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 ratios (SBRs) 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 SBRs, cell densities, labeling 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.

35 # Introduction

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

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

While super-resolution derivatives of confocal microscopy, known as Image Scanning Microscopy[16], can improve spatial resolution, 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 however 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 the cell surface. To overcome the drawbacks of traditional
92 mathematical image processing approaches, automated solutions based on supervised training of
93 deep convolutional neural networks (CNNs) have been used in recent years with outstanding
94 success for a wide range of problems in biomedical image analysis[31].

95 Here, we present Bacterial Cell Morphometry 3D (*BCM3D*)[32], a generally applicable
96 workflow for single-cell segmentation and shape determination in high-resolution 3D images of
97 bacterial biofilms. *BCM3D* uses CNNs, *in silico*-trained with computationally simulated biofilm
98 images, in combination with mathematical image analysis to achieve accurate single cell
99 segmentation in 3D. The CNNs employed in *BCM3D* are based on the 3D U-Net architecture and
100 training strategy, which has achieved excellent performance in biomedical data analysis
101 benchmark tests[31]. The mathematical image analysis modules of *BCM3D* enable post-processing
102 of the CNN results to further improve the segmentation accuracy. We establish that experimental
103 bacterial biofilms images, acquired by lattice light sheet microscopy, can be successfully

segmented using CNNs trained with computationally simulated biofilm 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. *BCM3D* also achieves marginally higher segmentation accuracy on 3D biofilm data than *Cellpose*[33], a state-of-the-art, CNN-based, generalist algorithm for cellular segmentation. 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[34,35]. 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

127 sheet reduces phototoxicity, because each excitation photon has multiple opportunities to be
128 absorbed by fluorophores in the excitation plane and produce in-focus fluorescence. Widefield
129 fluorescence images corresponding to each illuminated specimen plane are recorded on a sCMOS
130 detector (Hamamatsu ORCA Flash v2). In this work, 3D biofilm images were acquired by
131 translating the specimen through the light sheet in 200 nm steps using a piezo nanopositioning
132 stage (Physik Instrumente, P-621.1CD).

133 *E.coli* K12 Biofilm Imaging: Ampicillin resistant *E.coli* K12, constitutively expressing
134 GFP[36], were cultured at 37 degrees overnight in LB medium with 100 μg/ml ampicillin. Overnight
135 cultures were diluted 100 times into the same culture medium, grown to an optical density at 600
136 nm (OD600) of 0.6 – 1.0, and then diluted by an additional factor of 10. Round glass coverslips
137 with the diameter of 5 mm were put into a 24-well plate (Falcon) and 400 μL of cell culture was
138 added to the well. Cells were allowed to settle to the bottom of the well and adhere to the coverslip
139 for 1 hour. The round coverslips were then mounted onto a sample holder and placed into the
140 LLSM sample-basin filled with M9 medium. GFP fluorescence was excited using 488 nm light
141 sheet excitation. Biofilm growth was imaged at room temperature every 30 min for a total of 20
142 time points. At each time point, a single 3D image stack contained 400 images, each acquired with
143 a 15 ms exposure time to avoid motion blur.

144 *M. xanthus* Biofilm Imaging: Strain LS3908 expressing tdTomato under the control of the
145 IPTG-inducible promoter[37] and DK1622 (WT) were cultured in the nutrient rich CYE media at 30
146 degrees Celsius until it reached an OD600 of 0.6 - 1.0. Media was supplemented with 1 mM IPTG
147 for tdTomato expressing cells. Chitosan (Thermo Fisher)-coated 5 mm round glass coverslips were
148 prepared by incubating coverslips with 1% (w/v) chitosan (1.5 % glacial acetic acid (v/v)) at room
149 temperature for 1 hour. Coverslips were then rinsed with water and placed into a 24-well plate

150 (Falcon) with 350-400 µL of undiluted cell culture. WT cells were stained directly in the 24 well
151 plate with 5 ng/ml FM4-64 (Thermo Fisher) dye. Cells were allowed to settle and adhere to the
152 coverslip for 2 hours. After the settling period, the coverslip was gently rinsed with CYE media to
153 flush away unattached cells. The rinsed coverslip was then mounted onto a sample holder and
154 placed into the LLSM sample-basin filled with MC7 starvation buffer. tdTomato and FM 4-64
155 fluorescence was excited using 561 nm light sheet. The 3D image stack contained 400 2D images.
156 Each 2D slice was acquired with an exposure time of 30 ms.

157 Mixed Strain Biofilm Imaging: Ampicillin resistant *E.coli* K12, constitutively expressing
158 GFP[36], and ampicillin resistant *E.coli* K12, expressing mScarlet (pBAD vector, arabinose induce)
159 were cultured separately at 37 degrees overnight in LB medium with 100 µg/ml ampicillin.
160 Overnight cultures were diluted 100 times into the same culture medium, grown to an optical
161 density at 600 nm (OD600) of 0.6 – 1.0, and then diluted to an OD of 0.1. After dilution, the two
162 strains were mixed together. Round glass coverslips with the diameter of 5 mm were put into a 24-
163 well plate (Falcon) and 500 µL of cell culture was added to the well. Cells were allowed to settle
164 to the bottom of the well and adhere to the coverslip for 1 hour. The cell culture medium was then
165 removed and replaced by 500 uL M9 medium containing 0.2% (w/v) arabinose. The co-culture
166 was incubated at 30 degrees overnight. 10 mins before imaging, the co-culture was stained with
167 5 ng/ml FM4-64 (Thermo Fisher) dye. 3D image stacks of 20 planes with 5 ms exposure time per
168 frame were acquired using 488 nm excitation.

169

170 **Raw Data Processing**

171 Raw 3D images were background subtracted and then deskewed and deconvolved as
172 described previously[18,19]. The background was estimated by averaging intensity values of dark

173 areas (devoid of cells) in the field of view. Deconvolution was performed using the Richardson-
174 Lucy algorithm with 10 iterations using experimentally measured point spread functions (PSFs) as
175 the deconvolution kernel. The experimentally measured PSFs were obtained separately for each
176 color channel using fluorescent beads (200 nm FluoSpheres®, Thermo Fisher) coated on a
177 coverslip[38]. 3D images were rendered using the 3D Viewer plugin in Fiji[39] or ChimeraX[40].
178

179 **Generation of simulated biofilm images**

180 To generate data for training of CNNs, we computationally simulated fluorescence images
181 of 3D biofilms, for which spatial arrangements among individual cells are known precisely and
182 accurately. Growth and division of individual rod-shaped cells in a population were simulated
183 using CellModeller, an individual-based computational model of biofilm growth (**Figure 1a**)[41]. In
184 individual-based biofilm growth models, cells are the basic modeling units. Each cell is
185 characterized by a set of parameters, including its 3D position, volume, and spatial orientation. All
186 the cells in the simulated biofilm are then allowed to evolve in time according to predefined
187 biological, chemical, and mechanical rules. For example, cells grow at a defined rate and then
188 divide after reaching a certain volume threshold. Cellular collisions that are due to cell growth are
189 alleviated by imposing a minimum distance criterion between cells at each time point. For our
190 simulations, we chose cell diameter and cell length (*d, l*) parameters consistent with a given
191 bacterial species, namely (1 μm, 3 μm) for *E. coli*[42], (0.7 μm, 6 μm) for *M. xanthus*[43], and (1 μm,
192 1 μm) for spherically symmetric *S.aureus*[44].
193

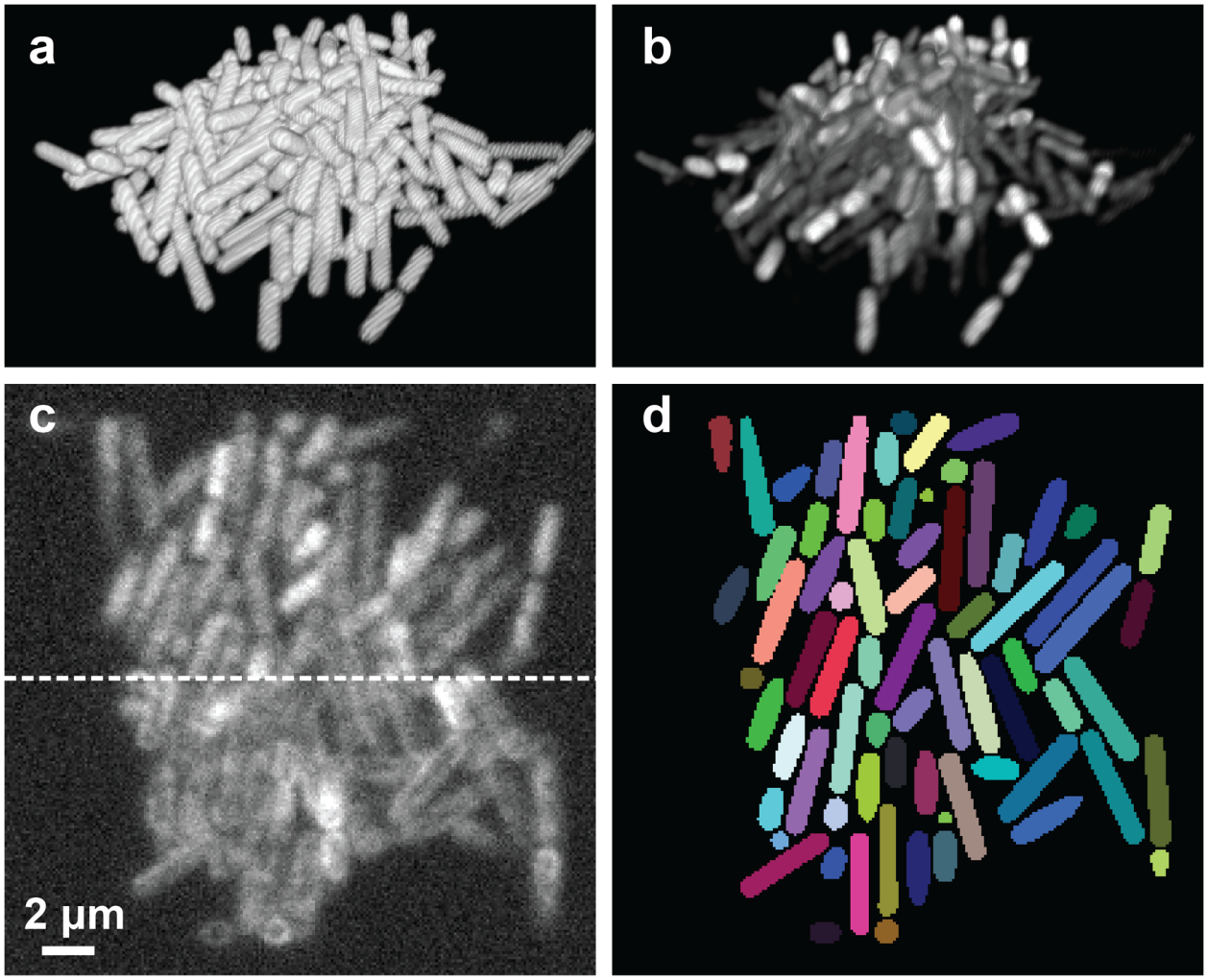


Figure 1. Simulation of fluorescent biofilms images and annotation maps. **(a)** Cell arrangements obtained by CellModeller. **(b)** Simulated 3D fluorescence image based on the cell arrangements in **a**. **(c)** XY slice through the 3D simulated fluorescence image in **b** (upper panel shows cells 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).

