

# Precise measurement of nanoscopic septin ring structures with deep learning-assisted quantitative superresolution microscopy

Amin Zehtabian<sup>\*†</sup>, Paul Markus Müller<sup>\*†</sup>, Maximilian Goisser, Leon Obendorf, Lea Jänisch, Nadja Hümpfer, Jakob Rentsch, and Helge Ewers

Institute for Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany

**ABSTRACT** The combination of image analysis and superresolution microscopy methods allows for unprecedented insight into the organization of macromolecular assemblies in cells. Advances in deep learning (DL)-based object recognition enable the automated processing of large amounts of data, resulting in high accuracy through averaging. However, while the analysis of highly symmetric structures of constant size allows for a resolution approaching the dimensions of structural biology, DL-based image recognition may introduce bias. This prohibits the development of readouts for processes that involve significant changes in size or shape of amorphous macromolecular complexes. Here we address this problem by using changes of septin ring structures in single molecule localization-based superresolution microscopy data as a paradigm. We identify potential sources of bias resulting from different training approaches by rigorous testing of trained models using real or simulated data covering a wide range of possible results. In a quantitative comparison of our models, we find that a trade-off exists between measurement accuracy and the range of recognized phenotypes. Using our thus verified models, we find that septin ring size can be explained by the number of subunits they are assembled from alone. Furthermore, we provide a new experimental system for the investigation of septin polymerization.

## Monitoring Editor

Melike Lakadamyali  
University of Pennsylvania

Received: Feb 10, 2022

Revised: Apr 29, 2022

Accepted: May 9, 2022

## INTRODUCTION

Light microscopy is a technique of fundamental importance in cell biology and, with the development of superresolution techniques, has allowed unprecedented insight into cellular processes on the brink of structural biology (Willig *et al.*, 2006; Gunzenhäuser *et al.*, 2012; Szymborska *et al.*, 2013; Balzarotti *et al.*, 2016; Sochacki *et al.*,

2017; Mund *et al.*, 2018; Sieben *et al.*, 2018; Gambarotto *et al.*, 2019; Zwettler *et al.*, 2020). Machine learning-based computational image processing furthermore allows for high-throughput analysis of even complex phenotypes in cells (Held *et al.*, 2010; Jones *et al.*, 2009; Sommer and Gerlich, 2013; Piccinini *et al.*, 2017; Morone *et al.*, 2020) and has accelerated progress in cell biological discovery. In recent years a number of open source deep learning (DL) platforms have been made available (Sommer *et al.*, 2017; Mathis *et al.*, 2018; Falk *et al.*, 2019; Moen *et al.*, 2019; Zhang *et al.*, 2019; Stringer *et al.*, 2020; Gómez-de-Mariscal *et al.*, 2021; Isensee *et al.*, 2021; Lucas *et al.*, 2021; von Chamier *et al.*, 2021) that require minimal know-how on the user side. At the same time, methods have been developed to reduce background (Weigert *et al.*, 2018; Fang *et al.*, 2021) and to accelerate image processing in superresolution microscopy (Ouyang *et al.*, 2018; Nehme *et al.*, 2020; Speiser *et al.*, 2021). Especially in superresolution microscopy, image processing allows for ultrahigh resolution of multiprotein complexes such as nuclear pores (Szymborska *et al.*, 2013; Heydarian *et al.*, 2021) or centrioles (Mennella *et al.*, 2012; Sieben *et al.*, 2018). However, these are highly symmetric structures of constant size that bear several symmetry

This article was published online ahead of print in MBoc in Press (<http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E22-02-0039>) on May 20, 2022.

<sup>†</sup>These authors contributed equally to this work.

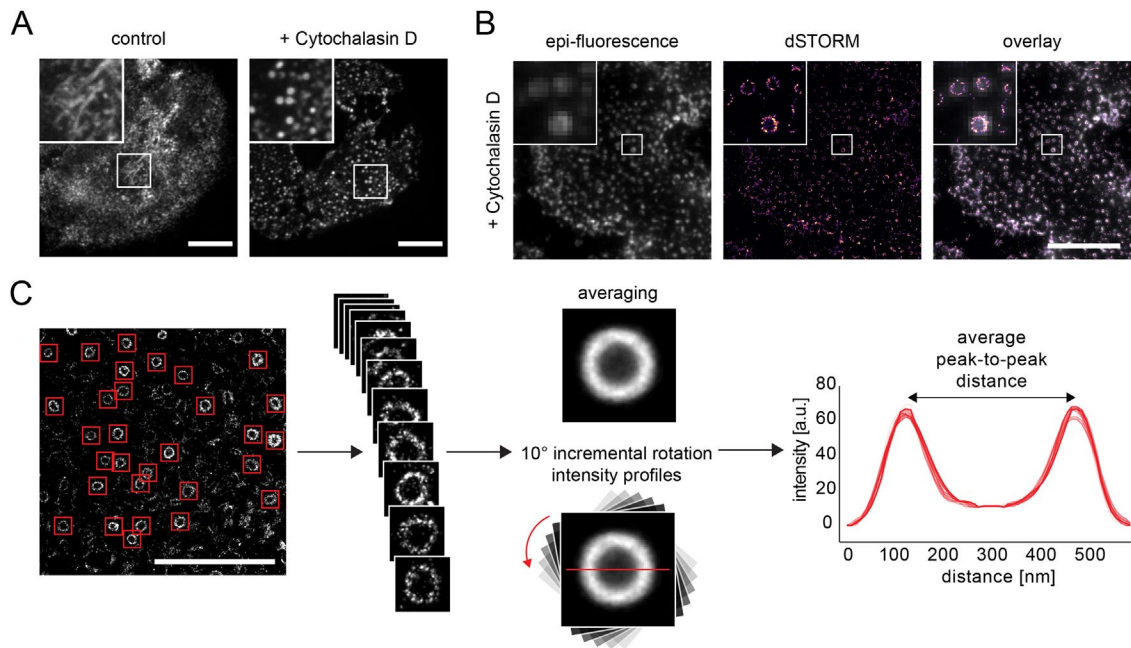
Competing interests: The authors declare no competing interests.

\*Address correspondence to: Amin Zehtabian ([amin.zehtabian@fu-berlin.de](mailto:amin.zehtabian@fu-berlin.de)); Paul Markus Müller ([paulmarkus.mueller@fu-berlin.de](mailto:paulmarkus.mueller@fu-berlin.de)).

Abbreviations used: DL, deep learning; dSTORM, direct stochastic optical reconstruction microscopy; GTP, guanosine-5'-triphosphate; ROI, region of interest; SMLM, single molecule localization-based superresolution microscopy; TIRF, total internal reflection fluorescence; wt, wildtype.

© 2022 Zehtabian *et al.* This article is distributed by The American Society for Cell Biology under license from the author(s). Two months after publication it is available to the public under an Attribution-NonCommercial-Share Alike 4.0 International Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/4.0>).

"ASCB®," "The American Society for Cell Biology®," and "Molecular Biology of the Cell®" are registered trademarks of The American Society for Cell Biology.



**FIGURE 1:** Superresolution microscopy assay for quantification of septin ring diameter. (A) Genome-edited NRK52E cells expressing mEGFP-Septin-2 from both alleles were treated with 5 μM cytochalasin D for 30 min (or DMSO as control) prior to fixation, resulting in the formation of ringlike structures. Images of mEGFP fluorescence were taken by spinning disk confocal microscopy. Scale bar: 10 μm. (B) Septin rings can be readily resolved by SMLM. NRK52E cells were treated with 5 μM cytochalasin D for 30 min and immunostained against Septin-7. Images were taken by conventional epifluorescence microscopy and by dSTORM. Scale bar: 10 μm. (C) Scheme of the procedure for the measurement of septin ring diameter. Left to right: Septin rings from superresolution microscopy images were cropped and averaged. The intensity profile of the averaged rings was measured at 18 successive 10° incremental rotations and the peak-to-peak distance of all of these profiles was averaged to return the ring diameter. Scale bar: 5 μm.

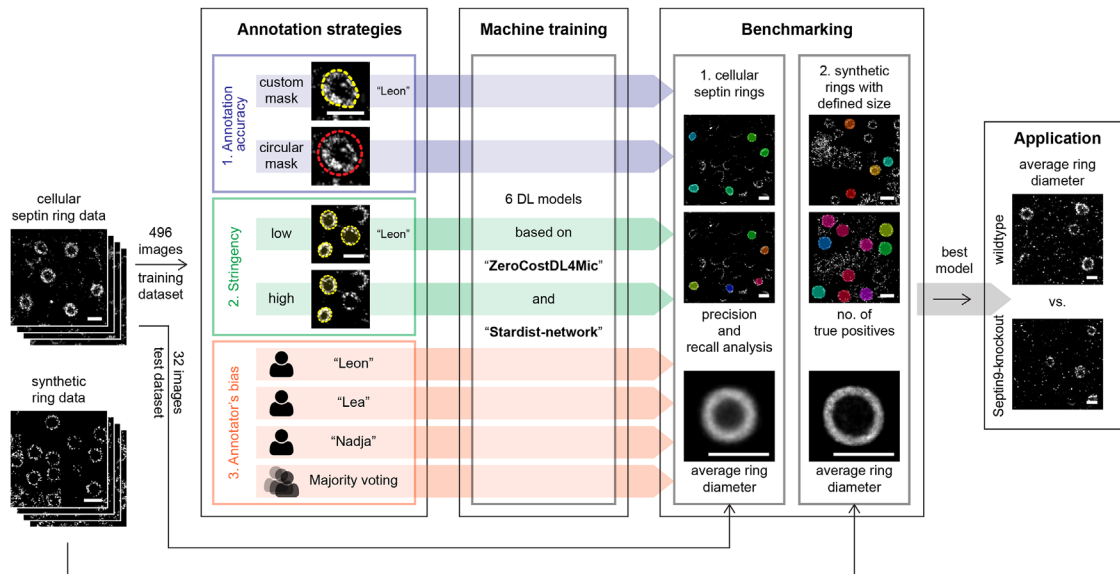
axes. As a result, they contain intrinsic means to allow for accurate averaging. For superresolution microscopy and averaging to harness its full power with the help of DL-based image processing to improve resolution of more amorphous structures and to develop quantitative readouts for structural changes, it remains to be determined to what extent DL methods can handle changes in the organization or even the size of target structures. Overfitting may make it impossible to detect even small changes in morphology or size; on the other hand, poor fitting may not allow for accurate measurements.

