Comparing Deep Learning Performance for Chronic Lymphocytic Leukaemia Cell Segmentation in Brightfield Microscopy Images

Markéta Vašinková[1] [iD], Vít Doleží[1], Michal Vašinek[1], Petr Gajdoš[1] and Eva Kriegová[2]

[1]Department of Computer Science, FEECS, VSB - Technical University of Ostrava, Ostrava, Czech Republic. [2]Department of Immunology, Faculty of Medicine and Dentistry, Palacky University & University Hospital, Olomouc, Czech Republic.

ABSTRACT

OBJECTIVES: This article focuses on the detection of cells in low-contrast brightfield microscopy images; in our case, it is chronic lymphocytic leukaemia cells. The automatic detection of cells from brightfield time-lapse microscopic images brings new opportunities in cell morphology and migration studies; to achieve the desired results, it is advisable to use state-of-the-art image segmentation methods that not only detect the cell but also detect its boundaries with the highest possible accuracy, thus defining its shape and dimensions.

METHODS: We compared eight state-of-the-art neural network architectures with different backbone encoders for image data segmentation, namely U-net, U-net++, the Pyramid Attention Network, the Multi-Attention Network, LinkNet, the Feature Pyramid Network, DeepLabV3, and DeepLabV3+. The training process involved training each of these networks for 1000 epochs using the PyTorch and PyTorch Lightning libraries. For instance segmentation, the watershed algorithm and three-class image semantic segmentation were used. We also used StarDist, a deep learning-based tool for object detection with star-convex shapes.

RESULTS: The optimal combination for semantic segmentation was the U-net++ architecture with a ResNeSt-269 background with a data set intersection over a union score of 0.8902. For the cell characteristics examined (area, circularity, solidity, perimeter, radius, and shape index), the difference in mean value using different chronic lymphocytic leukaemia cell segmentation approaches appeared to be statistically significant (Mann–Whitney U test, $P < .0001$).

CONCLUSION: We found that overall, the algorithms demonstrate equal agreement with ground truth, but with the comparison, it can be seen that the different approaches prefer different morphological features of the cells. Consequently, choosing the most suitable method for instance-based cell segmentation depends on the particular application, namely, the specific cellular traits being investigated.

KEYWORDS: Cell detection, cell segmentation, chronic lymphocytic leukaemia cells, image analysis, U-net++

Introduction

The combination of state-of-the-art instrumentation for the temporal surveillance of biological specimens, along with deep learning-based image processing techniques, offers promising opportunities to advance fundamental scientific research. We have used the latest advancements in image data processing using neural networks with different backbones and have proposed a method for segmenting chronic lymphocytic leukaemia (CLL) cells, which is an initial step towards further analyses of cell image data. Accurate image segmentation of cells is necessary to determine their morphological characteristics, which are important indicators of cell properties. Accurate cell segmentation is also a primary step towards successful cell tracking during *in vitro* migration analyses. Although most studies address the efficient detection of cancer cells from blood samples, our study addresses the need to segment cells in brightfield time-lapse microscopy images to study cell migration and changes in cell morphology using readily available computational technologies.

Our experiments answer the following questions:

- What are the most suitable new deep learning algorithms for cell detection in brightfield images?
- What is the best way to separate instances of individual objects? Comparison of three approaches. Which approach works best in our data?
- What parameters can be used to characterise cells? How do the values of these parameters vary depending on the different approaches?

To address this, we conducted a comprehensive comparative analysis using a variety of neural network architectures and tested their performance with different backbones.

*The need for precision cell segmentation
of CLL cells*

CLL is a chronic disease with significant genetic, morphological, and phenotypical heterogeneity. The condition exhibits considerable clinical variability among patients with respect to prognosis and clinical course, ranging from slow and stable to extremely aggressive.[1] CLL is characterised by the progressive accumulation of $CD5^+$, $CD23^+$, and $CD19^+$ B cells in the peripheral blood, bone marrow, lymph nodes and spleen.[2] Recent investigations have shown that variations in the size of CLL cells are indicative of their activation status, polarisation, and motility compared to the population of smaller CLL cells and indicate that the functional differences observed *in vitro* may reflect a different *in vivo* leukaemic potential, probably responsible for the heterogeneous characteristics of the disease and may contribute to different clinical manifestations in CLL patients.[3] In addition, a negative prognostic value was found to be associated with atypical CLL cell morphology.[4] Although CLL cells in peripheral blood are floating freely, their migration to and from the bone marrow is essential for their survival,[5] thus, a deeper understanding of the morphology and migratory behaviour of CLL cell subpopulations can be useful for both fundamental research and precision medicine.

Furthermore, state-of-the-art studies focus on automated detection and classification of cells from blood smear to perform accurate segmentation of tumour cells from normal cells,[6-8] classification of different types of blood cells,[9,10] or detection of specific properties of blood cells, such as deformability[11] and velocity.[12]

Cell segmentation methods

The investigation of cell segmentation has garnered growing interest in recent years due to the importance of cell morphology as a crucial phenotypic characteristic that provides insight into a cell's physiological condition. Furthermore, the delineation of cell boundaries is frequently necessary for the subsequent analysis of intracellular processes at the nanoscale or for studying cellular interactions at the millimetre scale. Therefore, cell segmentation is a major challenge in image data analysis due to its relevance in understanding cellular behaviour and enabling comprehensive investigations into various biological phenomena.[13]

Automated segmentation and detection of cells from optical microscopy and other imaging devices, particularly time-lapse microscopy, can be a valuable addition in addressing various biological questions that are heavily based on single-cell analysis. During the past decade, the field has evolved significantly, incorporating both conventional and machine learning-based methodologies.[14]

Traditional (conventional) methods for image segmentation. Conventional techniques rely on predefined rules and heuristics to separate foreground objects from the background. Currently, the field of conventional cell segmentation algorithms encompasses a range of methodologies. These include techniques such as thresholding,[15] watershed segmentation,[16-18] and active contour segmentation,[19-21] which are known to have an over-segmentation bias when the images are contaminated with noise. However, the advent of machine learning has revolutionised cell segmentation, enabling the development of sophisticated algorithms capable of learning and adapting to complex image patterns.[22]

Traditional methods are often combined and integrated into automated workflows,[23,24] not only to accelerate the segmentation procedure but also to complement the pre- and post-processing stages of neural networks.[25] They can also be effectively used in data annotation to streamline the generation of training data, which is essential for training neural networks.[26,27]

Deep learning–based segmentation and object detection. The current trend in microscopic cell image segmentation is based on machine learning and deep learning methods, which have gained dominance in cell segmentation and other fields in the last decade. The increasing complexity of microscopy images and the limitations of conventional approaches in cell segmentation have prompted substantial advancements in these techniques, which can be broadly categorised into supervised and unsupervised methodologies.[28,29] Supervised methods involve the development of mathematical functions or models through training data to enable the accurate prediction of new samples. Conversely, unsupervised machine learning algorithms operate independently of pre-assigned labels or scores in the training data. In particular, clustering methods have emerged as the predominant unsupervised techniques for addressing these challenges.[30-33]

In general, there are two main approaches to image segmentation, namely semantic and instance segmentation. In semantic segmentation, the objective is to assign a class label to each pixel in an image without explicitly distinguishing between individual instances of the same class. Conversely, the second category, instance segmentation, aims to differentiate between individual instances belonging to the same class. However, unlike semantic segmentation, instance segmentation does not provide predictions for all pixels in an image and focuses solely on predicting one particular type of object, namely, the foreground, while disregarding other objects present in the background.[34,35]

Cell segmentation requires instance segmentation because in microscopic images, individual cells touch, and using semantic segmentation alone results in the touching cells acting as a single object. Thus, a major challenge in cell segmentation is to design reliable feature representations that can recognise cell boundaries with high accuracy. To address this, there are several possible approaches. The first involves classifying each pixel

into a semantic class and grouping pixels of the same class into individual instances,[34] which can be achieved, for example, by applying connected component analysis.[36] This approach requires the most accurate semantic segmentation possible to achieve the precise categorisation of pixels in the cell-background transition region. Another approach involves localising each cell via rough shape representation using axis-aligned bounding boxes[37] or convex polygons for each pixel.[38] This approach is categorised as object detection rather than segmentation, in which object detection algorithms offer the capability to generate bounding boxes around individual cells. The process of drawing bounding boxes involves the labelling of two points in spatial coordinates and can exhibit greater efficiency compared with pixel-level annotation. In addition, the use of bounding boxes facilitates straightforward differentiation of overlapping objects.[39] In scenarios where pixel-level segmentation is necessary, object detection can be followed by post-processing techniques that establish accurate boundaries for each object.[40]

Consequently, a transition from pure semantic segmentation to instance-aware semantic segmentation is imperative. A viable approach involves incorporating a narrow artificial background ridge, measuring one pixel in width between the touching elements within the actual segmentation mask.[41]

For comparison, we selected different types of architectures developed for image segmentation and object detection. These networks differ in their structures on the desired results. The requirements of an optimal network vary with respect to the input data and the purpose of use. We defined the basic requirements as segmentation accuracy, ie, defining the boundary of an object as accurately as possible and distinguishing it from the background.