While the cell volume can be readily adjusted in CellModeller, the cellular volume density, which is determined by the intercellular spacing, is not directly adjustable. We therefore adjusted the cellular volume density after each simulation by scaling the cellular positions (cell centroids)

206 and thus the intercellular distances by a constant factor, while leaving cell sizes, shapes, and
207 orientations unchanged. This post-processing procedure enabled simulation of the exact same 3D
208 cell arrangements at adjustable cell volume densities.

209 We fluorescently labeled simulated cell volumes and surfaces according to two commonly
210 used labeling strategies in fluorescence microscopy. To simulate expression of intracellular
211 fluorescent proteins, the fluorescence emitters were placed at random positions within the cell
212 volume. To simulate membrane staining, the fluorescence emitters were placed at random
213 positions on the cell surface. Each cell contained between 500 - 1000 fluorophores to simulate
214 expression level variations between cells, which is often observed in experimental images. Once
215 the fluorophore spatial distributions were determined, a 3D fluorescence image (**Figure 1b**) was
216 computationally generated. Each fluorophore was treated as an isotropic point emitter, so that it
217 would produce a diffraction-limited point-spread-function (PSF) on the detector. Experimentally
218 measured 3D PSF shapes (see Raw Data Processing) were used as the convolution kernel. Next,
219 the fluorescence signal intensity was scaled by multiplying the image by a constant factor and then
220 a constant background intensity was added to the image at ~200 photons per pixel, as measured in
221 experimental data. This procedure enabled independent adjustments of the fluorescence signal and
222 background to obtain signal-to-background ratios (SBRs) consistent with experimental data. In a
223 final step, we introduced Poisson-distributed counting noise, based on the summed background
224 and signal intensities, as well as Gaussian-distributed camera read-out noise (experimentally
225 calibrated for our detector at 3.04 photons per pixel on average)[45]. This resulting image data
226 (**Figure 1c**) was then processed in the same manner as experimental data (see Raw Data
227 Processing). In contrast to experimental data, generation of the corresponding voxel-level

228 annotation maps is fast and error free, because the underlying ground truth cell arrangements are
229 known *a priori* (**Figure 1d**).

230 Differential labeling: To mimic imaging of reporter gene expression in a subset of cells,
231 we simulated biofilm images, in which all cells were stained at the cell surface (e.g. with a
232 membrane intercalating fluorescent dye) and a subset of cells additionally contained intracellular
233 fluorophores (e.g. through the expression of an intracellular fluorescent protein) (**Figure S1ab**).
234 The mixing ratios between membrane-labelled, and membrane and interior labelled cells were
235 10:90, 30:70, 50:50, 70:30 and 90:10. Ten different cell arrangements containing ~300 cells were
236 simulated for each ratio. To train the CNNs (see next section), six datasets were used, all with a
237 50:50 mixing ratio.

238 Mixed cell shapes: To mimic imaging of cells with different morphologies, we simulated
239 biofilms containing spherical and rod-shaped cells **(Figure S1cd)**. The spatial arrangements of the
240 mixed biofilm were generated by randomly positioning cells in a 3D volume. Collision among
241 cells was eliminated by adjusting cell positions and orientations. The size of the rod-shaped cells
242 is that of *E. coli* (~3 × 1 μm, length by diameter). The size of the spherical cells is that of *S. aureus*
243 (~1 μm in diameter)[46]. Both cell types were labelled by intracellular fluorophores, as described
244 above. The mixing ratios between rod-shaped and spherical cells were 10:90, 30:70, 50:50, 70:30
245 and 90:10. Ten different cell arrangements containing ~300 cells were simulated for each ratio. To
246 train the CNNs (see next section), we picked one image from each mixing ratio for a total of five
247 images.

248

249 **Training the convolutional neural networks**

250 We trained 3D U-Net CNNs for voxel-level classification tasks[47] within the NiftyNet platform[48]
251 (network architecture depth 4, convolution kernel size 3, ReLU activation function, 32 initial
252 feature maps, and random dropout of 0.5 during training). To achieve robust performance, we
253 trained these networks using five to ten simulated biofilm images with randomly selected cell
254 densities and signal-to-background ratios (see Generation of simulated biofilm images). The raw
255 data processing steps used for experimental data (see Raw Data Processing) were equally applied
256 to simulated data. 3D deconvolved simulated data and their corresponding voxel-level annotations
257 were used to train the CNNs. Each image used for training contained ~9 million voxels. We trained
258 CNNs by classifying each voxel as 'background', 'cell interior' or as 'cell boundary' based on the
259 underlying cell arrangements. For mixed-species biofilms, two additional classes, 'cell interior'
260 and 'cell boundary' of the second species, were used. This type of annotation scheme has been
261 shown to increase separation of bacterial cells in 2D[49]. For data augmentation, we applied
262 NiftyNet's built-in scaling, rotation, and elastic deformation functions. Instead of the original
263 cross-entropy loss function combined with uniform sampling, we used the Dice loss function and
264 'balanced sampler', so that every label has the same probability of occurrence in training. All
265 networks were trained for 2000 to 3600 iterations with a learning rate of 0.0001. Using these
266 parameters, it took approximately 24 hours to train the CNNs on a NVIDIA Tesla V100 GPU with
267 16 GB memory.

268

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

270 Voxel-level classification by CNNs generates different confidence maps (one confidence
271 map for each annotation class). The confidence values range between 0 and 1 and represent the

272 confidence of assigning individual voxels to a given class. After thresholding the 'cell interior'
273 confidence map to obtain a binary image (**Figure S2a-c**), connected voxel clusters can be isolated
274 and identified as single cell objects using 3D connected component labeling[50]. A conservative size-
275 exclusion filter was applied: small objects with a volume ~10 times less than the expected cell size
276 were considered background noise and filtered out using an area open operator[50]. Since the cell-
277 interior volumes do not contain the cell boundaries, we dilated each object by 1-2 voxels to
278 increase the cell volumes using standard morphological dilation[50]. The threshold value to segment
279 individual cell objects based on the 'cell interior' confidence map was determined by plotting the
280 overall voxel-level segmentation accuracy (see Performance Evaluation) versus the threshold
281 values (**Figure S2**). High segmentation accuracies were consistently obtained using confidence
282 thresholds between 0.88 and 0.94.
283

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

285 Thresholding of the 'cell interior' confidence map produces a binary segmentation result
286 (background = 0, cell interior =1), where groups of connected, non-zero voxels identify individual
287 cells in most cases (**Figure S2**). However, when cells are touching, they are often not segmented
288 as individuals, but remain part of the same voxel cluster (undersegmentation). On the other hand,
289 a single cell may be erroneously split into smaller sub-cellular objects (oversegmentation). Finally,
290 in datasets with low SBR, connected voxel clusters may be detected that do not correspond to cells
291 and thus produce false positive objects (**Figure 2a**). To address these errors and improve the
292 segmentation accuracy further, we included additional mathematical image analysis steps to post-
293 process the CNN results and reduce undersegmentation and oversegmentation errors.
294

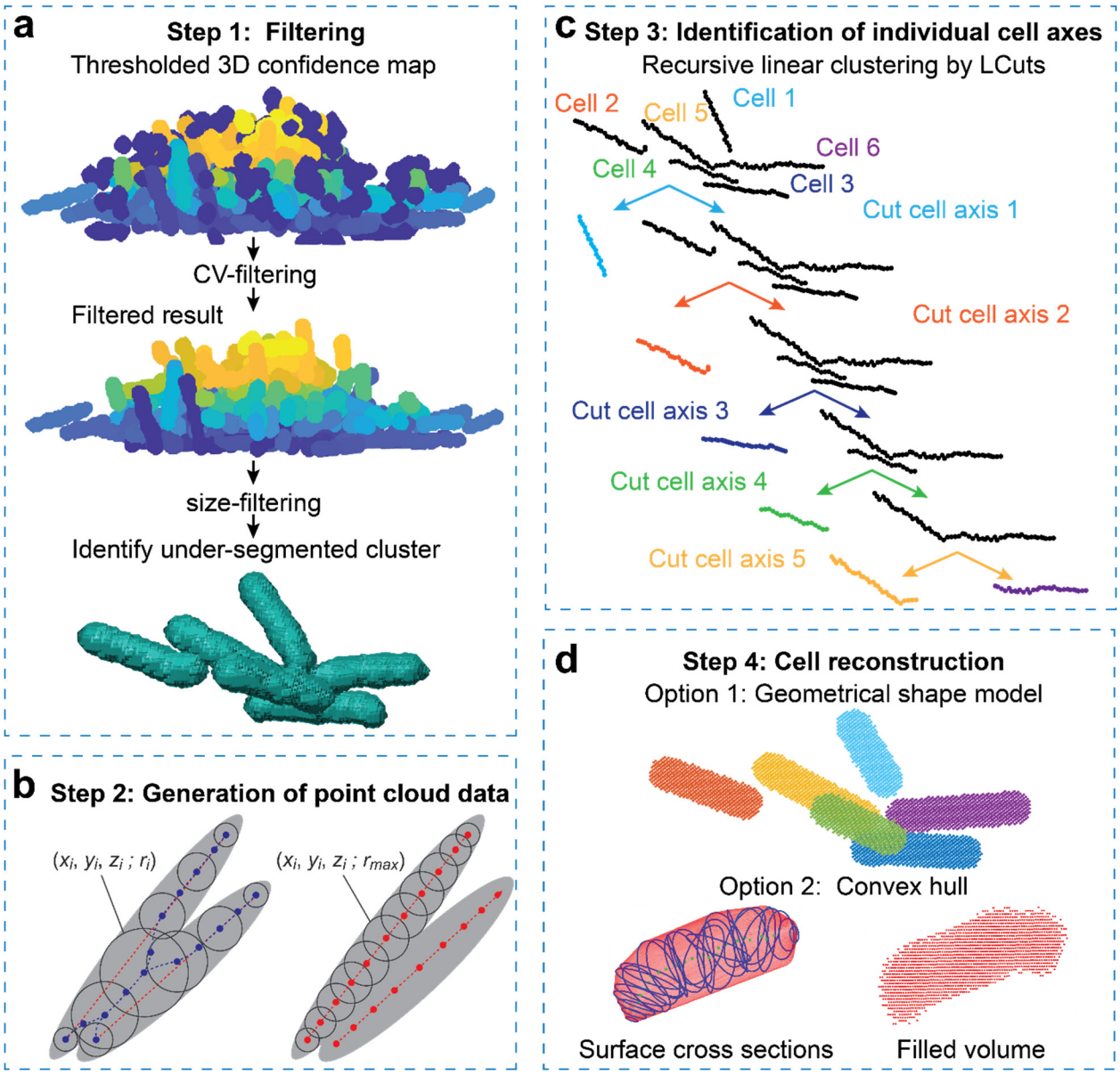


Figure 2. Post-processing of CNN-produced confidence maps using a refined *LCuts* processing pipeline. (**a**) False positive objects are detected and removed by CV- and size- filtering. Under-segmented clusters that are larger than single cells are selected for further splitting. **(b)** Illustration 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

