Non-Invasive Single-Cell Morphometry in Living Bacterial Biofilms

Mingxing Zhang[1,§], Ji Zhang[1,§], Yibo Wang[1,§], Jie Wang[3], Alecia M. Achimovich[2], Scott T. Acton[3], and Andreas Gahlmann[1,2,*]

[1]Department of Chemistry, University of Virginia, Charlottesville, Virginia, USA

[2]Department of Molecular Physiology & Biological Physics, University of Virginia School of Medicine, Charlottesville, Virginia, USA

[3]Department of Electrical & Computer Engineering, University of Virginia, Charlottesville, Virginia, USA

§These authors contribute equally to this paper

*Corresponding Author

Abstract

Fluorescence microscopy enables spatial and temporal measurements of live cells and cellular communities. However, this potential has not yet been fully realized for investigations of individual cell behaviors and phenotypic changes in dense, three-dimensional (3D) bacterial biofilms. Accurate cell detection and cellular shape measurement in densely packed biofilms are challenging because of the limited resolution and low signal to background ratio in fluorescence microscopy images. In this work, we present Bacterial Cell Morphometry 3D (*BCM3D*), an image analysis workflow that combines deep learning with mathematical image analysis to accurately segment and classify single bacterial cells in 3D fluorescence images. In *BCM3D*, deep convolutional neural networks (CNNs) are trained using simulated biofilm images with experimentally realistic signal-to-background ratios (SBRs), cell densities, labelling methods, and cell shapes. We systematically evaluated the segmentation accuracy of *BCM3D* using both simulated and experimental images. Compared to state-of-the-art bacterial cell segmentation approaches, *BCM3D* consistently achieves higher segmentation accuracy and further enables automated morphometric cell classifications in multi-population biofilms.

Introduction

Biofilms are multicellular communities of microorganisms that grow on biotic or abiotic surfaces[1-4]. In addition to cellular biomass, biofilms also contain an extracellular matrix (ECM) which is composed of polysaccharides, DNA, and proteins. Individual cells in biofilms interact with other cells, the ECM, or with the substrate surface, and the sum total of these interactions provide bacterial biofilms with emergent functional capabilities beyond those of individual cells. For example, biofilms are orders of magnitude more tolerant towards physical, chemical, and biological stressors, including antibiotic treatments and immune system clearance[1,2,5-8]. Understanding how such capabilities emerge from the coordination of individual cell behavior requires imaging technologies capable of resolving and simultaneous tracking of individual bacterial cells in 3D biofilms.

Live cell-compatible imaging technologies, such as light microscopy, can reveal the spatial and temporal context that affects cellular behaviors. However, conventional imaging modalities are not able to resolve individual cells within thick 3D biofilms over extended periods of time. For example, the diffraction-limited lateral *x,y*-resolution (~230 nm) of a confocal fluorescence microscope is barely sufficient to resolve bacterial cells positioned next to each other on flat glass coverslips. Even worse, the diffraction-limited axial *z*-resolution (570 nm) is comparable to the size of a single bacterial cell, so that densely-packed cells become unresolvable in the axial *z*-dimension[9,10]. Notable exceptions include loose biofilms (low cell density), spherical cell shapes[11,12], and mutant *Vibrio cholera* biofilms, in which cell-cell spacing is increased through the overproduction of ECM materials[13-15]. While single-cell resolved images have been obtained in such special situations, conventional optical microscopy modalities are not generally capable to accurately resolve and quantitatively track individual cells in dense 3D biofilms.

Super-resolution derivatives of confocal microscopy, known as Image Scanning Microscopy[16], can improve spatial resolution. However, a perhaps more important limitation for long-term live-cell imaging is photodamage to the specimen (phototoxicity) and to the fluorophores used for labeling (photobleaching)[17-19]. In confocal microscopy-based approaches, undesired out-of-focus fluorescence emission is filtered out by confocal pinholes to yield optically-sectioned images with high contrast, i.e. high signal-to-background ratios (SBRs). However, repeated illumination of out-of-focus regions during laser scanning and high light intensities at the focal volume result in rapid photobleaching of fluorophores and unacceptable phototoxicity for light sensitive specimens[17-19]. In fact, confocal fluorescence microscopy (as well as its super-resolution derivatives) uses illumination light intensities that are two to three orders of magnitude higher than the light intensities under which life has evolved[18]. The high rates of phototoxicity and photobleaching make confocal-based microscopy unsuitable for high frame-rate time-lapse imaging of living specimens over many hours and days[14,15,17,20,21].

In recent years, light sheet-based fluorescence excitation and imaging approaches have been developed to overcome the drawbacks of confocal microscopy. Among these, lattice light sheet microscopy (LLSM)[18,19] and field synthesis variants thereof[22], axially-swept light sheet microscopy (ASLM)[23,24], and single-objective oblique plane light sheet microscopes[25-30] now combine excellent 3D spatial resolution with fast temporal resolution and low phototoxicity at levels that cannot be matched by confocal microscopy. Specifically, light sheet-based microscopy approaches can operate at illumination intensities that are below the levels of cellular phototoxicity, even for notoriously light sensitive specimens, and reduce fluorophore photobleaching by 20-50 times compared to confocal microscopy, while maintaining comparable spatial resolution and contrast/SBR[18,26].

81 An additional challenge in high-resolution biofilm imaging is data quantification. Even if
82 sufficient resolution and high SBRs can be achieved to visually discern, i.e. qualitatively resolve
83 individual cells, robust computational algorithms are still needed for automated cell segmentation
84 and quantitative cell tracking. Towards this goal, image processing approaches based on the
85 watershed technique and intensity thresholding have been developed over the years for single-cell
86 segmentation in bacterial biofilms[13-15,21]. The broad applicability of watershed- and threshold-
87 based image processing algorithms is limited, because these algorithms require manual
88 optimization of many user-selected parameters. Even with optimal parameters, watershed- and
89 threshold-based image processing methods produce sub-optimal segmentation results, especially
90 when cell densities are high, when SBRs are low, and when cellular fluorescence intensities are
91 not uniform across the cytosol or on the cell surface. To overcome the drawbacks of watershed-
92 and threshold-based image processing approaches, deep learning by CNNs have been used in
93 recent years with outstanding success for a wide range of problems in biomedical image analysis[31].
94 Here, we present Bacterial Cell Morphometry 3D (*BCM3D*)[32], a generally applicable
95 workflow for single-cell segmentation and shape determination in high-resolution 3D images of
96 bacterial biofilms. *BCM3D* uses CNNs in combination with mathematical image analysis to
97 achieve accurate single cell segmentation in 3D. The CNNs employed in *BCM3D* are based on the
98 3D U-Net architecture and training strategy, which has achieved excellent performance in
99 biomedical data analysis benchmark tests[31]. The mathematical image analysis modules of *BCM3D*
100 enable post-processing of the CNN results to further improve the segmentation accuracy. We
101 establish that experimental bacterial biofilms images, acquired by lattice light sheet microscopy,
102 can be successfully segmented using CNNs trained with computationally simulated biofilm
103 images, for which the ground-truth voxel-level annotation maps are known accurately and

precisely. By systematically evaluating the performance of *BCM3D* for a range of SBRs, cell densities, and cell shapes, we find that voxel-level segmentation accuracies of >80%, as well as cell counting accuracies of >90%, can be robustly achieved. *BCM3D* consistently outperforms previously reported image segmentation approaches that rely exclusively on conventional image processing approaches. We expect that *BCM3D*, and CNN-based single-cell segmentation approaches in general, combined with non-invasive light sheet-based fluorescence microscopy will enable accurate cell tracking over time in dense 3D biofilms. This capability will launch a new era for bacterial biofilm research, in which the emergent properties of microbial populations can be studied in terms of the fully-resolved behavioral phenotypes of individual cells.

Materials and Methods

Lattice Light Sheet Imaging of Bacterial Biofilms

Fluorescence images of bacterial biofilms were acquired on a home-built lattice light sheet microscope (LLSM). LLSM enables specimen illumination with a thin light sheet derived from 2D optical lattice[33,34]. Here, a continuous illumination light sheet was produced by a time-averaged (dithered), square lattice pattern[18], and the illumination intensity at the sample was <1 W/cm^2. The submicrometer thickness of the excitation light sheet is maintained over long propagation distances (~30 μm), which enables optical sectioning, and thus high resolution, high contrast imaging of 3D specimens comparable to confocal microscopy. However, fluorophore excitation by a 2D light sheet reduces phototoxicity, because each excitation photon has multiple opportunities to be absorbed by fluorophores in the excitation plane and produce in-focus fluorescence. Widefield fluorescence images corresponding to each illuminated specimen plane are recorded on a sCMOS detector (Hamamatsu ORCA Flash v2). In this work, 3D biofilm images were acquired by

127 translating the specimen through the light sheet in 200 nm steps using a piezo nanopositioning
128 stage (Physik Instrumente, P-621.1CD).

129 *E.coli* K12 Biofilm Imaging: Ampicillin resistant *E.coli* K12, constitutively expressing
130 GFP[35], were cultured at 37 degrees overnight in LB medium with 100 μg/ml ampicillin. Overnight
131 cultures were diluted 100 times into the same culture medium, grown to an optical density at 600
132 nm (OD600) of 0.6 – 1.0, and then diluted by an additional factor of 10. Round glass coverslips
133 with the diameter of 5 mm were put into a 24-well plate (Falcon) and 400 μL of cell culture was
134 added to the well. Cells were allowed to settle to the bottom of the well and adhere to the coverslip
135 for 1 hour. The round coverslips were then mounted onto a sample holder and placed into the
136 LLSM sample-basin filled with M9 medium. GFP fluorescence was excited using 488 nm light
137 sheet excitation. Biofilm growth was imaged at room temperature every 30 min for a total of 20
138 time points. At each time point, a single 3D image stack contained 400 images, each acquired with
139 a 15 ms exposure time to avoid motion blur.

140 *M. xanthus* Biofilm Imaging: Strain LS3908 expressing tdTomato under the control of the
141 IPTG-inducible promoter[36] and DK1622 (WT) were cultured in the nutrient rich CYE media at 30
142 degrees Celsius until it reached an OD600 of 0.6 - 1.0. Media was supplemented with 1 mM IPTG
143 for tdTomato expressing cells. Chitosan (Thermo Fisher)-coated 5 mm round glass coverslips were
144 prepared by incubating coverslips with 1% (w/v) chitosan (1.5 % glacial acetic acid (v/v)) at room
145 temperature for 1 hour. Coverslips were then rinsed with water and placed into a 24-well plate
146 (Falcon) with 350-400 μL of undiluted cell culture. WT cells were stained directly in the 24 well
147 plate with 5 ng/ml FM4-64 (Thermo Fisher) dye. Cells were allowed to settle and adhere to the
148 coverslip for 2 hours. After the settling period, the coverslip was gently rinsed with CYE media to
149 flush away unattached cells. The rinsed coverslip was then mounted onto a sample holder and

150 placed into the LLSM sample-basin filled with MC7 starvation buffer. tdTomato and FM 4-64
151 fluorescence was excited using 561 nm light sheet. The 3D image stack contained 400 2D images.
152 Each 2D slice was acquired with an exposure time of 30 ms.

153 Mixed Strain Biofilm Imaging: Ampicillin resistant *E.coli* K12, constitutively expressing
154 GFP[35], and ampicillin resistant *E.coli* K12, expressing mScarlet (pBAD vector, arabinose induce)
155 were cultured separately at 37 degrees overnight in LB medium with 100 μg/ml ampicillin.
156 Overnight cultures were diluted 100 times into the same culture medium, grown to an optical
157 density at 600 nm (OD600) of 0.6 – 1.0, and then diluted to an OD of 0.1. After dilution, the two
158 strains were mixed together. Round glass coverslips with the diameter of 5 mm were put into a 24-
159 well plate (Falcon) and 500 μL of cell culture was added to the well. Cells were allowed to settle
160 to the bottom of the well and adhere to the coverslip for 1 hour. The cell culture medium was then
161 removed and replaced by 500 uL M9 medium containing 0.2% (w/v) arabinose. The co-culture
162 was incubated at 30 degrees overnight. 10 mins before imaging, the co-culture was stained with
163 5 ng/ml FM4-64 (Thermo Fisher) dye. 3D image stacks of 20 planes with 5 ms exposure time per
164 frame were acquired using 488 nm excitation.

165

166 **Raw Data Processing**

167 Raw 3D images were background subtracted and then deskewed and deconvolved as
168 described previously[18,19]. The background was estimated by averaging intensity values of dark
169 areas (devoid of cells) in the field of view. Deconvolution was performed using the Richardson-
170 Lucy algorithm with 10 iterations using experimentally measured point spread functions (PSFs) as
171 the deconvolution kernel. The experimentally measured PSFs were obtained separately for each

172 color channel using fluorescent beads (200 nm FluoSpheres®, Thermo Fisher) coated on a
173 coverslip[37]. 3D images were rendered using the 3D Viewer plugin in Fiji[38] or ChimeraX[39].
174

175 **Generation of simulated biofilm images**

176 To generate data for training of CNNs, we computationally simulated fluorescence images
177 of 3D biofilms, for which spatial arrangements among individual cells are known precisely and
178 accurately. Growth and division of individual rod-shaped cells in a population were simulated
179 using CellModeller, an individual-based computational model of biofilm growth (**Figure 1a**)[40]. In
180 individual-based biofilm growth models, cells are the basic modeling units. Each cell is
181 characterized by a set of parameters, including its 3D position, volume, and spatial orientation. All
182 the cells in the simulated biofilm are then allowed to evolve in time according to predefined
183 biological, chemical, and mechanical rules. For example, cells grow at a defined rate and then
184 divide after reaching a certain volume threshold. Cellular collisions that are due to cell growth are
185 alleviated by imposing a minimum distance criterion between cells at each time point. For our
186 simulations, we chose cell diameter and cell length (*d, l*) parameters consistent with a given
187 bacterial species, namely (1 μm, 3 μm) for *E. coli*[41], (0.7 μm, 6 μm) for *M. xanthus*[42], and (1 μm,
188 1 μm) for spherically symmetric *S.aureus*[43].

189 While the cell volume can be readily adjusted in CellModeller, the cellular volume density,
190 which is determined by the intercellular spacing, is not directly adjustable. We therefore adjusted
191 the cellular volume density after each simulation by scaling the cellular positions (cell centroids)
192 and thus the intercellular distances by a constant factor, while leaving cell sizes, shapes, and
193 orientations unchanged. This post-processing procedure enabled simulation of the exact same 3D
194 cell arrangements at adjustable cell volume densities.

We fluorescently labeled simulated cell volumes and surfaces according to two commonly used labeling strategies in fluorescence microscopy. To simulate expression of intracellular fluorescent proteins, the fluorescence emitters were placed at random positions within the cell volume. To simulate membrane staining, the fluorescence emitters were placed at random positions on the cell surface. Each cell contained between 500 - 1000 fluorophores to simulate expression level variations between cells, which is often observed in experimental images. Once the fluorophore spatial distributions were determined, a 3D fluorescence image (**Figure 1b**) was computationally generated. Each fluorophore was treated as an isotropic point emitter, so that it would produce a diffraction-limited point-spread-function (PSF) on the detector. Experimentally measured 3D PSF shapes (see Raw Data Processing) were used as the convolution kernel. Next, the fluorescence signal intensity was scaled by multiplying the image by a constant factor and then a constant background intensity was added to the image at ~200 photons per pixel, as measured in experimental data. This procedure enabled independent adjustments of the fluorescence signal and background to obtain signal-to-background ratios (SBRs) consistent with experimental data. In a final step, we introduced Poisson-distributed counting noise, based on the summed background and signal intensities, as well as Gaussian-distributed camera read-out noise (experimentally calibrated for our detector at 3.04 photons per pixel on average)[44]. This resulting image data (**Figure 1c**) was then processed in the same manner as experimental data (see Raw Data Processing). In contrast to experimental data, generation of the corresponding voxel-level annotation maps is fast and error free, because the underlying ground truth cell arrangements are known *a priori* (**Figure 1d**).