Here we combine DL-based image analysis based on an open-source platform with single molecule localization-based superresolution microscopy (SMLM; Betzig *et al.*, 2006; Rust *et al.*, 2006; Heilemann *et al.*, 2008) of subresolution septin ring structures. Septins are the fourth cytoskeleton (Weirich *et al.*, 2008; Mostowy and Cosart, 2012) and assemble from nonpolar heteromultimeric rod-shaped complexes into ringlike structures of subresolution size when not associated to tubulin or actin (Kinoshita *et al.*, 2002). These rings are free of actin and are thought to represent bona fide septin filamentous polymers (Sellin *et al.*, 2011). The composition of septin complexes plays an important role in septin function (Hu *et al.*, 2012; Kaplan *et al.*, 2017) and septin ring size depends on the composition of septin complexes (Kim *et al.*, 2011). Septin ring size may thus be a long sought-after readout for complex assembly and composition. Here we ask which factors in supervised DL affect ring recognition and, finally, accuracy of ring size measurements in native and perturbed cells with differing ring sizes. To do so, we train six different models on data annotated according to several different parameters and by three different experts and validate them on cellular and synthetic SMLM datasets. We find that while most models readily recognize ring structures in the size range expected in native

cells and allow for highly accurate determination of septin ring diameter from SMLM data, only a few models allow for the accurate recognition and measurement of rings of differing sizes. We furthermore show that septin ring size can be explained by septin complex composition alone, and we thus provide a new experimental paradigm for the investigation of septin complex assembly. Our results demonstrate that the combination of DL with SMLM can provide accurate readouts of changes in the size of amorphous multiprotein structures and that experimental design and careful validation are essential for the generation of reliable experimental pipelines.

## RESULTS

We chose septin rings as a model to test the robustness and accuracy of DL-assisted superresolution microscopy for measurements of subresolution cellular structures. The septins are a family of GTP-binding proteins that comprise the fourth cytoskeleton. Septins can assemble into filaments themselves or attach to actin stress fibers; however, when actin stress fibers are collapsed via the use of cytochalasin D, they form ring structures of around 350 nm diameter (Kim *et al.*, 2011; Kinoshita *et al.*, 2002) (Figure 1A). When imaged in SMLM, rings appear as homogeneously sized, symmetric rings with a thin perimeter (Figure 1B), but also incomplete rings, arcs, small filamentous structures and aggregates can be found. When we acquired many cells in SMLM imaging after cytochalasin D treatment and selected hundreds of rings from the stained septin structures, we could average them for accurate determination of their diameter of  $343.9 \pm 7.0$  nm ( $n = 181$ ; Figure 1C) with no apparent difference between antibody and nanobody-based immunostaining methods (Supplemental Figure S1). We concluded that we could accurately determine septin ring diameter using this method.



**FIGURE 2:** General annotation, training, and benchmarking strategy for DL-assisted septin ring recognition and measurement; 528 dSTORM images of septin rings in NRK52E cells were generated and used for annotation and training (496 images) or as a holdout test dataset (32 ground truth images). To benchmark annotation strategies for DL, images were annotated with different annotation accuracy, annotation stringency, and by three independent human experts as well as the majority voting thereof. Note that the model “Leon” served for comparison in different scenarios. For each annotation strategy we generated DL models based on the ZeroCostDL4Mic platform using the StarDist network. For benchmarking, we compared the ability of the models to recognize the rings and to return the correct ring diameter, both in cellular data and synthetic data with rings of known diameter. The best performing models were then applied to measure the diameter of septin rings in knockout cells. All scale bars: 500 nm.

We next aimed to use a widely available and versatile DL network to train it to recognize septin rings. We decided to use the ZeroCostDL4Mic (von Chamier *et al.*, 2021) platform which provides free cloud-based computation and the StarDist (Schmidt *et al.*, 2018) network that has been widely used for detection of cellular structures. However, since it had not been used for ring detection we decided to test if it could be trained from scratch to recognize septin rings in SMLM reconstructions. To do so, we used an annotated dataset of 496 images for training and adjusted parameters of the StarDist network until we could reliably reach convergence of our model. We next aimed to ask how annotation bias could affect model learning and the capability of the model to detect phenotypes strongly different from the wildtype (*wt*) situation in our images. In cellular SMLM images, septin rings are of variable circularity and completeness in labeling and occasionally touch. We hence decided to test if higher or lower stringency in determining a ring would be suitable for training and if masks precisely following the ring outline would lead to more accurate results. After training, we furthermore asked whether the DL models would also be able to accurately recognize and measure significantly smaller or bigger rings. For training, we generated a large amount of SMLM images of cytochalasin D-treated cells, which we separated into a training set of 496 images and a cellular holdout test dataset of 32 images. We then designed a framework to train DL models using six different annotation strategies and to test the models on the cellular holdout data, the synthetic data, and a biological test case that is rings devoid of Septin-9 (Figure 2). We asked one human expert to annotate in different ways in terms of accuracy of the masks (how does it fit to the ring) and annotation stringency (how circular and/or continuous do rings have to be), and after that we asked two more human experts to individually annotate with custom masks and low stringency. A final sixth model was trained using the consensus

masks. All trained models resulted in robust recognition of rings in the cellular test dataset of 32 SMLM images and we benchmarked the models in terms of precision as the fraction of true positives in all detection events (true positives over the sum of true positives and false positives), recall as the fraction of true positives compared with all ground truth rings (true positives over the sum of true positives and false negatives), and measured average ring size compared with ground truth (Figure 2). We then went on to run the models on synthetic SMLM data generated using FluoSim (Lagardère *et al.*, 2020) in which ring sizes from 125 to 500 nm were simulated. Finally, we tested the two best performing models on a cellular example of ring-size change, cells in which Septin-9 had been removed from the complex, leading to significantly smaller rings (Kim *et al.*, 2011).

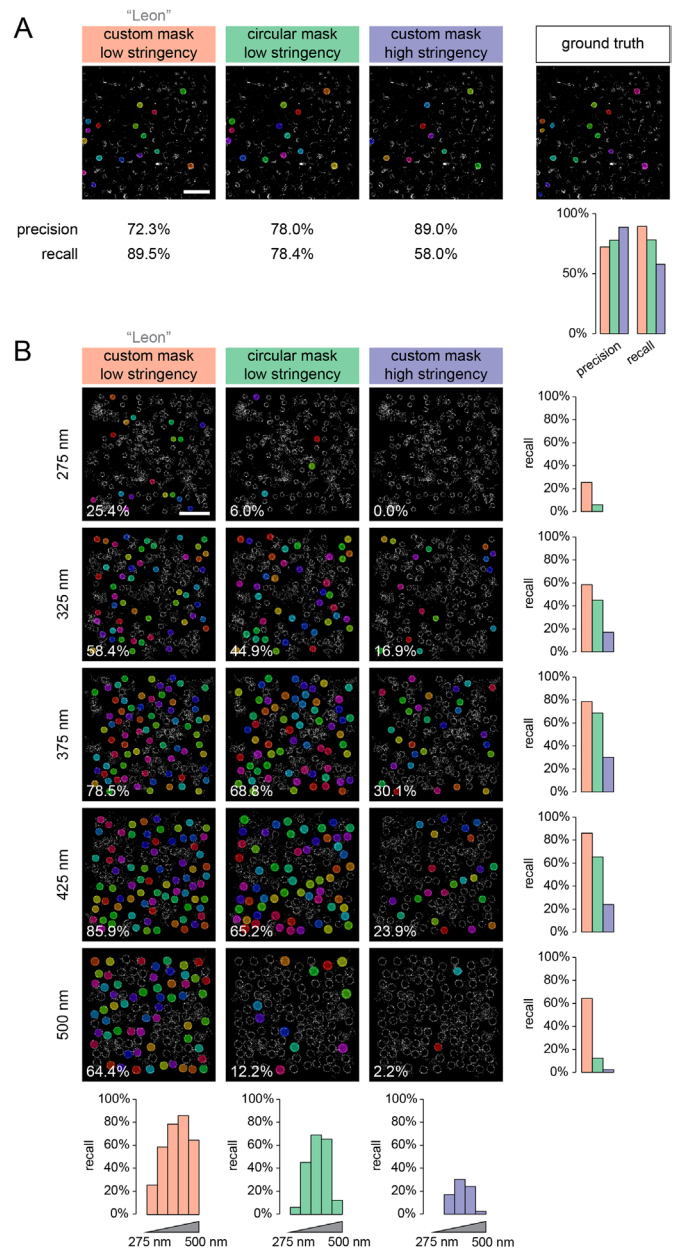
To evaluate the influence of annotation strategies on recognition by the DL models, a septin biologist recognized septin rings in the cellular test dataset of 32 SMLM images. These we accepted as bona fide septin rings, and the resulting dataset and the averaged diameter thereof are termed ground truth. The training dataset of SMLM data was first annotated by a human expert, “Leon,” according to three different strategies. The first strategy used low stringency in recognition as a septin ring and required the expert to draw a custom mask around each ring (Figure 2; Supplemental Figure S2). The second strategy required the expert to annotate the training dataset again with the same low stringency in septin ring recognition, but this time, instead of drawing a custom mask around each ring, used a circular mask or, to be more precise, a circular-shaped mask with a flexible size that covered the target ring, allowing a significant acceleration of the annotation procedure. The third strategy required the expert to apply high stringency in selecting rings but to use a custom-drawn mask, as in the first strategy. When the annotation-specific models were trained

and applied to the cellular test dataset, we found that all models recognized a large fraction of rings reliably when compared with ground truth albeit with varying precision and recall (Figure 3A). Specifically, the low stringency, custom mask-trained model led to the highest recall with 89.5% of rings found that were present in the ground truth. On the other hand, the high stringency, custom mask-trained model led to the highest precision with 89% of all recognized rings being true positives, but recall was low. The circular mask, low stringency model was of intermediate precision and recall, and no model showed precision or recall below 50%. Overall, we concluded that our models could reliably recognize septin rings in our cells.

