Eurographics Workshop on Visual Computing for Biology and Medicine (2021)
K. Nieselt and T. Schultz (Editors)

Short Paper

Polar space based shape averaging for star-shaped biological objects

K. Ruzaeva[1,2], K. Nöh[2] and B. Berkels [1]

[1]Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Aachen, Germany
[2]Institute of Bio- and Geosciences, IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany

Abstract

In this paper, we propose an averaging method for expert segmentation proposals of microbial organisms, resulting in a smooth, naturally looking segmentation ground truth. The approach exploits a geometrical property of the majority of the organisms – star-shapedness – and is based on contour averaging in polar space. It is robust and computationally efficient, where robustness is due to the absence of tuneable parameters. Moreover, the algorithm preserves the uncertainty (in terms of the standard deviation) of the experts' opinion, which allows to introduce an uncertainty-aware metric for estimation of the segmentation quality. This metric emphasizes the influence of ground truth regions with low variance. We study the performance of the proposed averaging method on time-lapse microscopy data of Corynebacterium glutamicum and the uncertainty-aware metric on synthetic data.

CCS Concepts

• ***Applied computing*** → *Imaging;* • ***Computing methodologies*** → *Image processing;*

1. Introduction

How bacteria maintain their shape and size is one of the big open questions of life [HT18]. Microfluidic single-cell analysis coupled with live-cell imaging microscopy is a versatile tool to study the diversity of cell shapes and sizes as well as their adaption with spatio-temporal resolution [LWK*19, GPP*12]. Many bacterial species, such as the gut bacterium *Escherichia coli* or the soil bacteria *Corynebacterium glutamicum* and *Bacillus subtilis* evolved into rod-shaped morphologies, with shapes ranging from spheric cocci to stretched, round-ended cylinders [SM15]. To quantitatively characterize every cell within each image frame captured in microfluidic single-cell experiments, multi-object segmentation is used. Here, Deep Learning (DL)-based segmentation methods, most prominently of U-Net type [RFB15], achieve state-of-the-art performance. The prediction accuracy of DL-based methods highly depends on the training data quality. For accurate predictions, high-quality (i.e. pixel-accurate) training data has to be labeled by the domain experts, which is, however, time-intense [LLS*19]. Despite their simplistic shape, in the abundant case of low-resolution and low-signal-to-noise ratio data, it is difficult to annotate images (i.e. to draw cell outlines), even for domain experts. In this case, repeated drawing of the desired segmentation by several raters (experts) is the only acceptable approach to come to a consensus ground truth (GT).

The variability in the annotations has two major origins:

- **Technical:** Different input devices (i.e. graphic tablet or mouse), which experts use to draw the cell outlines. Indeed, although leading to more coarse segmentation results, using a mouse as an input tool for a training data creation is the predominant approach, arguably due to being the cheaper and more abundant alternative.
- **Methodological:** Since there are no commonly established rules about the localization of cell borders, raters opinions may differ. In this case, the segmentation of one rater tends to be consistently wider/longer or narrow/shorter.

Therefore, a technique to derive a consensus or GT data from segmentation proposals of the experts is desired, which respects and preserves the uncertainty, induced by the inter-rater variation.

1.1. Related work

The averaging of the raters' proposals can be generalized to the average shape problem, which may be solved with several traditional approaches. Simple landmarks methods [Ray92, DTC*02] are based on a distance averaging between corresponding landmarks, important geometrical features, require manual laborious landmarks' specification, or an automatic derivation of the landmarks. Alternatively, methods based on the Fourier approximation of a closed contour and averaging of the corresponding Fourier coefficients for a given number of harmonics [KG82] find their application in [SKKN*19] and provide a smooth average outline of the average object, but highly depend on the number of the harmonics one approximates the contour with. As another option, variational-based averaging methods [BLR10] are dealing with sophisticated non-aligned objects, and may even provide statistical information in addition to the average shape [RW10]. However, these methods are computationally quite expensive and, due to the usually

DOI: 10.2312/vcbm.20211340

non-convex objective functionals involved, require careful initialization to avoid local minima in the optimization. A widely used algorithm, which is specifically aimed for a GT fusion given by different raters and has a user-friendly implementation [YLJB17], is the Simultaneous Truth And Performance Level Estimation (STAPLE) [WZW04]. STAPLE, however, does not produce uncertainty estimates.