Differential labeling: To mimic imaging of reporter gene expression in a subset of cells, we simulated biofilm images, in which all cells were stained at the cell surface (e.g. with a

membrane intercalating fluorescent dye) and a subset of cells additionally contained intracellular fluorophores (e.g. through the expression of an intracellular fluorescent protein) (**Figure S1ab**). The mixing ratios between membrane-labelled, and membrane + interior labelled cells were 10:90, 30:70, 50:50, 70:30 and 90:10. Ten different cell arrangements containing ~300 cells were simulated for each ratio. To train the CNNs (see next section), five datasets were used, all with a 50:50 mixing ratio.

Mixed cell shapes: To mimic imaging of cells with different morphologies, we simulated biofilms containing spherical and rod-shaped cells **(Figure S1cd)**. The spatial arrangements of the mixed biofilm were generated by randomly positioning cells in a 3D volume. Collision among cells was eliminated by adjusting cell positions and orientations. The size of the rod-shaped cells is that of *E. coli* (~3 × 1 μm, length by diameter). The size of the spherical cells is that of *S. aureus* (~1 μm in diameter)[45]. Both cell types were labelled by intracellular fluorophores, as described above. The mixing ratios between rod-shaped and spherical cells were 10:90, 30:70, 50:50, 70:30 and 90:10. Ten different cell arrangements containing ~300 cells were simulated for each ratio. To train the CNNs (see next section), we picked one image from each mixing ratio for a total of five images.

Training the convolutional neural networks

We trained 3D U-Net CNNs for voxel-level classification tasks[46] within the NiftyNet platform[47] (network architecture depth 4, convolution kernel size 3, ReLU activation function, 32 initial feature maps, and random dropout of 0.5 during training). To achieve robust performance, we trained these networks using five to ten simulated biofilm images with randomly selected cell densities and signal-to-background ratios (see Generation of simulated biofilm images). The raw

241 data processing steps used for experimental data (see Raw Data Processing) were equally applied
242 to simulated data. 3D deconvolved simulated data and their corresponding voxel-level annotations
243 were used to train the CNNs. Each image used for training contained ~9 million voxels. We trained
244 CNNs by classifying each voxel as 'background', 'cellular interior' or as 'cellular boundary' based
245 on the underlying cell arrangements. For mixed-species biofilms, two additional classes, 'cellular
246 interior' and 'cellular boundary' of the second species, were used. This type of annotation scheme
247 has been shown to increase separation of bacterial cells in 2D[48]. For data augmentation, we applied
248 NiftyNet's built-in scaling, rotation, and elastic deformation functions. Instead of the original
249 cross-entropy loss function combined with uniform sampling, we used the Dice loss function and
250 'balanced sampler', so that every label has the same probability of occurrence in training. All
251 networks were trained for 2000 to 3600 iterations with a learning rate of 0.0001. Using these
252 parameters, it took approximately 24 hours to train the CNNs on a NVIDIA Tesla V100 GPU with
253 16 GB memory.
254

255 **Thresholding of CNN-produced confidence maps**

256 Voxel-level classification by CNNs generates different confidence maps (one confidence
257 map for each annotation class). The confidence values range between 0 and 1 and represent the
258 confidence of assigning individual voxels to a given class. After thresholding the 'cell interior'
259 confidence map to obtain a binary image (**Figure S2a-c**), connected voxel clusters can be isolated
260 and identified as single cell objects using 3D connected component labeling[49]. A conservative size-
261 exclusion filter was applied: small objects with a volume ~10 times less than the expected cell size
262 were considered background noise and filtered out using an area open operator[49]. Since the cell-
263 interior volumes do not contain the cell boundaries, we dilated each object by 1-2 voxels to

increase the cell volumes using standard morphological dilation[49]. The threshold value to segment individual cell objects based on the 'cell interior' confidence map was determined by plotting the overall voxel-level segmentation accuracy (see Performance Evaluation) versus the threshold values (**Figure S2**). High segmentation accuracies were consistently obtained using confidence thresholds between 0.88 and 0.94.

Post-processing of U-Net result using a refined LCuts algorithm

Thresholding of the 'cell interior' confidence map produces a binary segmentation result (background = 0, cell interior =1), where groups of connected, non-zero voxels identify individual cells in most cases (**Figure S2**). However, when cells are touching, they are often not segmented as individuals, but remain part of the same voxel cluster (undersegmentation). On the other hand, a single cell may be erroneously split into smaller sub-cellular objects (oversegmentation). Finally, in datasets with low SBR, connected voxel clusters may be detected that do not correspond to cells and thus produce false positive objects (**Figure 2a**). To address these errors and improve the segmentation accuracy further, we included additional mathematical image analysis steps to post-process the CNN results and reduce undersegmentation and oversegmentation errors.

<u>Step 1. Filtering:</u> False positive objects are identified by evaluating the coefficient of variation[50,51] for each connected voxel cluster i:

$$CV_i = \frac{\sigma_i}{\mu_i}$$

where σ_i and μ_i denote the standard deviation and the mean of the intensity taken over all voxels contained in connected voxel cluster i. If the coefficient of variation is larger than ρ, then the current object will be classified as a false positive object and removed from the confidence map by setting all its voxels to zero. The removed objects will then no longer be counted when

evaluating the cell counting accuracy. The value of ρ is selected based on the coefficient of variation of the background. For the datasets analyzed here, this sample coefficient of variation was determined to be $\rho = 1.1$. After CV-filtering, objects smaller than 25% of the expected bacterial cell size are also removed by setting its voxels to zero. The remaining connected voxel clusters are then considered for further processing (**Figure 2a**).

Step 2. Identification of point coordinates along central bacterial cell axes: To identify and delineate individual cells in the connected voxel clusters identified in the previous step, we implemented medial axis extraction using the method of inscribed spheres[52], with the constraint that the sphere radii do not exceed the expected diameter of a single bacterial cell (e.g. $d = 0.8$ µm) (**Figure 2b** left). The set of N inscribed spheres are tangent to the object's surface and parameterized by $(x_i, y_i, z_i; r_i<d)$ for $i = 1, \ldots, N$. Determination of the $(x_i, y_i, z_i; r_i)$ coordinates is achieved using the Euclidean distance transform of the objects' boundary[53], so that the points with coordinates (x_i, y_i, z_i) reliably trace out the central cell axes of individual bacterial cells (**Figure 2b** right).

Step 3. Identification of individual cell axes: To separate different linear segments after cell axis extraction (**Figure 2c**), we used a refined version of the linear cuts (*LCuts*) algorithm[54,55]. *LCuts* is a graph-based data clustering method designed to detect linearly oriented groups of points with certain properties. The fundamental elements of a weighted mathematical graph are nodes, edges, and edge weights. Here, the points with coordinates (x_i, y_i, z_i) represent the graph nodes. Edges are the connections among nodes. Edges are assigned weights, for example, to reflect the confidence that two nodes belong to the same group. *LCuts* achieves grouping by assigning weights to edges in the fully connected graph to reflect the similarity between two nodes. The features of each node include its location and direction, where the location of each node is simply

its Cartesian coordinates. The direction of each node is found by first determining its 5-hop neighborhood, removing nodes at large relative angles, and evaluating the major direction of the outlier removed neighborhood (**Figure S3**).

The algorithm to separate the nodes into different groups is a recursive graph cutting method[54]. Graph cuts (*e.g.* nCut[56]) disconnect the edges between two groups of nodes when the combined weights of these edges are minimized. The weights, between node *i* and node *j*, are calculated as follows:

$$w_{ij} = w_D \cdot w_T$$

where

$$w_D = \begin{cases} e^{-D_{ij}^2/\sigma_D^2} & \text{if } D_{ij}^2 \leq r \\ 0 & \text{if } D_{ij}^2 > r \end{cases}$$

$$w_T = e^{-\left(\cos(\theta_{ij})-1\right)^2/\sigma_T^2}$$

w_D weighs the distance between two nodes and w_T weighs difference between node directions. D_{ij} is the Euclidean distance between node *i* and node *j*, and r is set to eliminate edges between two far away nodes. θ_{ij} is the relative angle between the directions of nodes *i* and *j*. σ_D and σ_T are adjustable parameters that control the rate of exponential decay. *LCuts* continues to separate groups of nodes until each group satisfies a stopping criterion. The stopping criterion is biologically inspired based on the expected length of a single cell and a group's linearity after each recursion. *LCuts* yields linearly oriented groups of points that trace out the central axes of individual cells (**Figure 2c**). Importantly, cell separation is achieved without having to specify the number of cells in the biofilm in advance. With this step, the clustering provided by the *LCuts* algorithm can fix under- and over-segmentation problems from the CNN-based segmentation.

Step 4. Cell reconstruction: The final output of linear clustering can provide length, location and orientation of each cell. Based on these linear clusters, the cellular architecture of the biofilms can be reconstructed by placing geometrical models of cells in space as shown in **Figure 2d**. For fast computation, spherocylinders are used as the geometrical model using a radius consistent with the known sizes of bacterial cells. To further refine the cell surfaces to better align with the CNN-segmented volumes, we enclosed the inscribed spheres found in Step 2 in a convex hull (**Figure 2d**).

Performance Evaluation

We quantified segmentation accuracy both at the cell-level (object counting) and at the voxel-level (cell shape estimation). To quantify the cell-level segmentation accuracy, we designated segmented objects as true positive (*TP*) if their voxel overlap with the ground truth or the manual annotation resulted in a Jaccard index[57] larger than 0.5. This criterion ensures one-to-one matching. If the segmented cell object could not be matched to a ground truth/manually annotated cell volume, then it was counted as a false positive (*FP*) and the Jaccard index of that segmented object was set to zero. If a ground truth/manually annotated cell volume was not identified in the image, then it was counted as false negative (*FN*). The cell counting accuracy was defined as $TP/(TP+FP+FN)$. The voxel-level segmentation accuracy quantifies the accuracy of cell shape estimation and is calculated as the average Jaccard index over all segmented objects in the image.

To evaluate the accuracy of cell segmentation on experimental data, three researchers separately traced the cell contours on experimental 2D slices by using freehand selections in Fiji ROI Manger[58]. Because human annotation is very time consuming (about 50 hours for a complete

354 3D dataset containing ~300 cells in a 22 x 32 x 12 um^3 volume), one to three single 2D slices were
355 selected for each dataset. One exception is the 3D *M. xanthus*, for which the cell outlines in all
356 available x, y and z slices were traced manually. For straight, rod-shaped cells, the centroids of the
357 resulting 2D cell contours all fall within the cell interior volume. To group together the contours
358 belonging to the same cells, the centroid of each contour was projected along the x, y and z
359 dimension. If the projected centroid was enclosed by any other contour in a different slice, then
360 the centroid of that contour was projected onto the plane of the initial contour. Two contours were
361 labeled as related if they contained each other's projected centroids. This process is repeated for
362 all possible contour pairs and their relationship is recorded in an adjacency matrix. Next, related
363 contours were assigned to individual cells. To separate incorrectly grouped contours, we
364 additionally identified clusters of centroids using the DBSCAN point clustering algorithm[59]. In a
365 final step, we manually removed incorrectly traced contours. This procedure determined the
366 approximate positions, shapes, and orientations of individual cells in the 3D biofilm.

367 To estimate the SBRs of both simulated and experimental images, we manually selected
368 and determined the intensities of approximately ten 'signal' and ten 'background' regions in the
369 images. We computed the SBR as the mean signal intensity divided by the mean background
370 intensity. To estimate the local density of a biofilm cluster of simulated images, we sliced the
371 considered several 3D tiles within the biofilm of size 64 by 64 by 8 voxels. We then estimated the
372 local density as total cell volume contained in each tile divided by the tile volume. We calculated
373 the mean density of the 10 densest tiles to define the 'local density' metric reported for each dataset
374 in the paper. To estimate the cell density in an experimentally acquired biofilm image, the same
375 calculations were performed on either 3D manual annotations (if available) or binary masks
376 obtained by CNN-processing.

Results and Discussion

Cell segmentation by thresholding CNN confidence maps

CNNs have been shown to perform well on pixel-level classification tasks for both 2D and 3D data[60,61]. Bacterial biofilms, however, present a unique challenge in this context. The cell shapes to be segmented are densely packed and barely resolvable even with the highest resolution optical microscopes. Additionally, living biofilms in fluorescence microscopes can only be excited with low laser intensities to ameliorate phototoxicity and photobleaching concerns. Low intensity excitation unfortunately also reduces the SBR in the acquired images. It remains unclear to what extent single-cell segmentation approaches can accurately identify and delineate cell shapes in bacterial biofilm images obtained under low intensity illumination conditions. To address this question, we implemented an *in silico* CNN training strategy for *BCM3D* (see Materials and Methods) and systematically evaluated its voxel-level classification (cell morphometry) and cell counting accuracies using simulated biofilm images with varying cell densities and SBRs similar to those encountered in experimental data.

We compared two commonly used cell labeling approaches, namely genetic labeling through the expression of cell-internal fluorescent proteins (**Figure 3a-c**) and staining of the cell membranes using fluorescent dyes (**Figure 3d-f**). For both labeling approaches, voxel-level segmentation and cell counting accuracies, obtained by thresholding CNN confidence maps (see Materials and Methods), depend mostly on cell density, whereas the SBR plays a less important role. For cell-internal labeling, SBRs of >1.7 and cell densities of <60% consistently produce voxel-level classification accuracies of >80% and cell counting accuracies of >95%. On the other hand, SBRs of <1.7 and cell densities of >60% lead to lower segmentation accuracies. While lower segmentation accuracies are expected for higher cell densities and lower SBRs, the sharp drop-

400 offs observed here may indicate a fundamental performance limitation of the CNNs employed.
401 Still, the voxel-level classification and cell counting accuracies consistently surpass previous
402 approaches for bacterial cell segmentation for commonly encountered cell densities and SBRs.
403 Specifically, the cell counting accuracies obtained by *BiofilmQ*[62], *Seg3D*[63], and Yan *et al.*[13] quickly
404 drop to zero as a function of increasing the Jaccard index threshold (a quantitative measure of cell
405 shape similarity relative to the ground truth, see Materials and Methods), indicating that cell shapes
406 are not accurately estimated by conventional image processing approaches (**Figure 3g-i**). By
407 contrast, the cell counting accuracy obtained by thresholding CNN confidence maps remain high
408 even for Jaccard index thresholds larger than 0.5, indicating that cell shapes are better estimated
409 by the CNNs.

410 The accuracies of single-cell shape estimation and cell counting are predominantly affected
411 by cell density. The variation is more pronounced for membrane-stained cells, because *inter-*
412 cellular fluorescence intensity minima are less pronounced when cell membranes are labeled and
413 cells physically contact each other (red arrow in **Figures 3c** and **3f**). By contrast, intracellular
414 fluorophores produce the highest intensities at the cell center, so that the gaps between cells are
415 more readily resolvable. Also noteworthy is the sharp drop-off in segmentation accuracies for
416 SBRs of <1.7 for all cases. In such low SBR regimes, fluorescence signals of the cells become too
417 difficult to be distinguished from the background. As a result, the CNNs falsely identify random
418 noisy patterns in the background as cells. Additionally, thresholding of the CNN confidence maps
419 often yields connected voxel clusters that contain multiple bacterial cells. False identification and
420 incomplete delineation of cells cause the pronounced decrease in segmentation accuracy for SBRs
421 of <1.7.
422

423 **Post-processing of CNN confidence maps**

424 To better identify individual cells in low SBR and high cell density datasets, we developed
425 a graph-based post-processing module (see Materials and Methods) that takes advantage of highly
426 conserved bacterial cell shape of a given species. Briefly, we transformed the CNN 'cell interior'
427 confidence maps into 3D point cloud data that trace out the central axes of individual cells. This
428 transformation was achieved by medial axis extraction using size-constrained inscribed spheres[52]
429 (**Figure 2**). Single-cell axes are then identified as linearly clustered data points by *LCuts* – a graph-
430 based data clustering method designed to detect linearly oriented groups of points[54]. The so-
431 identified single-cell axes are then mapped back onto the original segmentation volumes to obtain
432 estimates of the 3D positions, shapes, and orientations of the now separated cells.