When the composition of septin complexes is changed, the diameter of septin rings changes as well (Kim *et al.*, 2011), suggesting that ring size may provide a readout for septin complex assembly and composition. However, to accurately measure the diameter of rings with different sizes, our models must recognize rings of even strongly differing diameters for an accurate readout of ring-size changes. We thus aimed to test whether one of our training strategies would lead to a bias in the recognition of differently sized rings and only recognize rings in the size range found in our untreated cells. To do so, we analyzed a panel of synthetic SMLM test datasets in which rings of a predefined diameter from 275 to 500 nm, respectively, were present. When we then applied our models to recognize rings in these synthetic test datasets, we found striking differences in ring recognition (Figure 3B). Recall values varied both across the different diameters and across the different models, while precision was 100% for all the models, meaning that all recognized rings were true positives, as structures that were not rings but could still be recognized by the models were not generated in the synthetic SMLM test datasets. Recall values dropped for diameters significantly below or above the ring sizes present in our cells for all models. However, the low stringency, custom mask model (hereafter called “Leon”) clearly outperformed the other two models, with recall values between 25.4% at 275 nm and 85.9% at 425 nm. The other two models failed to reliably recognize rings of 275 or 500 nm. We concluded that the custom mask, high stringency model (hereafter called “high stringency”) and the circular mask, low stringency model (hereafter called “circular mask”) were biased for rings in the size range found in cells and that the training strategy significantly impacts recognition of septin rings toward superresolution measurements.

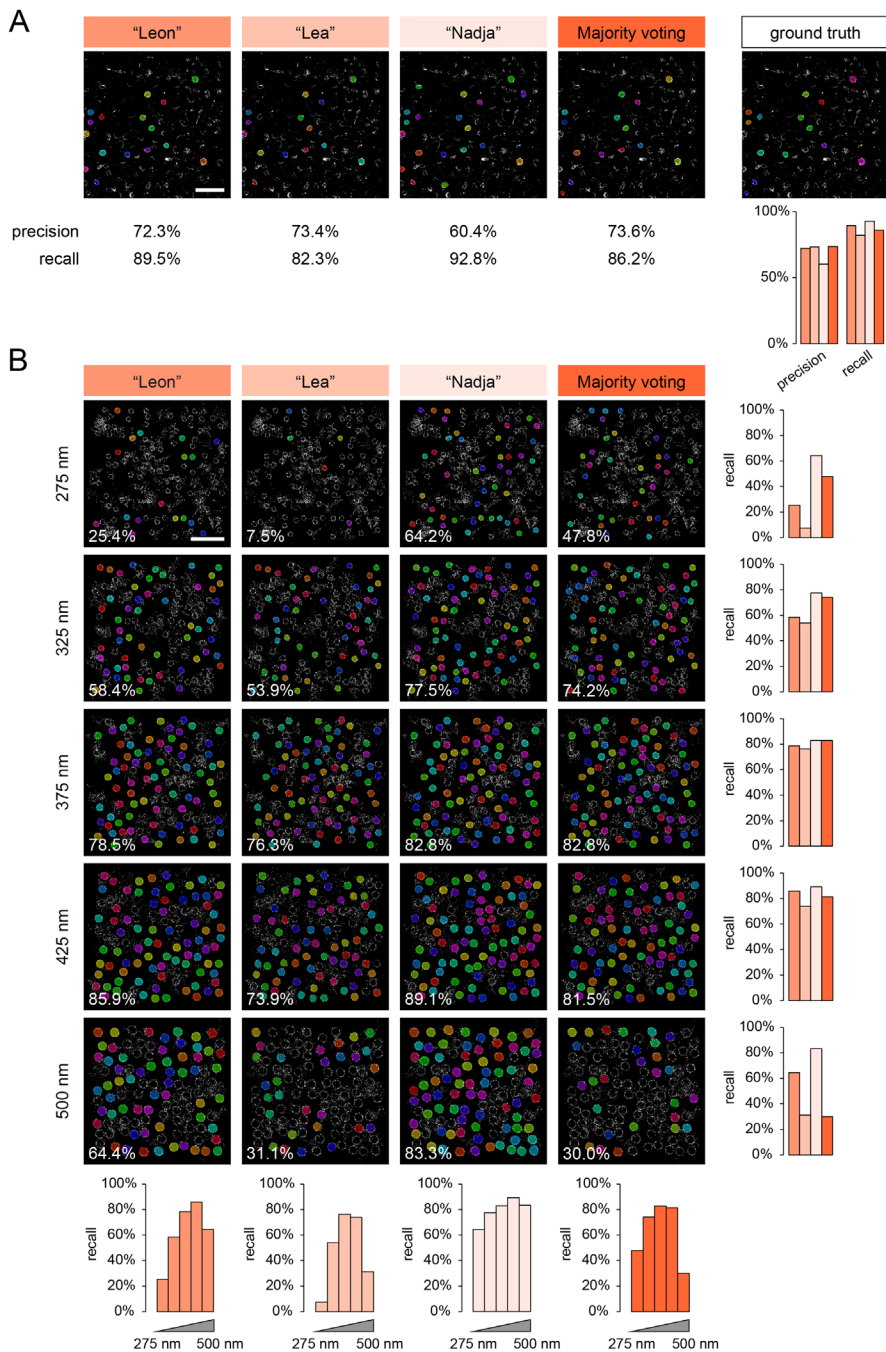
Next, we aimed to investigate how three different experts using the same annotation strategy would influence ring recognition by the resulting models. We asked two additional experts, “Lea” and “Nadja,” to annotate the cellular SMLM training dataset according to the low stringency, custom mask strategy. From the annotations of the three experts “Leon,” “Lea,” and “Nadja,” we furthermore created a “majority voting” annotation (Figure 2; Supplemental Figure S2). When we then trained additional models according to the “Lea,” “Nadja,” and “majority voting” annotations on our cellular test dataset, we found that all models exhibited a precision above 60% and a recall above 80% when compared with ground truth (Figure 4A).

We observed that precision and recall values of ring detection in the ground truth as compared over all six models showed an inverse relationship (Supplemental Figure S3). For example, the “Nadja” model exhibited an exceptional recall of 92.8%, however, at the price of only 60.4% precision, meaning that about 40% of the structures recognized by the “Nadja” model were false-positive objects that were not found in the ground truth. On the other hand, the “high stringency” model reached the highest precision of all models



**FIGURE 3:** Effect of annotation accuracy and stringency on the ability of DL models to detect rings of different size. (A) Overlay of randomly colored masks of recognized septin ring structures in dSTORM images by three different DL models: custom mask and low stringency model (aka “Leon”); circular mask and low stringency model; custom mask and high stringency model. The rings detected in a representative dSTORM reconstruction from the holdout data are shown together with the respective ground truth. Precision and recall calculations are shown below. Scale bar: 2 μm. (B) Overlay of randomly colored masks of recognized ring structures in synthetic dSTORM images with predetermined ring diameters from 275 to 500 nm as recognized by the DL models as in (A). Recall values of detected rings in each condition are shown on the bottom left of each image. Recall values of the models for each diameter are plotted on the right; recall values of a given model across all diameters are plotted at the bottom. Note: precision values were 100% for all models tested on synthetic data. Scale bar: 2 μm.

(89%) but the lowest recall (58%), meaning that it only recognized a bit more than half of the rings in the ground truth. The performance of the “majority voting” model was very similar to the “Leon”