1.2. This work

To alleviate the aforementioned limitations of existing approaches, we propose a new conceptually simple and fast averaging method. Our proposed method is based on the observation that most of the microbial cells are star-shaped. Therefore, we explore the use of the polar transform to find the average object outline to create ground truth for training and validation purposes of segmentation methods. The created ground truth retains information about the variation in segmentation proposals in terms of the standard deviation and thus allows to create an uncertainty-based metric for segmentation quality.

2. Methods

The gram-positive soil bacterium *C. glutamicum* is one of the major platform hosts for the biotechnological production of various amino acids, biochemicals, biofuels and proteins [EB05, BKD*18]. In the literature, it is described as a slightly bent rod (Fig. 1). One important geometrical aspect of *C. glutamicum* is that it is star-shaped: A set S in the Euclidean space $\mathbb{R}^n$ is called star-shaped if there exists $x_0 \in S$ such that for all $x \in S$ the line segment from x_0 to x is in S [HHMM20]. A point x_0 is called center of S.

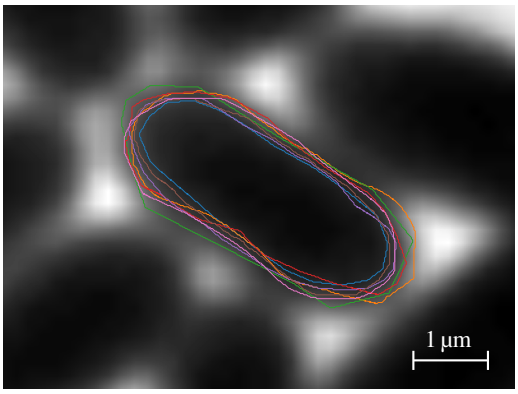

Figure 1: *Cutout from a phase-contrast microscopy image. Dark objects are* C. glutamicum *cells. Annotations from seven experts are shown for the central cell.*

Besides *C. glutamicum*, there are other important star-shaped organisms, i.e. *E. coli* and *S. cerevisiae* [Fel10] (mother cell and buds are considered separately), and also recent image processing methods explicitly exploit the star-shapedness [WSH*20, SWBM18]. Motivated by this, we consider star-shaped objects in this paper and use *C. glutamicum* to illustrate our proposed method. The method is not limited to this organism though, but applicable to the wide range of rod- and spherical-shaped bacteria.

2.1. Experimental setup

The test dataset consists of six images of *C. glutamicum*, where each image shows up to six cells (Table 1) and parts of neighboring cells (exemplified in Figure 1, which shows the image corresponding to Row 3 of Table 1). Nine experts were asked to annotate (manually segment) fully present cells with a mouse or stylus pen, where each cell is supposed to be segmented separately. The annotation was performed using `Hasty.ai`, a powerful, yet easy to use online annotation platform. To get a rough estimate of the inter-rater variability, we computed basic statistics of the segmentation proposals for some morphological features of the objects. Specifically, the resulting average length and width were calculated according to [FKCH19]. The length is measured along the longest middle line of the bounding (rotated) rectangle with minimum area and is $2.3037 \pm 0.1057\,\mu m$. The width is derived as the average of eleven equidistant width segments, parallel to the short middle line of the rectangle and is $0.8966 \pm 0.0505\,\mu m$. Finally, the area was computed as a sum of labeled pixels for each cell and determined to be $19.6530 \pm 2.6454\,\mu m^2$.