305 **Figure S3** for details). The graph (blue nodes) is then iteratively cut into smaller graphs (red nodes)
306 until the stopping criteria are reached (see text for details). **(d)** Post-processed graphs represented
307 in different color denoting different cells. The 3D surface of individual cells can be determined
308 using a geometrical cell shape model (e.g. a spherocylinder for rod shaped bacteria) or by
309 calculating the convex hull around the inscribed spheres found in step 2.
310
311
312 Step 1. Filtering: False positive objects are identified by evaluating the coefficient of
313 variation[51,52] for each connected voxel cluster *i*:

314 $$CV_i = \frac{\sigma_i}{\mu_i}$$

315 where σ_i and μ_i denote the standard deviation and the mean of the intensity taken over all voxels
316 contained in connected voxel cluster *i*. If the coefficient of variation is larger than ρ, then the
317 current object will be classified as a false positive object and removed from the confidence map
318 by setting all its voxels to zero. The removed objects will then no longer be counted when
319 evaluating the cell counting accuracy. The value of ρ is selected based on the coefficient of
320 variation of the background. For the datasets analyzed here, this sample coefficient of variation
321 was determined to be ρ = 1.1. After CV-filtering, objects smaller than 25% of the expected
322 bacterial cell size are also removed by setting its voxels to zero. The remaining connected voxel
323 clusters are then considered for further processing (**Figure 2a**).
324 Step 2. Identification of point coordinates along central bacterial cell axes: To identify and
325 delineate individual cells in the connected voxel clusters identified in the previous step, we
326 implemented medial axis extraction using the method of inscribed spheres[53], with the constraint
327 that the sphere radii do not exceed the expected diameter of a single bacterial cell (e.g. d = 0.8 μm)
328 (**Figure 2b** left). The set of N inscribed spheres are tangent to the object's surface and

329 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
330 achieved using the Euclidean distance transform of the objects' boundary[54], so that the points with
331 coordinates (x_i, y_i, z_i) reliably trace out the central cell axes of individual bacterial cells (**Figure 2b**
332 right).

333 Step 3. Identification of individual cell axes: To separate different linear segments after
334 cell axis extraction (**Figure 2c**), we used a refined version of the linear cuts (*LCuts*) algorithm[55,56].
335 *LCuts* is a graph-based data clustering method designed to detect linearly oriented groups of points
336 with certain properties. The fundamental elements of a weighted mathematical graph are nodes,
337 edges, and edge weights. Here, the points with coordinates (x_i, y_i, z_i) represent the graph nodes.
338 Edges are the connections among nodes. Edges are assigned weights, for example, to reflect the
339 confidence that two nodes belong to the same group. *LCuts* achieves grouping by assigning
340 weights to edges in the fully connected graph to reflect the similarity between two nodes. The
341 features of each node include its location and direction, where the location of each node is simply
342 its Cartesian coordinates. The direction of each node is found by first determining its 5-hop
343 neighborhood, removing nodes at large relative angles, and evaluating the major direction of the
344 outlier removed neighborhood (**Figure S3**).

345 The algorithm to separate the nodes into different groups is a recursive graph cutting
346 method[55]. Graph cuts (*e.g.* nCut[57]) disconnect the edges between two groups of nodes when the
347 combined weights of these edges are minimized. The weights, between node i and node j, are
348 calculated as follows:

349
$$w_{ij} = w_D \cdot w_T$$

350 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.

<u>Step 4. Cell reconstruction:</u> 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**).

371 **Performance Evaluation**

372 We quantified segmentation accuracy both at the cell-level (object counting) and at the
373 voxel-level (cell shape estimation). To quantify the cell-level segmentation accuracy, we
374 designated segmented objects as true positive (*TP*) if their voxel overlap with the ground truth or
375 the manual annotation resulted in a Intersect of Union value (IoU value, aka Jaccard index[58]) larger
376 than 0.5. This criterion ensures one-to-one matching. If the segmented cell object could not be
377 matched to a ground truth/manually annotated cell volume, then it was counted as a false positive
378 (*FP*) and the IoU value of that segmented object was set to zero. If a ground truth/manually
379 annotated cell volume was not identified in the image, then it was counted as false negative (*FN*).
380 The cell counting accuracy was defined as $TP/(TP+FP+FN)$. The voxel-level segmentation
381 accuracy quantifies the accuracy of cell shape estimation and is calculated as the average IoU value
382 over all segmented objects in the image.

383 To evaluate the accuracy of cell segmentation on experimental data, three researchers
384 separately traced the cell contours on experimental 2D slices by using freehand selections in Fiji
385 ROI Manger[59]. Because human annotation is very time consuming (about 50 hours for a complete
386 3D dataset containing ~300 cells in a 22 x 32 x 12 um^3 volume), one to three single 2D slices were
387 selected for each dataset. One exception is the 3D *M. xanthus*, for which the cell outlines in all
388 available x, y and z slices were traced manually (**Figure S4a**). For straight, rod-shaped cells, the
389 centroids of the resulting 2D cell contours all fall within the cell interior volume. To group together
390 the contours belonging to the same cells, the centroid of each contour was projected along the x, y
391 and z dimension. If the projected centroid was enclosed by any other contour in a different slice,
392 then the centroid of that contour was projected onto the plane of the initial contour. Two contours
393 were labeled as related if they contained each other's projected centroids (**Figure S4b**). This

process is repeated for all possible contour pairs and their relationship is recorded in an adjacency matrix. Next, related contours were assigned to individual cells (**Figure S4c**). To separate incorrectly grouped contours, we additionally identified clusters of centroids using the DBSCAN point clustering algorithm[60] (**Figure S4d**). In a final step, we manually removed incorrectly traced contours (**Figure S4e**). Cells are reconstructed by creating convex hulls with the grouped contours (**Figure S4f** and **S4g**). This procedure determined the approximate positions, shapes, and orientations of individual cells in the 3D biofilm.

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

411 # Results and Discussion

412 **Cell segmentation by thresholding CNN confidence maps**

413 CNNs have been shown to perform well on pixel-level classification tasks for both 2D and
414 3D data[61,62]. Bacterial biofilms, however, present a unique challenge in this context. The cell
415 shapes to be segmented are densely packed and barely resolvable even with the highest resolution
416 optical microscopes. Additionally, living biofilms in fluorescence microscopes can only be imaged
417 with low laser intensities to ameliorate phototoxicity and photobleaching concerns. Unfortunately,
418 low intensity fluorescence excitation also reduces the SBR in the acquired images. So far, it
419 remains unclear to what extent single-cell segmentation approaches can accurately identify and
420 delineate cell shapes in bacterial biofilm images obtained under low intensity illumination
421 conditions. To address this question, we implemented an *in silico* CNN training strategy (see
422 Materials and Methods) and systematically evaluated its voxel-level classification (cell
423 morphometry) and cell counting accuracies using simulated biofilm images with varying cell
424 densities and SBRs similar to those encountered in experimental data.

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

434 offs observed here may indicate a fundamental performance limitation of the CNNs employed.
435 Still, the voxel-level classification and cell counting accuracies consistently surpass previous
436 approaches for bacterial cell segmentation for commonly encountered cell densities and SBRs.
437 Specifically, the cell counting accuracies obtained by *BiofilmQ*[63], *Seg3D*[64], and Yan *et al.*[13] quickly
438 drop to zero as a function of increasing IoU matching threshold (a quantitative measure of cell
439 shape similarity relative to the ground truth, aka the Intersect-over Union (IoU); Materials and
440 Methods), indicating that cell shapes are not accurately estimated by conventional image
441 processing approaches (**Figure 3g-i**). We also evaluated the segmentation accuracy of *Cellpose*, a
442 recently developed, CNN-based cellular segmentation algorithm[33]. The segmentation accuracy of
443 *Cellpose* is superior to conventional image processing approaches – a considerable achievement
444 given that *Cellpose* was trained primarily on images of eukaryotic cells. However, being a pre-
445 trained generalist model, the segmentation accuracy of *Cellpose* is lower than the accuracy
446 achieved by the specialist *in silico*-trained CNNs of *BCM3D*, which were trained specifically for
447 3D bacterial biofilm segmentation. Overall, the cell counting accuracy obtained by *BCM3D* are
448 higher than other methods and remain higher even for IoU matching thresholds larger than 0.5,
449 indicating that cell shapes are better estimated by the *in silico*-trained CNNs.

450 The accuracies of single-cell shape estimation and cell counting are predominantly affected
451 by cell density. The variation is more pronounced for membrane-stained cells, because *inter*-
452 cellular fluorescence intensity minima are less pronounced when cell membranes are labeled and
453 cells physically contact each other (red arrow in **Figures 3c** and **3f**). By contrast, intracellular
454 fluorophores produce the highest intensities at the cell center, so that the gaps between cells are
455 more readily resolvable. Also noteworthy is the sharp drop-off in segmentation accuracies for
456 SBRs of <1.7 for all cases. In such low SBR regimes, fluorescence signals of the cells become too

difficult to be distinguished from the background. As a result, the CNNs falsely identify random noisy patterns in the background as cells. Additionally, thresholding of the CNN confidence maps often yields connected voxel clusters that contain multiple bacterial cells. False identification and incomplete delineation of cells cause the pronounced decrease in segmentation accuracy for SBRs of <1.7.