433 Post-processing with *LCuts* takes advantage of *a priori* knowledge about expected bacterial
434 cell sizes (**Figure S4**) by removing erroneously segmented volumes that are significantly smaller
435 than the expected value and by splitting incompletely segmented volumes representing fused cells.
436 Improvements of up to 15% and 36% are observed for cells labeled with cytosolic fluorophores
437 (**Figure 4a-c**) and membrane-localized fluorophores (**Figure 4d-f**), respectively. The more
438 substantial improvement for membrane-stained cells is due to fact that CNNs trained on
439 membrane-stained cells are more prone to erroneously identifying speckled background noise as
440 fluorescence signals in low SBR images. In addition, membrane-intercalating fluorophores of two
441 adjacent cells are in close proximity, making it difficult to resolve fluorescence signals from two
442 separate cells due to spatial signal overlap (see the red arrow, **Figure 3c** and **3f**).

443 **Segmentation of experimental biofilm images**

444 To test the performance of *BCM3D* on experimentally acquired biofilm images, we
445 acquired time-lapse images of GFP expressing *E. coli* biofilms every thirty minutes for ten hours.
446 We then manually annotated one 2D slice in the 3D images at the t = 5, 6, and 10-hour time points.
447 When referenced to these manual segmentation results, both the CNN outputs and the *LCuts*-
448 processed CNN outputs consistently achieved better cell counting accuracies than conventional
449 segmentation methods (**Figure 5, Figure S5**). We note however that the cell counting accuracy is
450 much lower (45% at a 0.5 Jaccard index threshold) when referencing manual annotation. The more
451 rapid drop-off of the cell counting accuracy as a function of increasing Jaccard index threshold can
452 be attributed to the following factors. First, human annotation of experimentally acquired biofilm
453 images is not as consistent as simulated data, and therefore, it is less likely to produce good
454 matches between segmented and annotated cell shapes. Second, the experimentally acquired
455 biofilm images differ from the simulated training and annotation datasets in several ways. Real
456 cells may appear blurred due to partially immobilized/wiggling cells. Some cells are in the process
457 of cell division, so that different human annotators may count these cells either as one cell or two
458 cells. Finally, optical aberrations and scattering effects were not included in training data
459 simulations. We note, however that, at a lower Jaccard index threshold, the cell counting accuracies
460 is in the range of 60% to 70%, which is substantially higher than the current state-of-the-art in
461 biofilm imaging (**Figure 5**).

462 To demonstrate that *BCM3D* can achieve similarly high segmentation accuracies for
463 membrane-stained cells, we analyzed a small patch of a *M. xanthus* biofilm, which was stained
464 with the membrane intercalating dye FM4-64 (**Figure 6a**). For this dataset, we manually annotated
465 each *xy*, *xz*, and *yz* slice through a single 3D image stack (**Figure 6b**). When referenced to this 3D

466 manual segmentation result, *BCM3D* (**Figure 6c**) again produced cell counting accuracies in the
467 range of 60% to 70% at low Jaccard index thresholds (**Figure 6d**). We note that comparison with
468 the state-of-the-art cannot be made here, because previous biofilm segmentation approaches were
469 not designed to work with membrane-stained cells.

470

471 **Morphological separation of mixed cell populations**

472 Given the much-improved segmentation results obtained using *BCM3D*, we reasoned that
473 the same CNNs may also be able to assign segmented objects to different cell types based on subtle
474 morphological differences in the acquired images. Differences in the imaged cell morphologies
475 arise due to physical differences in cell shapes (e.g. spherical vs. rod-shaped cells) or due to
476 differences in the fluorescent labeling protocols (e.g. intracellular vs. cell membrane labeling),
477 because fluorescence microscopes simply measure the spatial distributions of fluorophores in the
478 sample. The ability to separate different cell morphologies is important for the study of
479 multispecies biofilms, where interspecies cooperation and competition dictate population-level
480 outcomes[3,64-71]. Separation of differentially labeled cells is also important for the study of gene
481 activation in response to cell-to-cell signaling[72]. Expression of cytosolic fluorescent proteins by
482 transcriptional reporter strains is a widely-used technique to visualize activation of a specific gene
483 or genetic pathway in living cells. Such genetic labeling approaches can be complemented by
484 chemical labeling approaches, e.g. using membrane intercalating chemical dyes that help visualize
485 cells non-specifically or environmentally-sensitive membrane dyes that provide physiological
486 information, including membrane composition[73,74], membrane organization/integrity[75-77], and
487 membrane potential[78,79]. Chemical and genetic labeling approaches are traditionally implemented
488 in two different color channels. However, there are important drawbacks to using multiple colors.

489 First and foremost, the amount of excitation light delivered is increased to excite differently
490 colored fluorophores, raising phototoxicity and photobleaching concerns. Second, it takes *N* times
491 as along to acquire *N*-color images (unless different color channels can be acquired
492 simultaneously), making it challenging to achieve high temporal sampling in time-lapse
493 acquisition. For these reasons, methods that extract complementary physiological information
494 from a single-color image are preferable.

495 We evaluated the ability of *BCM3D* to automatically segment and separate rod-shaped and
496 spherical bacterial cells consistent with sizes of *E. coli* and *S. aureus*, respectively, in simulated
497 images (**Figure S6**). Similar to a single-population segmentation problem, we trained CNNs that
498 classify pixels into five different classes: 'background', 'cellular interior of population 1', 'cellular
499 boundary of population 1', 'cellular interior of population 2' and 'cellular boundary of population
500 2'. Thresholding the CNNs confidence maps can achieve cell counting accuracies larger than 90%
501 for both cell types, independent of their population fractions (**Figure 7a**). Post-processing of this
502 result using *LCuts* improved the cell counting accuracy by less than 0.5% on average, indicating
503 that under-segmented cell clusters are not prevalent in this dataset.

504 We next evaluated the ability of *BCM3D* to automatically segment and separate membrane-
505 stained cells that express cytosolic fluorescent proteins from those that do not (**Figure S7**). Again,
506 the cell counting accuracy is consistently above 80% for all tested mixing ratios (**Figure 7b**).
507 Finally, we applied *BCM3D* to experimentally acquired biofilm images of two different *E. coli*
508 strains. Both strains were stained by the membrane intercalating dye FM4-64, but the second strain
509 additionally expressed GFP (**Figure S8**). The cells were homogeneously mixed prior to mounting
510 to randomize the spatial distribution of different cell types in the biofilm (see Materials and
511 Methods). Multiple 2D slices from the 3D image stack were manually annotated and compared

512 with the results obtained by *BCM3D*. Consistent with the single-species experimental data, a cell
513 counting accuracy of 50% is achieved for each cell type at a 0.5 Jaccard index threshold and, at a
514 lower Jaccard index thresholds, the counting accuracies increased to 60% to 70%, (**Figure 7cd**).
515 Thus, *BCM3D* enables automated and accurate cell type assignments based on subtle differences
516 in cell morphologies in mixed population biofilms – a capability not available using conventional
517 image processing methods.
518

519 **Conclusions**

520 CNNs have been successful applied to many different problems in biological image
521 analysis, but their ability to segment individual cells in 3D and time-lapse 3D bacterial biofilm
522 images has not yet been fully explored. Here, we demonstrated a new CNN-based image analysis
523 workflow, termed *BCM3D*, for single-cell segmentation and shape classification (morphometry)
524 in 3D images of bacterial biofilms. We applied *BCM3D* to 3D images acquired by lattice light
525 sheet microscopy. However, *BCM3D* can also be applied to 3D images acquired by confocal
526 microscopy or more advanced super-resolution microscopy modalities, provided that realistic
527 image formation models are used to simulate the training datasets. The use of simulated training
528 data is a major advantage of *BCM3D*, because it overcomes inconsistencies inherent in manual
529 dataset annotation and provides needed flexibility not only in terms of the microscope platform
530 used, but also in terms of the bacterial cell shapes that are to be segmented.

531 We systematically investigated the advantages and limitations of *BCM3D* by evaluating its
532 voxel- and cell-level segmentation accuracy using simulated and experimental datasets of different
533 cell densities and SBRs. CNN-based image analysis enabled accurate segmentation of individual
534 cells in crowded environments and automatic assignments of individual cells to specific cell

535 populations for most of the tested parameter space. Such capabilities are not readily available when
536 using previously established segmentation methods that rely exclusively on conventional image
537 and signal processing algorithms.

538 While *BCM3D* surpasses the performance of previous approaches, we stress that further
539 improvements are still possible and, for certain experiments, absolutely needed. Our systematic
540 analysis revealed that high cell density and low SBR datasets are particularly challenging for the
541 CNNs used in this work. Future work will therefore focus on increasing the contrast and resolution
542 in bacterial biofilm images. The use of optical super-resolution modalities can certainly provide
543 higher spatial resolution, but such resolution improvements often come at a cost of reduced image
544 contrast and faster photobleaching/phototoxicity, especially for fast frame-rate time-lapse
545 imaging. Software solutions will therefore play an increasingly important role in biological
546 imaging. Indeed, CNN-based processing modules developed for contrast enhancement and image
547 denoising have also surpassed the performance of conventional methods based on mathematical
548 signal processing[80-83]. Incorporation of these tools into *BCM3D* may yield further improvements
549 in the single-cell segmentation accuracies. We anticipate that the ability to accurately identify and
550 delineate individual cells in dense 3D biofilms will in turn enable accurate cell tracking over long
551 periods of time. Detailed measurements of behavioral single-cell phenotypes in larger bacterial
552 communities will help determine how macroscopic biofilm properties, such as its mechanical
553 cohesion/adhesion and its biochemical metabolism, emerge from the collective actions of
554 *individual* bacteria.

555 **Author Contributions**

556 M.Z., J.Z., Y.W., J.W., S.T.A., and A. G. designed research;

557 M.Z., J.Z., Y.W., J.W., and A.M.A. performed research;

558 M.Z., J.Z., Y.W., J.W., S.T.A., and A. G. analyzed data;

559 and M.Z., J.Z., Y.W., J.W., S.T.A., and A. G. wrote the paper.

560

561 **Acknowledgements**

562 This work was supported by a Jeffress Memorial Trust Awards in Interdisciplinary Research (to

563 A.G and S.T.A.). M.Z. and J.W. were supported by a University of Virginia Presidential

564 Fellowship in Data Science. Y.W. was supported in part by a Dean's M.S.-Ph.D. Fellowship in

565 Data Science. We thank Lotte Søgaard-Andersen, Larry Shimkets, Huiwang Ai and Ingmar

566 Riedel-Kruse for providing the bacterial strains used in this work. We thank Karsten Siller,

567 computational research consultant at the University of Virginia, for help with utilizing the

568 Universities super-computing resources.

569

Figure 1.

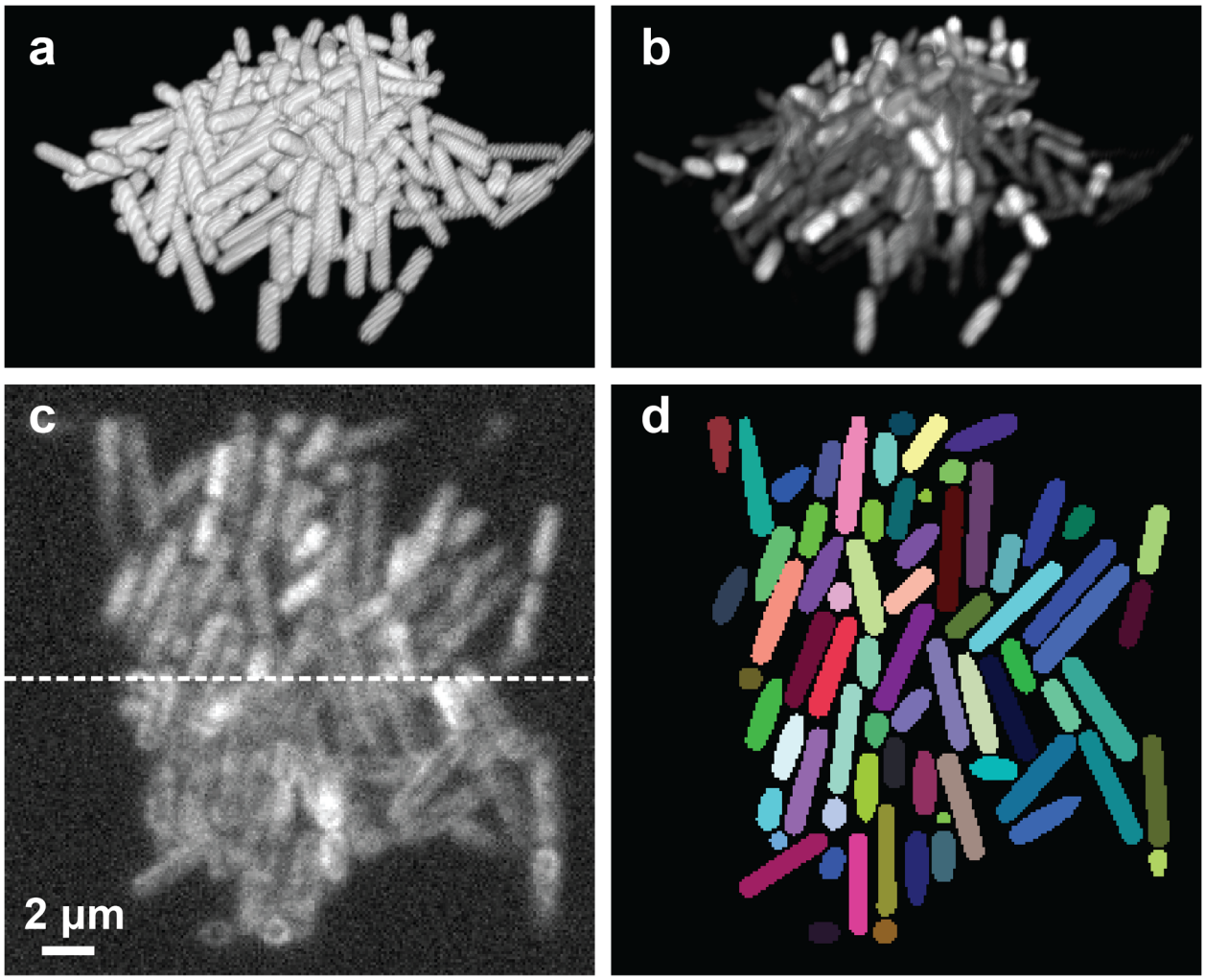


Figure 1. Simulation of fluorescent biofilms images and annotation maps. **(a)** Cell arrangements obtained by CellModeller. **(b)** Simulated 3D fluorescence image based on the cell arrangements in **a**. **(c)** XY slice through the 3D simulated fluorescence image in **b** (upper panel shows cells expressing cytosolic fluorescent proteins, lower panel shows cells stained with membrane-intercalating dyes). **(d)** Ground truth cell arrangements giving rise to the image shown in **c**. Voxels are displayed as black (background), or in different colors (indicating different cells).

Figure 2.