2.2. The polar average of star-shaped sets

An important property of star-shaped objects is that their boundary or contour can be expressed as a graph in polar coordinates. Let $R \subset \mathbb{R}^2$ be star-shaped with center $(x_0, y_0) \in R$ and boundary $C = \partial R$. The corresponding transformation to polar coordinates is

$$\begin{aligned} r &: \mathbb{R}^2 \to [0,\infty),\ (x,y) \mapsto r(x,y) = \sqrt{(x-x_0)^2 + (y-y_0)^2}, \\ \theta &: \mathbb{R}^2 \to (-\pi,\pi],\ (x,y) \mapsto \theta(x,y) = \operatorname{arctan2}(x-x_0, y-y_0). \end{aligned} \tag{1}$$

Since R is star-shaped, the polar transformed boundary

$$\mathcal{P} = \{(\theta(x,y), r(x,y)) : (x,y) \in C\} \tag{2}$$

is a graph over $(-\pi,\pi]$, i.e. there is a function $f : (-\pi,\pi] \to [0,\infty)$ such that $\mathcal{P} = \{(\theta, f(\theta)) : \theta \in (-\pi,\pi]\}$. This polar representation of star-shaped objects is the key ingredient for our contour averaging approach.

Let $R_1, \ldots, R_N \subset \mathbb{R}^2$ be N star-shaped sets, each representing a rater proposal for the segmentation of a given object, and all with a common center (x_0, y_0). Assuming that there is no consistent bias of the rater proposals, a suitable average of the proposals should be a good estimate of the GT segmentation. For $i = 1, \ldots, N$, let f_i be the graph function from the polar representation of ∂R_i. Then, the average of these f_i, i.e. $f := \frac{1}{N}\sum_{i=1}^{N} f_i$, is the graph function of the polar representation $\mathcal{P}$ of the contour of the averaged rater proposals. Using the inverse transform of (1), i.e.

$$[0,\infty) \times (-\pi,\pi],\ (r,\theta) \mapsto (r\cos(\theta) + x_0, r\sin(\theta) + y_0), \tag{3}$$

on $\mathcal{P}$, we get the contour of the rater proposal average and thus our GT estimate.

2.3. Preprocessing and contour averaging

For real data, rater proposal are given as binary pixel images $B_1, \ldots B_N$, i.e. matrices whose entries only take the values zero or one. From each such mask B_i, we extract the contour of the corresponding proposal as a list of (x,y) coordinates by applying the

discrete Laplace operator to the binary image and then extracting the positions of the zero crossings.

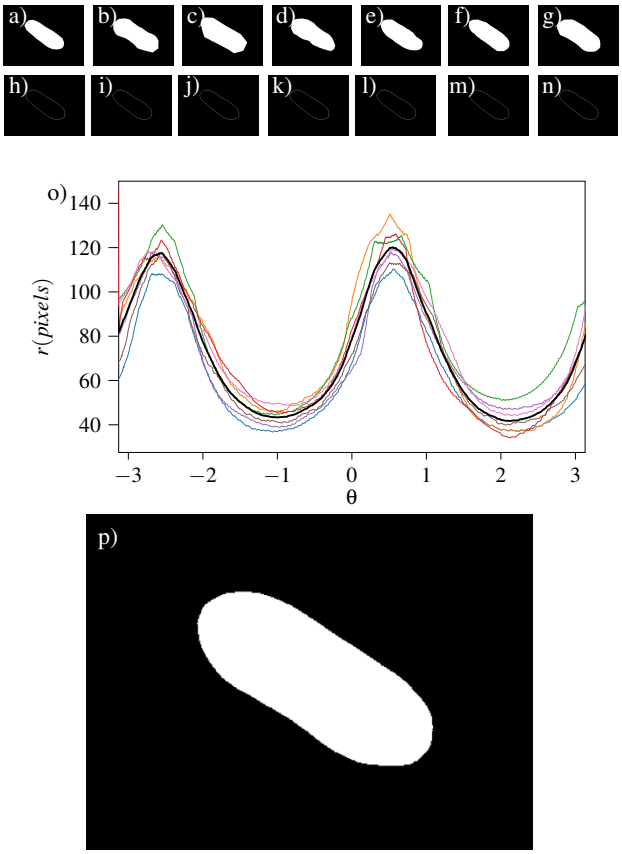

Figure 2: *Averaging GT pipeline. Segmentation proposals (a)-(g), their contours (h)-(n). Contours of the segmentation proposals in polar space and their average (black) (o). The resulting average of segmentation proposals (p).*