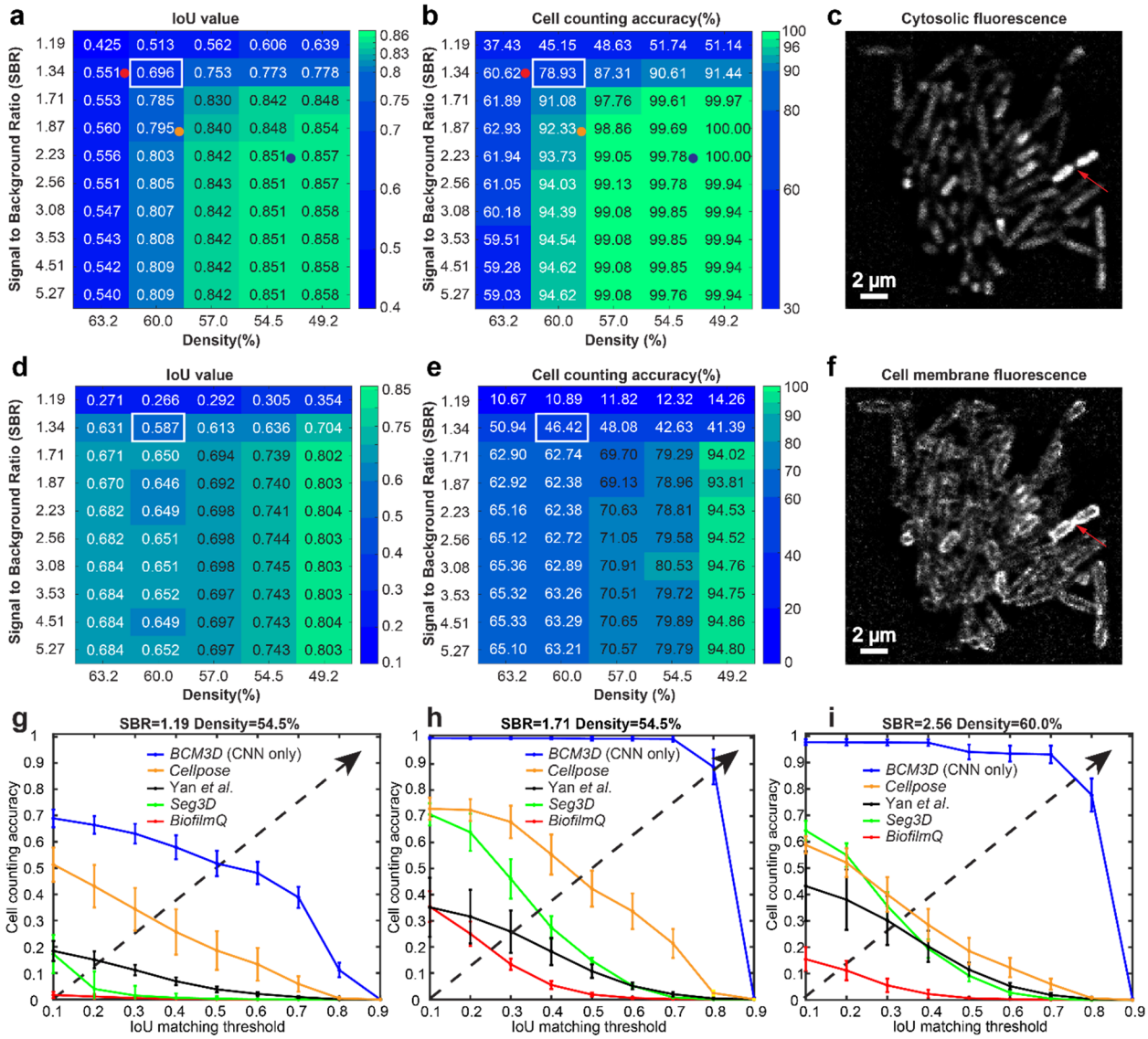


Figure 3. Performance of *BCM3D* using *in silico*-trained CNNs only on previously unseen simulated biofilm images. **(a)** The IoU value quantifies whether each voxel has been assigned to

466 the correct class ('cell interior', cell surface', or 'background'). Solid circles represent the
467 maximum local density and average SBRs encountered in experimental datasets (red, orange and
468 blue: *E. coli* expressing GFP). **(b)** The cell counting accuracy (using a IoU matching threshold of
469 0.5 for each segmented object, see Materials and Methods) averaged over n=10 replicate datasets
470 for cells labeled with cytosolic fluorophores. **(c)** Example image of cells labeled with cytosolic
471 fluorophores (Cell density = 60.0%, SBR = 1.34, indicated by white rectangle in panels a and b).
472 **(d)** IoU value and **(e)** cell counting accuracy averaged over n=10 replicate datasets for cells labeled
473 with membrane-localized fluorophores. Purple solid circles represent experimental datasets of *M.*
474 *xanthus* stained with the membrane intercalating dye FM4-64. **(f)** Example image of cells labeled
475 with membrane-localized fluorophores (Cell density = 60.0%, SBR = 1.34, indicated by white
476 rectangles in panels d and e). The red arrows indicate a close cell-to-cell contact point. **(g), (h) and**
477 **(i)** Comparison of segmentation accuracies achieved by conventional segmentation approaches
478 (*BiofilmQ*, *Seg3D*, Yan *et al.*), *Cellpose*, and *BCM3D* (only using *in silico*-trained CNNs). Three
479 simulated datasets (cytosolic fluorophores) with different SBRs and cell densities are shown.
480 Segmentation accuracy is parameterized in terms of cell counting accuracy (*y* axis) and IoU
481 matching threshold (*x* axis, a measure of cell shape estimation accuracy). Each data point is the
482 average of ten independent biofilm images. Error bars represent ± one standard deviation. Curves
483 approaching the upper right-hand corner indicate higher overall segmentation accuracy.
484

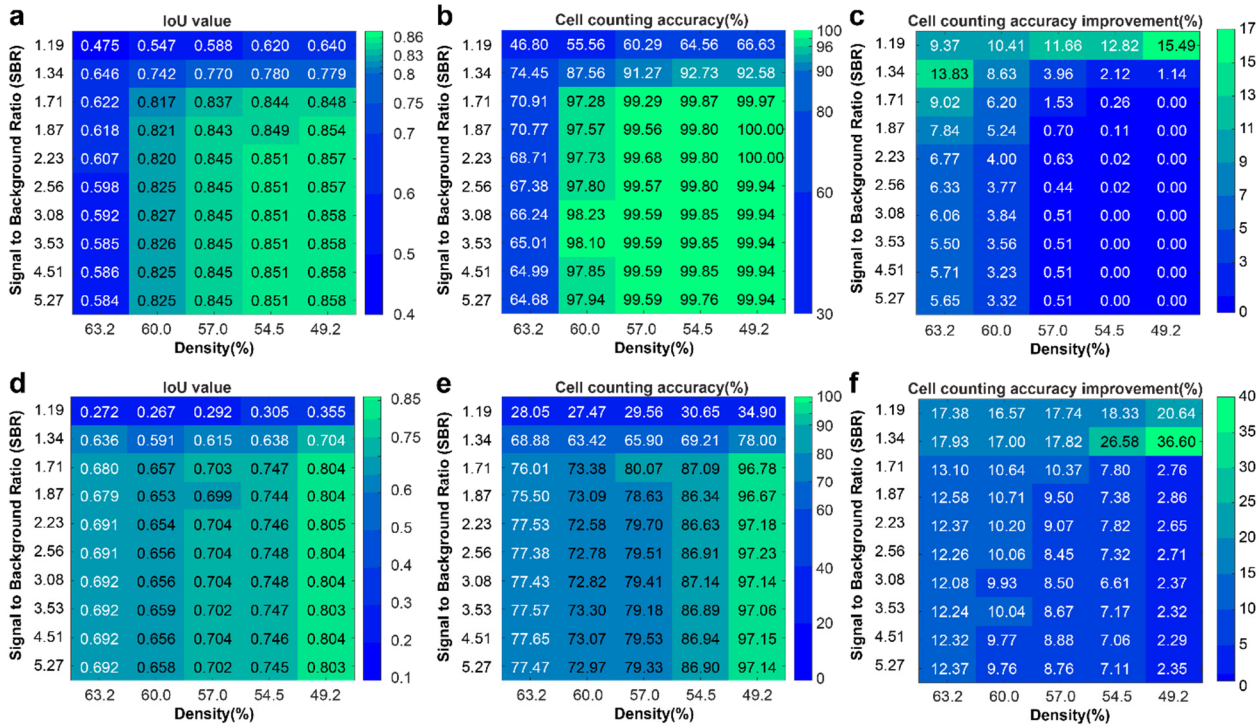
485

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

487 To better identify individual cells in low SBR and high cell density datasets, we developed
488 a graph-based post-processing module (see Materials and Methods) that takes advantage of highly
489 conserved bacterial cell shape of a given species. Briefly, we transformed the CNN 'cell interior'
490 confidence maps into 3D point cloud data that trace out the central axes of individual cells. This
491 transformation was achieved by medial axis extraction using size-constrained inscribed spheres[53]
492 (**Figure 2**). Single-cell axes are then identified as linearly clustered data points by *LCuts* – a graph-
493 based data clustering method designed to detect linearly oriented groups of points[55]. The so-

identified single-cell axes are then mapped back onto the original segmentation volumes to obtain estimates of the 3D positions, shapes, and orientations of the now separated cells.

Post-processing with *LCuts* takes advantage of *a priori* knowledge about expected bacterial cell sizes (**Figure S5**) by removing erroneously segmented volumes that are significantly smaller than the expected value and by splitting incompletely segmented volumes representing fused cells. Improvements in cell counting accuracy of up to 15% and 36% are observed for cells labeled with cytosolic fluorophores (**Figure 4a-c**) and membrane-localized fluorophores (**Figure 4d-f**), respectively. The more substantial improvement for membrane-stained cells is due to fact that CNNs trained on membrane-stained cells are more prone to erroneously identifying speckled background noise as fluorescence signals in low SBR images. In addition, membrane-intercalating fluorophores of two adjacent cells are in close proximity, making it difficult to resolve fluorescence signals from two separate cells due to spatial signal overlap (see the red arrow, **Figure 3c** and **3f**).



508 **Figure 4.** Performance of *BCM3D* (*in silico*-trained CNNs *and* additional post-processing by
509 *LCuts*) on previously unseen simulated data. **(a)** IoU value and **(b)** cell counting accuracy (using
510 a IoU matching threshold of 0.5 for each segmented object) averaged over n=10 replicate datasets
511 for cells labeled with cytosolic fluorophores. **(c)** Improvement relative to *silico*-trained
512 convolutional neural networks without post-processing. **(d)** IoU value and **(e)** cell counting
513 accuracy averaged over n=10 replicate datasets for cells labeled with membrane-localized
514 fluorophores. **(f)** Improvements relative to *silico*-trained convolutional neural networks without
515 post-processing.
516
517
518 **Segmentation of experimental biofilm images**
519 To test the performance of *BCM3D* on experimentally acquired biofilm images, we
520 acquired time-lapse images of GFP expressing *E. coli* biofilms every thirty minutes for ten hours.
521 We then manually annotated one 2D slice in the 3D images at the t = 5, 6, and 10-hour time points.
522 When referenced to these manual segmentation results, both the CNN outputs and the *LCuts*-
523 processed CNN outputs consistently achieved better cell counting accuracies than conventional
524 segmentation methods (**Figure 5, Figure S6**). On the other hand, *Cellpose* outperforms the *in*
525 *silico*-trained CNNs on two out of three of the test images, for which *in silico*-trained CNNs
526 struggle with undersegmentation problems. However, mathematical post-processing of the CNN
527 outputs by *LCuts* corrects some of these errors, so that the integrated *BCM3D* workflow achieves
528 comparable or slightly better results than *Cellpose* on average. Visual inspection of the
529 segmentation results showed that *Cellpose* accurately segments individual cells in low density
530 regions, but suffers from oversegmentation errors in high density regions (**Figure S6e**). On the
531 other hand, the integrated *BCM3D* workflow (CNN + *LCuts*) gives biological accurate
532 segmentation results even in dense biofilm regions (**Figure 5**).