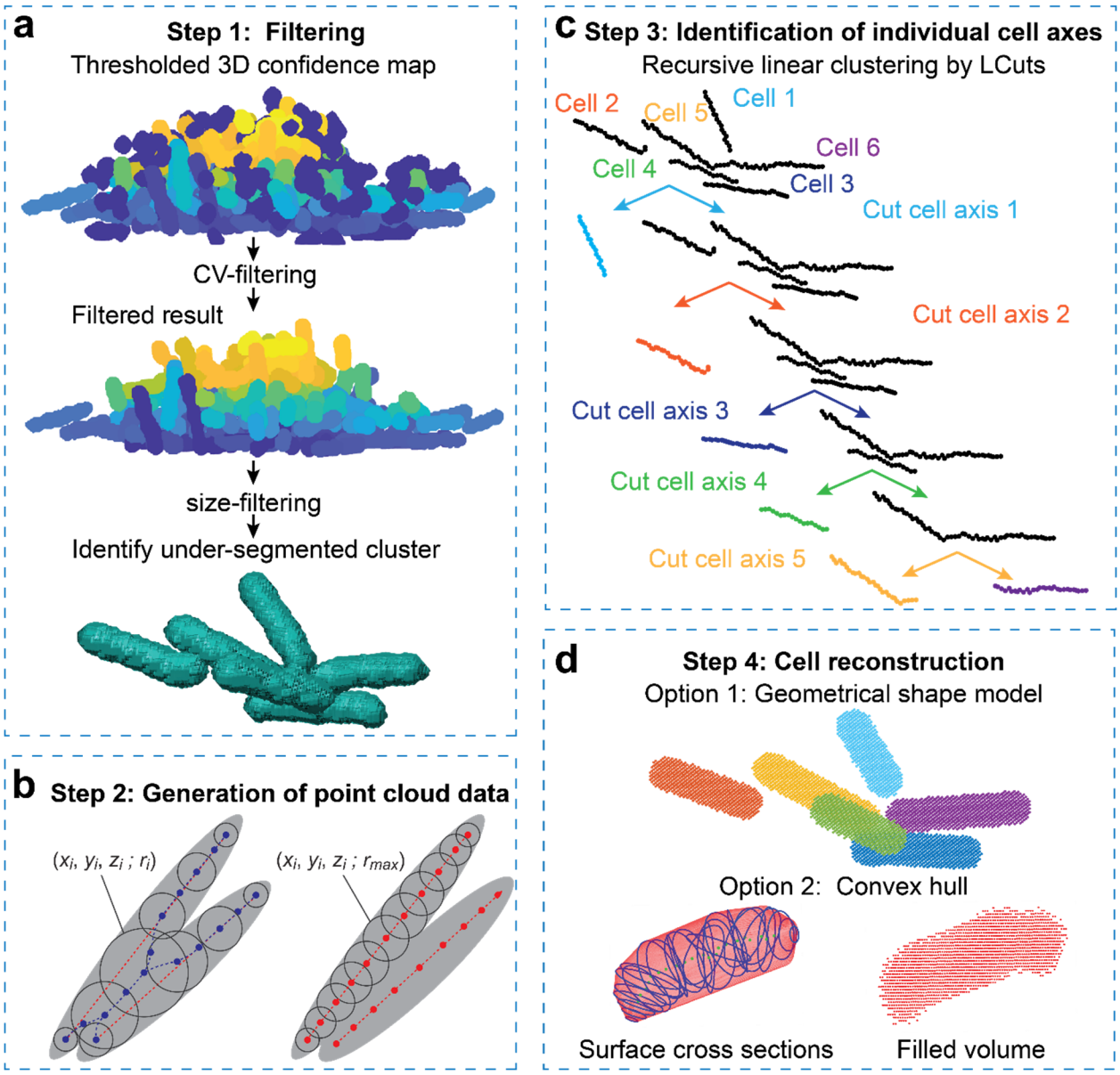


Figure 2. Post-processing of CNN-produced confidence maps using a refined *LCuts* processing pipeline. (**a**) False positive objects are detected and removed by CV- and size- filtering. Under-segmented clusters that are larger than single cells are selected for further splitting. **(b)** Illustration of modified medial axis (red dashed lines) extraction to generate point cloud data from fused clusters of rod-shaped cells using the method of inscribed spheres. When cells are touching, the traditional medial axis extraction process fails to align with the actual cell central axis (left). To

overcome this drawback, we limited the size of the inscribed spheres based on prior knowledge of bacterial cell diameters (right). **(c)** The set of inscribed sphere centers are then treated as a fully-connected, undirected graph in 3D with two node features: location and direction (see text and **Figure S3** for details). The graph (blue nodes) is then iteratively cut into smaller graphs (red nodes) until the stopping criteria are reached (see text for details). **(d)** Post-processed graphs represented in different color denoting different cells. The 3D surface of individual cells can be determined using a geometrical cell shape model (e.g. a spherocylinder for rod shaped bacteria) or by calculating the convex hull around the inscribed spheres found in step 2.

Figure 3.

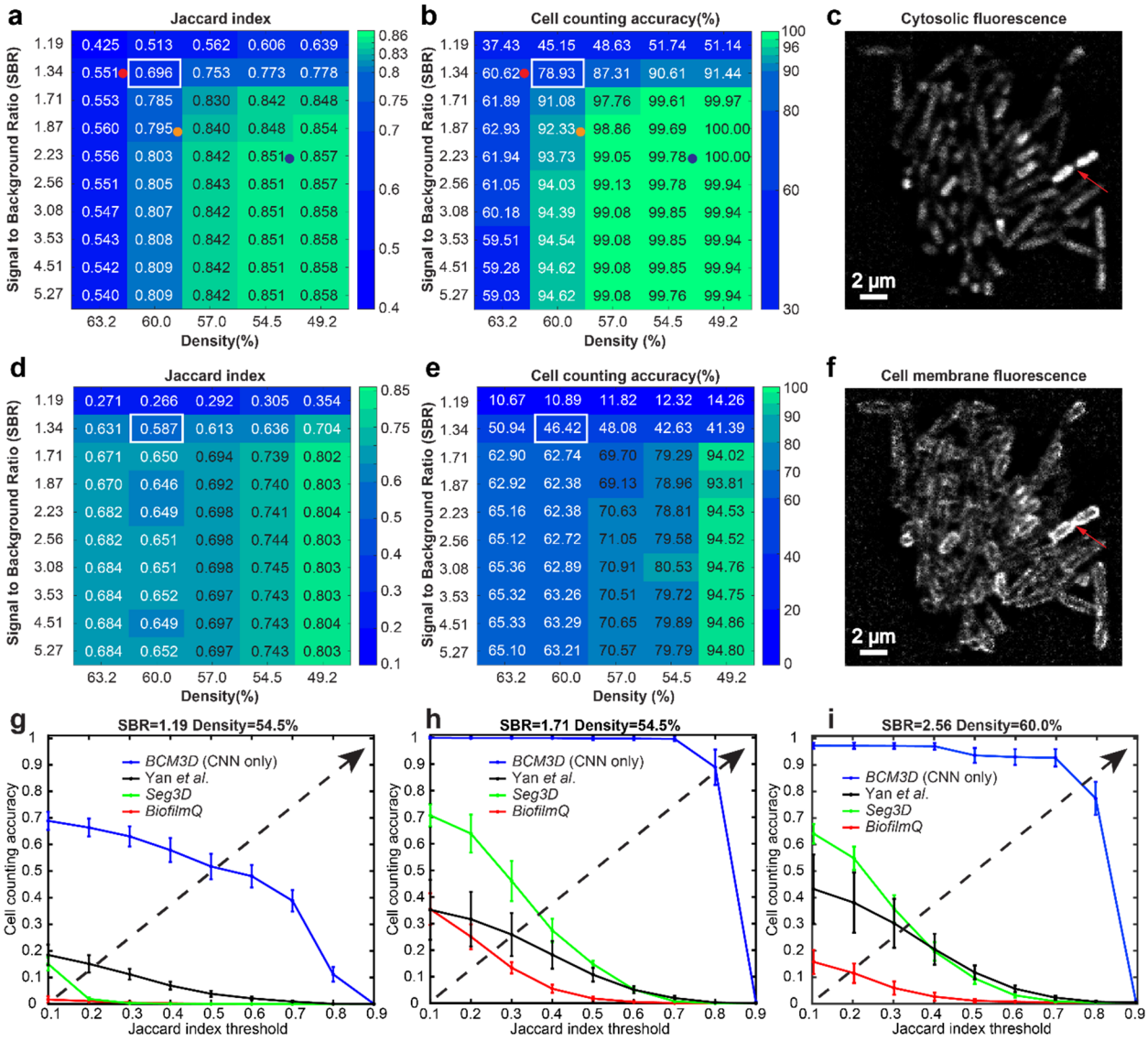


Figure 3. Performance of *BCM3D* using *in silico*-trained CNNs only on previously unseen simulated biofilm images. **(a)** The Jaccard index quantifies whether each voxel has been assigned to the correct class ('cell interior', cell surface', or 'background'). Solid circles represent the maximum local density and average SBRs encountered in experimental datasets (red, orange and blue: *E. coli* expressing GFP). **(b)** The cell counting accuracy using a Jaccard Index threshold of 0.5 for each segmented object (Materials and Methods) averaged over n=10 replicate datasets for cells labeled with cytosolic fluorophores. **(c)** Example image of cells labeled with cytosolic

606 fluorophores (Cell density = 60.0%, SBR = 1.34, indicated by white rectangle in panels a and b).
607 **(d)** Jaccard index and **(e)** Cell counting accuracy averaged over n=10 replicate datasets for cells
608 labeled with membrane-localized fluorophores. Purple solid circles represent experimental
609 datasets of *M. xanthus* stained with the membrane intercalating dye FM4-64. **(f)** Example image
610 of cells labeled with membrane-localized fluorophores (Cell density = 60.0%, SBR = 1.34,
611 indicated by white rectangles in panels d and e). The red arrows indicate a close cell-to-cell contact
612 point. **(g), (h) and (i)** Comparison of segmentation accuracies achieved by conventional (*BiofilmQ*,
613 *Seg3D*, and Yan *et al.*) and *BCM3D* using *in silico*-trained CNNs only. Three simulated datasets
614 (cytosolic fluorophores) with different SBRs and cell densities are shown. Segmentation accuracy
615 is parameterized in terms of cell counting accuracy (*y* axis) and accuracy of single-cell shape
616 estimation (*x* axis, quantified as the single-cell Jaccard index). Curves approaching the upper right-
617 hand corner indicate higher overall segmentation accuracy.
618
619

Figure 4.

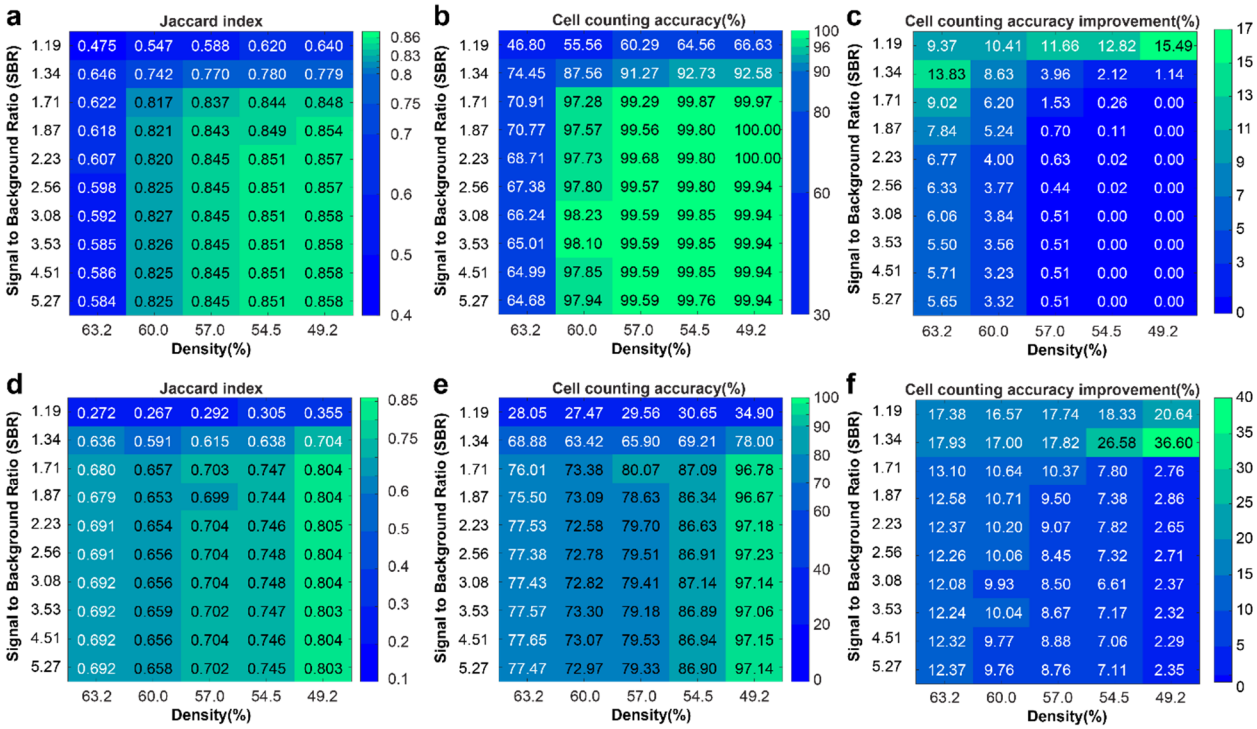
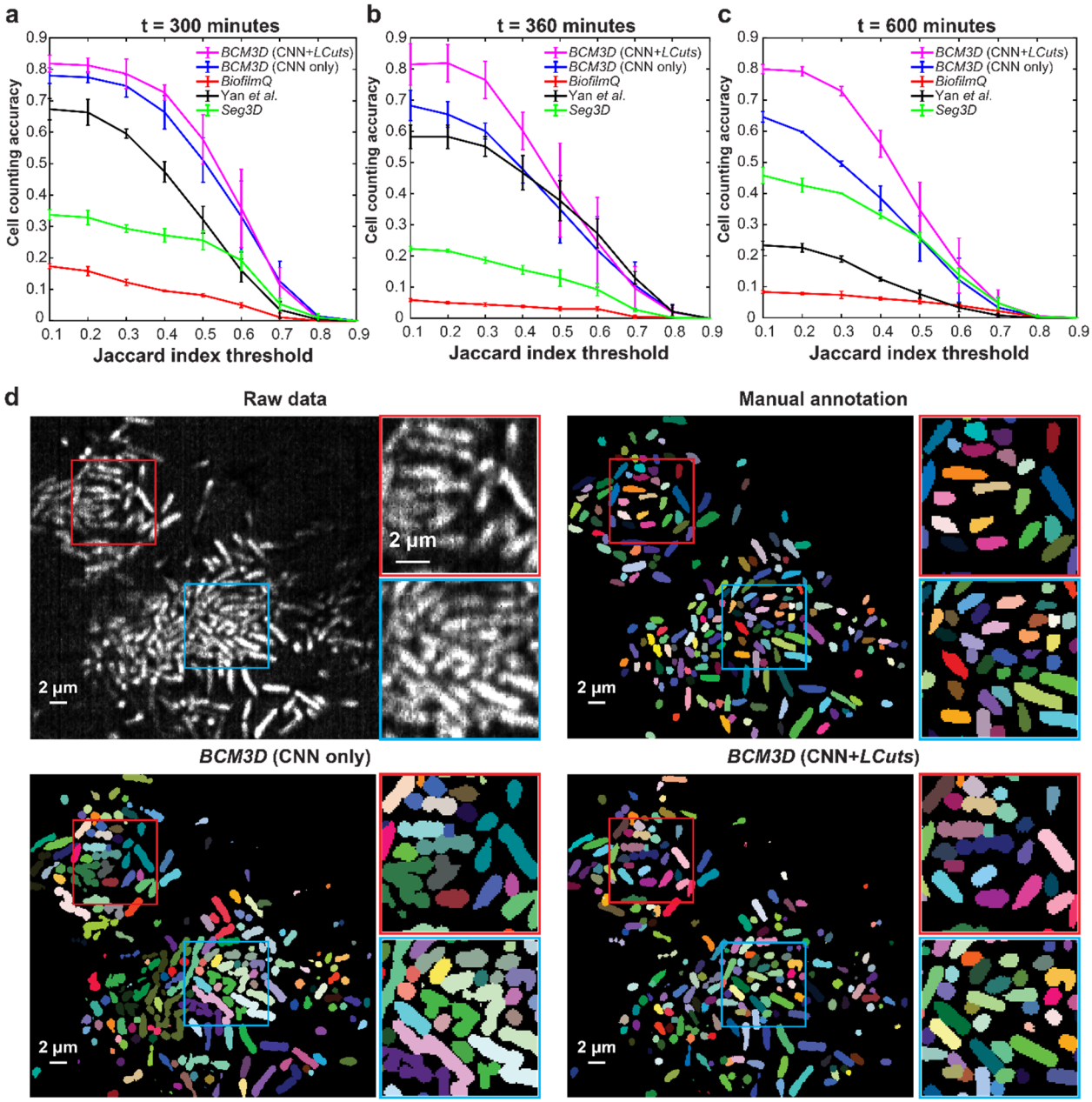


Figure 4. Performance of *BCM3D* (*in silico*-trained CNNs *and* additional post-processing by *LCuts*) on previously unseen simulated data. **(a)** Jaccard index and **(b)** The cell counting accuracy using a Jaccard Index threshold of 0.5 for each segmented object averaged over n=10 replicate datasets for cells labeled with cytosolic fluorophores. **(c)** Improvement relative to *silico*-trained convolutional neural networks without post-processing. **(d)** Jaccard index and **(e)** Cell counting accuracy averaged over n=10 replicate datasets for cells labeled with membrane-localized fluorophores. **(f)** Improvements relative to *silico*-trained convolutional neural networks without post-processing.