U-net and U-net++. U-net is a neural network architecture known for its encoder-decoder network structure, which enables end-to-end solving of semantic segmentation tasks, even when the training data set is limited. The U-net network, proposed by Ronneberger et al,[42] builds on the foundation established by Shelhamer et al[43] and their use of fully convolutional networks.

An important common feature of encoder-decoder networks used in image segmentation tasks is the integration of skip connections. In the U-net architecture, these skip connections are achieved by concatenating up-sampled feature maps and incorporating convolutions and non-linearities between each up-sampling step. These skip connections improve gradient flow and enable the transfer of information between the down-sampling and the up-sampling paths by connecting each pair of encoder-decoder layers, thus proving its effectiveness in recovering full spatial resolution at the network output and making fully convolutional methods suitable for accurate semantic segmentation.[44]

Since its introduction in 2015, the U-net network has experienced a rapid rise in the field of medical and biological image data evaluation; naturally, various versions of it are emerging, namely 3D U-net,[45] Residual U-net,[46] Attention U-net,[47] Inception U-net,[48] Dense U-net,[49] R2 U-net,[50] U-net++,[51] *SE* U-net,[52] and many other networks that make use of an ensemble of U-nets.[53-56]

In our study, we employed the U-net++ architecture, which incorporates nested and dense skip connections. The fundamental premise underlying this architecture is that the model can capture intricate details of foreground objects by progressively enhancing high-resolution feature maps from the encoder network before merging them with semantically enriched feature maps from the decoder network. U-net++ comprises an encoder and decoder linked by a sequence of nested dense convolutional blocks. The primary purpose behind U-net++ is to close the semantic gap between the feature maps of the encoder and decoder prior to fusion.[51]

The Feature Pyramid Network (FPN). Feature pyramids are a fundamental element within recognition systems designed to detect objects across various scales. The concept of feature pyramids was originally introduced by Lin et al,[57] with the aim of effectively leveraging multiscale features for tasks involving object detection and recognition. FPN is constructed on top of a backbone network, typically a pre-trained convolutional neural network (CNN) like ResNet or VGGNet, and is responsible for extracting features from the input image. Feature pyramid networks enable these CNNs to capture features at distinct levels of abstraction and spatial resolution. The uppermost levels of the pyramid encompass high-level features that are semantically rich but possess a lower spatial resolution compared to the lower levels of the pyramid. On the contrary, the lower levels of the pyramid exhibit high spatial resolution but contain less semantic information in comparison with the higher levels. By integrating features from multiple scales, the FPN facilitates robust and precise object detection and recognition.

The Multiscale Attention Net (MA-Net). Attention is an important mechanism that can be used for a variety of deep learning models in many different domains.[58] Attention mechanisms have been introduced into computer vision to simulate the human ability to detect salient regions in complex scenes and have been successfully applied in many areas, including semantic segmentation.[59-62] MA-Net was first proposed by Fan et al[63] for the segmentation of the liver and tumours. Apply self-attention modules in the bottleneck of the U-net architecture. The first module, the Position-wise Attention Block, is used to obtain spatial dependencies between pixels in feature maps, while the second module, the Multiscale Fusion Attention Block, considers the channel dependencies between any feature maps.

The Pyramid Attention Network (PAN). The PAN is an architectural design that aims to enhance the feature representation capabilities of CNNs by incorporating attention mechanisms. It was inspired by the attention mechanism to provide precise pixel-level attention for high-level features extracted

from CNNs and was designed to capture both local and global contextual information in an image. This is achieved by using two modules: The Feature Pyramid Attention module provides pixel-level attention information and increases the receptive field by utilising a pyramid structure, and the Global Attention Up-sample module exploits high-level feature maps to guide low-level features recovering pixel localisation. By effectively integrating attention mechanisms at different scales, the PAN architecture enables the network to capture meaningful context and enhance its power for image classification, object detection, and semantic segmentation.[64]

LinkNet. LinkNet was initially designed to address the requirements of real-time semantic segmentation of visual scenes. It follows an architectural paradigm to that of U-net and other segmentation networks, featuring an encoder on the left and a decoder on the right. The achievement of temporal efficiency in LinkNet is due to the specific interconnections established between the encoder and the decoder. This entails selectively directing the input of each encoder layer to its corresponding decoder output, thereby facilitating the segmentation process. Because of this strategy, the decoder can function with reduced parametrisation in each layer, contributing to improved temporal efficiency.[65] LinkNet was employed in Araújo et al[66] for the segmentation of melanoma skin cancer in combination with transfer learning and fine tuning. Although LinkNet makes efficient use of parameters and a number of operations in real-time applications, it lags behind U-net in accuracy.

DeepLabV3 and DeepLabV3+. DeepLabV3 is a state-of-the-art model in the domain of semantic segmentation; it utilises atrous convolution and post-processing via fully connected conditional random fields.[67] Using atrous convolution, Deep-LabV3 achieves an expansion of the receptive field of filters without increasing the number of parameters, facilitating the acquisition of local and global contextual information. This is accomplished through the introduction of voids within convolutional kernels, consequently regulating the spacing between the values utilised for convolutional computation. Dilated convolutions support exponential expansion of the receptive field without loss of resolution or coverage and provide explicit control over the resolution of features generated by CNNs.[68]

DeepLabV3+ expands on the foundations of DeepLabV3 by incorporating an encoder-decoder architecture. This architecture employs DeepLabV3 as an encoder module and a simple yet effective decoder module. The encoder module handles multiscale contextual information by employing atrous convolutions across various scales. In contrast, the decoder module further enhances the precision of segmentation outcomes in relation to object boundaries.[69]

Methods

For our experiments, we selected image sequences featuring CLL cells isolated from a single patient, which were captured using brightfield time-lapse microscopy. A detailed description
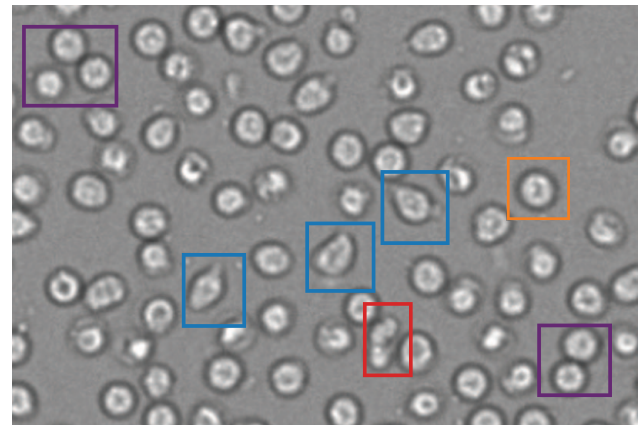

Figure 1. Brief description of the input image data. The appearance of typical chronic lymphocytic leukaemia cells: (orange) rounded cell with a blurry dark edge, brighter inner ring part, and darker middle; (blue) nontypical: non-rounded shape (red).

of the conducted experiments and the resulting image data are presented below.

Sample preparation and image acquisition

Peripheral blood mononuclear cells were isolated by density gradient centrifugation from a patient with a very high percentage ($>95\%$) of lymphocytes. CLL cells were identified and then sorted by gating for CD19$^+$, with $>99\%$ purity. Small (s-CLL) and large (l-CLL) cells were sorted by gating by the size of the smallest and largest portions of CLL cells in the CLL population (20% each). From 2.5 million CLL cells, approximately 350 000 s-CLL and 250 000 l-CLL cells were sorted and cultured in complete foetal bovine serum, 2 mM L-glutamine, supplemented with 10% heat-inactivated foetal bovine serum, 2 mM L-glutamine, 100 U mL^{-1} penicillin and 100 µg mL^{-1} streptomycin. Cells s-CLL and l-CLL were treated separately with CpG (3 µg mL^{-1}), cultured in a climate-controlled chamber (37 °C and 5% CO_2) for 24h and then placed on a heated microscopy stage (37 °C and 5% CO_2). Cell polarisation and motility were monitored by time-lapse video microscopy of cells for 60 min and recorded using Gen5 software and a BioTec Cytation 5 reader at a final magnification of $200\times$ (Figure 1). The camera used was a Blackfly BFLY-U3-23S6M, with an objective Olympus lens (size: 20; numerical aperture: 0.45; objective PSF Sigma: 0.806) and an image size of 1224 × 904 pixels.