The center of mass of the sum of the masks is used as estimate of the common center, (x_0, y_0), and each list of contour coordinates is transformed to polar coordinates using (1) (note that, in a brief sensitivity analysis, we found that the influence of changing (x_0, y_0) was insignificant for our experiments). Thus, for each B_i, we get a list of polar contour coordinates $(\theta_{i,1}, r_{i,1}), \ldots, (\theta_{i,K_i}, r_{i,K_i})$, sorted by θ in ascending order. Here, the number of entries $K_i \in \mathbb{N}$ depends on the mask and the spacing of the $\theta_{i,j}$ is non-uniform. Using linear interpolation, as a trade-off between simplicity and computational complexity, of the polar contour coordinate vector, we get a continuous function $R_i : (-\pi, \pi] \to [0, \infty)$ such that $R_i(\theta_{i,j}) = r_{i,j}$ for $j = 1, \ldots, K_i$. The average of these R_i gives us the polar representation of the average contour, which is then sampled at angles $\theta_i, \ldots, \theta_K \in (-\pi, \pi]$, i.e.

$$R_{a,i} = \frac{1}{N} \sum_{n=0}^{N-1} R_n(\theta_i) \tag{4}$$

We select θ as a vector of uniformly distributed points in $(-\pi, \pi]$, which should not be shorter than the longest vector of proposed object boundaries, i.e. $K \geq \max_{i=1,\ldots,N} K_i$. For the sake of simplicity, we used $K = 1000$ for all experiments, to exceed the length of the longest proposed contour in our test dataset. Finally, the polar space coordinate vector is transformed back to real space using (3), rounding the results to the nearest pixel position. To obtain the average mask I_a, the resulting outline is filled. To fill the outline, we use the `fillPoly` function from [Bra00], which fills an area bounded by a polygonal (composed by the outline coordinates) contour.

2.4. Uncertainty-aware segmentation quality metric

One byproduct of the polar average (4) is that each R_{a_i} is associated with a standard deviation

$$\sigma_{a,i} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (R_n(\theta_i) - R_{a,i})^2} \tag{5}$$

encoding the inter-rater variability. Usually, inter-rater deviation in segmentation is inevitable. Moreover, unevenly distributed deviation may indicate varying difficulty to locate the boundary of an object in different regions (e.g. it is easier to segment cells in non-crowded colonies with no other cells around than in regions with densely packed cells). Most of the available metrics for the segmentation quality estimation are based on spatial overlap (e.g. Dice or Jacquard scores), and do not take this kind of uncertainty into account. However, to avoid bias, the estimation of the segmentation quality of a given algorithm using a "fuzzy" GT segmentation, should take into account this uncertainty. In other words, an uncertainty-aware metric is desired.

We propose to use a weighted root-mean-square error (WRMSE), where the weights are inversely proportional to the standard deviation of the object boundary to emphasize the influence of the contour points with no (or low) variation. The suggested metric takes into account every boundary pixel, while respecting the uncertainty.

The root-mean-square error (RMSE) is a common distance-based metric to evaluate the performance of an algorithm [JdC18] and illustrates the average distance between each of the n predicted contour points in polar space and the corresponding GT (average) contour points. In our case, the RMSE is:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (R_{\text{pred}}(\theta_i) - R_a(\theta_i))^2} \tag{6}$$

where R_{pred} is the predicted contour in polar space and R_a is the GT contour. Based on this, we define the WRMSE as follows:

$$WRMSE = \sqrt{\frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=0}^{n-1} w_i (R_{\text{pred}}(\theta_i) - R_a(\theta_i))^2} \tag{7}$$

where w is a vector of positive weights. To avoid division by zero and the extreme influence of low deviation regions, we use an exponential weighting of the form $w = e^{-\sigma}$, i.e.

$$w_i = e^{-\sigma_{a,i}} \tag{8}$$

3. Results

3.1. Ground truth averaging results

The results of the averaging method are examined in terms of the similarity to the segmentation proposals, using Dice score, and visual contour smoothness and compared with the STAPLE algorithm for different thresholds and the median of the segmentation proposals. The Dice score is defined as:

$$\mathrm{Dice}(I_a, (I_i)_{i=1}^N) = \frac{1}{N}\sum_{i=1}^{N}\frac{2|I_a \cap I_i|}{|I_a| + |I_i|} \tag{9}$$

where I_i is the i-th segmentation proposal for the given image and I_a is the average mask.