632 **Figure 5.**



633

634

635 **Figure 5**. Comparison of segmentation accuracies achieved by conventional (*BiofilmQ*, *Seg3D*,

636 and Yan *et al.*), and *BCM3D* (this work). The estimated SBRs are 2.2, 1.8, and 1.3, respectively.

637 The estimated cell densities are 54.8%, 59.0%, and 64.6%, respectively. **(a-c)** Three experimental

638 *E. coli* datasets (cytosolic expression of GFP) acquired at different time points after inoculation of

cells. Segmentation accuracy is parameterized in terms of cell counting accuracy (*y* axis) and accuracy of single-cell shape estimation (*x* axis, quantified as the single-cell Jaccard index). Curves approaching the upper right-hand corner indicate higher overall segmentation accuracy. **(d)** Comparison of segmentation results achieved at the t= 600 minutes time point by manual annotation, and by BCM3D using *in silico*-trained CNNs only and after further refinement of CNN outputs using *LCuts*. Segmentation results of the other methods are shown in **Figure S5**.

Figure 6.

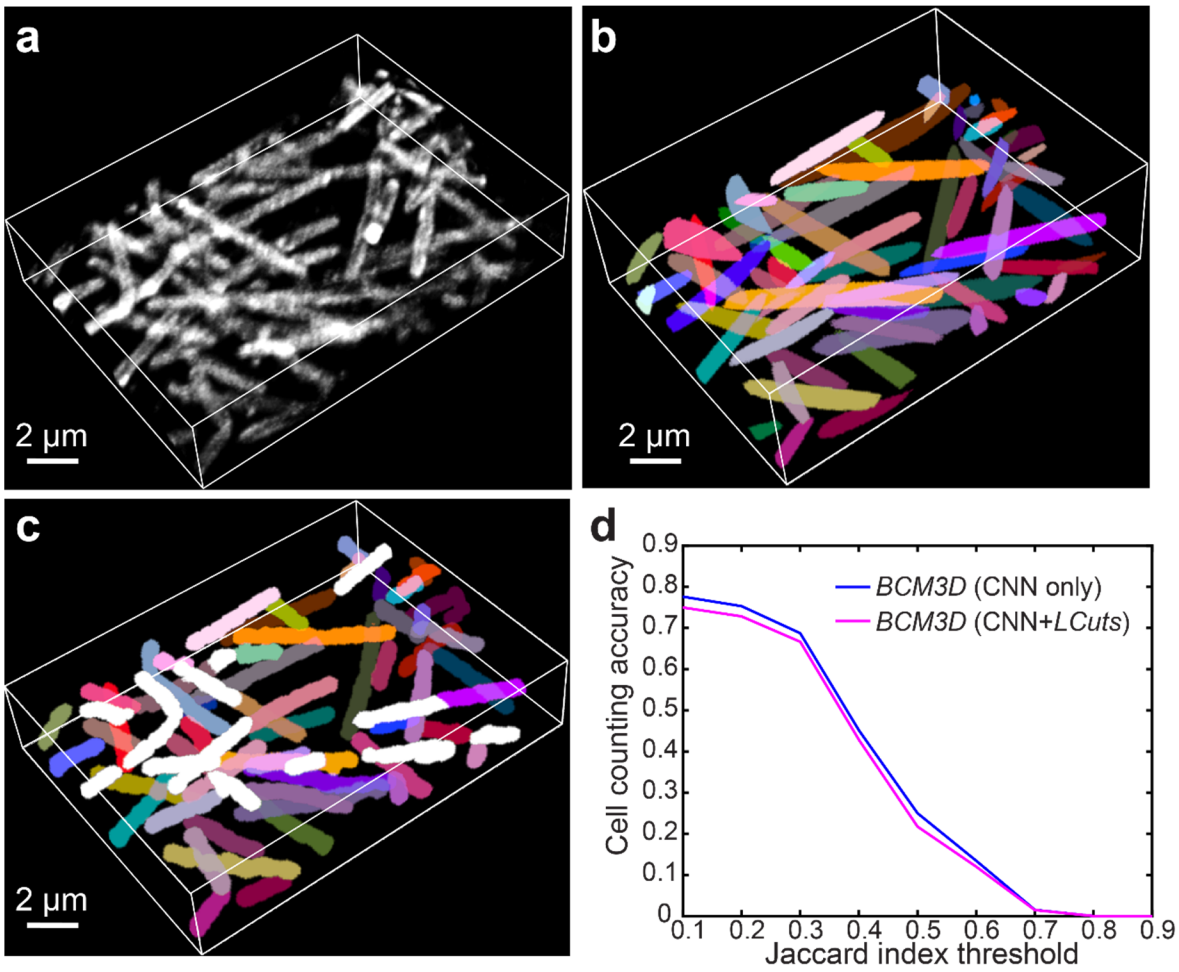


Figure 6. 3D Segmentation accuracy evaluation using *M. xanthus* biofilm images (cell density = 36.2%, and SBR = 1.58) using *in silico*-trained CNN processing. **(a)** Maximum intensity projection of a 3D *M. xanthus* fluorescence image. Cells were labeled with membrane-intercalating dye, FM4-64. **(b)** Maximum intensity projection of the manually obtained 3D segmentation result. **(c)** Maximum intensity projection of a CNN-based 3D segmentation result after *LCuts* post-processing. Cells that can be matched with the GT are displayed in the same colors as GT or otherwise colored in white. **(d)** Segmentation accuracy parameterized in terms of cell counting accuracy (*y* axis) and accuracy of single-cell shape estimation (*x* axis, quantified as the single-cell Jaccard index).

Figure 7.

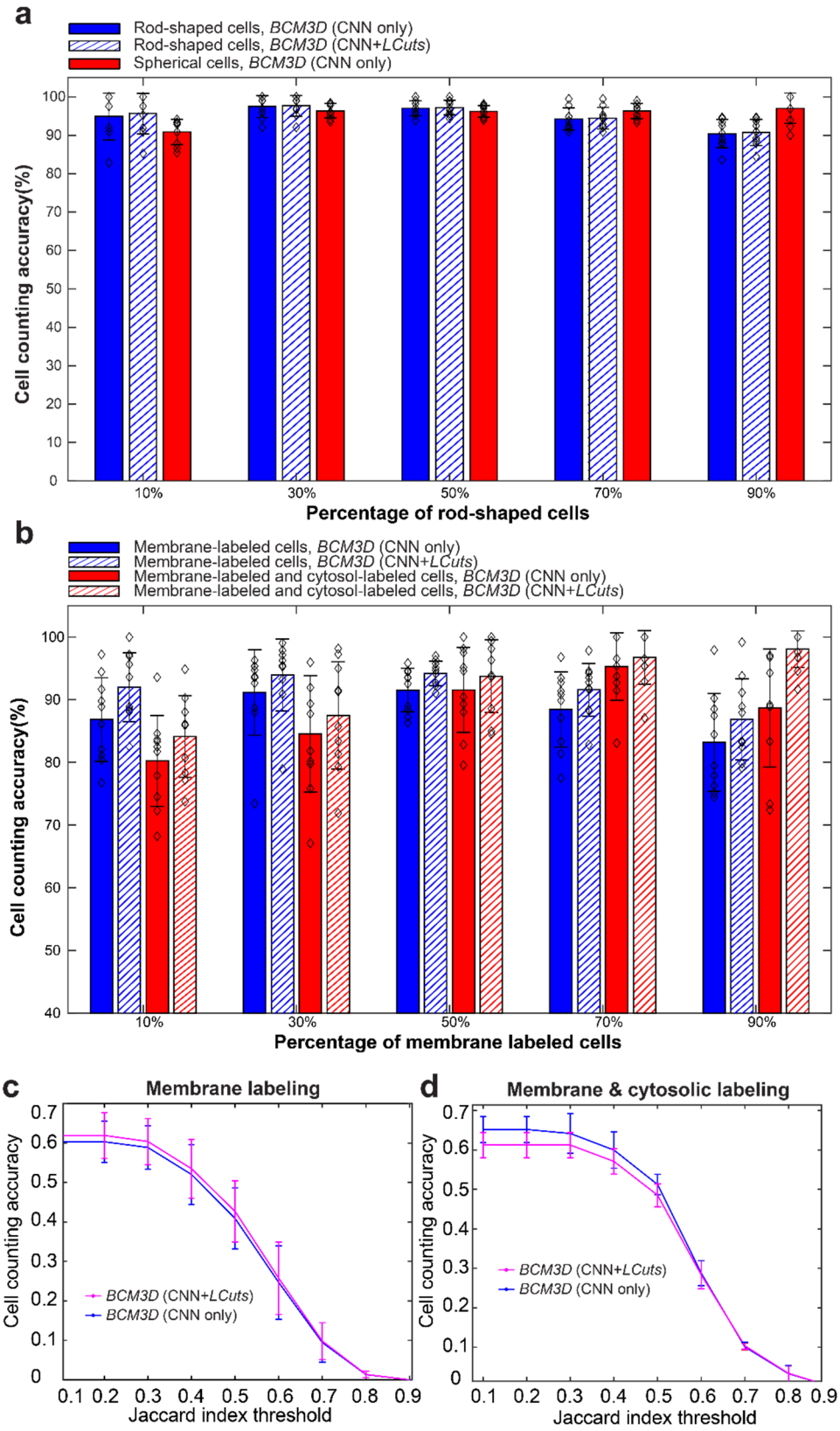

a
Rod-shaped cells, BCM3D (CNN only)
Rod-shaped cells, BCM3D (CNN+LCuts)
Spherical cells, BCM3D (CNN only)
Cell counting accuracy(%)
Percentage of rod-shaped cells
10%
30%
50%
70%
90%
b
Membrane-labeled cells, BCM3D (CNN only)
Membrane-labeled cells, BCM3D (CNN+LCuts)
Membrane-labeled and cytosol-labeled cells, BCM3D (CNN only)
Membrane-labeled and cytosol-labeled cells, BCM3D (CNN+LCuts)
Cell counting accuracy(%)
Percentage of membrane labeled cells
c
Membrane labeling
Cell counting accuracy
BCM3D (CNN+LCuts)
BCM3D (CNN only)
Jaccard index threshold
d
Membrane & cytosolic labeling
Cell counting accuracy
BCM3D (CNN+LCuts)
BCM3D (CNN only)
Jaccard index threshold

Figure 7. (a) Cell counting accuracy of *BCM3D* on simulated images containing different ratios of rod-shaped and spherical cells. Black diamonds represent the counting accuracy for each independent dataset. **(b)** Cell counting accuracy of *BCM3D* on simulated images with different ratios of membrane-labeled, and membrane-labeled and interior fluorescent protein expressing cells. Black diamonds represent the counting accuracy for 10 independent datasets. **(c** and **d)** Cell counting accuracy of *BCM3D* on experimental images of (c) membrane-labeled, and (d) membrane-labeled and interior fluorescent protein expressing *E. coli* cells (mixing ratio ~ 1:1).

666 **References**

667 1. Hall-Stoodley, L., Costerton, J.W. & Stoodley, P. Bacterial biofilms: from the natural
668 environment to infectious diseases. *Nat Rev Microbiol* **2**, 95-108 (2004).

669 2. O'Toole, G., Kaplan, H.B. & Kolter, R. Biofilm formation as microbial development. *Annual*
670 *Review of Microbiology* **54**, 49-79 (2000).

671 3. Nadell, C.D., Drescher, K. & Foster, K.R. Spatial structure, cooperation and competition in
672 biofilms. *Nat Rev Microbiol* **14**, 589-600 (2016).

673 4. Persat, A., Inclan, Y.F., Engel, J.N., Stone, H.A. & Gitai, Z. Type IV pili mechanochemically
674 regulate virulence factors in Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A* **112**, 7563-
675 7568 (2015).

676 5. Zhang, Q., Lambert, G., Liao, D., Kim, H., Robin, K., Tung, C.-k., Pourmand, N. & Austin,
677 R.H. Acceleration of Emergence of Bacterial Antibiotic Resistance in Connected
678 Microenvironments. *Science* **333**, 1764-1767 (2011).

679 6. Donlan, R.M. Biofilms: microbial life on surfaces. *Emerg Infect Dis* **8**, 881-890 (2002).

680 7. Yan, J. & Bassler, B.L. Surviving as a Community: Antibiotic Tolerance and Persistence in
681 Bacterial Biofilms. *Cell host & microbe* **26**, 15-21 (2019).

682 8. Kovach, K., Davis-Fields, M., Irie, Y., Jain, K., Doorwar, S., Vuong, K., Dhamani, N.,
683 Mohanty, K., Touhami, A. & Gordon, V.D. Evolutionary adaptations of biofilms infecting
684 cystic fibrosis lungs promote mechanical toughness by adjusting polysaccharide production.
685 *NPJ Biofilms Microbiomes* **3**, 1 (2017).

686 9. Berk, V., Fong, J.C., Dempsey, G.T., Develioglu, O.N., Zhuang, X., Liphardt, J., Yildiz, F.H.
687 & Chu, S. Molecular architecture and assembly principles of *Vibrio cholerae* biofilms.
688 *Science* **337**, 236-239 (2012).

689 10. Vidakovic, L., Singh, P.K., Hartmann, R., Nadell, C.D. & Drescher, K. Dynamic biofilm
690 architecture confers individual and collective mechanisms of viral protection. *Nat Microbiol*
691 **3**, 26-31 (2018).

692 11. Stewart, E.J., Ganesan, M., Younger, J.G. & Solomon, M.J. Artificial biofilms establish the
693 role of matrix interactions in staphylococcal biofilm assembly and disassembly. *Sci Rep* **5**,
694 13081 (2015).

695 12. Stewart, E.J., Satorius, A.E., Younger, J.G. & Solomon, M.J. Role of environmental and
696 antibiotic stress on Staphylococcus epidermidis biofilm microstructure. *Langmuir* **29**, 7017-
697 7024 (2013).