Comparison of state-of-the-art neural network architectures for semantic segmentation of CLL cells

In this study, our primary aim was to detect CLL cells (rounded, touching low-contrast objects) from brightfield time-lapse microscopy images. To accomplish this, we conducted a comprehensive analysis utilising a variety of neural network

architectures. Specifically, the neural networks evaluated in our study included the ones described in Section 3; namely U-net, U-net++, the Multiscale attention Net (MA-Net), LinkNet, the FPN, DeepLabV3, DeepLabV3+ and the Pyramid Attention Network (PAN). The purpose of utilising these architectures was to facilitate the accurate identification and localisation of CLL cells within the acquired BF images without distinguishing individual cells in the case of their contact.

The training process involved training each of these networks for 1000 epochs using PyTorch and PyTorch Lightning libraries. Using segmentation models,[70] we incorporated various backbone encoders into neural networks, namely InceptionResNetV2,[71] ResNeSt-269,[72] ResNet-152,[46] ResNeXt-101_32x32d,[73] and VGG19.[74]

Compared variables

Loss function. The Dice coefficient, or the Dice-Sørensen coefficient, is a common metric for pixel segmentation that can also be modified to act as a loss function:

$$DCS = \frac{2\left|X \cap Y\right|}{\left|X\right| + \left|Y\right|}.$$

The Dice-Sørensen coefficient is calculated by taking twice the count of elements shared by both sets and dividing it by the sum of the elements in each set. In general, the purpose of loss functions is to compute a metric that the model should try to minimise throughout the training process.

Intersection over union metric. In image segmentation, intersection over union (IoU) is a primary metric for evaluating model accuracy. The IoU is the intersection ratio of the area between the ground truth (GT) and the segmentation mask (TP = true positive) to the union of TP, the predicted area outside the ground truth (FP = false positive), and the number of pixels in the GT area not predicted by the model (FN = false negative):

$$IoU = \frac{TP}{TP + FP + FN}.$$

We calculate IoU in two ways: the per-image IoU and the data set IoU. The per-image IoU means that we first calculate the IoU score for each image and then compute the mean of these scores. IoU of the data set means that we aggregate the intersection and union over the whole data set and then calculate the IoU score. The difference between IoU per data set and the IoU per image can be observed in data sets with 'empty' images (ie, images without a target class). Empty images have a large influence on the IoU per image and much less influence on the IoU of the data set.

Training. The graphics processing unit hardware used for this test was an NVIDIA V100 with 32 GB of VRAM.[75] The original CLL images had a resolution of 1224 × 904 pixels with a colour depth of 16 bits; for labelling and training, we chose 224 × 224 pixel cutouts. Each network was run for 1000 epochs, and there was no restriction on the time it took to train.

Proposed workflow for instance segmentation of CLL cells and other lymphoblast–like cells

The preceding investigation aimed to identify the optimal neural network to effectively detect CLL cells in brightfield microscopy. The process involved semantic segmentation, where each pixel within the images was assigned specific values based on the probability of being part of the target object or the background. To achieve instance segmentation, which is particularly advantageous in this context (as it enables differentiation between touching and overlapping objects), requires another step in the post-processing stage. Since the methods for achieving instance-based cell segmentation depend on the annotation of the data itself, an overview of three types of workflows is given below (Figure 2).

First step–data annotation. There are several procedures available to separate individual segmented cells, some of which require specific preprocessing steps. These steps influence cell segmentation during the annotation of image data, the initial phase of image segmentation. In this study, we used three types of data annotation, each tailored to the subsequent cell segmentation procedure. Basic annotation involves creating a binary mask, distinguishing between two classes: black representing the background and white representing cells. By slightly modifying this mask (adding a 1-pixel border between cells and then algorithmically outlining each cell), a mask of three classes was created (green – cell, red – background, blue – boundary). The output of a neural network trained on such data comprises predictions for these three classes, including the cell-background boundary. The last type is the annotation of each cell separately, that is, each cell corresponds to a different pixel value. To create these masks, it is possible to use freely available tools and plugins, eg, Labkit in ImageJ,[76] AnnotatorJ,[77] or QuPath.[78]

Instance segmentation methods. In this study, we used three types of approaches to perform instance-based cell image segmentation. The first is the application of traditional image analysis algorithms to the segmented image through 2-class segmentation (Section 4 Materials and methods), the second is the application of these algorithms to the data obtained through 3-class segmentation, and the third is the use of the StarDist detector,[38] which uses a lightweight neural network based on U-net and works similarly to most detection algorithms; however, instead of axis-aligned bounding boxes, it predicts a star-convex polygon for each pixel to represent the shapes. StarDist is widely used to detect convex-shaped objects, such as cell nuclei. It is available as a library written in Python and as a tool in Fiji/ImageJ,[79] where it is also part of the TrackMate[80] cell tracking tool.
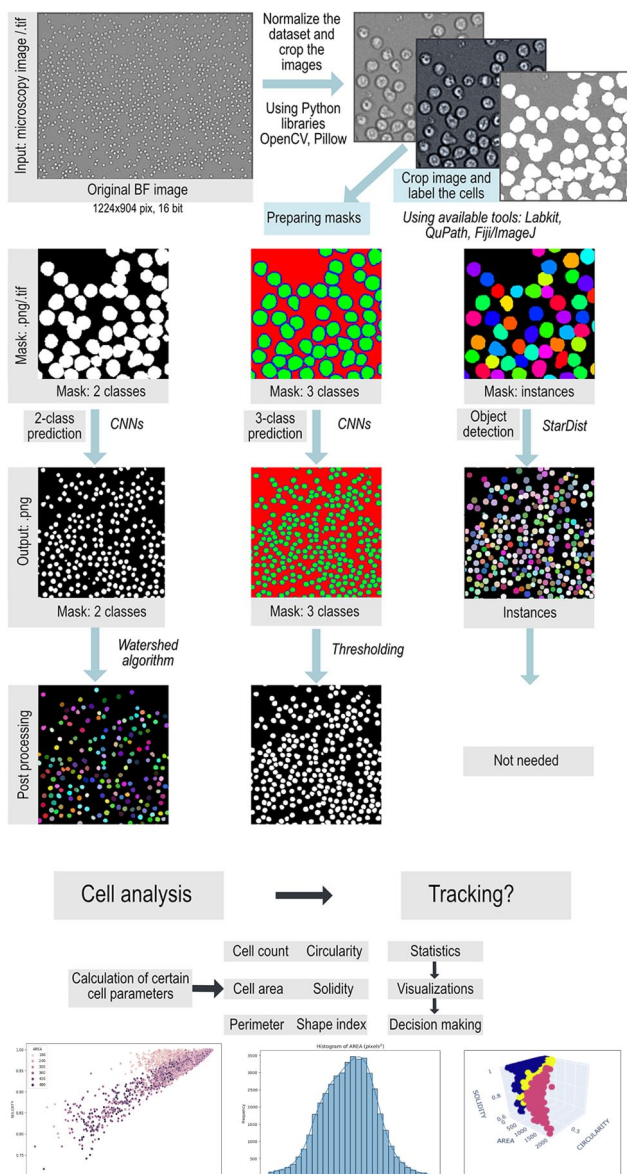
Figure 2. The workflow was tailored to processing and evaluating brightfield microscopy images of CLL cells and cells with similar morphology. The input comprises microscopy images or image sequences, the dimensions of which require adjustment depending on the semantic segmentation algorithm employed. Image cutouts with dimensions of 224 × 224 pixels and 256 × 256 pixels (StarDist) were utilised for neural network training. Preprocessing entails image normalisation and augmentation. Masks are prepared based on the segmentation type: 2-class, 3-class, instance. Post-processing involves employing traditional image processing methods such as thresholding and the watershed algorithm.

A typical algorithm used in post-processing to separate objects is the watershed algorithm, which interprets pixel values as local topography or elevation, facilitating the demarcation of object boundaries within images. Although the output of a 2-class segmentation framework can be post-processed using this algorithm, more precise results were obtained through 3-class segmentation (Figure 3).

The resulting segmented images were then processed using Fiji/ImageJ software. Since the data set contained time series

content, ie, images taken through a device designed for time-lapse microscopy, the Fiji/ImageJ plugin TrackMate with integrated threshold object detector was used to obtain cellular characteristics (area, circularity, solidity, perimeter, radius, and shape index). The input was a sequence of 47 images in .png format of 1280 × 928 pixels, converted to 8-bit depth in binary format, and the output is a .csv table with the positions of detected cells and their characteristic. The post-processing pipeline of U-net++ 2-class and U-net++ 3-class segmentation outputs is depicted in Figure 4.

Results

The following section summarises the results of both parts of the study, the first is the selection of the most appropriate architecture for semantic segmentation, followed by a comparison of approaches to creating instance-based cell segmentation and a comparison with the widely used StarDist tool.