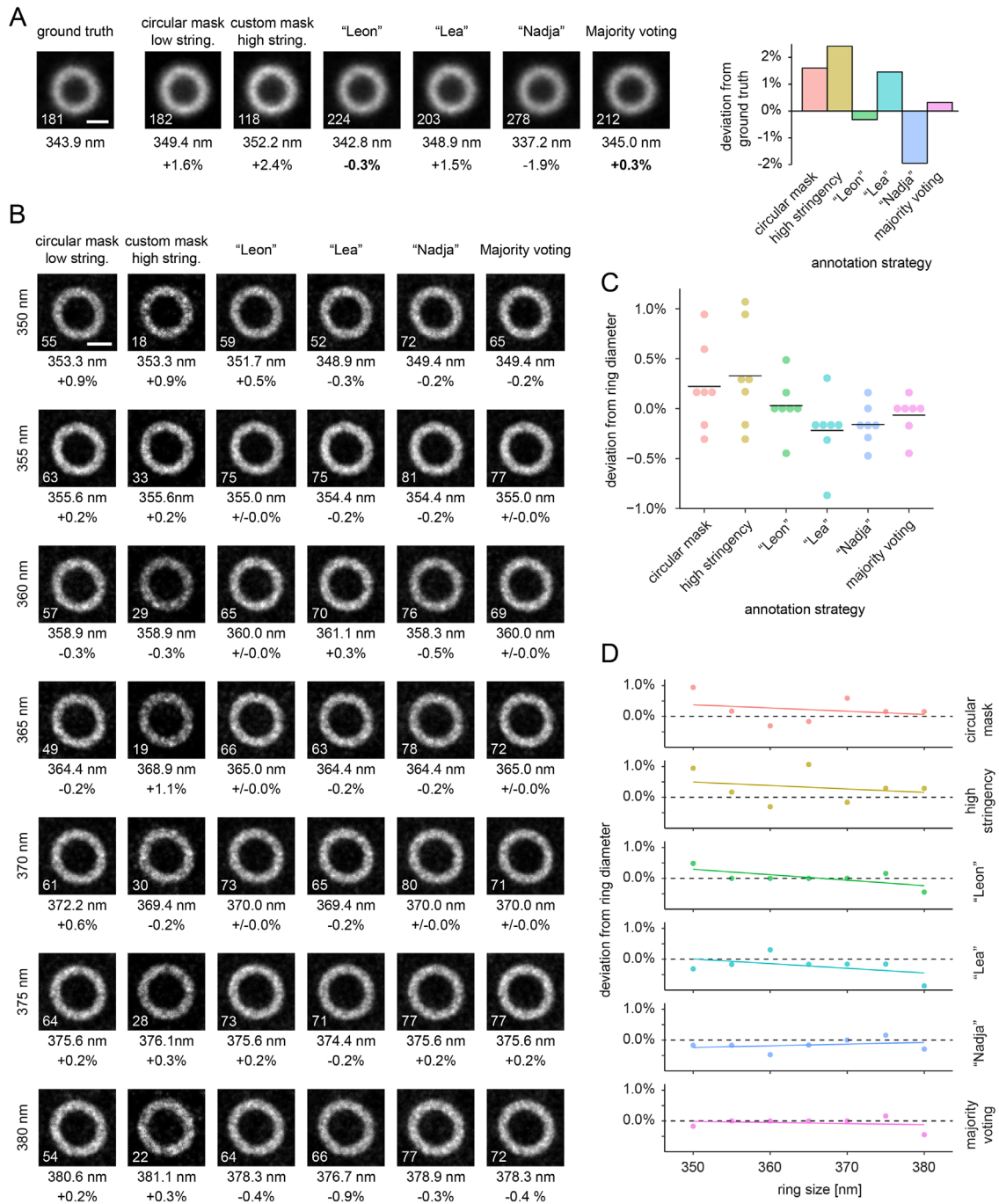


**FIGURE 4:** Effect of annotator's bias on the ability of DL models to detect rings of different sizes. (A) Overlay of randomly colored masks of recognized septin ring structures in dSTORM images by four different DL models: custom mask and low stringency model each for individual annotators "Leon," "Lea," and "Nadja" and for the "majority voting" model. The rings detected in a representative dSTORM reconstruction from the holdout data are shown together with the respective ground truth. Precision and recall calculations are shown below. Scale bar: 2  $\mu$ m. (B) Overlay of randomly colored masks of recognized ring structures in synthetic dSTORM images with predetermined ring diameters from 275 to 500 nm as recognized by the DL models as in (A). Recall values of detected rings in each condition are shown on the bottom left of each image. Recall values of the models for each diameter are plotted on the right; recall values of a given model across all diameters are plotted at the bottom. Note: precision values were 100% for all models tested on synthetic data. Scale bar: 2  $\mu$ m.

model. We concluded that indeed the specific annotation style of different experts influences ring recognition significantly.

When we next applied the "Lea," "Nadja," and "majority voting" models to the panel of synthetic test datasets and compared them with the "Leon" model, we found significant discrepancies in performance (Figure 4B). Both the "Nadja" and the "majority voting" models reliably outperformed the other models for smaller rings, while for rings in the range of those found in cells, all models performed equally well. We also tested rings below 275 nm in size (Supplemental Figure S4; Supplemental Figure S5). For the smallest size of 125 nm, no model could recognize any rings in synthetic SMLM test datasets. However, the "Nadja" and the "majority voting" models could reliably recognize rings from a diameter of 255 nm upward, making them clearly stand out from the other models regarding the recognition of small-sized rings. Large (500 nm) rings, on the other hand, were readily detected by the "Nadja" and "Leon" models, while the "Lea" and "majority voting" models recognized less than half of the rings detected by either the "Nadja" or the "Leon" models. We concluded that the bias introduced by individual annotators has significant consequences on the recognition of septin rings in SMLM data.

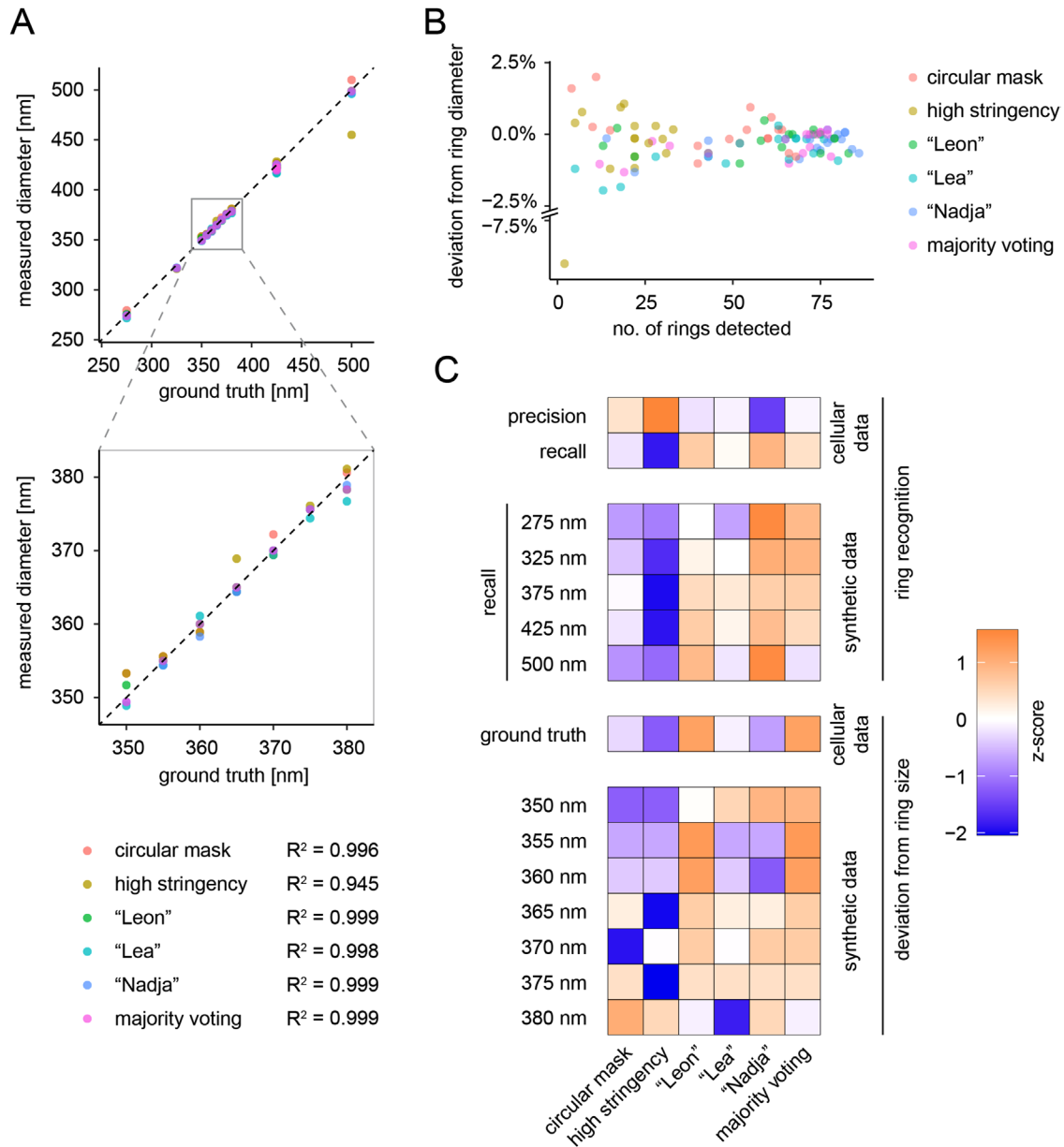
We next aimed to ask to what extent these biases in septin ring annotation would impact the accurate measurement of septin ring diameter in datasets analyzed by our models. To do so, we averaged the recognized septin rings for every model and condition and measured the diameter in nm as described in Figure 1C. When we compared the results of our models on the cellular SMLM test dataset, we found that all models resulted in a measurement with an accuracy within 3% of the 343.9 nm found in the ground truth (Figure 5A). Next, we tested if we could measure differences as small as 5 nm in ring diameter. To do so, we generated synthetic SMLM test datasets with diameters in 5-nm steps from 350 to 380 nm. When we ran our models on these datasets and measured the average diameters of the respectively recognized rings, we found that all models resulted in very accurate measurements of the respective ring sizes, with 3–4 nm errors maximally (Figure 5B). Overall, the highest spread was found for the "circular mask" and "high stringency" models (Figure 5C), while the residuals were lowest for the "majority voting" model (Figure 5D). Importantly, we did not observe a ring diameter dependence of recall values in the synthetic data of rings from 350 to



**FIGURE 5: Quantification of the precision of DL-based ring measurements.** (A) Results for septin ring diameter measurements in cellular dSTORM holdout data by the respective DL models. Averaged recognized ring images are shown for each model and average ring diameter and deviation from ground truth are shown below the images. The number of rings averaged is shown on the bottom left of each image. The plot on the right shows the deviation of measured septin ring diameter from ground truth for each model in percentage. Scale bar: 200 nm. (B) Results for septin ring diameter measurements in synthetic dSTORM data of known ring size by respective DL models. The underlying ground truth ring diameter is given on the left. Averaged recognized ring images are shown for each model and diameter; average ring diameter and deviation from ground truth are shown below the images. The number of rings averaged is shown on the bottom left of each image. Scale bar: 200 nm. (C) Deviation from respective ground truth ring diameter in percentage points measured for all models and for all seven ground truth ring diameters shown in (B). Black lines indicate mean deviation. (D) Deviation from respective ground truth ring diameter in percentage plotted against ground truth diameter for all models. Dashed line marks 0% deviation; colored lines show a linear fit.

380 nm (Supplemental Figure S6). Together, we concluded that all models were suitable to automatically extract and measure ring size in native cellular SMLM datasets.

When we plotted the measured ring diameter against ground truth ring diameter in synthetic SMLM test data, we found that all models exhibited excellent linearity over the range from 275 to



**FIGURE 6:** DL models accurately and robustly measure septin ring size in dSTORM data. (A) Measured ring diameter in synthetic dSTORM data plotted against ground truth ring diameter for all models used in this study. The magnified inset below shows the ring size range exemplified in Figure 5(B). (B) Deviation from respective ground truth ring diameter for synthetic rings from 275 to 500 nm in percentage points measured for all models plotted against the respective number of detected rings in each case. Note that deviation from ground truth decreases with a larger number of rings detected. (C) Global analysis of DL model performance. From top to bottom: precision and recall values in cellular data, recall values in synthetic data, ring size deviation from ground truth in cellular data, and deviation from ground truth ring size in synthetic data. Heat map shows horizontally z-scored values.