698 13. Yan, J., Sharo, A.G., Stone, H.A., Wingreen, N.S. & Bassler, B.L. Vibrio cholerae biofilm
699 growth program and architecture revealed by single-cell live imaging. *Proc Natl Acad Sci U*
700 *S A* **113**, E5337-5343 (2016).

14. Drescher, K., Dunkel, J., Nadell, C.D., van Teeffelen, S., Grnja, I., Wingreen, N.S., Stone, H.A. & Bassler, B.L. Architectural transitions in Vibrio cholerae biofilms at single-cell resolution. *Proceedings of the National Academy of Sciences* **113**, E2066-E2072 (2016).

15. Hartmann, R., Singh, P.K., Pearce, P., Mok, R., Song, B., Díaz-Pascual, F., Dunkel, J. & Drescher, K. Emergence of three-dimensional order and structure in growing biofilms. *Nature Physics* **15**, 251-256 (2019).

16. Gregor, I. & Enderlein, J. Image scanning microscopy. *Curr Opin Chem Biol* **51**, 74-83 (2019).

17. Waldchen, S., Lehmann, J., Klein, T., van de Linde, S. & Sauer, M. Light-induced cell damage in live-cell super-resolution microscopy. *Sci Rep* **5**, 15348 (2015).

18. Chen, B.C., Legant, W.R., Wang, K., Shao, L., Milkie, D.E., Davidson, M.W., Janetopoulos, C., Wu, X.S., Hammer, J.A., 3rd, Liu, Z., English, B.P., Mimori-Kiyosue, Y., Romero, D.P., Ritter, A.T., Lippincott-Schwartz, J., Fritz-Laylin, L., Mullins, R.D., Mitchell, D.M., Bembenek, J.N., Reymann, A.C., Bohme, R., Grill, S.W., Wang, J.T., Seydoux, G., Tulu, U.S., Kiehart, D.P. & Betzig, E. Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).

19. Li, D., Shao, L., Chen, B.C., Zhang, X., Zhang, M., Moses, B., Milkie, D.E., Beach, J.R., Hammer, J.A., 3rd, Pasham, M., Kirchhausen, T., Baird, M.A., Davidson, M.W., Xu, P. & Betzig, E. Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science* **349**, aab3500 (2015).

20. Nadell, C.D., Ricaurte, D., Yan, J., Drescher, K. & Bassler, B.L. Flow environment and matrix structure interact to determine spatial competition in Pseudomonas aeruginosa biofilms. *Elife* **6**(2017).

21. Diaz-Pascual, F., Hartmann, R., Lempp, M., Vidakovic, L., Song, B., Jeckel, H., Thormann, K.M., Yildiz, F.H., Dunkel, J., Link, H., Nadell, C.D. & Drescher, K. Breakdown of Vibrio cholerae biofilm architecture induced by antibiotics disrupts community barrier function. *Nat Microbiol* (2019).

22. Chang, B.J., Kittisopikul, M., Dean, K.M., Roudot, P., Welf, E.S. & Fiolka, R. Universal light-sheet generation with field synthesis. *Nat Methods* **16**, 235-238 (2019).

23. Dean, K.M., Roudot, P., Welf, E.S., Danuser, G. & Fiolka, R. Deconvolution-free Subcellular Imaging with Axially Swept Light Sheet Microscopy. *Biophysical journal* **108**, 2807-2815 (2015).

24. Chakraborty, T., Driscoll, M.K., Jeffery, E., Murphy, M.M., Roudot, P., Chang, B.J., Vora, S., Wong, W.M., Nielson, C.D., Zhang, H., Zhemkov, V., Hiremath, C., De La Cruz, E.D., Yi, Y., Bezprozvanny, I., Zhao, H., Tomer, R., Heintzmann, R., Meeks, J.P., Marciano, D.K., Morrison, S.J., Danuser, G., Dean, K.M. & Fiolka, R. Light-sheet microscopy of cleared tissues with isotropic, subcellular resolution. *Nat Methods* **16**, 1109-1113 (2019).

738 25. Dunsby, C. Optically sectioned imaging by oblique plane microscopy. *Optics Express* **16**,
739 20306-20316 (2008).

740 26. Voleti, V., Patel, K.B., Li, W., Perez Campos, C., Bharadwaj, S., Yu, H., Ford, C., Casper,
741 M.J., Yan, R.W., Liang, W., Wen, C., Kimura, K.D., Targoff, K.L. & Hillman, E.M.C. Real-
742 time volumetric microscopy of in vivo dynamics and large-scale samples with SCAPE 2.0.
743 *Nat Methods* **16**, 1054-1062 (2019).

744 27. Flemming, H.C. & Wuertz, S. Bacteria and archaea on Earth and their abundance in biofilms.
745 *Nat Rev Microbiol* **17**, 247-260 (2019).

746 28. York, A. High NA single-objective light-sheet. in
747 *https://andrewgyork.github.io/high_na_single_objective_lightsheet/* (Github, 2020).

748 29. Voleti, V., Patel, K.B., Li, W., Campos, C.P., Bharadwaj, S., Yu, H., Ford, C., Casper, M.J.,
749 Yan, R.W. & Liang, W. Real-time volumetric microscopy of in vivo dynamics and large-
750 scale samples with SCAPE 2.0. *Nature methods* **16**, 1054-1062 (2019).

751 30. Bouchard, M.B., Voleti, V., Mendes, C.S., Lacefield, C., Grueber, W.B., Mann, R.S., Bruno,
752 R.M. & Hillman, E.M. Swept confocally-aligned planar excitation (SCAPE) microscopy for
753 high-speed volumetric imaging of behaving organisms. *Nature photonics* **9**, 113 (2015).

754 31. Ulman, V., Maska, M., Magnusson, K.E.G., Ronneberger, O., Haubold, C., Harder, N.,
755 Matula, P., Matula, P., Svoboda, D., Radojevic, M., Smal, I., Rohr, K., Jalden, J., Blau, H.M.,
756 Dzyubachyk, O., Lelieveldt, B., Xiao, P., Li, Y., Cho, S.Y., Dufour, A.C., Olivo-Marin, J.C.,
757 Reyes-Aldasoro, C.C., Solis-Lemus, J.A., Bensch, R., Brox, T., Stegmaier, J., Mikut, R.,
758 Wolf, S., Hamprecht, F.A., Esteves, T., Quelhas, P., Demirel, O., Malmstrom, L., Jug, F.,
759 Tomancak, P., Meijering, E., Munoz-Barrutia, A., Kozubek, M. & Ortiz-de-Solorzano, C. An
760 objective comparison of cell-tracking algorithms. *Nat Methods* **14**, 1141-1152 (2017).

761 32. Source code available at: https://github.com/GahlmannLab/BCM3D.git.

762 33. Chen, B.C., Legant, W.R., Wang, K., Shao, L., Milkie, D.E., Davidson, M.W., Janetopoulos,
763 C., Wu, X.S., Hammer, J.A. & Liu, Z. Lattice light-sheet microscopy: imaging molecules to
764 embryos at high spatiotemporal resolution. *Science* **346**, 1257998 (2014).

765 34. Valm, A.M., Cohen, S., Legant, W.R., Melunis, J., Hershberg, U., Wait, E., Cohen, A.R.,
766 Davidson, M.W., Betzig, E. & Lippincott-Schwartz, J. Applying systems-level spectral
767 imaging and analysis to reveal the organelle interactome. *Nature* **546**, 162 (2017).

768 35. Glass, D.S. & Riedel-Kruse, I.H. A synthetic bacterial cell-cell adhesion toolbox for
769 programming multicellular morphologies and patterns. *Cell* **174**, 649-658. e616 (2018).

770 36. Cotter, C.R., Schuttler, H.B., Igoshin, O.A. & Shimkets, L.J. Data-driven modeling reveals
771 cell behaviors controlling self-organization during Myxococcus xanthus development. *Proc*
772 *Natl Acad Sci U S A* **114**, E4592-E4601 (2017).

773 37. Gao, R., Asano, S.M., Upadhyayula, S., Pisarev, I., Milkie, D.E., Liu, T.-L., Singh, V.,
774 Graves, A., Huynh, G.H. & Zhao, Y. Cortical column and whole-brain imaging with
775 molecular contrast and nanoscale resolution. *Science* **363**, eaau8302 (2019).

776 38. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
777 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein,
778 V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-
779 image analysis. *Nature Methods* **9**, 676-682 (2012).

780 39. Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H. & Ferrin,
781 T.E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci*
782 **27**, 14-25 (2018).

783 40. Rudge, T.J., Steiner, P.J., Phillips, A. & Haseloff, J. Computational modeling of synthetic
784 microbial biofilms. *ACS Synth Biol* **1**, 345-352 (2012).

785 41. Reshes, G., Vanounou, S., Fishov, I. & Feingold, M. Cell shape dynamics in Escherichia
786 coli. *Biophysical journal* **94**, 251-264 (2008).

787 42. Starruß, J., Peruani, F., Jakovljevic, V., Søgaard-Andersen, L., Deutsch, A. & Bär, M.
788 Pattern-formation mechanisms in motility mutants of Myxococcus xanthus. *Interface focus* **2**,
789 774-785 (2012).

790 43. Harris, L.G., Foster, S. & Richards, R.G. An introduction to Staphylococcus aureus, and
791 techniques for identifying and quantifying S. aureus adhesins in relation to adhesion to
792 biomaterials: review. *Eur Cell Mater* **4**, 39-60 (2002).

793 44. Yan, T., Richardson, C.J., Zhang, M. & Gahlmann, A. Computational correction of spatially
794 variant optical aberrations in 3D single-molecule localization microscopy. *Optics Express* **27**,
795 12582-12599 (2019).

796 45. Muller, F.D., Schink, C.W., Hoiczyk, E., Cserti, E. & Higgs, P.I. Spore formation in
797 Myxococcus xanthus is tied to cytoskeleton functions and polysaccharide spore coat
798 deposition. *Mol Microbiol* **83**, 486-505 (2012).

799 46. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T. & Ronneberger, O. 3D U-Net: Learning
800 Dense Volumetric Segmentation from Sparse Annotation. 424-432 (Springer International
801 Publishing, Cham, 2016).

802 47. Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D.I., Wang, G., Eaton-Rosen, Z., Gray, R.,
803 Doel, T., Hu, Y., Whyntie, T., Nachev, P., Modat, M., Barratt, D.C., Ourselin, S., Cardoso,
804 M.J. & Vercauteren, T. NiftyNet: a deep-learning platform for medical imaging. *Computer*
805 *Methods and Programs in Biomedicine* **158**, 113-122 (2018).

806 48. Van Valen, D.A., Kudo, T., Lane, K.M., Macklin, D.N., Quach, N.T., DeFelice, M.M.,
807 Maayan, I., Tanouchi, Y., Ashley, E.A. & Covert, M.W. Deep Learning Automates the
808 Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments. *PLoS Comput*
809 *Biol* **12**, e1005177 (2016).

49. Acton, S.T. Fast Algorithms for Area Morphology. *Digit Signal Process* **11**, 187-203 (2001).

50. Yu, Y. & Acton, S.T. Speckle reducing anisotropic diffusion. *Trans. Img. Proc.* **11**, 1260-1270 (2002).

51. Tabassum, N., Vaccari, A. & Acton, S. Speckle removal and change preservation by distance-driven anisotropic diffusion of synthetic aperture radar temporal stacks. *Digit Signal Process* **74**, 43-55 (2018).

52. Choi, H.I., Choi, S.W. & Moon, H.P. Mathematical theory of medial axis transform. *Pacific Journal of Mathematics* **181**, 57-88 (1997).

53. Yan, T.Q. & Zhou, C.X. A Continuous Skeletonization Method Based on Distance Transform. *Comm Com Inf Sc* **304**, 251-258 (2012).

54. Wang, J., Batabyal, T., Zhang, M.X., Zhang, J., Aziz, A., Gahlmann, A. & Acton, S.T. LCuts: Linear Clustering of Bacteria Using Recursive Graph Cuts. *2019 IEEE International Conference on Image Processing (ICIP)*, 1575-1579 (2019).

55. Source code available at: https://github.com/jwang-c/Postprocessing-using-LCuts.

56. Shi, J.B. & Malik, J. Normalized cuts and image segmentation. *Ieee Transactions on Pattern Analysis and Machine Intelligence* **22**, 888-905 (2000).

57. Jaccard, P. The distribution of the flora in the alpine zone. *New Phytologist* **11**, 37-50 (1912).

58. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S. & Schmid, B. Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682 (2012).

59. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* 226-231 (AAAI Press, Portland, Oregon, 1996).

60. Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M. & Van Valen, D. Deep learning for cellular image analysis. *Nat Methods* (2019).

61. Caicedo, J.C., Roth, J., Goodman, A., Becker, T., Karhohs, K.W., Broisin, M., Molnar, C., McQuin, C., Singh, S., Theis, F.J. & Carpenter, A.E. Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry Part A* **95**, 952-965 (2019).

62. Hartmann, R., Jeckel, H., Jelli, E., Singh, P.K., Vaidya, S., Bayer, M., Vidakovic, L., Díaz-Pascual, F., Fong, J.C.N., Dragoš, A., Besharova, O., Nadell, C.D., Sourjik, V., Kovács, Á.T., Yildiz, F.H. & Drescher, K. BiofilmQ, a software tool for quantitative image analysis of microbial biofilm communities. *bioRxiv* (2019).

843 63. Reyer, M.A., McLean, E.L., Chennakesavalu, S. & Fei, J. An Automated Image Analysis
844 Method for Segmenting Fluorescent Bacteria in Three Dimensions. *Biochemistry* **57**, 209-
845 215 (2018).

846 64. Liu, J., Prindle, A., Humphries, J., Gabalda-Sagarra, M., Asally, M., Lee, D.Y., Ly, S.,
847 Garcia-Ojalvo, J. & Suel, G.M. Metabolic co-dependence gives rise to collective oscillations
848 within biofilms. *Nature* **523**, 550-554 (2015).

849 65. Prindle, A., Liu, J., Asally, M., Ly, S., Garcia-Ojalvo, J. & Suel, G.M. Ion channels enable
850 electrical communication in bacterial communities. *Nature* **527**, 59-63 (2015).

851 66. Humphries, J., Xiong, L., Liu, J., Prindle, A., Yuan, F., Arjes, H.A., Tsimring, L. & Suel,
852 G.M. Species-Independent Attraction to Biofilms through Electrical Signaling. *Cell* **168**,
853 200-209 e212 (2017).

854 67. Liu, J., Martinez-Corral, R., Prindle, A., Lee, D.-y.D., Larkin, J., Gabalda-Sagarra, M.,
855 Garcia-Ojalvo, J. & Süel, G.M. Coupling between distant biofilms and emergence of nutrient
856 time-sharing. *Science* **356**, 638-642 (2017).

857 68. Mitri, S. & Foster, K.R. The genotypic view of social interactions in microbial communities.
858 *Annu Rev Genet* **47**, 247-273 (2013).

859 69. Drescher, K., Nadell, C.D., Stone, H.A., Wingreen, N.S. & Bassler, B.L. Solutions to the
860 public goods dilemma in bacterial biofilms. *Current biology : CB* **24**, 50-55 (2014).

861 70. Persat, A., Nadell, C.D., Kim, M.K., Ingremeau, F., Siryaporn, A., Drescher, K., Wingreen,
862 N.S., Bassler, B.L., Gitai, Z. & Stone, H.A. The mechanical world of bacteria. *Cell* **161**, 988-
863 997 (2015).

864 71. Papenfort, K. & Bassler, B.L. Quorum sensing signal-response systems in Gram-negative
865 bacteria. *Nat Rev Microbiol* **14**, 576-588 (2016).

866 72. Kroos, L. Highly Signal-Responsive Gene Regulatory Network Governing Myxococcus
867 Development. *Trends Genet* **33**, 3-15 (2017).