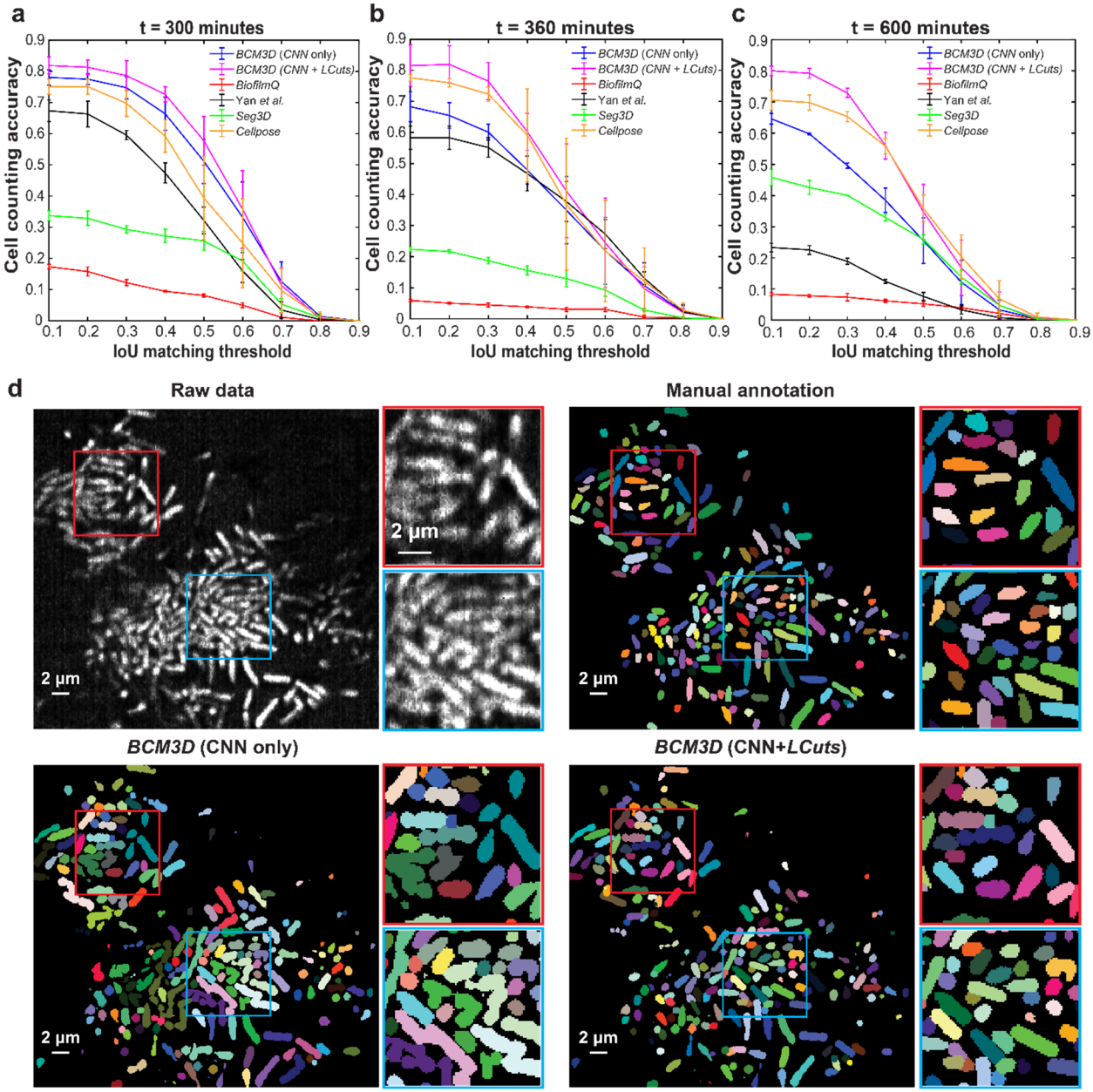


Figure 5. Comparison of segmentation accuracies achieved by conventional segmentation approaches (*BiofilmQ*, *Seg3D*, Yan *et al.*), *Cellpose*, and *BCM3D*. The estimated SBRs are 2.2, 1.8, and 1.3, respectively. The estimated cell densities are 54.8%, 59.0%, and 64.6%, respectively. **(a-c)** Three experimental *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 IoU matching threshold (*x* axis). Each data point is the average of the cell counting accuracies calculated using annotation maps traced by three different researchers. Error bars represent ± one standard deviation. 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**.

We attribute the more rapid drop-off of the cell counting accuracy as a function of increasing IoU matching threshold to the following factors. First, human annotation of experimentally acquired biofilm images is not as consistent as simulated data, and therefore, it is less likely to produce close matches between segmented and annotated cell shapes. Second, the experimentally acquired biofilm images differ from the simulated training and annotation datasets in several ways. Real bacterial cells may appear blurred because they are only partially immobilized and therefore wiggle during frame acquisition. Furthermore, some cells are in the process of cell division and different human annotators may count these cells either as one cell or two cells. Finally, optical aberrations and scattering effects were not included in training data simulations. However, at a lower IoU matching threshold, the cell counting accuracy remains above 70%, which represents a substantial improvement over traditional bacterial cell segmentation approaches (**Figure 5**).

To demonstrate that *BCM3D* can achieve similarly high segmentation accuracies for membrane-stained cells and different cellular arrangements, we analyzed a small patch of a *M. xanthus* biofilm, which was stained with the membrane intercalating dye FM4-64 (**Figure 6a**).

565 In contrast to *E. coli* biofilms, the submerged *M. xanthus* biofilm imaged here features cells in a
566 mesh-like arrangement with close cell-to-cell contacts, which presents a unique challenge for 3D
567 single-cell segmentation. To obtain reference data 3D segmentation accuracy determination, we
568 manually annotated each *xy*, *xz*, and *yz* slice of an entire 3D image stack (**Figure 6b**). When
569 referenced to this 3D manual segmentation result, *BCM3D* (**Figure 6c**) again produced cell
570 counting accuracies in the range of 60% to 70% at low IoU matching thresholds (**Figure 6d**). We
571 note that comparison with conventional segmentation methods mentioned above cannot be made
572 here, because those methods were not designed to work with membrane-stained cells. The
573 performance of *Cellpose* on this type of biofilm architecture is inferior to the results achieved using
574 the *in silico*-trained CNNs of *BCM3D*. One reason might be that the pre-trained, generalist
575 *Cellpose* model has not been trained sufficiently on long, thin, and highly interlaced rod-shaped
576 cells, such as those contained in a *M. xanthus* biofilm.

577

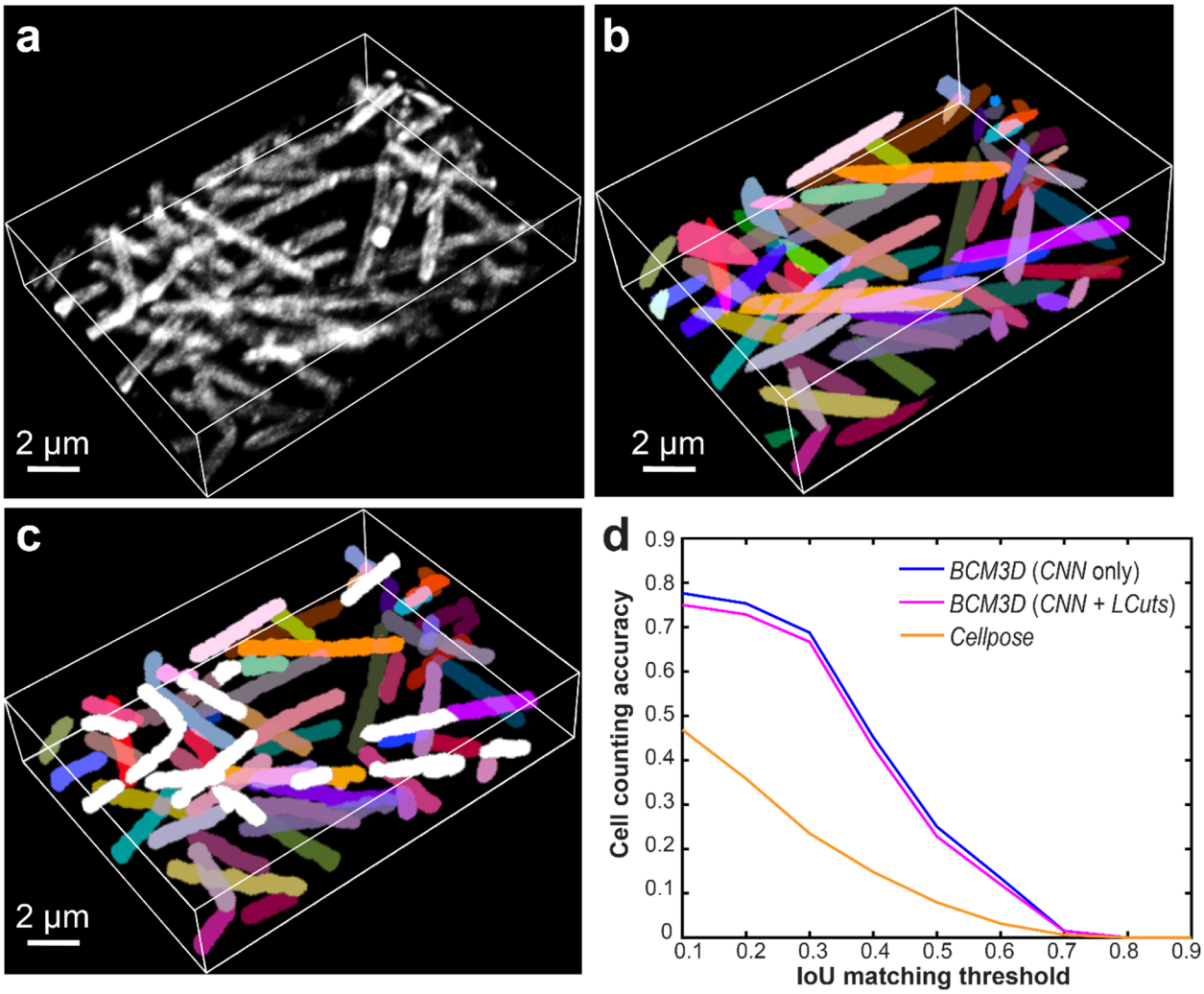


Figure 6. 3D Segmentation accuracy evaluation using *M. xanthus* biofilm images (cell density = 36.2%, and SBR = 1.58) using *in silico*-trained CNN processing. **(a)** Maximum intensity projection of a 3D *M. xanthus* fluorescence image. Cells were labeled with membrane-intercalating dye, FM4-64. **(b)** 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 IoU matching threshold (*x* axis).