Comparison of different neural network architectures for semantic segmentation of CLL cells

Although some of the network architectures converged to the solutions faster than others, as seen in Figure 5, where the FPN network has a lower loss function value faster than other networks, and the same can be observed in Figure 6.

Finally, U-net++ had the best results in terms of the loss function and all IoU metrics. From the loss graphs, the IoU metric chart (Figure 7) and the results in Table 1, U-net++ provides the best results, followed by U-net and MA-Net. U-net, U-net++, and MA-Net were primarily designed for medical image segmentation. It was found that the attention mechanisms applied in the MA-Net do not improve the final score in the segmentation of CLL cells. However, MA-Net outperformed U-net in the Test Dataset-IoU metric with the ResNet-152 backbone, where it provided the second-best result. It also has the advantage of having shorter training times than U-net++, but this metric does not play a significant role in the case of this application. The optimal combination was U-net++ architecture with ResNeSt-269 backbone, with an IoU score of 0.8902.

An example of the input and neural network output can be seen in Figure 8, where outputs from both DeepLab architectures have blurry outputs, and LinkNet has some undesired artefacts, which can be corrected in post-processing. A single image slice of several cells was selected and then compared across all outputs. The highlighted values represent the best result in each measured category.

Comparison of different approaches for instance segmentation of CLL cells

For the characteristics examined (area, circularity, solidity, perimeter, radius, and shape index), the difference in mean value using different CLL cell segmentation approaches appeared to be statistically significant (Mann–Whitney U test, $P < .0001$). Figure 9 shows the distribution of the values of the
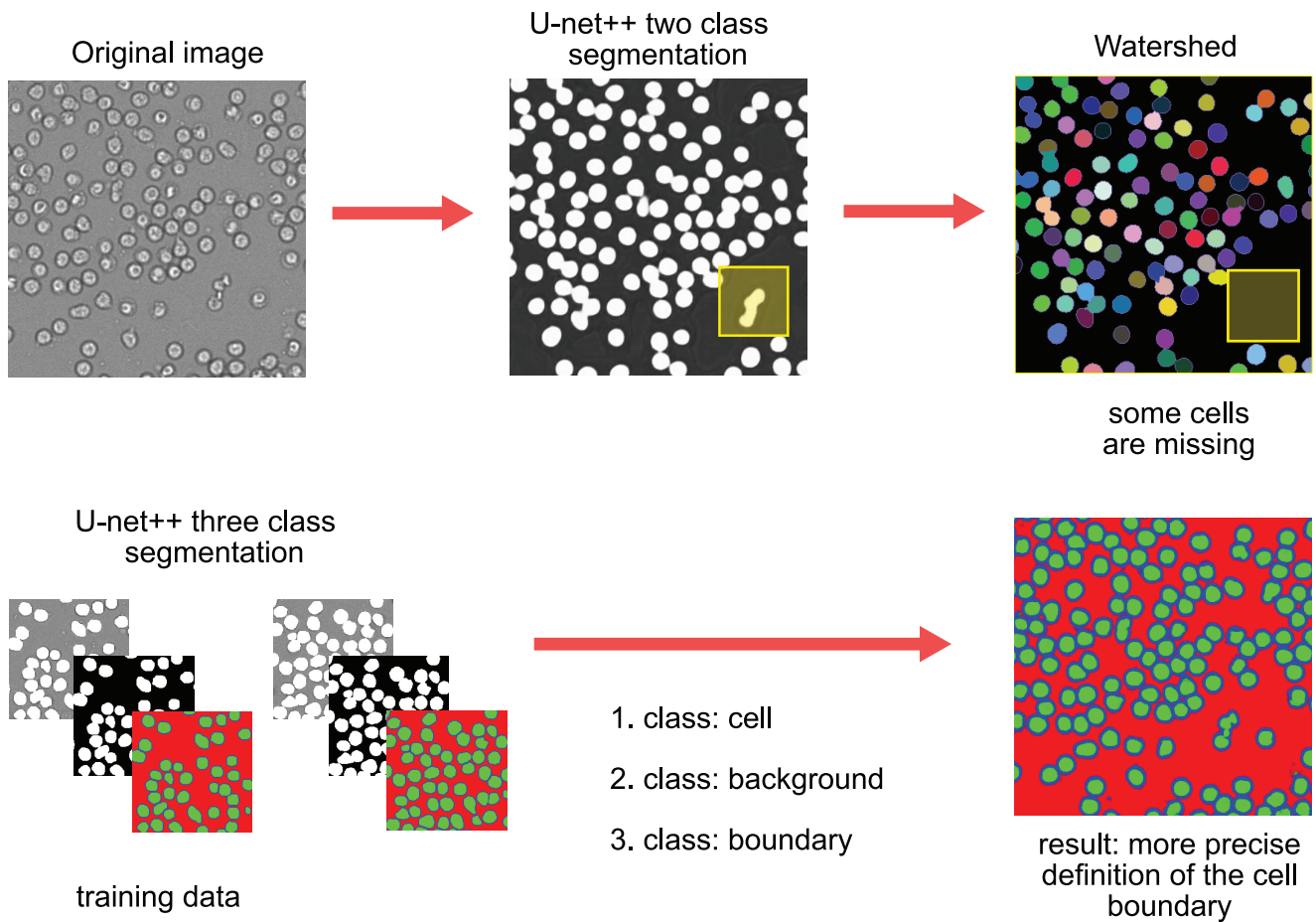
Figure 3. Proposed workflow containing post-processing steps leading to instance segmentation of CLL cells with utilising U-net++. The watershed algorithm was used on the output of the U-net++ model. The result revealed that certain cell boundaries were not detected with the necessary precision, potentially impacting subsequent analyses. Consequently, we decided to adopt an alternative approach: the three-class segmentation method. In this scenario, we found that the boundaries of touching cells were more effectively resolved.
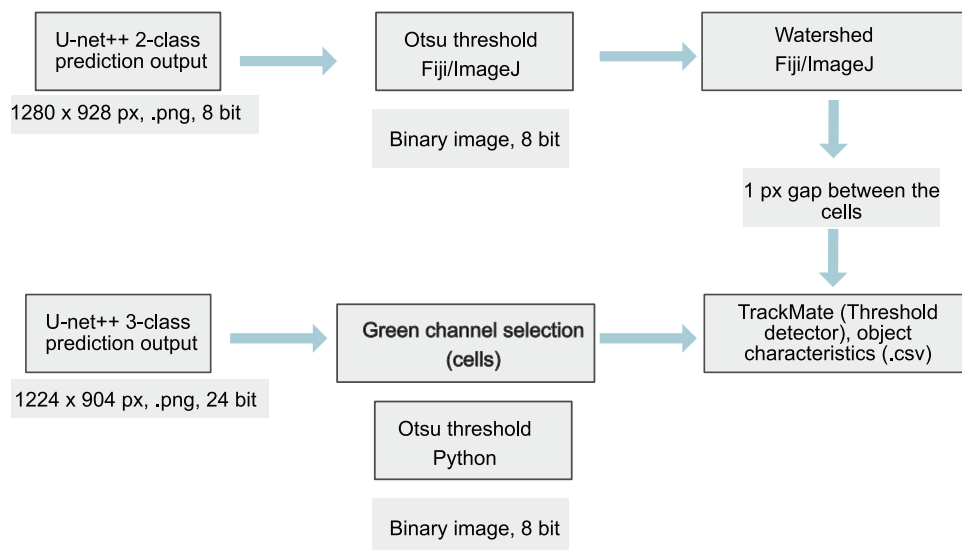