868 73. Moon, S., Yan, R., Kenny, S.J., Shyu, Y., Xiang, L., Li, W. & Xu, K. Spectrally resolved,
869 functional super-resolution microscopy reveals nanoscale compositional heterogeneity in
870 live-cell membranes. *Journal of the American Chemical Society* **139**, 10944-10947 (2017).

871 74. Bramkamp, M. & Lopez, D. Exploring the existence of lipid rafts in bacteria. *Microbiol.*
872 *Mol. Biol. Rev.* **79**, 81-100 (2015).

873 75. Zou, S.B., Hersch, S.J., Roy, H., Wiggers, J.B., Leung, A.S., Buranyi, S., Xie, J.L., Dare, K.,
874 Ibba, M. & Navarre, W.W. Loss of elongation factor P disrupts bacterial outer membrane
875 integrity. *Journal of bacteriology* **194**, 413-425 (2012).

76. Gonelimali, F.D., Lin, J., Miao, W., Xuan, J., Charles, F., Chen, M. & Hatab, S.R. Antimicrobial properties and mechanism of action of some plant extracts against food pathogens and spoilage microorganisms. *Frontiers in microbiology* **9**, 1639 (2018).

77. Parasassi, T., De Stasio, G., d'Ubaldo, A. & Gratton, E. Phase fluctuation in phospholipid membranes revealed by Laurdan fluorescence. *Biophysical journal* **57**, 1179-1186 (1990).

78. Strahl, H. & Hamoen, L.W. Membrane potential is important for bacterial cell division. *Proceedings of the National Academy of Sciences* **107**, 12281-12286 (2010).

79. Prindle, A., Liu, J., Asally, M., Ly, S., Garcia-Ojalvo, J. & Süel, G.M. Ion channels enable electrical communication in bacterial communities. *Nature* **527**, 59-63 (2015).

80. Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C. & Culley, S. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature methods* **15**, 1090-1097 (2018).

81. Krull, A., Vicar, T. & Jug, F. Probabilistic Noise2Void: Unsupervised content-aware denoising. *arXiv preprint arXiv:1906.00651* (2019).

82. Krull, A., Buchholz, T.-O. & Jug, F. Noise2void-learning denoising from single noisy images. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2129-2137 (2019).

83. Buchholz, T.-O., Prakash, M., Krull, A. & Jug, F. DenoiSeg: Joint Denoising and Segmentation. *arXiv preprint arXiv:2005.02987* (2020).

Figure S1

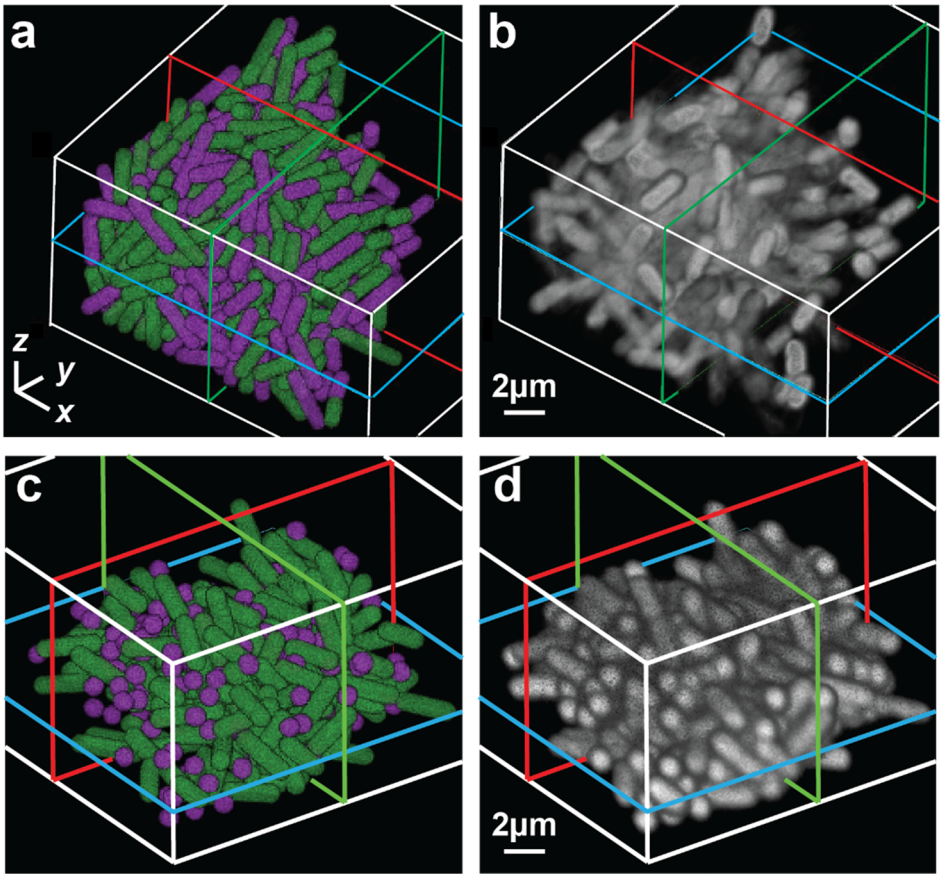


Figure S1. Simulation of mixed labeling and mixed cell shape biofilms. (**a**) Cell arrangements (green indicates membrane labeled cells, purple indicates membrane labeled cells that simultaneously express interior fluorescence protein). (**b**) Fluorescence image based on the cell arrangements in (a) as displayed by the volume viewer plugin of Fiji[1]. (**c**) Cell arrangements (green indicates rod-shaped cells, magenta indicates spherical shaped cells). (**d**) Fluorescence image based on the cell arrangements in (c) as displayed by the volume viewer plugin of Fiji[1].

Figure S2

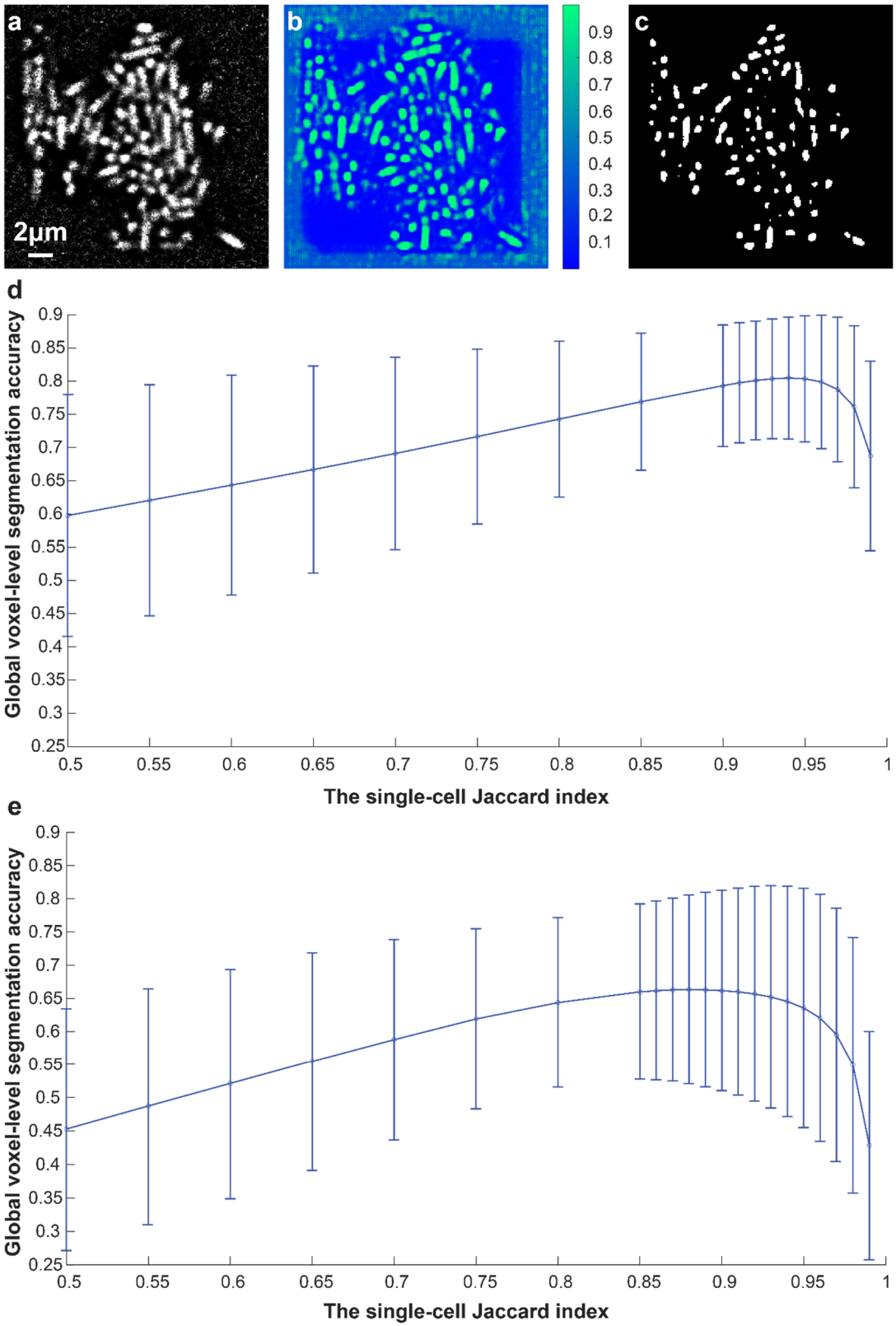

a
2μm
b
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
c
d
Global voxel-level segmentation accuracy
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
The single-cell Jaccard index
e
Global voxel-level segmentation accuracy
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
The single-cell Jaccard index

Figure S2. Binary segmentation result produced by thresholding the 'cell interior' confidence map at a high value (0.88-0.94). (**a**) Deconvolved fluorescence image. (**b**) 'Cell interior' confidence map. (**c**) Binary segmentation result (confidence threshold = 0.94). (**d** and **e**) Global voxel-level segmentation accuracy (y axis) versus the threshold of single-cell shape estimation (*x* axis, quantified as the single-cell Jaccard index) for cells labeled with cytosolic fluorophores (d) and cells labeled with membrane-localized fluorophores (e).

Figure S3

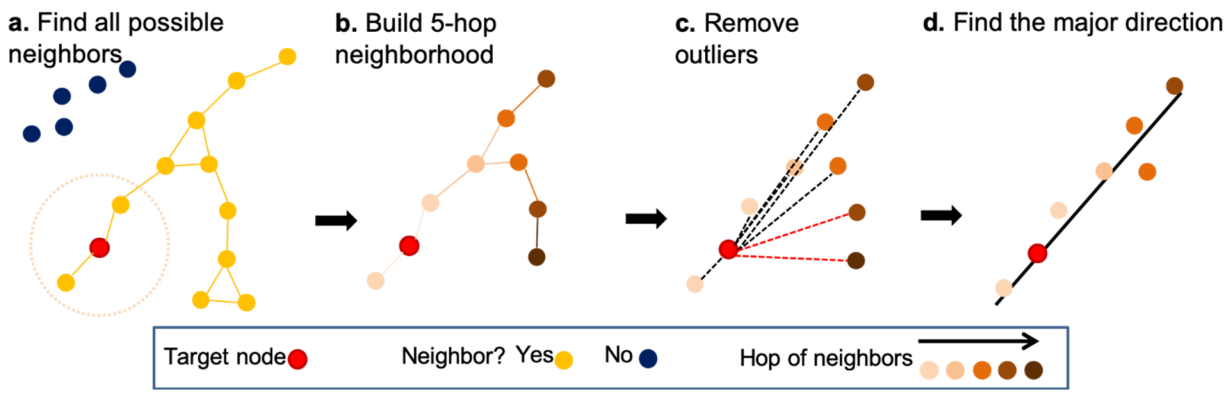


Figure S3. Determination of node direction in an outlier-removed neighborhood[2]. (**a**) A neighborhood of the target node (in red) is a sub-graph, where all adjacent nodes (in yellow) are connected via edges to the target node. Here, if the distance of two nodes is less than a chosen value (indicated by the dashed circle), these nodes are adjacent to each other. The blue dots are not part of the neighborhood. (**b**) A hop is defined as the number of edges that one has to traverse from one node to the other node in the graph. Here, the 5-hop neighborhood of the target node is shown. (**c**) The directional vectors are found from the target node to all the other nodes within in the 5-hop neighborhood (dashed lines). The nodes are classified as outliers if they have large relative angles compared to all the other directional vectors (red dashed lines). (**d**) Finally, the direction feature of the current node is evaluated as the major direction of the outlier removed neighborhood using principle component analysis.

Figure S4

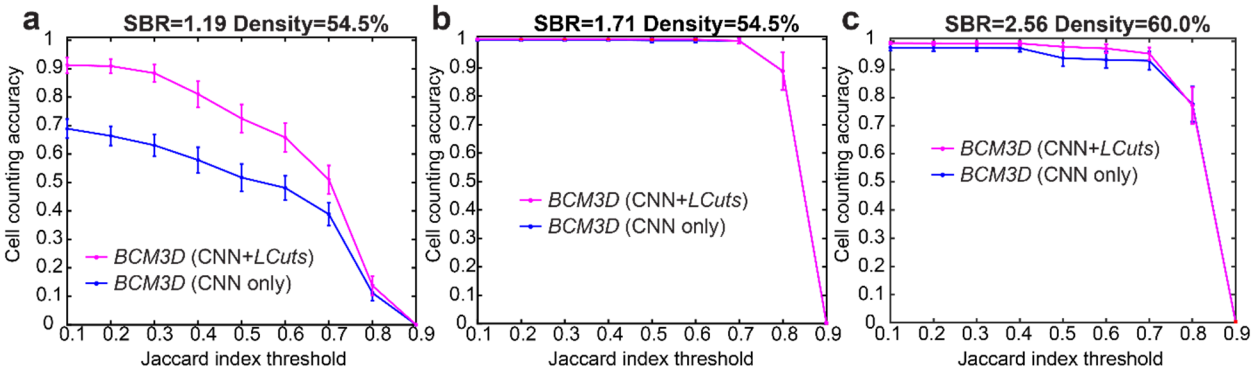


Figure S4. Segmentation accuracies achieved for biofilm images with different SBR and cell densities. (**a**), (**b**), and (**c**) represent results for datasets with SBR 1.19, Density 54.5%, SBR 1.71, Density 54.5% and SBR 2.56, Density 54.5%, respectively. Each curve is plotted by averaging 10 different datasets.