587 **Morphological separation of mixed cell populations**

588 Given the improved segmentation results obtained using *BCM3D*, we reasoned that the
589 same CNNs may have additional capacity to assign segmented objects to different cell types based
590 on subtle morphological differences in the acquired images. Differences in the imaged cell
591 morphologies arise due to physical differences in cell shapes (e.g. spherical vs. rod-shaped cells)
592 or due to differences in the fluorescent labeling protocols (e.g. intracellular vs. cell membrane
593 labeling), because fluorescence microscopes simply measure the spatial distributions of
594 fluorophores in the sample. The ability to separate different cell morphologies is important for the
595 study of multispecies biofilms, where interspecies cooperation and competition dictate population-
596 level outcomes[3,65-72]. Separation of differentially labeled cells is also important for the study of
597 gene activation in response to cell-to-cell signaling[73]. Expression of cytosolic fluorescent proteins
598 by transcriptional reporter strains is a widely-used technique to visualize activation of a specific
599 gene or genetic pathway in living cells. Such genetic labeling approaches can be complemented
600 by chemical labeling approaches, e.g. using membrane intercalating chemical dyes that help
601 visualize cells non-specifically or environmentally-sensitive membrane dyes that provide
602 physiological information, including membrane composition[74,75], membrane organization and
603 integrity[76-78], and membrane potential[79,80]. Chemical and genetic labeling approaches are
604 traditionally implemented in two different color channels. However, there are important drawbacks
605 to using multiple colors. First and foremost, the amount of excitation light delivered is increased
606 to excite differently colored fluorophores, raising phototoxicity and photobleaching concerns.
607 Second, it takes *N* times as along to acquire *N*-color images (unless different color channels can
608 be acquired simultaneously), making it challenging to achieve high temporal sampling in time-

609 lapse acquisition. For these reasons, methods that extract complementary physiological
610 information from a single-color image are preferable.

611 We evaluated the ability of *BCM3D* to automatically segment and separate rod-shaped and
612 spherical bacterial cells consistent with sizes of *E. coli* and *S. aureus*, respectively, in simulated
613 images (**Figure S7**). For segment cells in two-population biofilms, we trained CNNs that classify
614 pixels into five different classes: 'background', 'cell interior of population 1', 'cell boundary of
615 population 1', 'cell interior of population 2' and 'cell boundary of population 2'. Thresholding the
616 CNNs confidence maps can achieve cell counting accuracies larger than 90% for both cell types,
617 independent of their population fractions (**Figure 7a**). Post-processing of this result using *LCuts*
618 improved the cell counting accuracy by less than 0.5% on average, indicating that under-
619 segmented cell clusters are not prevalent in this dataset.

620 We next evaluated the ability of *BCM3D* to automatically segment and separate membrane-
621 stained cells that express cytosolic fluorescent proteins from those that do not (**Figure S8**). Again,
622 the cell counting accuracy is consistently above 80% for all tested mixing ratios (**Figure 7b**).
623 Finally, we applied *BCM3D* to experimentally acquired biofilm images of two different *E. coli*
624 strains. Both strains were stained by the membrane intercalating dye FM4-64, but the second strain
625 additionally expressed GFP (**Figure S9**). The cells were homogeneously mixed prior to mounting
626 to randomize the spatial distribution of different cell types in the biofilm (see Materials and
627 Methods). Multiple 2D slices from the 3D image stack were manually annotated and compared
628 with the results obtained by *BCM3D*. Consistent with the single-species experimental data, a cell
629 counting accuracy of 50% is achieved for each cell type at a 0.5 IoU matching threshold and, at
630 lower IoU matching thresholds, the counting accuracies increased to 60% to 70%, (**Figure 7cd**).
631 Thus, using appropriately trained CNNs in *BCM3D* enables automated and accurate cell type

assignments based on subtle differences in cell morphologies in mixed population biofilms – a capability not available using conventional image processing methods.

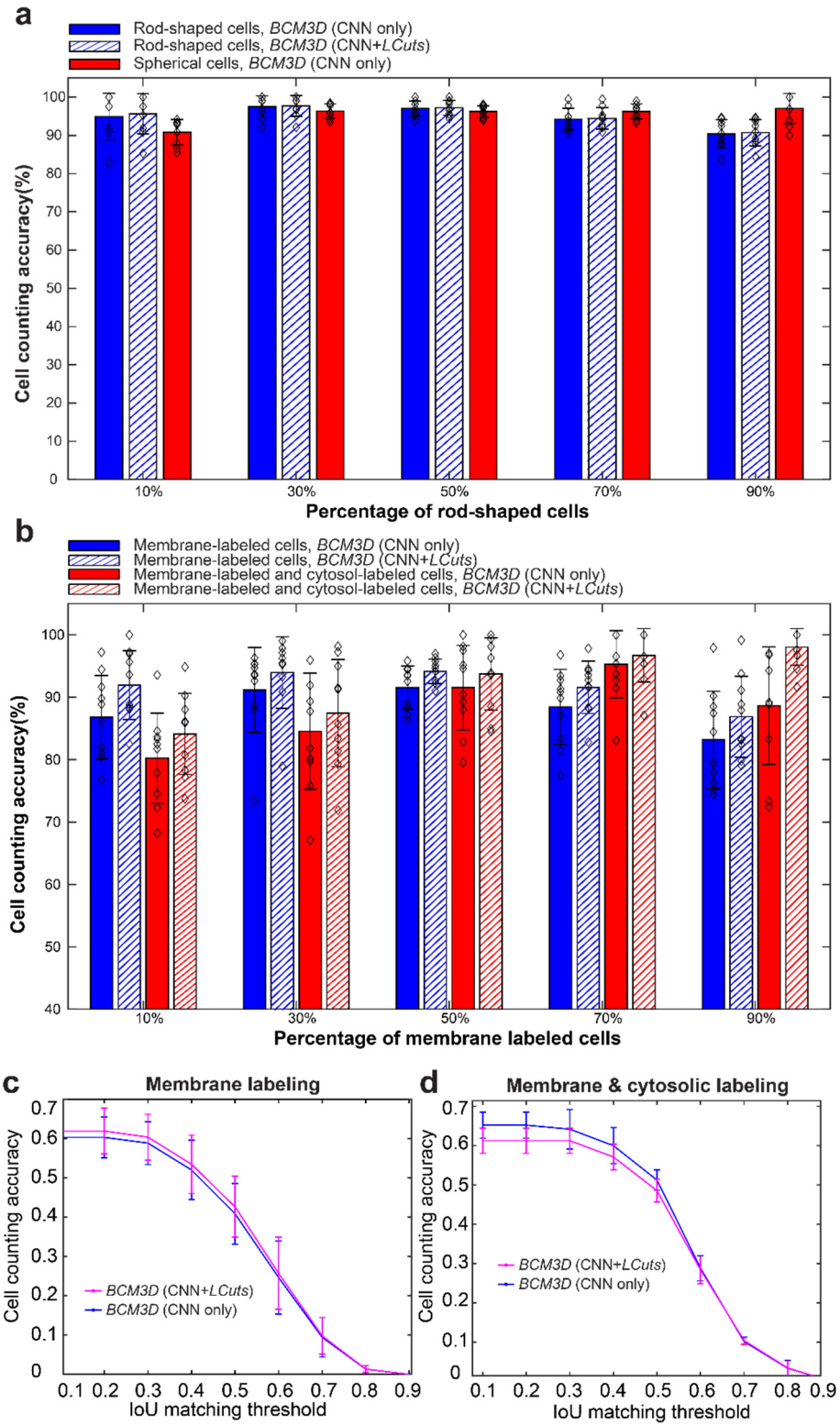


Figure 7. (a) Cell counting accuracy of *BCM3D* on simulated images containing different ratios of rod-shaped and spherical cells. Black diamonds represent the counting accuracy for each independent dataset. Error bars represent ± one standard deviation. **(b)** Cell counting accuracy of *BCM3D* on simulated 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. Error bars represent ± one standard deviation. **(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). Each data point is the average of the cell counting accuracies calculated using annotation maps traced by three different researchers. Error bars represent ± one standard deviation.

Conclusions

CNNs have been successful applied to many different problems in biological image analysis, but their ability to segment individual cells in 3D and time-lapse 3D bacterial biofilm images has not yet been fully explored. Here, we demonstrated a new CNN-based image analysis workflow, termed *BCM3D*, for single-cell segmentation and shape classification (morphometry) in 3D images of bacterial biofilms. In this work, we applied *BCM3D* to 3D images acquired by lattice light sheet microscopy. However, *BCM3D* readily generalizes to 3D images acquired by confocal microscopy or advanced super-resolution microscopy modalities, provided that realistic image formation models are used to simulate the training datasets. The use of simulated training data is a major advantage of *BCM3D*, because it overcomes inconsistencies inherent in manual dataset annotation and thus solves the problem of obtaining sufficient amounts of accurately annotated 3D image data. Thus, the ability of use simulated training data provides the needed flexibility not only in terms of the microscope platform used, but also in terms of the bacterial cell shapes that are to be segmented.

661 We systematically investigated the advantages and limitations of *BCM3D* by evaluating
662 both voxel- and cell-level segmentation accuracies using simulated and experimental datasets of
663 different cell densities and SBRs. *BCM3D* enabled accurate segmentation of individual cells in
664 crowded environments and automatic assignments of individual cells to specific cell populations
665 for most of the tested parameter space. Such capabilities are not readily available when using
666 previously established segmentation methods that rely exclusively on conventional image and
667 signal processing algorithms.