Figure 4. Post-processing pipeline. Difference in procedure when using 2-class and 3-class prediction. In the case of 3-class prediction, the output is an RGB image in which the cell boundaries are already detected during prediction, which facilitates post-processing because it is not necessary to use cell separation algorithms, in this case the watershed algorithm, which creates a 1-pixel gap between cells.
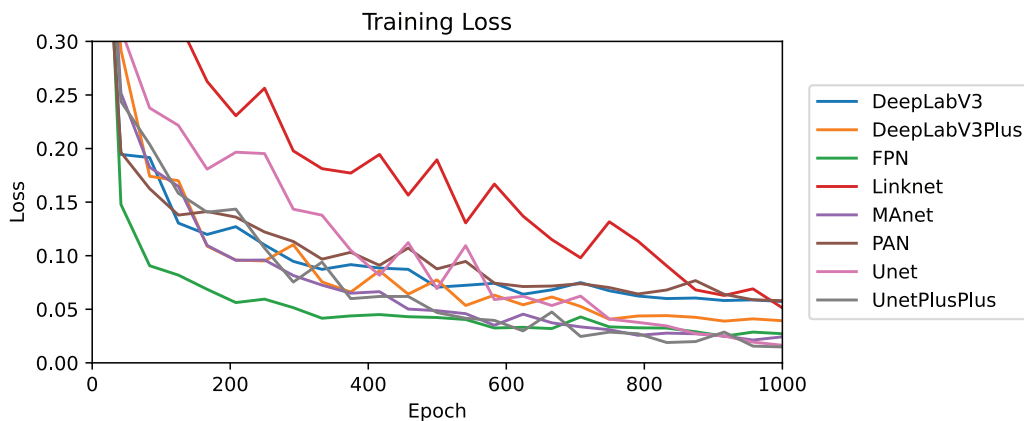
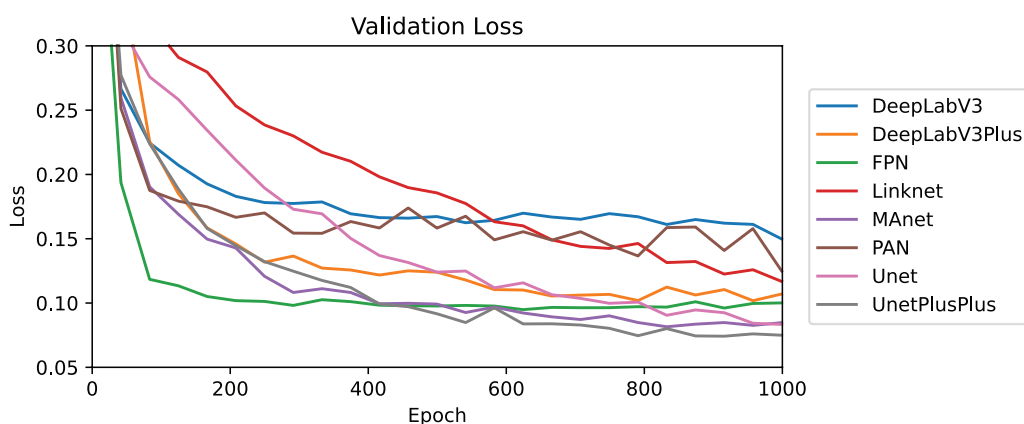**Figure 5.** Training loss per epoch for each network.
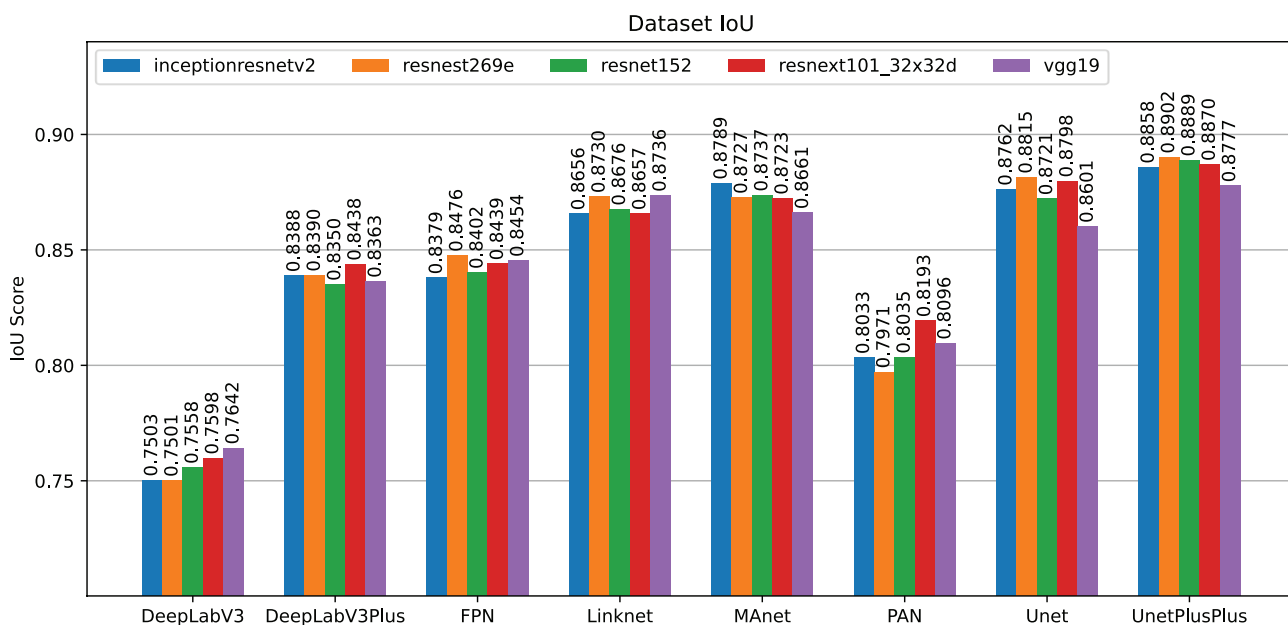

Figure 6. Validation loss for each epoch.

Figure 7. Best dataset-IoU score for architecture and backbone.

Table 1. The results for the backbone resnet152.

NETWORK	PARAMETERS	TRAINING TIME	LOSS	TEST DATA SET
	(MILLIONS)	(MM: SS)		IOU (%)
U-net	67.2	21:41	0.0165	87.21
U-net++	83.6	30:10	**0.0149**	**88.88**
MA-Net	182	25:31	0.0242	87.37
LinkNet	65.8	21:02	0.0517	86.76
FPN	60.8	**15:39**	0.0272	84.01
DeepLabV3	74.3	18:49	0.0572	75.57
DeepLabV3+	61.3	48:33	0.0394	83.50
PAN	**58.9**	16:46	0.0581	80.35

different cell characteristics using three different approaches, 2-class U-net++, 3-class U-net++, StarDist (custom model). From these graphs, the differences are mainly in the distribution of outliers. Another finding is the characteristic 'CIRCULARITY', which acquires a small variance when using the StarDist detector, while its variance is larger when using other post-processing procedures.

Subsequently, the effectiveness of each approach to fit the manually labelled data was compared (GT). Individual approaches were found to differ primarily in the nature of cell segmentation, not in the overall IoU score (Figure 10). Individual approaches were also compared with each other.

Discussion

Deep learning has exerted substantial influence across multiple domains in the realm of image processing, including computer vision, image classification, semantic segmentation, instance segmentation, and panoptic segmentation. In the domain of medical imaging, as well as cell segmentation and tracking, CNN and deep learning models have demonstrated their ability to achieve state-of-the-art accuracy, primarily by effectively harnessing both local and global image features.[71]

Our methodology for segmenting time-lapse microscopy images of CLL cells involved a comprehensive evaluation of state-of-the-art neural network architectures, resulting in the selection of the optimal model for segmenting CLL cells. This chosen methodology also outlines procedures applicable to segmenting time-lapse images of other cell types in brightfield microscopy, with the primary focus being on accurately defining cell boundaries, a prerequisite for successfully determining their characteristics. Furthermore, we investigated different network backbones to identify the optimal configuration, ultimately finding that the combination of the U-net++ network with the ResNeSt-269 backbone was the most effective for our specific application. The choice of the appropriate method for cell segmentation depends on the quality and quantity of the input data but also on the experience of the evaluator. In this case, the cells have a relatively regular, circular to ellipsoidal shape but vary in size and contain irregularities. However, irregularities may have a significant effect on the evaluation of the experiments performed. For these reasons, accurate identification of object shapes in microscopic images is essential. The workflows mentioned demonstrate the possibilities to segment this type of image data, ie, to separate individual cells as accurately as possible, which is also important for subsequent monitoring of cell motility and migration.

The effectiveness of the proposed methods is based on the following points.

(a) By comparing the combinations of backbone and advanced deep learning architectures for image semantic segmentation, the most efficient was selected and used in the two proposed methods.

(b) Then 3-class segmentation was applied to the proposed method, where the pixels of the training data were segmented into object (cell), cell boundary and background classes. The image data labelled in this way was then used for training.

(c) 3-class segmentation, in contrast to 2-class segmentation, shows a more precise separation of individual cells, ie, creating instances without the need to apply an additional post-processing algorithm, for example, watershed, which is not always efficient.