500 nm with  $R^2$  values above 0.94 (Figure 6A). Thus while all models were very precise in their measurement of ring diameter, a clear dependence of measurement accuracy on the number of recognized rings was found (Figure 6B), suggesting that the generation of higher amounts of data could further ameliorate measurement accuracy for the less well-performing models. In Figure 6C, we summarize the data from Figures 3 to 5 in an accessible overview according to the z scores for ring recognition (top) and ring diameter measurement (bottom). Found values are horizontally z scored, thus emphasizing differences between models even if they may be small in absolute size and thus describe relative performance. However,

this overview allows for general observations. It seems that beyond the inverse relation of precision and recall of the ground truth (Supplemental Figure S3), the z scores for precision and recall, respectively, are correlated with specific outcomes. The model with the lowest precision in the cellular data, "Nadja," performed best in recognizing rings of sizes significantly different from those measured in native cells, suggesting that a less stringent selection of cellular structures as rings allow the model to recognize rings of different sizes more easily. In contrast, the model with the highest precision in the cellular data, "high stringency," did not perform well in recognition and measurement of rings in synthetic data. On the other hand,

the models that most reliably recognized and most accurately measured rings in the expected size in both cells and synthetic data showed high recall values (“Leon” and “majority voting”), whereas the models with low recall (“high stringency” and “circular mask”) exhibited the worst performance. We concluded that precision and recall are important indicators of model performance, but they do not suffice to predict experimental outcome.

Having found models that can recognize septin rings of different sizes and accurately measure their diameter, we decided to test our models on an experimental system where septin ring size is known to vary. Septin complexes in mammalian cells are nonpolar, palindromic rod-shaped hetero-octamers of around 34 nm length (Bertin *et al.*, 2010, 2008; DeMay *et al.*, 2011; Frazier *et al.*, 1998; Kaplan *et al.*, 2015) that bear two Septin-9 subunits in their middle (McMurray and Thorner, 2019; Mendonça *et al.*, 2019; Soroor *et al.*, 2021). When Septin-9 is removed from cells through RNA interference, septin rings are known to be smaller (Kim *et al.*, 2011). This suggests that ring size is connected to complex length as the septin complexes created in the absence of Septin-9 are known to be hexamers (Sellin *et al.*, 2012) of around 26 nm length. If that is the case, septin ring size should reflect complex size and ring size should be dependent on the number of complexes in the ring. To ask whether we can thus estimate the number of septin complexes in rings from SMLM measurements, we performed SMLM on mouse embryonic fibroblast cells from *wt* and Septin-9<sup>-/-</sup> mice (Füchtbauer *et al.*, 2011). We found septin rings to be in the size range of 300–400 nm and decided to analyze ring diameter using the “Leon” and the “majority voting” models as they most accurately measured the average ring diameter in that range. Indeed, both models readily recognized many rings in measured SMLM data (Figure 7A; Supplemental Figure S7). In *wt* cells, we measured an average ring diameter of 413.3 nm (“majority voting”) or 425.5 nm (“Leon”). Geometrically, this would translate into 39 septin octamers with a 9.2° association angle (“Leon”) or 38 octamers with a 9.4° association angle (“majority voting”). On the other hand, in Septin-9<sup>-/-</sup> cells, we measured ring diameters of 334.4 nm (“Leon”) or 335.0 nm (“majority voting”), which would translate into 40 septin hexamers with an 8.9° association angle for both models (Figure 7B). Strikingly, we could thus determine that septin ring size in our cells must be determined by the propagation angle of about 9° between polymerizing septin complexes that would then lead to 40 septin complexes required to form a ring. We concluded that machine-learning-supported SMLM superresolution microscopy could reveal information on the ultrastructural arrangement of septin complexes in rings.

Finally, we asked if our model would allow the recognition of septin ring structures in other systems. We previously established SMLM for yeast cells and provided quantitative measurements of septin ring structures during the cell cycle by staining Cdc11-GFP via Alexa Fluor 647-coupled nanobodies in dividing yeast cells (Ries *et al.*, 2012). When we applied the “Nadja” model to the original data from this publication, we found that this model readily recognized the ring-shaped septin structure at the bud neck of *Saccharomyces cerevisiae* cells (Supplemental Figure S8) in cells that presented the septin ring in the imaging plane. At the same time, the larger and more amorphous bud scars were not recognized as rings by our model. This suggests that the combination of machine learning and SMLM superresolution microscopy of septin rings can be used to investigate septin assembly in different experimental settings. Taken together, our deep-learning model here provides a mature solution for the automated recognition of septin ring structures in several systems and provides very accurate measurements.

## DISCUSSION

We here developed a DL-based assay for the recognition and measurement of septin ring size in superresolution images of mammalian cells. We trained DL models in several different ways to investigate possible bias. We demonstrate that the outline of the annotation mask, the stringency in annotation, and the generation of a consensus between experts influence the outcome of both the accuracy of measurements and robustness in the detection of target structures. For example, we found that some of the models were more sensitive to changes in ring size than others. Especially if the models are supposed to recognize structures with a different phenotype in cells, such bias can be decisive. This suggests that benchmarking on simulated data covering the whole phenotypic range provides an excellent means to control the reliability of the DL models.

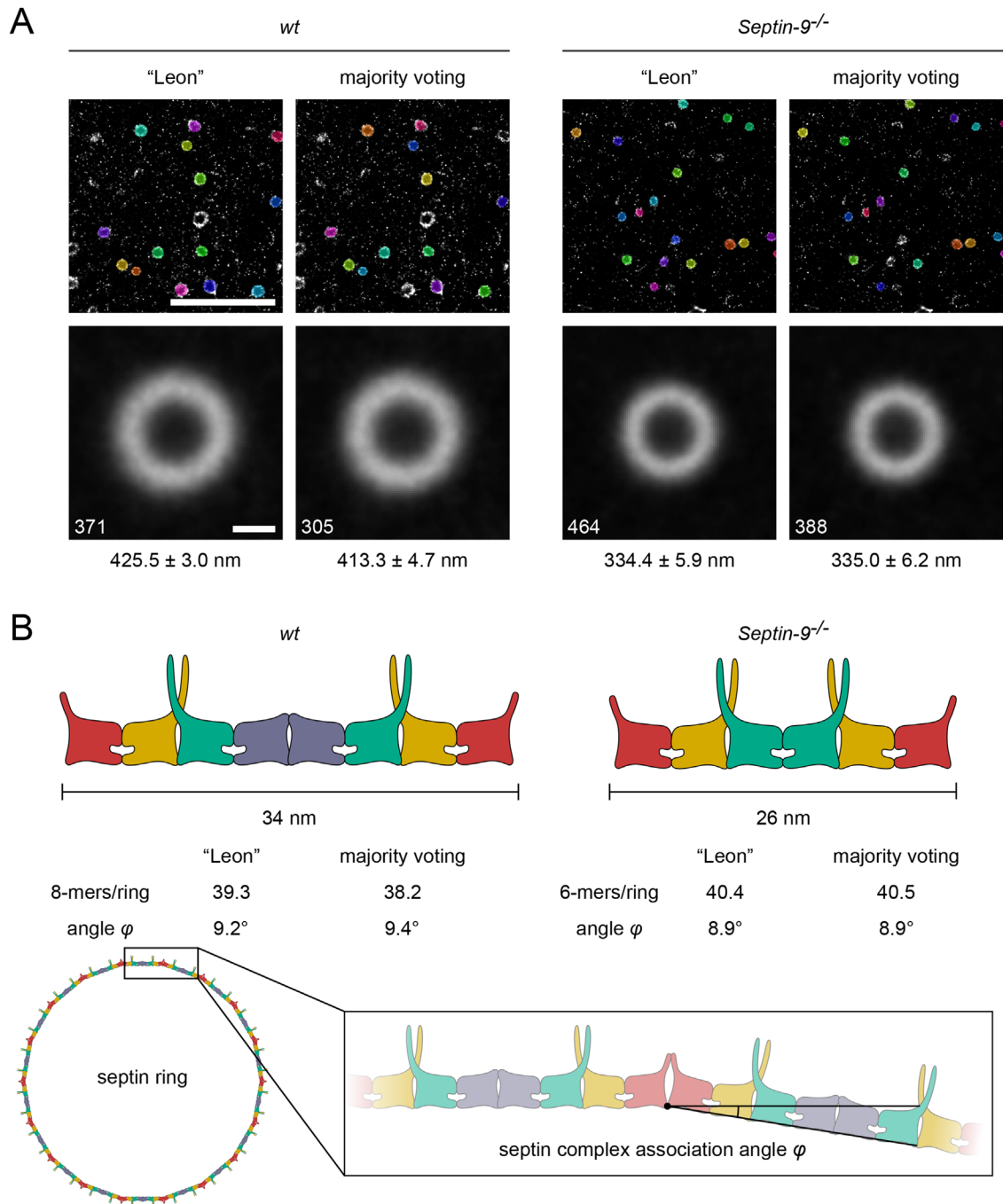
So far, DL models for the recognition of fluorescence-labeled structures in cells have mostly been used to recognize structures of low or highly predictable variability such as cellular nuclei (Carpenter *et al.*, 2006; Schmidt *et al.*, 2018; Weigert *et al.*, 2018; Falk *et al.*, 2019). The correct recognition of different phenotypes is, however, an essential property of data analysis models in experimental biology. To ensure that this is possible for our models, we use a wide array of synthetic SMLM data to test our models on the capability to recognize phenotypes that are different from what is observed in *wt* cells. Our data suggest that very stringent criteria in the selection of native rings may lead to overfitting and thus be detrimental to the recognition of rings of different size in cells, as observed for the “high stringency” and likely also the “Lea” model. On the other hand, less stringent criteria lead to a slightly lower accuracy in the measurement of ring diameter, as observed for the “Nadja” model. This model also readily recognized septin rings in a different model, *S. cerevisiae*, even though it was not trained with any such data, demonstrating excellent portability of the model. From all of our models, the “majority voting” model belonged to the best performing models both for recognizing *wt* cellular rings and for recognizing rings of significantly different sizes. It also resulted in very low error in ring size determination. Our data agree with previous observations that consensus or majority voting models often perform better than models based on individual annotators (Segebarth *et al.*, 2020).

In our study we retrieve the accurate ring diameter through averaging. In a more dynamic scenario where rings of different sizes and shapes are present in the same sample, such as the different organizational states of septins in the budding yeast *S. cerevisiae* during the cell cycle (Ries *et al.*, 2012), this would not be useful. However, the combination of ring segmentation and the calibration of the segmented area to known diameters by means of synthetic data would also allow the measurement of individual rings and thus their distribution in a heterogenous mixture of rings. This suggests that the use of simulated data is - beyond benchmarking - a useful tool that might allow for the accurate measurement of individual rings in future studies.