668 While *BCM3D* surpasses the performance of previous approaches, we stress that further
669 improvements are possible and, for certain experiments, absolutely needed. Our systematic
670 analysis revealed that high cell density and low SBR datasets are particularly challenging for the
671 CNNs used in this work. Future work will therefore focus on increasing the contrast and resolution
672 in bacterial biofilm images. While, the use of optical super-resolution modalities can provide
673 higher spatial resolution, such resolution improvements often come at a cost of reduced image
674 contrast and faster photobleaching/phototoxicity, especially for high frame-rate time-lapse
675 imaging. Software solutions that can process images with limited resolution and low SBRs will
676 therefore play a tremendously important role in biological imaging. *BCM3D* is a general workflow
677 that integrates computational simulation of training data, *in silico*-training of CNNs for a *specific*
678 task or a *specific* cell type, and mathematical post-processing of the CNN outputs. Incorporating
679 different training strategies and different CNNs, such as the generalist CNN used in *Cellpose*[33],
680 into the *BCM3D* workflow will enable automated cross-validation of segmentation results when a
681 ground truth or manual annotation map is not available. Furthermore, CNN-based image
682 processing modules developed for contrast enhancement and denoising have also surpassed the
683 performance of conventional methods based on mathematical signal processing[81-84]. Incorporating

these tools into the *BCM3D* workflow promises to further improve the single-cell segmentation accuracies. We anticipate that the ability to accurately identify and delineate individual cells in dense 3D biofilms will in turn enable accurate cell tracking over long periods of time. Detailed measurements of behavioral single-cell phenotypes in larger bacterial communities will help determine how macroscopic biofilm properties, such as its mechanical cohesion/adhesion and its biochemical metabolism, emerge from the collective actions of *individual* bacteria.

Author Contributions

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

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

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

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

Acknowledgements

This work was supported by a Jeffress Memorial Trust Awards in Interdisciplinary Research (to A.G and S.T.A.). M.Z. and J.W. were supported by a University of Virginia Presidential Fellowship in Data Science. Y.W. was supported in part by a Dean's M.S.-Ph.D. Fellowship in Data Science. We thank Lotte Søgaard-Andersen, Larry Shimkets, Huiwang Ai and Ingmar Riedel-Kruse for providing bacterial strains used in this work. We thank Karsten Siller, computational research consultant at the University of Virginia, for help with utilizing the Universities super-computing resources.

Biological materials availability

All strains used in this work are available from their original sources upon request.

Data and code availability

The code for running all the modules of BCM3D as well as training and test data is available at https://github.com/GahlmannLab/BCM3D.git.

714 **References**

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

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

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

721 4. Persat, A., Inclan, Y.F., Engel, J.N., Stone, H.A. & Gitai, Z. Type IV pili mechanochemically
722 regulate virulence factors in Pseudomonas aeruginosa. *Proc Natl Acad Sci U S A* **112**, 7563-
723 7568 (2015).

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

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

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

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

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

737 10. Vidakovic, L., Singh, P.K., Hartmann, R., Nadell, C.D. & Drescher, K. Dynamic biofilm
738 architecture confers individual and collective mechanisms of viral protection. *Nat Microbiol*
739 **3**, 26-31 (2018).

740 11. Stewart, E.J., Ganesan, M., Younger, J.G. & Solomon, M.J. Artificial biofilms establish the
741 role of matrix interactions in staphylococcal biofilm assembly and disassembly. *Sci Rep* **5**,
742 13081 (2015).

743 12. Stewart, E.J., Satorius, A.E., Younger, J.G. & Solomon, M.J. Role of environmental and
744 antibiotic stress on Staphylococcus epidermidis biofilm microstructure. *Langmuir* **29**, 7017-
745 7024 (2013).

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

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

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

755 16. Gregor, I. & Enderlein, J. Image scanning microscopy. *Curr Opin Chem Biol* **51**, 74-83
756 (2019).

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

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

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

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

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

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

778 23. Dean, K.M., Roudot, P., Welf, E.S., Danuser, G. & Fiolka, R. Deconvolution-free
779 Subcellular Imaging with Axially Swept Light Sheet Microscopy. *Biophysical journal* **108**,
780 2807-2815 (2015).

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

786 25. Dunsby, C. Optically sectioned imaging by oblique plane microscopy. *Optics Express* **16**,
787 20306-20316 (2008).

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

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

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

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

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

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

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

810 33. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for
811 cellular segmentation. *bioRxiv*, 2020.2002.2002.931238 (2020).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

861 51. Yu, Y. & Acton, S.T. Speckle reducing anisotropic diffusion. *Trans. Img. Proc.* **11**, 1260-
862 1270 (2002).

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

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

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

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

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

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

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

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

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

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

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

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

893 64. Reyer, M.A., McLean, E.L., Chennakesavalu, S. & Fei, J. An Automated Image Analysis
894 Method for Segmenting Fluorescent Bacteria in Three Dimensions. *Biochemistry* **57**, 209-
895 215 (2018).

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

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

901 67. Humphries, J., Xiong, L., Liu, J., Prindle, A., Yuan, F., Arjes, H.A., Tsimring, L. & Suel,
902 G.M. Species-Independent Attraction to Biofilms through Electrical Signaling. *Cell* **168**,
903 200-209 e212 (2017).

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

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

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

911 71. Persat, A., Nadell, C.D., Kim, M.K., Ingremeau, F., Siryaporn, A., Drescher, K., Wingreen,
912 N.S., Bassler, B.L., Gitai, Z. & Stone, H.A. The mechanical world of bacteria. *Cell* **161**, 988-
913 997 (2015).

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

916 73. Kroos, L. Highly Signal-Responsive Gene Regulatory Network Governing Myxococcus
917 Development. *Trends Genet* **33**, 3-15 (2017).

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

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

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

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

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

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

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

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

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

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

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

945

Figure S1

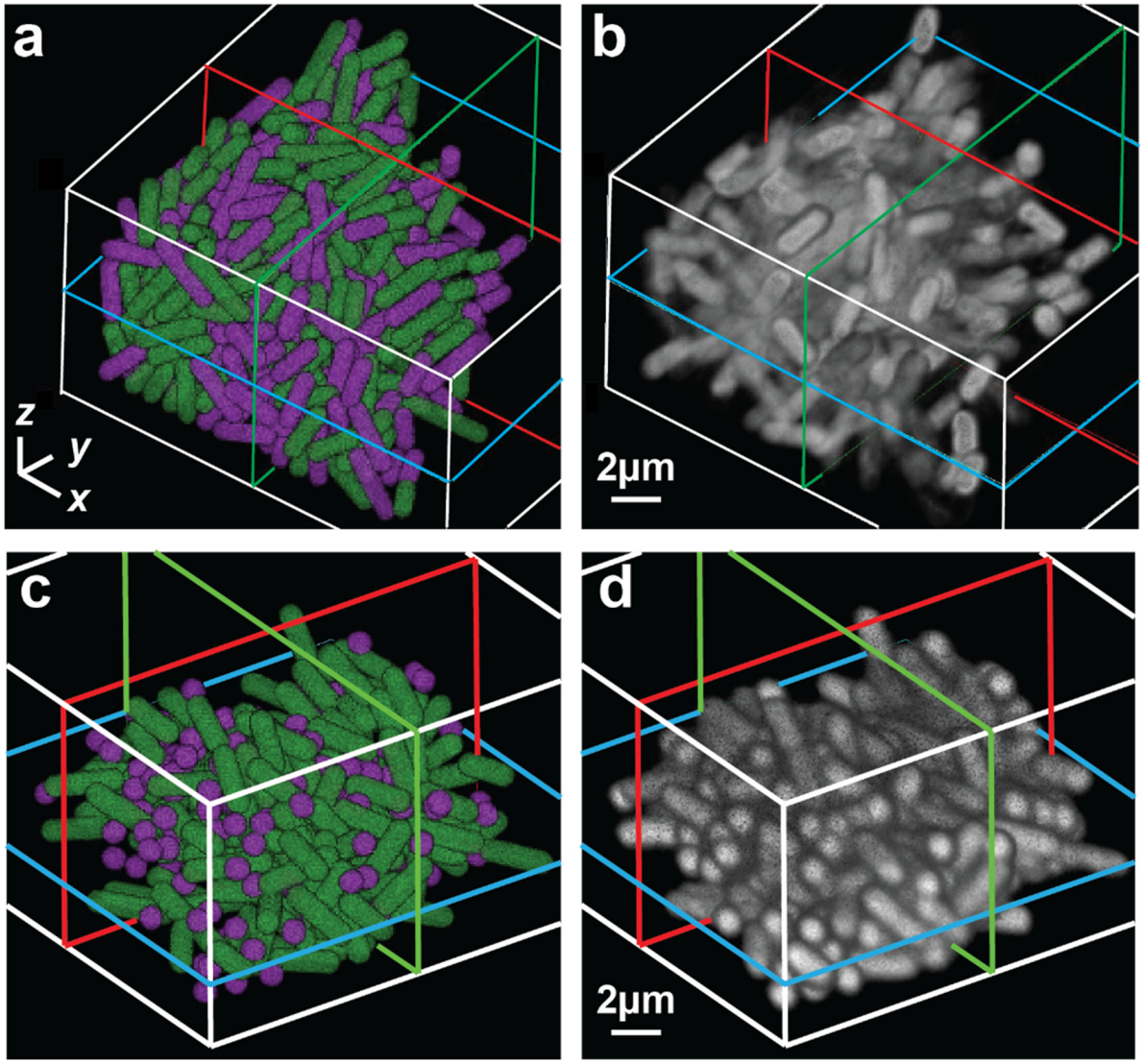


Figure S1. Simulation of mixed labeling and mixed cell shape biofilms. (**a**) Cell arrangements (green indicates membrane labeled cells, purple indicates membrane labeled cells that simultaneously express interior fluorescence protein). (**b**) Fluorescence image based on the cell arrangements in (a) 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