Conclusion

The result of the whole study is a comparison of methods that can assign characteristics to image data/cells, which can then be used in the study of cell migration. The main objective of the study is to show differences in cell segmentation using different approaches and how they can affect the resulting measurements. Our next objective is to develop an optimal
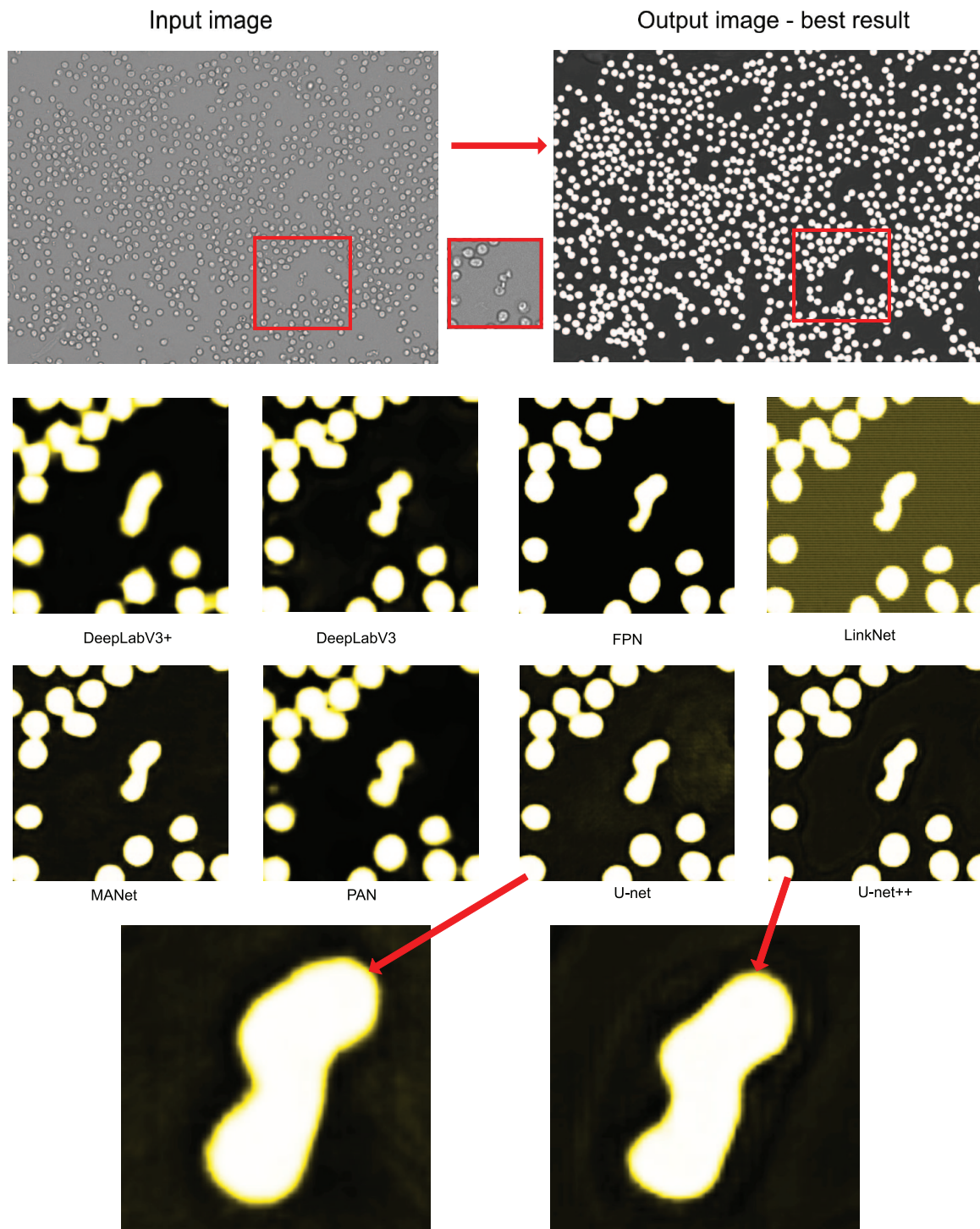
Figure 8. Comparison of the detail of the image in the output from the neural networks used to detect CLL cells in the brightfield time-lapse microscopy image. In the magnified images, the different segmentations of this detail can be observed, as well as the cell boundaries; yellow was added to make the details more visible. It is desirable that the neural network detects pixels belonging to cells with the highest possible probability, that there is the most obvious transition between the cell and the background.
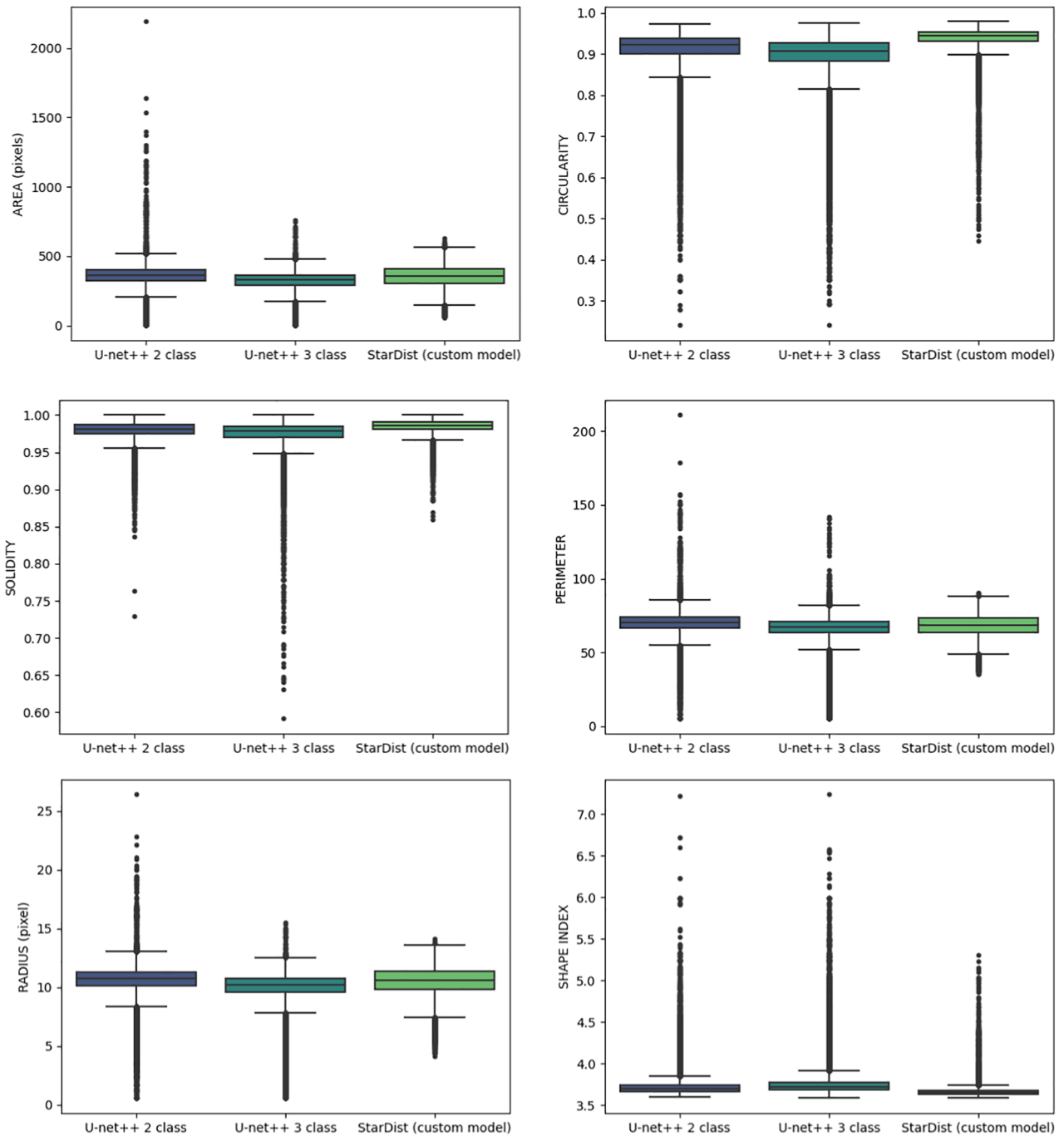
Figure 9. Boxplots showing the variability of a given cell parameter (area, circularity, solidity, perimeter, radius, shape index) when using a specific method, namely U-net++ 2-class segmentation, U-net++ 3-class segmentation and StarDist (custom model).

algorithm for tracking CLL cells. Subsequent research will investigate the potential of CLL cellular characteristics to predict disease progression, which could be beneficial for precision medicine. The difference in the mean value of some cellular parameters using the different approaches to segmentation was found to be statistically significant; however, the different approaches differed in the nature of cell segmentation, not in the overall IoU score.
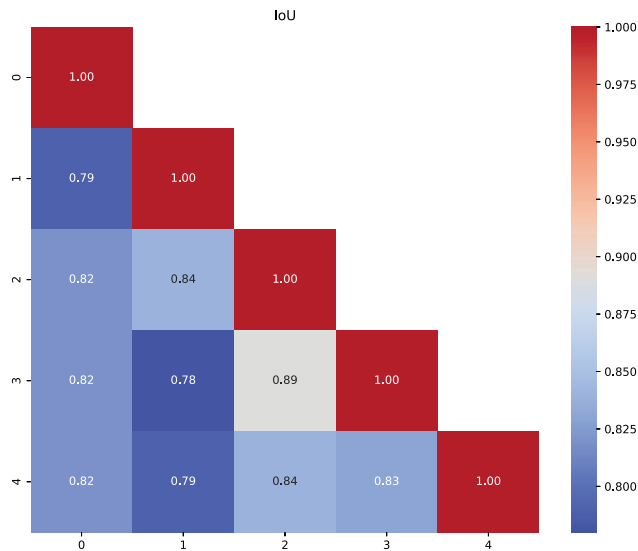
Figure 10. IoU visualisation comparing predictions against manually labelled cells (ground truth). When compared, we found that overall the algorithms demonstrate equal agreement with GT, but with the comparison it can be seen that the different approaches prefer different morphological features of the cells, ie, StarDist due to its primary purpose of detecting round (convex) shapes preferred round cells, while 3-class segmentation better defined irregular shapes. The difference between 2-class segmentation with watershed and 3-class segmentation is largely due to the fact that watershed separates cells from each other using a 1-pixel gap, which however does not always reflect the actual cell shape, unlike 3-class segmentation, where the gaps are already present in the training data.