Our data provide insight into the requirements in experimental design of DL-based data analysis pipelines of superresolution microscopy data specifically and into the generation of robust DL models in general. We furthermore provide a robust quantitative assay for the investigation of subresolution structures assembled from an essential cytoskeletal component, the septins.

Our finding that septin ring diameter both in Septin-9 *wt* and in Septin-9<sup>-/-</sup> cells can be explained by 40 complexes, each adding an angle of 9° toward ring closure, suggests that superresolution microscopy of septin complexes may be useful as a readout to investigate septin complex assembly in the future. The finding that the





**FIGURE 7:** Accurate measurement of cellular septin ring size reveals the geometric architecture of septin rings.

(A) Septin ring recognition in *wt* and *Septin-9<sup>-/-</sup>* mouse embryonic fibroblasts. Septin ring recognition was performed with the two DL models "Leon" and "majority voting." Representative overlay of randomly colored masks of recognized septin ring structures in dSTORM images and averaged ring images are shown for each case. Average ring diameter is shown below. The number of rings averaged is shown on the bottom left of each image. Scale bars: upper row, 5  $\mu$ m; bottom row, 200 nm. (B) Model of septin ring assembly. Shown are schematic representations of septin complexes. In mouse fibroblasts the knockout of Septin-9 is thought to remove the central Septin-9 dimer (blue) from the palindromic, hetero-octameric complex, resulting in a heterohexamer. Since the polymerization interface for septin filament assembly is the same for hexamer and octamer, the rules for assembly of rings should be the same for both complexes and resulting rings thus smaller. Septin rings measured via DL-assisted superresolution microscopy data analysis in *Septin-9* knockout cells are smaller than *wt* rings, consistent with a constant mode of assembly of rings from  $\sim 40$  complexes with an interface angle  $\phi$  of  $\sim 9^\circ$ .

angle in *Septin-9* containing complexes is the same as in *Septin-9*-free complexes suggests that the  $9^\circ$  propagation angle is not introduced at the center of the hexamer but rather at the polymerization

interface between terminal Septin-2-group septins in adjacent complexes. When the observation of smaller Septin rings in shRNA-mediated *Septin-9* knockdown was first made (Kim *et al.*, 2011) this

conclusion could not be made, as at the time it was assumed SEPT9 would occupy the termini of the complex and a switch between Septin-9 to Septin-7 capped septin complexes would take place on Septin-9 depletion. Only recently it was clarified that Septin9 indeed occupies the central position in the complex and thus the organization of the complex would be similar to that observed in yeast (Mendonça *et al.*, 2019; Soroor *et al.*, 2021). In baker's yeast *S. cerevisiae*, it is known that the terminal septin in the heterooctamer, here Cdc11 or Shs1, determines the degree to which septin filaments bend. Straight filaments form from Cdc11-capped complexes and ringlike filaments form from Shs1-capped complexes (Garcia *et al.*, 2011). It is tempting to speculate that the Septin-2-group septins, which occupy the termini of the mammalian octamer may play a similar role. It is known that the Septin-2 group members can replace one another at position X in a Septin-X/6/7/7/6/X heterohexameric biochemically (Kinoshita, 2003) and that they exhibit nonoverlapping functions (Ihara *et al.*, 2005; Asada *et al.*, 2010; Kaplan *et al.*, 2017; Song *et al.*, 2019) that may be a result of their mode of filament assembly. In the future, ring size as measured with our assay may allow for the investigation of how complex-terminal subunits control septin polymerization.

## MATERIALS AND METHODS

[Request a protocol](#) through *Bio-protocol*.

### Cell culture and immunostaining

The NRK52E-Septin-2-GFP<sup>EN/EN</sup> genome edited cell line expressing mEGFP-Septin-2 from both alleles was generated as described (Banko *et al.*, 2019). Septin-9<sup>c/c</sup> (*wt*) or Septin-9<sup>-/-</sup> (knockout) mouse embryonic fibroblasts were a generous gift from Ernst-Martin Füchtbauer (Department of Molecular Biology and Genetics, Aarhus University; Füchtbauer *et al.*, 2011). All steps in the following were carried out at room temperature unless otherwise stated. NRK52E-Septin-2-GFP<sup>EN/EN</sup>, Septin-9<sup>c/c</sup> (*wt*), or Septin-9<sup>-/-</sup> (knockout) mouse embryonic fibroblasts were seeded on round coverslips (thickness no. 1.5) coated with 0.01% (wt/vol) poly-L-Lysine solution (Sigma) and grown in DMEM (Life Technologies, Thermo Fisher Scientific) supplemented with 10% (vol/vol) fetal calf serum, 2 mM L-glutamine (Life Technologies, Thermo Fisher Scientific), and 100 U/ml penicillin–streptomycin (Sigma) at 37°C and 5% CO<sub>2</sub> in a humidified incubator. The next day, cells were treated with 5 μM cytochalasin D (Cayman Chemical, Cat# 11330) for 30 min at 37°C in DMEM without supplements to induce septin ring formation. Cells were fixed in 4% (wt/vol) paraformaldehyde solution in PHEM buffer (60 mM PIPES-KOH, pH 6.9, 25 mM HEPES, 10 mM EGTA, 1 mM MgCl<sub>2</sub>) for 15 min at 37°C. Fixation was quenched with 50 mM NH<sub>4</sub>Cl in PHEM buffer for 7 min. Cells were permeabilized using 0.25% (vol/vol) Triton X-100 in PHEM buffer for 5 min and then blocked with Image-iT FX signal enhancer (Invitrogen, Thermo Fisher Scientific) for 30 min before blocking for 1 h with 4% (vol/vol) horse serum, 1% (wt/vol) bovine serum albumin (BSA), and 0.1% (vol/vol) Triton X-100 in PHEM buffer. The sample was then stained with 250 ng/ml anti-GFP nanobody coupled to Alexa Fluor 647 (Platonova *et al.*, 2015) at 4°C overnight or with rabbit anti-human-Septin-7 IgG (used at 0.5 μg/ml, Cat# 18991, IBL America) overnight and subsequently with goat anti-rabbit Alexa Fluor 647–conjugate IgG-Fab (used at 4 μg/ml, Cat# A-21246, Thermo Fisher Scientific) for 45 min in PHEM buffer supplemented with 1% (wt/vol) BSA and 0.1% (vol/vol) Triton X-100. Samples were postfixed in 4% (wt/vol) paraformaldehyde in PHEM buffer for 10 min and quenched as described above. For the cellular septin ring dataset 13 samples were prepared out of which 6 were stained with anti-GFP nanobody and

7 were stained with anti-Septin-7 antibody. Images were taken from 18 (anti-GFP nanobody) and 15 cells (anti-Septin-7 antibody) resulting in 288 and 240 images, respectively. We used different immunostaining techniques for this dataset so that our models later would be tolerant to data generated by different labeling methods. Of these 528 images 32 were randomly selected to be used as a test dataset while the remaining 496 images were used for training of the DL models. For the Septin-9<sup>c/c</sup> or Septin-9<sup>-/-</sup> mouse embryonic fibroblasts 6 cells were imaged from one sample preparation per condition resulting in 96 images per condition.

### Spinning disk confocal microscopy

Images were acquired on an inverted IX71 microscope (Olympus) equipped with a CSU-X1 spinning disk unit (Yokogawa) and an iLas laser illumination system (Gataca Systems) with a 491-nm laser for illumination. A 60× NA 1.42 oil objective (Olympus) was used, and images were taken with an ORCA Flash 4.0LT sCMOS camera (Hamamatsu). The System was operated using the software MetaMorph.

### dSTORM of cellular septin rings

Coverslips with immunostained samples were mounted in a mounting chamber and filled with GLOX + BME buffer (4% [wt/vol] glucose, 140 mM β-mercaptoethanol, 10 mM NaCl, 200 mM Tris-HCl, pH 8, 500 μg/ml glucose oxidase, 40 μg/ml glucose catalase and 10% [vol/vol] glycerol). dSTORM images were acquired on a Vutara 352 superresolution microscope (Bruker) equipped with an ORCA Flash4.0 sCMOS camera (Hamamatsu) and a 60× NA 1.49 ApoN oil immersion total internal reflection fluorescence (TIRF) objective (Olympus), yielding a pixel size of 98 nm. Data were acquired in TIRF illumination at a laser power density of ~35 kW/cm<sup>2</sup> using a 639-nm laser. dSTORM images were reconstructed from 10,000 images taken with an exposure of 20 ms per image. Reference epifluorescence images were taken with an exposure of 500 ms and ~0.5 kW/cm<sup>2</sup> laser power density. Images were reconstructed with a pixel size of 10 nm using the software Vutara SRX v.6.04.14.

### dSTORM of yeast septin rings

Experiments were described previously (Ries *et al.*, 2012). For a detailed instruction of yeast dSTORM staining, see Kaplan and Ewers (2015).

### Generation of synthetic dSTORM ring data

Synthetic storm data were generated using FluoSim (Lagardère *et al.*, 2020). Geometry files were created using a python script, which generates nonoverlapping randomly spaced rings with a thickness of 150 nm. Each geometry file contains 100 rings of the same radius. To generate background geometry files, a python script was used which creates 100 overlapping irregular icosahedrons per geometry with random side lengths of up to 500 nm. The geometry files were imported into FluoSim. To mimic the immunostaining process 10,000 dyes were distributed randomly onto the geometries. Single-molecule localization reconstructions were then generated using parameters which resemble the experimental STORM imaging parameters (pixel size 10 nm, switch-on rate 0.027 Hz, switch-off rate 10 Hz, simulation time step 20 ms, localization precision 10 nm, 10,000 frames).