The median of the segmentation proposals is defined as pixel-wise median across the proposals and thus equivalent to the shape average with respect to the L^2-norm. To obtain a binary mask with an even number of raters and an equal amount of votes, the value 1 is chosen (instead of 0.5). The Dice score in Table 1 was calculated as the average of the Dice score for every cell (not image). Thus, for the pictures 4-6, every cell was processed separately.

I n.	Mean	Median	S 0.85	S 0.95	PA
1						
2						
3						
4						
5						
6						
Dice Score		0.9357	0.9348	0.9342	0.9331

Table 1: *The result of different averaging methods: Median, STAPLE with different thresholds (S 0.85 and S 0.95), and the proposed method based on the contour averaging in polar space (PA).*

The result in the Table 1 shows that, while all methods are almost indistinguishable in terms of the Dice score, the proposed method provides smooth, naturally looking average results (Figure 3) that do not inherit the coarse edges proposed by experts and which stem from the labeling with a mouse. This is an important property, for instance for training data generation, since ML-based algorithms could learn the artifacts of the provided training data.

3.2. Uncertainty-aware metric

Since the deviation of the expert proposals used to create the ground truth above is distributed equally over the cell boundary, the metric

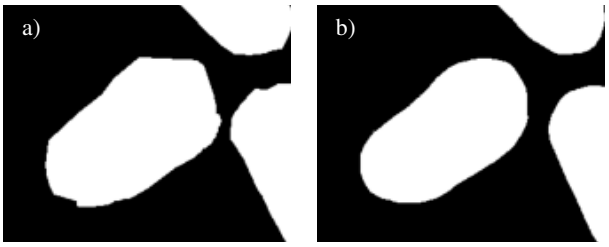

Figure 3: *Comparison of the results obtained by STAPLE(85) (S 85 of Table 1) (a) and the proposed method (b), which shows the more natural smoothness of the proposed method compared to STAPLE.*

will be examined with synthetic segmentation proposals. Five binary rod-shaped masks (segmentation proposals) were generated, where the width of the rod is well-defined, and the length deviates over the proposals (Figure 4 (a)-(e)).

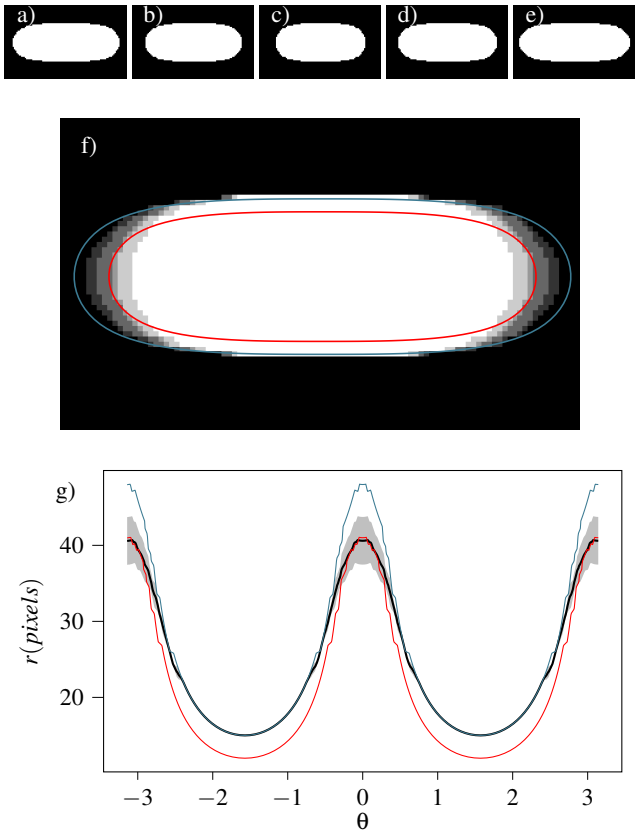

Figure 4: *Synthetic segmentation proposals (a)-(e). Contours to examine (red and blue) plotted over a sum of the segmentation proposals (f). Average cell shape contour (black) with a standard deviation, and the contours of interest (g).*