Acknowledgements
The authors thank Arootin Gharibian for his contribution to to the section Methods.

Author Contributions
Markéta V and VD conceived of the presented idea. Markéta V created a framework and wrote the manuscript. VD and Michal V designed the experiment and perform the computations. PG and EK reviewed and approved the manuscript. All the authors are accountable for the integrity of this work. All the authors read and approved the final manuscript.

Consent
All patients involved in this study provided written informed consent for the use of biological material for research purposes, in accordance with the Declaration of Helsinki.

Ethics
The Ethics Committee of the University Hospital and Palacký University Olomouc approved the study.

Data Availability Statement
The data sets analysed during the current study are available from the corresponding author on reasonable request.

ORCID iD
Markéta Vašinková [iD] https://orcid.org/0009-0003-9321-0751

REFERENCES
1. Scarfò L, Ferreri AJ, Ghia P. Chronic lymphocytic leukaemia. *Crit Rev Oncol Hematol*. 2016;104:169-182. doi:10.1016/j.critrevonc.2016.06.003
2. Bosch F, Dalla-Favera R. Chronic lymphocytic leukaemia: from genetics to treatment. *Nat Rev Clin Oncol*. 2019;16:684-701. doi:10.1038/s41571-019-0239-8
3. Manukyan G, Mikulkova Z, Turcsanyi P, et al. Towards a better characterisation of leukemic cells in chronic lymphocytic leukaemia: cell-size heterogeneity reflects their activation status and migratory abilities. *Cancers (Basel)*. 2021;13:4922. doi:10.3390/cancers13194922
4. Oscier D, Else M, Matutes E, Morilla R, Strefford JC, Catovsky D. The morphology of CLL revisited: the clinical significance of prolymphocytes and correlations with prognostic/molecular markers in the LRF CLL4 trial. *Br J Haematol*. 2016;174:767-775. doi:10.1111/bjh.14132
5. Davids MS, Burger JA. Cell trafficking in chronic lymphocytic leukemia. *Open J Hematol*. 2012;3:-3. doi:10.13055/ojhmt_3_s1_03.120221
6. Ashwini A, Sriram SR, Sheela JJJ. Detection of chronic lymphocytic leukemia using Deep Neural Eagle Perch Fuzzy Segmentation – a novel comparative approach. *Biomed Signal Process Control*. 2024;90:105905. doi:10.1016/j.bspc.2023.105905
7. Chatap N, Shibu S. Analysis of blood samples for counting leukemia cells using Support Vector Machine and Nearest Neighbor. *IOSR J Comp Eng*. 2014;16:79-87.
8. Acharya V, Kumar P. Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms. *Med Biol Eng Comput*. 2019;57:1783-1811. doi:10.1007/s11517-019-01984-1
9. Gu W, Sun K. AYOLOv5: improved YOLOv5 based on attention mechanism for blood cell detection. *Biomed Signal Process Control*. 2024;88:105034. doi:10.1016/j.bspc.2023.105034
10. Fırat H. Classification of microscopic peripheral blood cell images using multi-branch lightweight CNN-based model. *Neural Comput Appl*. 2024;36:1599-1620. doi:10.1007/s00521-023-09158-9
11. Liang M, Zhong J, Shannon CS, Agrawal R, Ai Y. Intelligent image-based deformability assessment of red blood cells via dynamic shape classification. *Sens Actuators B Chem*. 2024;401:135056. doi:10.1016/j.snb.2023.135056
12. Ali N, Liu X, Wang W, et al. Blood cell characterization based on deep learning and diffraction phase microscopy. *Opt Commun*. 2024;561:130522. doi:10.1016/j.optcom.2024.130522
13. Meijering E, Dzyubachyk O, Smal I. Methods for cell and particle tracking. *Methods Enzymol*. 2012;504:183-200. doi:10.1016/B978-0-12-391857-4.00009-4
14. Greenwald NF, Miller G, Moen E, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat Biotechnol*. 2022;40:555-565. doi:10.1038/s41587-021-01094-0
15. Zhang J, Hu J. Image segmentation based on 2D Otsu method with histogram analysis. Paper presented at: 2008 International Conference on Computer Science and Software Engineering, 2008. doi:10.1109/CSSE.2008.206
16. Malpica N, de Solórzano CO, Vaquero JJ, et al. Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry*. 1997;28:289-297. doi:10.1002/(sici)1097-0320(19970801)28
17. Zhou X, Li F, Yan J, Wong ST. A novel cell segmentation method and cell phase identification using Markov model. *IEEE Trans Inf Technol Biomed*. 2009;13:152-157. doi:10.1109/TITB.2008.2007098
18. Buggenthin F, Marr C, Schwarzfischer M, et al. An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinformatics*. 2013;14:297. doi:10.1186/1471-2105-14-297
19. Molnar C, Jermyn IH, Kato Z, et al. Accurate morphology preserving segmentation of overlapping cells based on active contours. *Sci Rep*. 2016;6:32412. doi:10.1038/srep32412
20. Braiki M, Benzinou A, Nasreddine K, Mouelhi A, Labidi S, Hymery N. Human dendritic cells segmentation based on K-means and active contour. Paper presented at: Image and Signal Processing: 8th International Conference, ICISP 2018, July 2-4, 2018; Cherbourg. Springer-Verlag. doi:10.1007/978-3-319-94211-7_3
21. Niaz A, Iqbal E, Akram F, Kim J, Choi KN. Self-initialized active contours for microscopic cell image segmentation. *Sci Rep*. 2022;12:14947. doi:10.1038/s41598-022-18708-5

22. Maška M, Ulman V, Delgado-Rodriguez P, et al. The cell tracking challenge: 10 years of objective benchmarking. *Nat Methods.* 2023;20:1010-1020. doi:10.1038/s41592-023-01879-y

23. Voigt SP, Ravikumar K, Basu B, et al. Automated image processing workflow for morphological analysis of fluorescence microscopy cell images. *JOM.* 2021;73:2356-2365. doi:10.1007/s11837-021-04707-w

24. Salvi M, Morbiducci U, Amadeo F, et al. Automated segmentation of fluorescence microscopy images for 3D cell detection in human-derived cardiospheres. *Sci Rep.* 2019;9:6644. doi:10.1038/s41598-019-43137-2

25. Kowal M, Żejmo M, Skobel M, Korbicz J, Monczak R. Cell nuclei segmentation in cytological images using convolutional neural network and seeded watershed algorithm. *J Digit Imaging.* 2020;33:231-242. doi:10.1007/s10278-019-00200-8

26. Englbrecht F, Ruider IE, Bausch AR. Automatic image annotation for fluorescent cell nuclei segmentation. *PLoS ONE.* 2021;16:e0250093. doi:10.1371/journal.pone.0250093

27. Gudla PR, Zaki G, Shachar S, Misteli T, Pegoraro G. Deep learning based segmentation of nuclei from fluorescence microscopy images. *Microsc Microanal.* 2019;25:1376-1377. doi:10.1017/S143192761900761X

28. de Bruijne M. Machine learning approaches in medical image analysis: from detection to diagnosis. *Med Image Anal.* 2016;33:94-97. doi:10.1016/j.media.2016.06.032

29. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol.* 2019;19:64. doi:10.1186/s12874-019-0681-4

30. Bradbury L, Wan JW. A spectral k-means approach to bright-field cell image segmentation. *Annu Int Conf IEEE Eng Med Biol Soc.* 2010;2010:4748-4751. doi:10.1109/IEMBS.2010.5626380

31. Nasir AA, Mashor MY, Mohamed Z. Segmentation based approach for detection of malaria parasites using moving k-means clustering. Paper presented at: 2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences, 2012. doi:10.1109/IECBES.2012.6498073

32. Abbas N, Mohamad DB. Microscopic RGB color images enhancement for blood cells segmentation in YCbCr color space for k-means clustering. *J Theor Appl Inf Technol.* 2013;55:117-125.

