.-H., Elfassy, A.-Y., Tropp, N., Parnis, A., Ayalon, B., Jhe, W., Cohen, S. "A semiautomated measurement of muscle fiber size using the Imaris software", American Journal of Physiology-Cell Physiology, 321(3), pp. C615–C631, 2021. https://doi.org/10.1152/ajpcell.00206.2021

[18] Ulman, V., Maška, M., Magnusson, K. E. G., Ronneberger, O., Haubold, C., ..., Ortiz-de-Solorzano, C. "An objective comparison of cell-tracking algorithms", Nature Methods, 14(12), pp. 1141–1152, 2017. https://doi.org/10.1038/nmeth.4473

[19] Bhattarai, A., Meyer, J., Petersilie, L., Shah, S. I., Neu, L. A., Rose, C. R., Ullah, G. "Deep-Learning-Based Segmentation of Cells and Analysis (DL-SCAN)", Biomolecules, 14(11), 1348, 2024. https://doi.org/10.3390/biom14111348

[20] Lv, Z., Cao, X., Jin, X., Xu, S., Deng, H. "High-accuracy morphological identification of bone marrow cells using deep learning-based Morphogo system", Scientific Reports, 13(1), 13364, 2023. https://doi.org/10.1038/s41598-023-40424-x

[21] Kovács, K. D., Szittner, Z., Magyaródi, B., Péter, B., Szabó, B., Vörös, A., Kanyó, N., Székács, I., Horvath, R. "Optical sensor reveals the hidden influence of cell dissociation on adhesion measurements", Scientific Reports, 14(1), 11719, 2024. https://doi.org/10.1038/s41598-024-61485-6

[22] Sacks, B., Mody, J., Huang, Y., Jung, S., Kimelman, ..., Kara, E. "Protocol for automated analysis of biological images using Python code", STAR Protocols, 7(1), 104368, 2026. https://doi.org/10.1016/j.xpro.2026.104368

[23] Xu, X., Xu, S., Jin, L., Song, E. "Characteristic analysis of Otsu threshold and its applications", Pattern Recognition Letters, 32(7), pp. 956–961, 2011. https://doi.org/10.1016/j.patrec.2011.01.021

[24] Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W. "LoRA: Low-Rank Adaptation of Large Language Models", [preprint] arXiv, 16 October 2021. https://doi.org/10.48550/arXiv.2106.09685

[25] Lalit, M., Funke, J. "An Investigation of Unsupervised Cell Tracking and Interactive Fine-Tuning", In: 2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Honolulu, HI, USA, 2025, pp. 5851–5859. ISBN 979-8-3315-8988-2 https://doi.org/10.1109/ICCVW69036.2025.00610

[26] Passerotti, G., Alberello, A., Vichi, M., Bennetts, L. G., Bailey, J., Toffoli, A. "Sea Ice Floe Segmentation in Close-Range Optical Imagery Using Active Contour and Foundation Models", Earth and Space Science, 13(2), e2025EA004453, 2026. https://doi.org/10.1029/2025EA004453

[27] Azad, R., Aghdam, E. K., Rauland, A., Jia, Y., Avval, A. H., Bozorgpour, A., Karimijafarbigloo, S., Cohen, J. P., Adeli, E., Merhof, D. "Medical Image Segmentation Review: The Success of U-Net", IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(12), pp. 10076–10095, 2024. https://doi.org/10.1109/TPAMI.2024.3435571

[28] Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., Maier-Hein, K. H. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation", Nature Methods, 18(2), pp. 203–211, 2021.
https://doi.org/10.1038/s41592-020-01008-z.

[29] Stringer, C., Wang, T., Michaelos, M., Pachitariu, M. "Cellpose: a generalist algorithm for cellular segmentation", Nature Methods, 18(1), pp. 100–106, 2021.
https://doi.org/10.1038/s41592-020-01018-x

[30] Selzer, G. J., Rueden, C. T., Hiner, M. C., Evans, E. L., Harrington, K. I. S., Eliceiri, K. W. "napari-imagej: ImageJ ecosystem access from napari", Nature Methods, 20(10), pp. 1443–1444, 2023.
https://doi.org/10.1038/s41592-023-01990-0

[31] Lam, V. K., Byers, J. M., Robitaille, M. C., Kaler, L., Christodoulides, J. A., Raphael, M. P. "A self-supervised learning approach for high throughput and high content cell segmentation", Communications Biology, 8(1), 780, 2025.
https://doi.org/10.1038/s42003-025-08190-w

[32] Sheikh, E. M., Tharwat, A., Schwan, C., Schenck, W. "CGMD: Centrality-Guided Maximum Diversity for Annotation-Efficient Fine-Tuning of Pretrained Cell Segmentation Models", [preprint] arXiv, 07 March 2024.
https://doi.org/10.48550/arXiv.2403.04567

[33] Schmidt, U., Weigert, M., Broaddus, C., Myers, G. "Cell Detection with Star-Convex Polygons", In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, Granada, Spain, 2018, 11071, pp. 265–273. ISBN 978-3-030-00934-2
https://doi.org/10.1007/978-3-030-00934-2_30

[34] Shi, C., Li, Y., Guo, J., Chen, Q., Cao, T., Liao, S., Chen, A., Li, M., Zhang, Y. "CSRefiner: a lightweight framework for fine-tuning cell segmentation models with small datasets", Briefings in Bioinformatics, 27(1), bbaf718, 2026.
https://doi.org/10.1093/bib/bbaf718

[35] Mustafa, S., Li, G., Iqbal, Y., Akhtar, S., Iqbal, A., Lin, L. "Improving blood cell classification in imbalanced medical imaging using Mixup-augmented graph-attention encoder-decoder network", Biomedical Signal Processing and Control, 120, 110116, 2026.
https://doi.org/10.1016/j.bspc.2026.110116

[36] Asadov, C., Shirinova, A., Alimirzoyeva, Z., Hasanova, A. "Digital Medicine and Artificial Intelligence in Chronic Myeloid Leukemia: Current Applications, Challenges, and Future Directions", Cureus, 18(2), e103951, 2026.
https://doi.org/10.7759/cureus.103951