Considering two contours (blue and red), where one is matching the length of the GT and another is matching the width, as the result of the segmentation algorithm, the WRMSE and the Dice score are

Table 2: *Dice score and RMSE for* 2 *contours*

	Blue	Red
Dice	0.9044	0.9172
WRMSE	1.0482	3.4881

calculated, cf. Table 2. Despite the fact that the red contour shows a better Dice score (bigger overlap with the average mask), it does not match the regions with the highest confidence (low variation), which is reflected in the WRMSE. On the other hand, the blue contour perfectly matches low variation regions, explaining the lower (better) WRMSE in this case.

4. Conclusions

We here introduce an averaging method for segmentation proposals of star-shaped objects provided by experts. The proposed method is based on averaging in polar space, it is simple, computationally efficient, provides smooth, naturally looking outlines, and preserves the uncertainty derived from the variation of the segmentation proposals in terms of the standard deviation. The obtained uncertainty information is used to construct a weighted root-mean-square error, which is useful as a metric for segmentation quality when the segmentation GT has uncertainty that noticeably varies along the boundary.

Acknowledgements The authors gratefully acknowledge the contributions of the experts, members of the departments of Micro Scale Bioengineering (FZ Juelich) and Multiscale Bioengineering (Bielefeld University), who performed the labelling of the images.

Funding This work was performed as part of the Helmholtz School for Data Science in Life, Earth and Energy (HDS-LEE) and received funding from the Helmholtz Association of GRC.

References

[BKD*18] BARITUGO K.-A., KIM H. T., DAVID Y., CHOI J.-I., HONG S. H., JEONG K. J., CHOI J. H., JOO J. C., PARK S. J.: Metabolic engineering of Corynebacterium glutamicum for fermentative production of chemicals in biorefinery. *Applied Microbiology and Biotechnology 102*, 9 (2018), 3915–3937. doi:10.1007/s00253-018-8896-6. 2

[BLR10] BERKELS B., LINKMANN G., RUMPF M.: An SL(2) invariant shape median. *Journal of Mathematical Imaging and Vision 37*, 2 (Feb. 2010), 85–97. doi:10.1007/s10851-010-0194-6. 1

[Bra00] BRADSKI G.: The OpenCV Library. *Dr. Dobb's Journal of Software Tools* (2000). 3

[DTC*02] DAVIES R., TWINING C., COOTES T., WATERTON J., TAYLOR C.: A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging 21*, 5 (May 2002), 525–537. doi:10.1109/tmi.2002.1009388. 1

[EB05] EGGELING L., BOTT M. (Eds.): *Handbook of Corynebacterium glutamicum*. CRC Press, Mar. 2005. 2

[Fel10] FELDMANN H.: *Yeast : molecular and cell biology*. Wiley-VCH, Weinheim, 2010. 2

[FKCH19] FACCHETTI G., KNAPP B., CHANG F., HOWARD M.: Reassessment of the basis of cell size control based on analysis of cell-to-cell variability. *Biophysical Journal 117*, 9 (Nov. 2019), 1728–1738. doi:10.1016/j.bpj.2019.09.031. 2

[GPP*12] GRÜNBERGER A., PACZIA N., PROBST C., SCHENDZIELORZ G., EGGELING L., NOACK S., WIECHERT W., KOHLHEYER D.: A disposable picolitre bioreactor for cultivation and investigation of industrially relevant bacteria on the single cell level. *Lab on a Chip 12*, 11 (2012), 2060. doi:10.1039/c2lc40156h. 1