33. Savkare SS, Narote SP. Blood cell segmentation from microscopic blood images. Paper presented at: 2015 International Conference on Information Processing (ICIP), 2015. doi:10.1109/INFOP.2015.7489435

34. Jung S, Heo H, Park S, Jung SU, Lee K. Benchmarking deep learning models for instance segmentation. *Appl Sci.* 2022;12:8856. doi:10.3390/app12178856

35. Gu W, Bai S, Kong L. A review on 2D instance segmentation based on deep neural networks. *Image and Vision Comp.* 2022;120:104401. doi:10.1016/j.imavis.2022.104401

36. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods.* 2021;18:100-106. doi:10.1038/s41592-020-01018-x

37. Mahani GK, Li R, Evangelou N, et al. Bounding box based weakly supervised deep convolutional neural network for medical image segmentation using an uncertainty guided and spatially constrained loss. Paper presented at: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), 2022. doi:10.1109/ISBI52829.2022.9761558

38. Schmidt U, Weigert M, Broaddus C, Myers G. Cell detection with star-convex polygons. Paper presented at: Medical Image Computing and Computer Assisted Intervention–MICCAI, 2018. doi:10.1007/978-3-030-00934-2_30

39. Chen J, Zhang B. Segmentation of overlapping cervical cells with mask region convolutional neural network. *Comput Math Methods Med.* 2021;2021:3890988. doi:10.1155/2021/3890988

40. Hung J, Goodman A, Ravel D, et al. Keras R-CNN: library for cell detection in biological images using deep neural networks. *BMC Bioinform.* 2020;21:300. doi:10.1186/s12859-020-03635-x

41. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry [published correction appears in *Nat Methods.* 2019;16(4):351]. *Nat Methods.* 2019;16:67-70. doi:10.1038/s41592-018-0261-2

42. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. Paper presented at: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015. doi:10.1007/978-3-319-24574-4_28

43. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2017;39:640-651. doi:10.1109/TPAMI.2016.2572683

44. Siddique N, Paheding S, Elkin CP, Devabhaktuni V. U-Net and its variants for medical image segmentation: a review of theory and applications. *IEEE Access.* 2021;9:82031-82057. doi:10.1109/ACCESS.2021.3086020

45. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-net: learning dense volumetric segmentation from sparse annotation. Paper presented at: Medical Image Computing and Computer Assisted Intervention–MICCAI 2016. Lecture Notes in Computer Science, 2016. doi:10.1007/978-3-319-46723-8_49

46. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Paper presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. doi:10.1109/CVPR.2016.90

47. Oktay O, Schlemper J, Folgoc LL, et al. Attention u-net: learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018.

48. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. Paper presented at: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. doi:10.1109/CVPR.2015.7298594

49. Cai S, Tian Y, Lui H, Zeng H, Wu Y, Chen G. Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network. *Quant Imaging Med Surg.* 2020;10:1275-1285. doi:10.21037/qims-19-1090

50. Alom MZ, Yakopcic C, Hasan M, Taha TM, Asari VK. Recurrent residual U-Net for medical image segmentation. *J Med Imaging (Bellingham).* 2019;6:014006. doi:10.1117/1.JMI.6.1.014006

51. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: a nested U-Net architecture for medical image segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2018).* 2018;11045:3-11. doi:10.1007/978-3-030-00889-5_1

52. Hu J, Shen L, Albanie S, Sun G, Wu E. Squeeze-and-excitation networks. *IEEE Trans Pattern Anal Mach Intell.* 2020;42:2011-2023. doi:10.1109/TPAMI.2019.2913372

53. Abd-Ellah MK, Khalaf AAM, Awad AI, Hamed HFA. TPUAR-Net: two parallel U-Net with asymmetric residual-based deep convolutional neural network for brain tumor segmentation. Paper presented at: Image Analysis and Recognition, 2019. doi:10.1007/978-3-030-27272-2_9

54. Wu Y, Shen H, Tan Y, Shi Y. Automatic liver tumor segmentation used the cascade multi-scale attention architecture method based on 3D U-Net. *Int J Comput Assist Radiol Surg.* 2022;17:1915-1922. doi:10.1007/s11548-022-02653-9

55. Li H, Li A, Wang M. A novel end-to-end brain tumor segmentation method using improved fully convolutional networks. *Comput Biol Med.* 2019;108:150-160. doi:10.1016/j.compbiomed.2019.03.014

56. Piantadosi G, Sansone M, Fusco R, Sansone C. Multi-planar 3D breast segmentation in MRI via deep convolutional neural networks. *Artif Intell Med.* 2020;103:101781. doi:10.1016/j.artmed.2019.101781

57. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. Paper presented at: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. doi:10.1109/CVPR.2017.106

58. Brauwers G, Frasincar F. A general survey on attention mechanisms in deep learning. *IEEE Trans Knowl Data Eng.* 2023;35:3279-3298. doi:10.1109/TKDE.2021.3126456

59. Li H, Luo H, Huan W, et al. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network. *Neural Comput Appl.* 2021;33:11589-11602. doi:10.1007/s00521-021-05856-4

60. Gao C, Ye H, Cao F, Wen C, Zhang Q, Zhang F. Multiscale fused network with additive channel–spatial attention for image segmentation. *Knowledge-Based Syst.* 2021;214:106754. doi:10.1016/j.knosys.2021.106754

61. Guo MH, Xu TX, Liu JJ, et al. Attention mechanisms in computer vision: a survey. *Comp Visual Media.* 2022;8:331-368. doi:10.1007/s41095-022-0271-y

62. Fan Z, Hu G, Sun X, Wang G, Dong J, Su C. Self-attention neural architecture search for semantic image segmentation. *Knowledge-Based Syst.* 2022;239:107968. doi:10.1016/j.knosys.2021.107968

63. Fan T, Wang G, Li Y, Wang H. MA-Net: a multi-scale attention network for liver and tumor segmentation. *IEEE Access.* 2020;8:179656-179665. doi:10.1109/ACCESS.2020.3025372

64. Li H, Xiong P, An J, Wang L. Pyramid attention network for semantic segmentation. *arXiv preprint arXiv:1805.10180*, 2018.

65. Chaurasia A, Culurciello E. LinkNet: exploiting encoder representations for efficient semantic segmentation. Paper presented at: 2017 IEEE Visual Communications and Image Processing (VCIP), 2017. doi:10.1109/VCIP.2017.8305148

66. Araújo RL, de Araújo FHD, e Silva RRV. Automatic segmentation of melanoma skin cancer using transfer learning and fine-tuning. *Multimedia Syst.* 2022;28:1239-1250. doi:10.1007/s00530-021-00840-3

67. Chen LC, Papandreou G, Schroff F, Adam H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv*:1706.05587, 2017.

68. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. *arXiv preprint arXiv*:1511.07122, 2016.

69. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. Paper presented at: Computer Vision – ECCV 2018, 2018. doi:10.1007/978-3-030-01234-2_49

70. Segmentation Models Pytorch. https://github.com/qubvel/segmentation_models.pytorch. Accessed February 19, 2024.

71. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception-ResNet and the impact of residual connections on learning. arXiv preprint arXiv:1602.07261, 2016.

72. Zhang H, Wu C, Zhang Z, et al. ResNeSt: split-attention networks. Paper presented at: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022. doi:10.1109/CVPRW56347.2022.00309

73. Mahajan D, Girshick R, Ramanathan V, et al. Exploring the limits of weakly supervised pretraining. Paper presented at: Computer Vision – ECCV 2018. doi:10.1007/978-3-030-01216-8_12.

74. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv*:1409.1556, 2015.

75. Jia Z, Maggioni M, Staiger B, Scarpazza DP. Dissecting the NVIDIA Volta GPU architecture via microbenchmarking. *arXiv preprint arXiv*:1804.06826, 2018.

76. Arzt M, Deschamps J, Schmied C, et al. LABKIT: labeling and segmentation toolkit for big image data. *Front Comput Sci*. 2022;4:777728.

77. Hollandi R, Diósdi Á, Hollandi G, Moshkov N, Horváth P. AnnotatorJ: an ImageJ plugin to ease hand annotation of cellular compartments. *Mol Biol Cell*. 2020;31:2179-2186. doi:10.1091/mbc.E20-02-0156

78. Bankhead P, Loughrey MB, Fernández JA, et al. QuPath: open source software for digital pathology image analysis. *Sci Rep*. 2017;7:16878. doi:10.1038/s41598-017-17204-5

79. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676-682. doi:10.1038/nmeth.2019

80. Ershov D, Phan MS, Pylvänäinen JW, et al. TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat Methods*. 2022;19:829-832. doi:10.1038/s41592-022-01507-1