### Data preparation for DL

The DL trainings have been carried out on 496 cellular SMLM reconstructed images of size 1024 × 1024 pixels rendered at a pixel size of 10 nm, where the number of septin rings per image ranged from 0 to ~30 rings. Each training image has been manually annotated

using each of the six different annotation strategies resulting in six different paired sets of raw and corresponding mask images. The total number of annotated rings varied from ~1700 rings (in “high stringency” model) to ~4000 rings (in the “Nadja” model). The manual annotation was performed using the Fiji plugin Labkit. This way, our annotators manually drew masks around each potential septin ring resulting in mask images in which all pixels within an annotated ring were labeled with a distinct integer, while background pixels were labeled with zero. Based on the model as well as the number of septinlike rings in the image, a few to tens of seconds are spent on manual annotation of a single image. Moreover, the “majority voting” annotation—where the certainty level of the annotation is high—is created from the annotations made by the three expert annotators. A given pixel in the “majority voting” annotation image is considered as a part of a mask only if it has been annotated by at least two individual annotators; otherwise it is considered background. The unseen test dataset comprising 32 cellular SMLM images was manually annotated by a septin biologist and the result was used as ground truth for further benchmarking of the models.

### DL model training and parameter settings

For septin ring recognition, we trained each of the six models from scratch with the same cellular SMLM image data paired with the corresponding set of model specific mask images. To do so, we selected the 2D variant of the StarDist (Schmidt *et al.*, 2018) machine-learning network with the U-Net structure as its core. We trained the StarDist network on the ZeroCostDL4Mic platform (von Chamier *et al.*, 2021) as it provided simple access to a range of popular DL networks including StarDist. The hyperparameters of the StarDist structure, originally designed for cell and nuclei segmentation problems, were tuned in our implementations for the recognition of rings in SMLM image data. To do so, we adjusted the ZeroCostDL4Mic StarDist notebook to enable a long-running grid search for hyperparameter optimization as well as offline data analysis. The grid search yielded different sets of hyperparameters when optimizing for recall and precision, of which we chose the one that ensured a reasonable balance between the two criteria in all the six trained models. Supplemental Table S1 summarizes the values of the StarDist network’s hyperparameters used in our experiments. To avoid adding another cause for bias, no augmentation was used for training. Furthermore, transfer learning is not considered in our work. The batch size was set to 8 and the mean absolute error function was chosen as the loss function. Of the training dataset, 10% were used as a validation set to monitor possible overfitting during training.

To guarantee a fair comparison of the six different annotation strategies, we trained each pair of raw images and corresponding mask images on the same StarDist network with fixed hyperparameters. To train the StarDist network, the key adopted python packages included tensorflow (v 0.1.12), Keras (v 2.3.1), csbdeep (v 0.6.3), numpy (v 1.19.5), and cuda (v 11.1.105). The training process has been accelerated using a Tesla P100 GPU and depending on the model it took between 123 and 161 min.

### Ring recognition by means of trained models

For benchmarking, we ran the six trained models on the cellular test dataset and the synthetic ring dataset using the ZeroCostDL4Mic platform. For the ring recognition in the Septin-9<sup>c/c</sup> and Septin-9<sup>-/-</sup> cell data we exported the models from the ZeroCostDL4Mic platform and ran them using the StarDist Fiji plugin. The outputs generated by the models were mask images of the same size of the corresponding test image.

### Ring averaging and diameter measurement

To generate average ring images, the model output mask images were converted into binary images using Fiji (Schindelin *et al.*, 2012). The centers of the individual ring masks were calculated and fixed-size bounding boxes were created around the center points. Bounding boxes touching the image borders were excluded from further measurements. The bounding box ROIs were then together with the corresponding images fed to CellProfiler (Carpenter *et al.*, 2006) to extract per-ring crops from the corresponding raw images. Using a custom-written Fiji macro, we measured the average ring diameter as follows: the ring crops per condition were averaged and the diameter was determined by measuring horizontal line profiles with a line thickness of 1 pixel on 18 successive 10° rotations of the image (see Figure 1C). In each of these 18 intensity plots the peak-to-peak distance was measured and then averaged to calculate the ring diameter.

### Calculation of septin ring geometry

The septin ring circumference  $C$  was calculated based on the formula  $C = \pi * d$  where  $d$  is the measured ring diameter. The circumference was then divided by the approximate periodicity length of hetero-octameric or heterohexameric complexes as can be found in filaments including the Septin-2-NC interface gap in the open conformation as predicted by pdb structure data (pdb codes: 2QAG, 7M6J, 6UQQ, 5CYO). This was in line with the observed periodicity length of the hetero-octameric complex as found in filamentous septin assemblies in yeast (Bertin *et al.*, 2010, 2008; Byers and Goetsch, 1976; DeMay *et al.*, 2011; Frazier *et al.*, 1998; Kaplan *et al.*, 2015). The septin complex association angle was calculated by dividing 360° by the number of septin complexes found in the ring.

### Statistics

We used the precision metric to show how relevant the recognized objects are in comparison with the septin rings in the ground truth. It was calculated as  $TP / (TP + FP)$ , where  $TP$  stands for the number of true positives and  $FP$  is the number of falsely recognized objects. Recall, on the other hand, was used to show how well the recognized objects match the septin rings in the ground truth. Thus, it was calculated as  $TP / (TP + FN)$ , where  $FN$  is the number of relevant septin rings in the ground truth that were not recognized by the models. Both measures fall in the [0%, 100%] interval, with higher values corresponding to better septin ring recognition performance.  $R^2$  values were calculated in R (R Project) using the “lm” function; z-score values were calculated across the six models using the formula  $z = (x - \mu) / \sigma$ , where  $x$  is the variable (recall, precision, ring diameter, no of rings recognized),  $\mu$  is the mean over the group of variables of all six models, and  $\sigma$  is the SD from the mean.

### Data availability

All data generated in this study are available on request. All codes used to generate data in this paper can be found here: <https://github.com/AG-Ewers/SeptinRecognition>.

### ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of TRR 186 (Project Number 278001972) and SFB 958 and core funding through Freie Universität Berlin. SMLM imaging was performed in the Core Facility BioSupraMol facility of Freie Universität Berlin. The authors thank Richard C. Garratt for help with the measurement of the septin hetero-octamer dimensions based on pdb structures. Furthermore, the authors thank all members of the Ewers laboratory for helpful discussions.