Figure S5

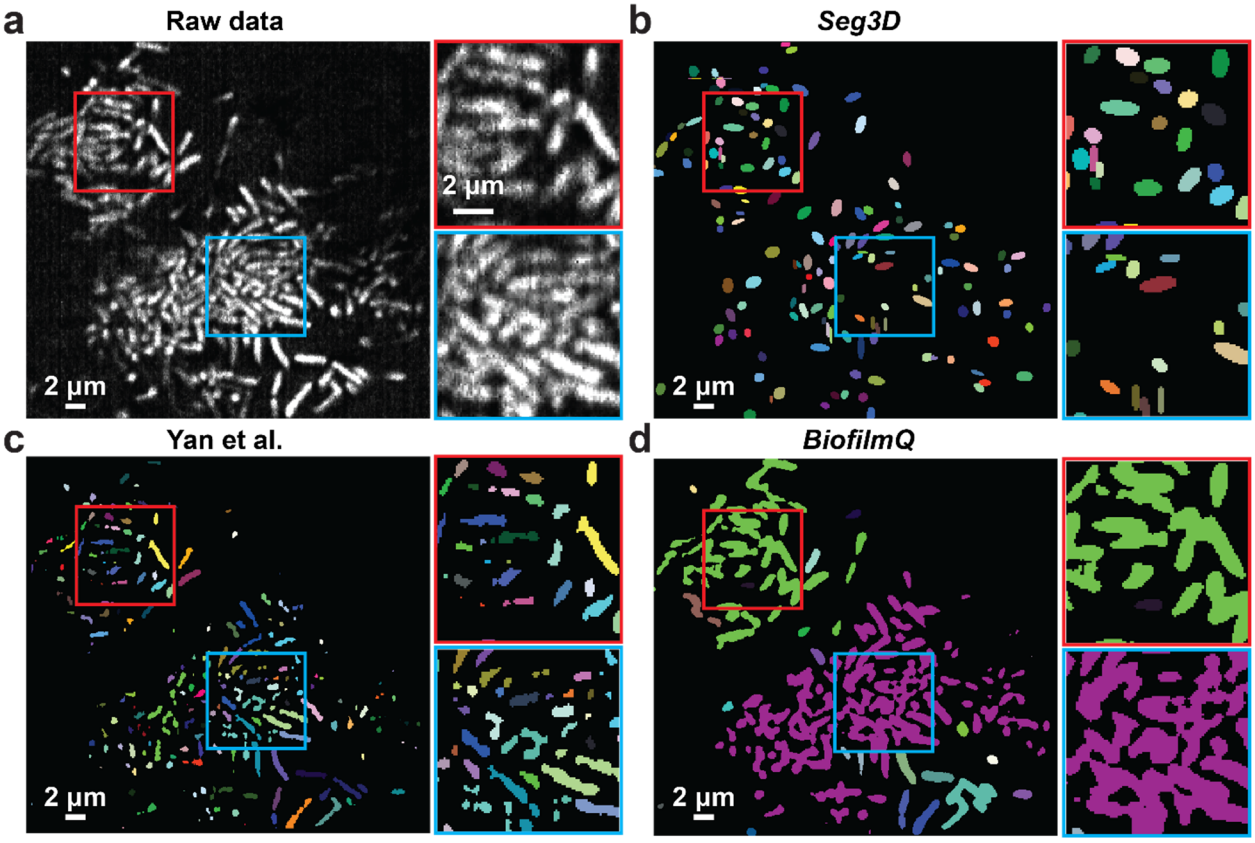


Figure S5. Visual comparison of segmentation results achieved by previous segmentation approaches that rely solely on mathematical image processing. (**a**) Experimental dataset is a *E. coli* biofilm containing GFP expressing cells 10 hours after the inoculation. (**b**) Segmentation result obtained using *Seg3D*[3]. (**c**) Segmentation result obtained using the algorithm in Yan *et al.*[4]. (**d**) Segmentation result obtained using *BiofilmQ*[5].

Figure S6

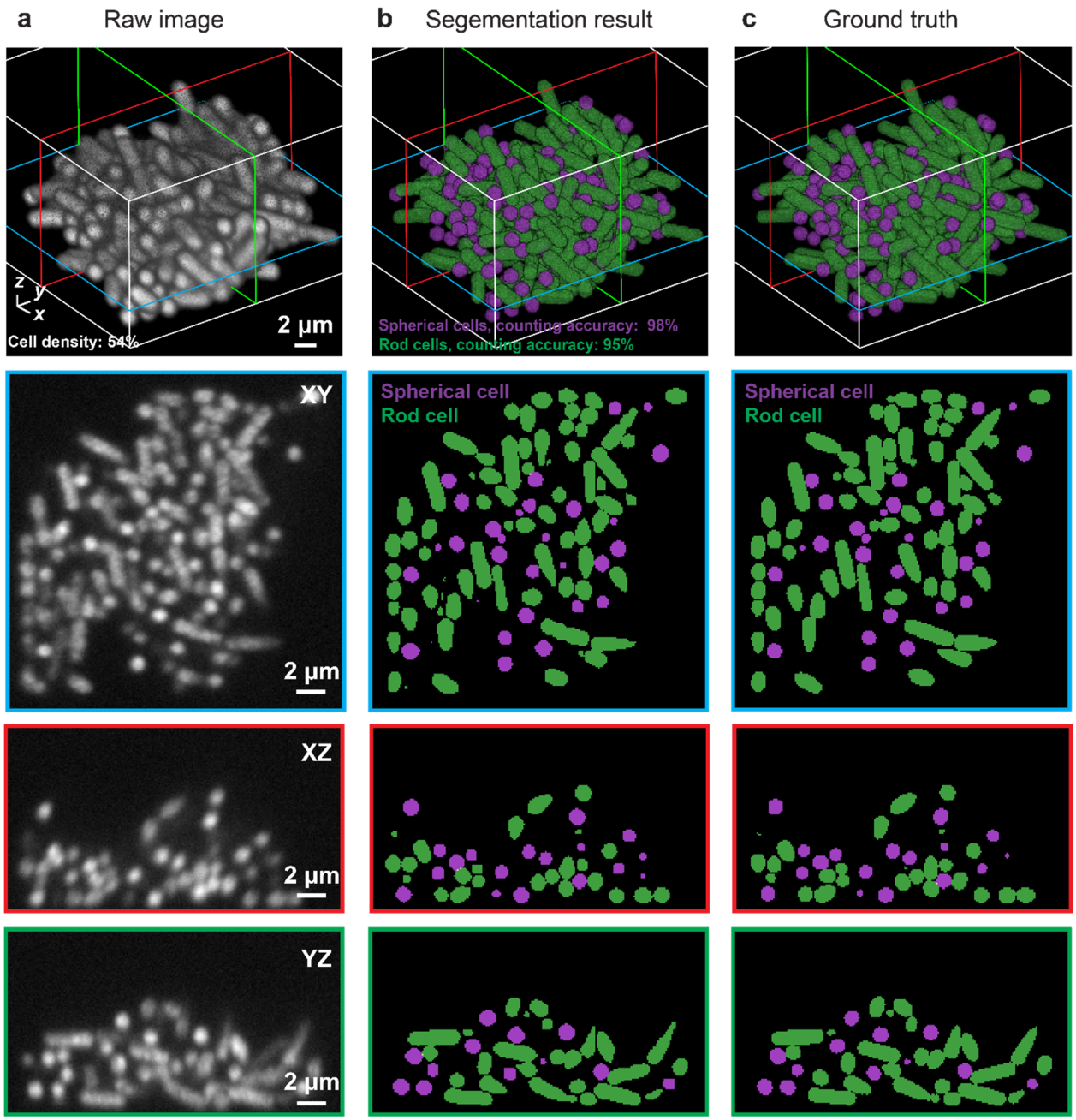


Figure S6. Segmentation of mixed population biofilms containing spherical cells and rod-shaped cells. (**a**) Simulated fluorescence image of a mixture of spherical cells and rod-shaped cells. The mixing ratio is 50:50. (**b**) Segmentation result obtained by *BCM3D* and (**c**) ground truth. Three orthogonal planes are shown below each 3D image.

Figure S7

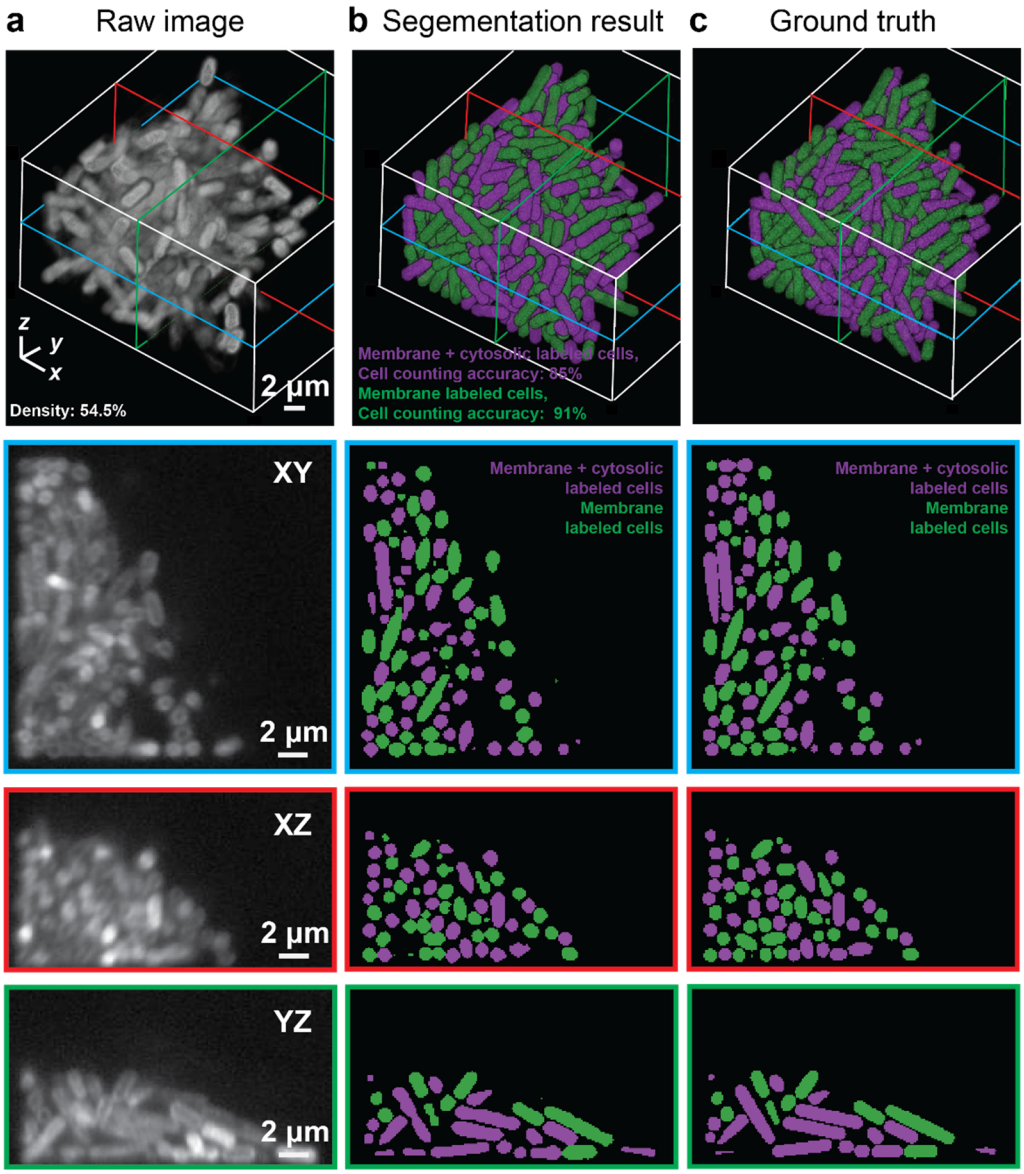


Figure S7. Segmentation of mixed population biofilms containing membrane-stained cells and membrane-stained cells that additionally express an intracellular fluorescent protein. (**a**) Simulated fluorescence image of a mixture of membrane-stained cells and membrane-stained cells that additionally express an intracellular fluorescent protein. The mixing ratio is 50:50. (**b**) Segmentation result obtained by *BCM3D* and (**c**) ground truth. Three orthogonal planes are shown below each 3D image.

Figure S8

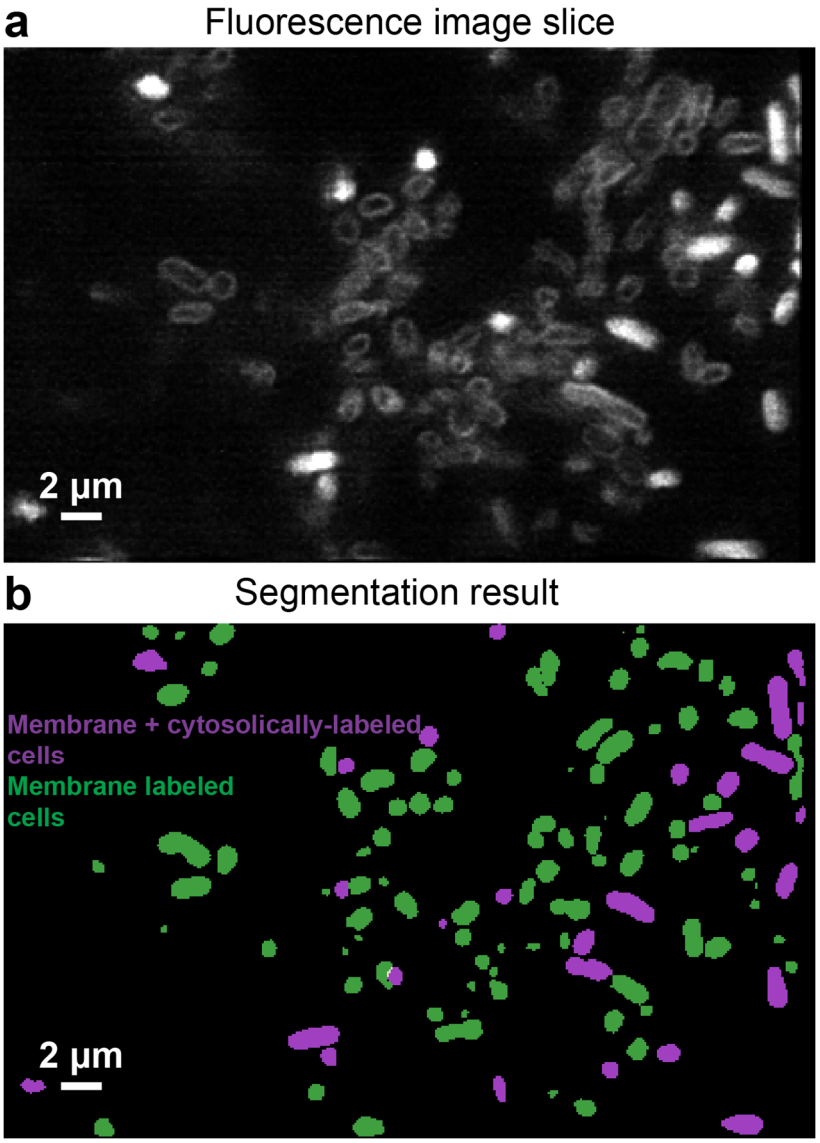


Figure S8. (**a**) Experimental 2D slice of a mixed *E. coli* population containing membrane-stained cells and membrane-stained cells that additionally express an intracellular fluorescent protein. The mixing ratio at the time of inoculation was 50:50. All cells were labeled by the FM4-64 membrane-intercalating dye. **(b)** *BCM3D* segmentation result corresponding to the image shown in (a). Membrane-stained cells are displayed in green, and cells that were both membrane-stained and cytosolically-labeled are displayed in magenta.

References

1. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S. & Schmid, B. Fiji: an open-source platform for biological-image analysis. *Nature methods* **9**, 676-682 (2012).

2. Wang, J., Batabyal, T., Zhang, M.X., Zhang, J., Aziz, A., Gahlmann, A. & Acton, S.T. LCuts: Linear Clustering of Bacteria Using Recursive Graph Cuts. *2019 IEEE International Conference on Image Processing (ICIP)*, 1575-1579 (2019).

3. Reyer, M.A., McLean, E.L., Chennakesavalu, S. & Fei, J. An Automated Image Analysis Method for Segmenting Fluorescent Bacteria in Three Dimensions. *Biochemistry* **57**, 209-215 (2018).

4. Yan, J., Sharo, A.G., Stone, H.A., Wingreen, N.S. & Bassler, B.L. Vibrio cholerae biofilm growth program and architecture revealed by single-cell live imaging. *Proc Natl Acad Sci U S A* **113**, E5337-5343 (2016).

5. Hartmann, R., Jeckel, H., Jelli, E., Singh, P.K., Vaidya, S., Bayer, M., Vidakovic, L., Díaz-Pascual, F., Fong, J.C.N., Dragoš, A., Besharova, O., Nadell, C.D., Sourjik, V., Kovács, Á.T., Yildiz, F.H. & Drescher, K. BiofilmQ, a software tool for quantitative image analysis of microbial biofilm communities. BioRxiv 735423 (2019).