[HHMM20] HANSEN G., HERBURT I., MARTINI H., MOSZYŃSKA M.: Starshaped sets. *Aequationes mathematicae 94*, 6 (May 2020), 1001–1092. doi:10.1007/s00010-020-00720-7. 2

[HT18] HARRIS L. K., THERIOT J. A.: Surface area to volume ratio: A natural variable for bacterial morphogenesis. *Trends in Microbiology 26*, 10 (Oct. 2018), 815–832. doi:10.1016/j.tim.2018.04.008. 1

[JdC18] JOHNSTON B., DE CHAZAL P.: A review of image-based automatic facial landmark identification techniques. *EURASIP Journal on Image and Video Processing 2018*, 1 (Sept. 2018). doi:10.1186/s13640-018-0324-4. 3

[KG82] KUHL F. P., GIARDINA C. R.: Elliptic fourier features of a closed contour. *Computer Graphics and Image Processing 18*, 3 (Mar. 1982), 236–258. doi:10.1016/0146-664x(82)90034-x. 1

[LLS*19] LEYGEBER M., LINDEMANN D., SACHS C. C., KAGANOVITCH E., WIECHERT W., NÖH K., KOHLHEYER D.: Analyzing microbial population heterogeneity—expanding the toolbox of microfluidic single-cell cultivations. *Journal of Molecular Biology 431*, 23 (Nov. 2019), 4569–4588. doi:10.1016/j.jmb.2019.04.025. 1

[LWK*19] LINDEMANN D., WESTERWALBESLOH C., KOHLHEYER D., GRÜNBERGER A., VON LIERES E.: Microbial single-cell growth response at defined carbon limiting conditions. *RSC Advances 9*, 25 (2019), 14040–14050. doi:10.1039/c9ra02454a. 1

[Ray92] RAY T. S.: Landmark eigenshape analysis: Homologous contours: Leaf shape in syngonium (araceae). *American Journal of Botany 79*, 1 (Jan. 1992), 69. doi:10.2307/2445199. 1

[RFB15] RONNEBERGER O., FISCHER P., BROX T.: U-net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science*. Springer International Publishing, 2015, pp. 234–241. doi:10.1007/978-3-319-24574-4_28. 1

[RW10] RUMPF M., WIRTH B.: An elasticity-based covariance analysis of shapes. *International Journal of Computer Vision 92*, 3 (June 2010), 281–295. doi:10.1007/s11263-010-0358-2. 1

[SKKN*19] SAKAMOTO L., KAJIYA-KANEGAE H., NOSHITA K., TAKANASHI H., KOBAYASHI M., KUDO T., YANO K., TOKUNAGA T., TSUTSUMI N., IWATA H.: Comparison of shape quantification methods for genomic prediction, and genome-wide association study of sorghum seed morphology. *PLOS ONE 14*, 11 (Nov. 2019), e0224695. doi:10.1371/journal.pone.0224695. 1

[SM15] SATTLEY W., MADIGAN M.: Microbiology. *Encyclopedia of Life Sciences (eLS)* (Aug. 2015). doi:10.1002/9780470015902.a0000459.pub2. 1

[SWBM18] SCHMIDT U., WEIGERT M., BROADDUS C., MYERS G.: Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Springer International Publishing, 2018, pp. 265–273. doi:10.1007/978-3-030-00934-2_30. 2

[WSH*20] WEIGERT M., SCHMIDT U., HAASE R., SUGAWARA K., MYERS G.: Star-convex polyhedra for 3d object detection and segmentation in microscopy. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)* (Mar. 2020), IEEE. doi:10.1109/wacv45572.2020.9093435. 2

[WZW04] WARFIELD S., ZOU K., WELLS W.: Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. *IEEE Transactions on Medical Imaging 23*, 7 (July 2004), 903–921. doi:10.1109/tmi.2004.828354. 2

[YLJB17] YANIV Z., LOWEKAMP B. C., JOHNSON H. J., BEARE R.: SimpleITK image-analysis notebooks: a collaborative environment for education and reproducible research. *Journal of Digital Imaging 31*, 3 (Nov. 2017), 290–303. doi:10.1007/s10278-017-0037-8. 2