## REFERENCES

- Asada A, Takahashi J, Taniguchi M, Yamamoto H, Kimura T, Saito T, Hisanaga S-I (2010). Neuronal expression of two isoforms of mouse Septin 5. *J Neurosci Res* 88, 1309–1316.
- Balzarotti F, Eilers Y, Gwosch KC, Gynnå AH, Westphal V, Stefani FD, Elf J, Hell SW (2016). Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science*. doi:10.1126/science.aak9913
- Banko M, Mucha Kruczynska I, Weise C, Heyd F, Ewers H (2019). A homozygous genome-edited Sept2-EGFP fibroblast cell line. *Cytoskeleton (Hoboken)* 76, 73–82.
- Bertin A, McMurray MA, Grob P, Park S-S, Garcia G, Patanwala I, Ng H-L, Alber T, Thorner J, Nogales E (2008). Saccharomyces cerevisiae septins: supramolecular organization of heterooligomers and the mechanism of filament assembly. *Proc Natl Acad Sci USA* 105, 8274–8279.
- Bertin A, McMurray MA, Thai L, Garcia G, Votin V, Grob P, Allyn T, Thorner J, Nogales E (2010). Phosphatidylinositol-4,5-bisphosphate promotes budding yeast septin filament assembly and organization. *J Mol Biol* 404, 711–731.
- Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, Davidson MW, Lippincott-Schwartz J, Hess HF (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642–1645.
- Byers B, Goetsch L (1976). A highly ordered ring of membrane-associated filaments in budding yeast. *J Cell Biol* 69, 717–721.
- Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, Guertin DA, Chang JH, Lindquist RA, Moffat J, et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7, R100.
- DeMay BS, Bai X, Howard L, Occhipinti P, Meseroll RA, Spiliotis ET, Oldenbourg R, Gladfelter AS (2011). Septin filaments exhibit a dynamic, paired organization that is conserved from yeast to mammals. *J Cell Biol* 193, 1065–1081.
- Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, Böhm A, Deubner J, Jäckel Z, Seiwald K, et al. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods* 16, 67–70.
- Fang L, Monroe F, Novak SW, Kirk L, Schiavon CR, Yu SB, Zhang T, Wu M, Kastner K, Latif AA, et al. (2021). Deep learning-based point-scanning super-resolution imaging. *Nat Methods* 18, 406–416.
- Frazier JA, Wong ML, Longtine MS, Pringle JR, Mann M, Mitchison TJ, Field C (1998). Polymerization of purified yeast septins: evidence that organized filament arrays may not be required for septin function. *J Cell Biol* 143, 737–749.
- Füchtbauer A, Lassen LB, Jensen AB, Howard J, Quiroga A de S, Warming S, Sørensen AB, Pedersen FS, Füchtbauer EM (2011). Septin9 is involved in septin filament formation and cellular stability. *Biol Chem* 392, 769–777.
- Gambarotto D, Zwettler FU, Le Guennec M, Schmidt-Cernohorska M, Fortun D, Borgers S, Heine J, Schloetel J-G, Reuss M, Unser M, et al. (2019). Imaging cellular ultrastructures using expansion microscopy (U-ExM). *Nat Methods* 16, 71–74.
- Garcia G, Bertin A, Li Z, Song Y, McMurray MA, Thorner J, Nogales E (2011). Subunit-dependent modulation of septin assembly: Budding yeast septin Shs1 promotes ring and gauze formation. *J Cell Biol*. doi:10.1083/jcb.201107123
- Gómez-de-Mariscal E, García-López-de-Haro C, Ouyang W, Donati L, Lundberg E, Unser M, Muñoz-Barrutia A, Sage D (2021). DeepImageJ: A user-friendly environment to run deep learning models in ImageJ. *Nat Methods* 18, 1192–1195.
- Gunzenhäuser J, Olivier N, Pengo T, Manley S (2012). Quantitative Super-Resolution Imaging Reveals Protein Stoichiometry and Nanoscale Morphology of Assembling HIV-Gag Virions. *Nano Lett*. doi:10.1021/nl3021076
- Heilemann M, van de Linde S, Schüttelpelz M, Kasper R, Seefeldt B, Mukherjee A, Tinnefeld P, Sauer M (2008). Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed Engl* 47, 6172–6176.
- Held M, Schmitz MHA, Fischer B, Walter T, Neumann B, Olma MH, Peter M, Ellenberg J, Gerlich DW (2010). CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nat Methods* 7, 747–754.
- Heydarian H, Joosten M, Przybylski A, Schueder F, Jungmann R, van Werkhoven B, Keller-Findeisen J, Ries J, Stallinga S, Bates M, Rieger B (2021). 3D particle averaging and detection of macromolecular symmetry in localization microscopy. *Nat Commun* 12, 2847.
- Hu J, Bai X, Bowen JR, Dolat L, Korobova F, Yu W, Baas PW, Svitkina T, Gallo G, Spiliotis ET (2012). Septin-driven coordination of actin and microtubule remodeling regulates the collateral branching of axons. *Curr Biol* 22, 1109–1115.
- Ihara M, Kinoshita A, Yamada S, Tanaka H, Tanigaki A, Kitano A, Goto M, Okubo K, Nishiyama H, Ogawa O, et al. (2005). Cortical organization by the septin cytoskeleton is essential for structural and mechanical integrity of mammalian spermatozoa. *Dev Cell* 8, 343–352.
- Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat Methods* 18, 203–211.
- Jones TR, Carpenter AE, Lamprecht MR, Moffat J, Silver SJ, Grenier JK, Castoreno AB, Eggert US, Root DE, Golland P, Sabatini DM (2009). Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proc Natl Acad Sci USA* 106, 1826–1831.
- Kaplan C, Ewers H (2015). Optimized sample preparation for single-molecule localization-based superresolution microscopy in yeast. *Nat Protoc* 10, 1007–1021.
- Kaplan C, Jing B, Winterlood CM, Bridges AA, Occhipinti P, Schmied J, Grinhagens S, Gronemeyer T, Tinnefeld P, Gladfelter AS, et al. (2015). The absolute arrangement of subunits in cytoskeletal septin filaments in cells measured by fluorescence microscopy. *Nano Lett*. doi:10.1021/acs.nanolett.5b00693
- Kaplan C, Steinmann M, Zapiorkowska NA, Ewers H (2017). Functional Redundancy of Septin Homologs in Dendritic Branching. *Front Cell Dev Biol* 5, 11.
- Kim MS, Froese CD, Estey MP, Trimble WS (2011). SEPT9 occupies the terminal positions in septin octamers and mediates polymerization-dependent functions in abscission. *J Cell Biol* 195, 815–826.
- Kinoshita M (2003). Assembly of mammalian septins. *J Biochem* 134, 491–496.
- Kinoshita M, Field CM, Coughlin ML, Straight AF, Mitchison TJ (2002). Self- and actin-templated assembly of Mammalian septins. *Dev Cell* 3, 791–802.
- Lagardère M, Chamma I, Bouilhol E, Nikolski M, Thoumine O (2020). FluoSIm: simulator of single molecule dynamics for fluorescence live-cell and super-resolution imaging of membrane proteins. *Sci Rep* 1–14. doi:10.1038/s41598-020-75814-y
- Lucas AM, Ryder PV, Li B, Cimini BA, Eliceiri KW, Carpenter AE (2021). Open-source deep-learning software for bioimage segmentation. *MBoC* 32, 823–829.
- Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* 1–12. doi:10.1038/s41593-018-0209-y
- McMurray MA, Thorner J (2019). Turning it inside out: The organization of human septin heterooligomers. *Cytoskeleton (Hoboken)* 76, 449–456.
- Mendonça DC, Macedo JN, Guimarães SL, Barroso da Silva FL, Cassago A, Garratt RC, Portugal RV, Araujo APU (2019). A revised order of subunits in mammalian septin complexes. *Cytoskeleton (Hoboken)* 76, 457–466.
- Mennella V, Keszthelyi B, McDonald KL, Chhun B, Kan F, Rogers GC, Huang B, Agard DA (2012). Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. *Nat Cell Biol* 14, 1159–1168.
- Moën E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D (2019). Deep learning for cellular image analysis. *Nat Methods* 1–14. doi:10.1038/s41592-019-0403-1
- Morone D, Marazza A, Bergmann TJ, Molinari M (2020). Deep learning approach for quantification of organelles and misfolded polypeptide delivery within degradative compartments. *MBoC* 31, 1512–1524.
- Mostowy S, Cossart P (2012). Septins: the fourth component of the cytoskeleton. *Nat Rev Mol Cell Biol* 13, 183–194.
- Mund M, van der Beek JA, Deschamps J, Dmitrieff S, Hoess P, Monster JL, Picco A, Nédélec F, Kaksonen M, Ries J (2018). Systematic nanoscale analysis of endocytosis links efficient vesicle formation to patterned actin nucleation. *Cell* 174, 884–896.e17.
- Nehme E, Freedman D, Gordon R, Ferdman B, Weiss LE, Alalouf O, Naor T, Orange R, Michaeli T, Shechtman Y (2020). DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning. *Nat Methods* 1–10. doi:10.1038/s41592-020-0853-5
- Ouyang W, Aristov A, Lelek M, Hao X, Zimmer C (2018). Deep learning massively accelerates super-resolution localization microscopy. *Nat Biotechnol* 36, 460–468.
- Piccinini F, Balassa T, Szkalitsy A, Molnar C, Paavolaainen L, Kujala K, Buzas K, Sarazova M, Pietiäinen V, Kutay U, et al. (2017). Advanced

- cell classifier: user-friendly machine-learning-based software for discovering phenotypes in high-content imaging data. *Cell Syst* 4, 651–655.e5.
- Platonova E, Winterflood CM, Ewers H (2015). A simple method for GFP- and RFP-based dual color single-molecule localization microscopy. *ACS Chem Biol* 10, 1411–1416.
- Ries J, Kaplan C, Platonova E, Eghlidi H, Ewers H (2012). A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. *Nature Methods* 9, 582–584.
- Rust MJ, Bates M, Zhuang X (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* 3, 793–795.
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682.
- Schmidt U, Weigert M, Broaddus C, Myers G (2018). Cell detection with Star-convex polygons. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. Lecture Notes in Computer Science()*, eds. A Frangi, J Schnabel, C Davatzikos, C Alberola-López and G Fichtinger, Cham, Switzerland: Springer, vol 11071, [https://doi.org/10.1007/978-3-030-00934-2\\_30](https://doi.org/10.1007/978-3-030-00934-2_30).
- Segebarth D, Griebel M, Stein N, von Collenberg CR, Martin C, Fiedler D, Comeras LB, Sah A, Schoeffler V, Lüffe T, et al. (2020). On the objectivity, reliability, and validity of deep learning enabled bioimage analyses. *eLife* 9, e59780.
- Sellin ME, Holmfeldt P, Stenmark S, Gullberg M (2011). Microtubules support a disc-like septin arrangement at the plasma membrane of mammalian cells. *Mol Biol Cell*. doi:10.1091/mbc.E11-09-0754
- Sellin ME, Stenmark S, Gullberg M (2012). Mammalian SEPT9 isoforms direct microtubule-dependent arrangements of septin core heteromers. *Mol Biol Cell* 23, 4242–4255.
- Sieben C, Banterle N, Douglass KM, Gönczy P, Manley S (2018). Multicolor single-particle reconstruction of protein complexes. *Nat Methods* 1–9. doi:10.1038/s41592-018-0140-x
- Sochacki KA, Dickey AM, Strub M-P, Taraska JW (2017). Endocytic proteins are partitioned at the edge of the clathrin lattice in mammalian cells. *Nat Cell Biol*. doi:10.1038/ncb3498
- Sommer C, Gerlich DW (2013). Machine learning in cell biology - teaching computers to recognize phenotypes. *J Cell Sci* 126, 5529–5539.
- Sommer C, Hoefler R, Samwer M, Gerlich DW (2017). A deep learning and novelty detection framework for rapid phenotyping in high-content screening. *Mol Biol Cell* 28, 3428–3436.
- Song K, Gras C, Capin G, Gimber N, Lehmann M, Mohd S, Puchkov D, Rödiger M, Wilhelmi I, Daumke O, et al. (2019). A SEPT1-based scaffold is required for Golgi integrity and function. *J Cell Sci* 132, jcs225557.
- Soroosh F, Kim MS, Palander O, Balachandran Y, Collins RF, Benlekbir S, Rubinstein JL, Trimble WS (2021). Revised subunit order of mammalian septin complexes explains their in vitro polymerization properties. *Mol Biol Cell* 32, 289–300.
- Speiser A, Müller L-R, Hoess P, Matti U, Obara CJ, Legant WR, Kreshuk A, Macke JH, Ries J, Turaga SC (2021). Deep learning enables fast and dense single-molecule localization with high accuracy. *Nat Methods* 18, 1082–1090.
- Stringer C, Wang T, Michaelos M, Pachitariu M (2020). Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* 1–21. doi:10.1038/s41592-020-01018-x
- Szymborska A, de Marco A, Daigle N, Cordes VC, Briggs JAG, Ellenberg J (2013). Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. *Science*. doi:10.1126/science.1240672
- von Chamier L, Laine RF, Jukkala J, Spahn C, Krentzel D, Nehme E, Lerche M, Hernández-Pérez S, Mattila PK, Karinou E, et al. (2021). Democratizing deep learning for microscopy with ZeroCostDL4Mic. *Nat Commun* 12, 2276.
- Weigert M, Schmidt U, Boothe T, Müller A, Dibrov A, Jain A, Wilhelm B, Schmidt D, Broaddus C, Culley S, et al. (2018). Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat Methods* 15, 1090–1097.
- Weirich CS, Erzberger JP, Barral Y (2008). The septin family of GTPases: architecture and dynamics. *Nat Rev Mol Cell Biol* 9, 478–489.
- Willig KI, OO RS, Westphal V, Jahn R, Hell SW (2006). STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis. *Nature* 440, 935–939.
- Zhang J, Wang Z, Chen Y, Han R, Liu Z, Sun F, Zhang F (2019). PIXER: an automated particle-selection method based on segmentation using a deep neural network. *BMC Bioinformatics* 20, 41.
- Zwettler FU, Spindler MC, Reinhard S, Klein T, Kurz A, Benavente R, Sauer M (2020). Tracking down the molecular architecture of the synaptonemal complex by expansion microscopy. *Nat Commun* 11, 3222.