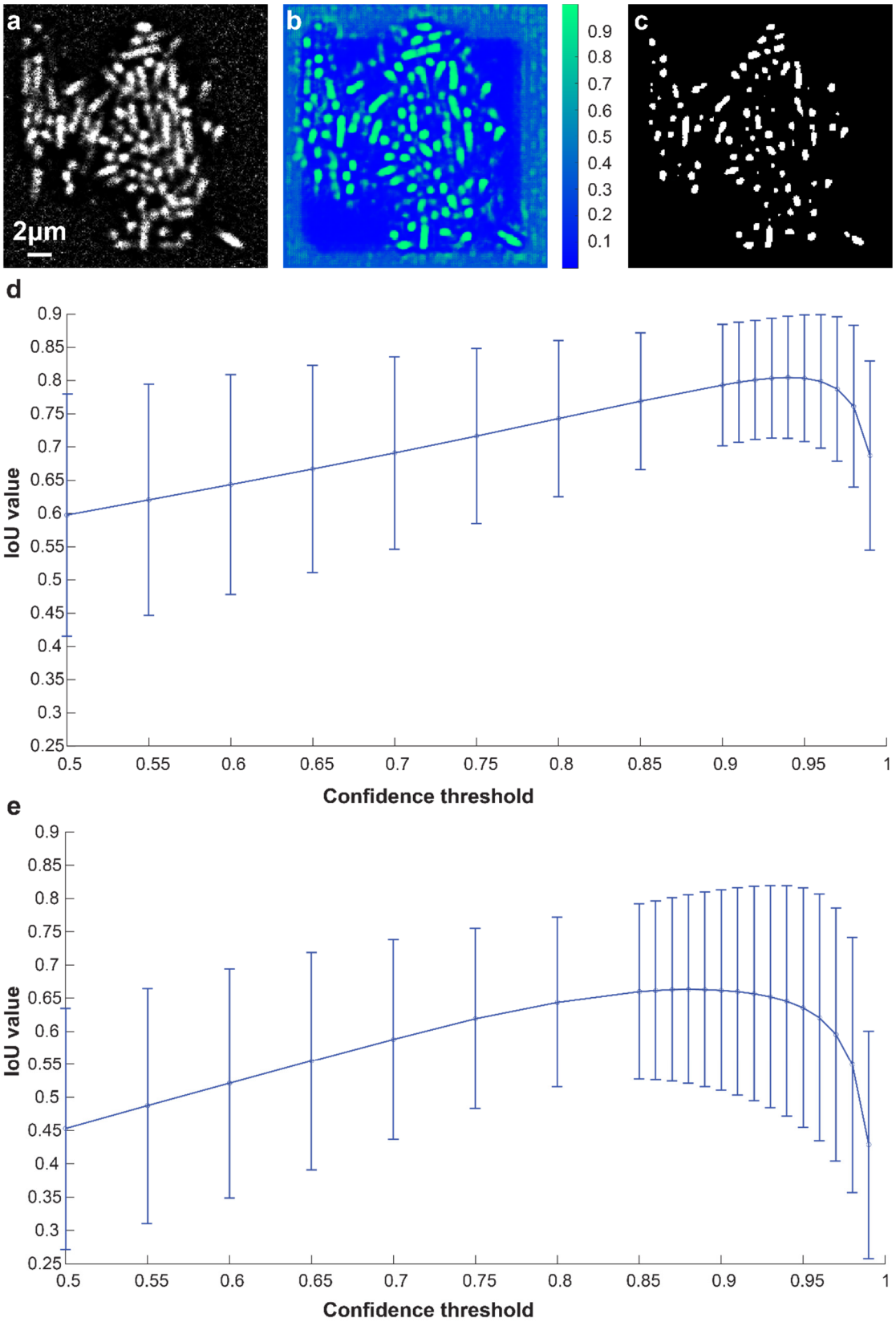

a
2μm
b
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
c
d
IoU value
Confidence threshold
e
IoU value
Confidence threshold

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**) IoU value (y axis) versus the confidence map threshold (*x* axis) for cells labeled with cytosolic fluorophores (d) and cells labeled with membrane-localized fluorophores (e). Each curve is plotted by averaging 500 different datasets. Error bars represent ± one standard deviation.

Figure S3

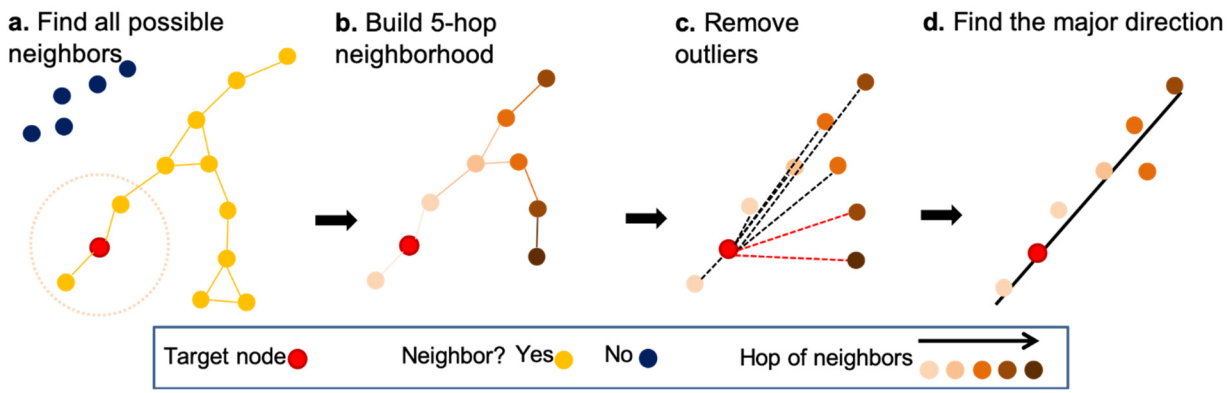


Figure S3. Determination of node direction in an outlier-removed neighborhood[2]. (**a**) A neighborhood of the target node (in red) is a sub-graph, where all adjacent nodes (in yellow) are connected via edges to the target node. Here, if the distance of two nodes is less than a chosen value (indicated 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

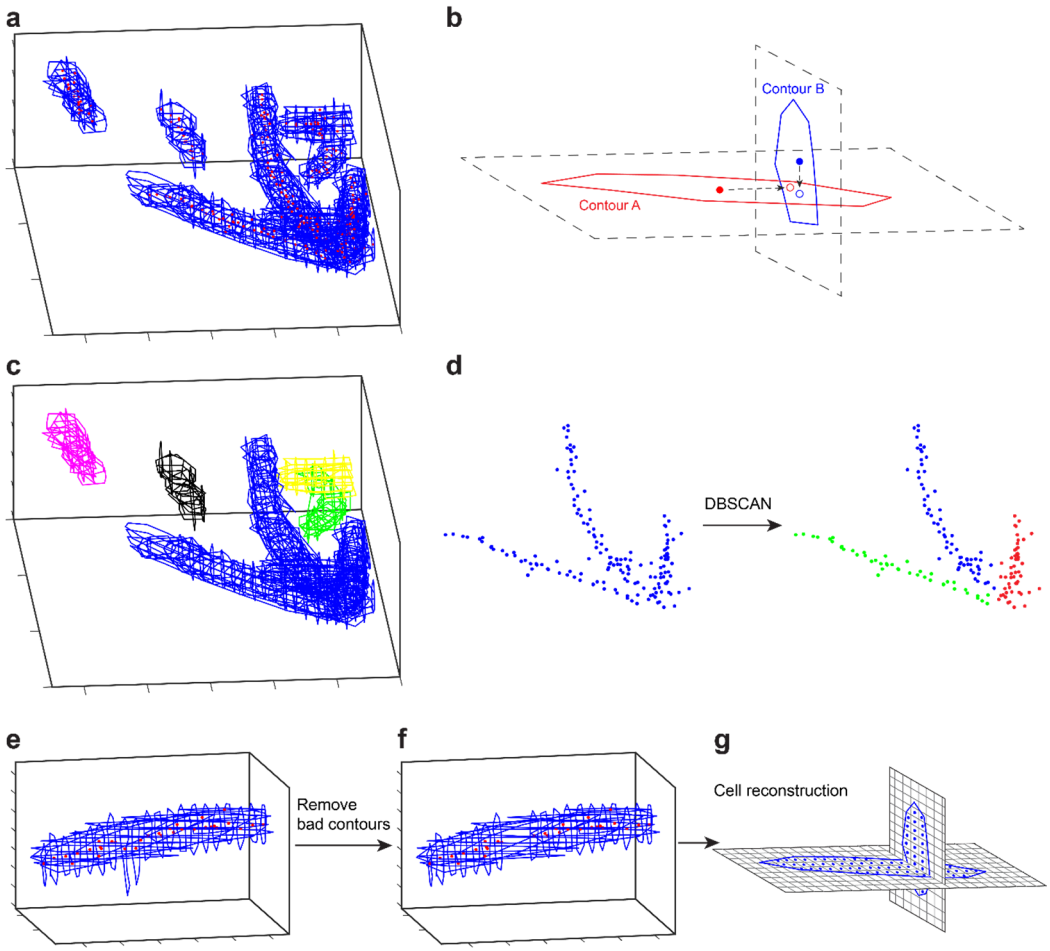


Figure S4. Manually trace cell outline. (**a**) Cell outlines in all available x, y and z slices were traced manually. (**b**) Find the contours belonging to the same cell. The centroid of each contour was projected along the x, y and z dimension. If the projected centroid was enclosed by any other contour in a different slice, then the centroid of that contour was projected onto the plane of the initial contour. Two contours were labeled as related if they contained each other's projected centroids. This process is repeated for all possible contour pairs and their relationship is recorded in an adjacency matrix. (**c**) The related contours are grouped as cells. Different colors represent different cells. (**d**) Segment big clusters that contain more than one cell by grouping the centroids of the contours. This step will run manually and iteratively to segment all single cells from the big cluster. (**e**) Manually check all contours for each cell. (**f**) Remove the bad contours, such as unreasonably large ones. (**g**) A convex hull is built based on the contours for each cell. The convex hull is then used as the mask to extract cell volume from the raw data.

Figure S5

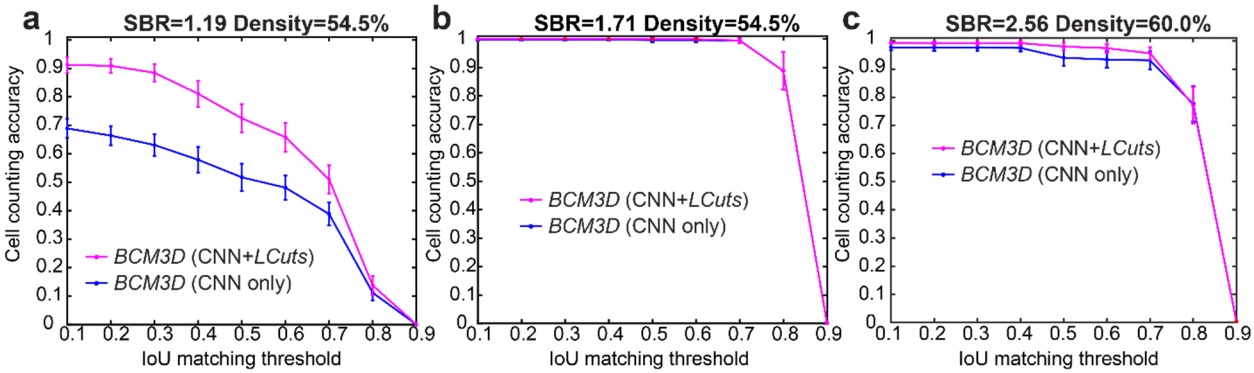


Figure S5. Segmentation accuracies achieved for biofilm images with different SBR and cell densities. (**a**), (**b**), and (**c**) represent results for datasets with SBR 1.19, Density 54.5%, SBR 1.71, Density 54.5% and SBR 2.56, Density 54.5%, respectively. Each curve is plotted by averaging 10 different datasets. Error bars represent ± one standard deviation.

Figure S6

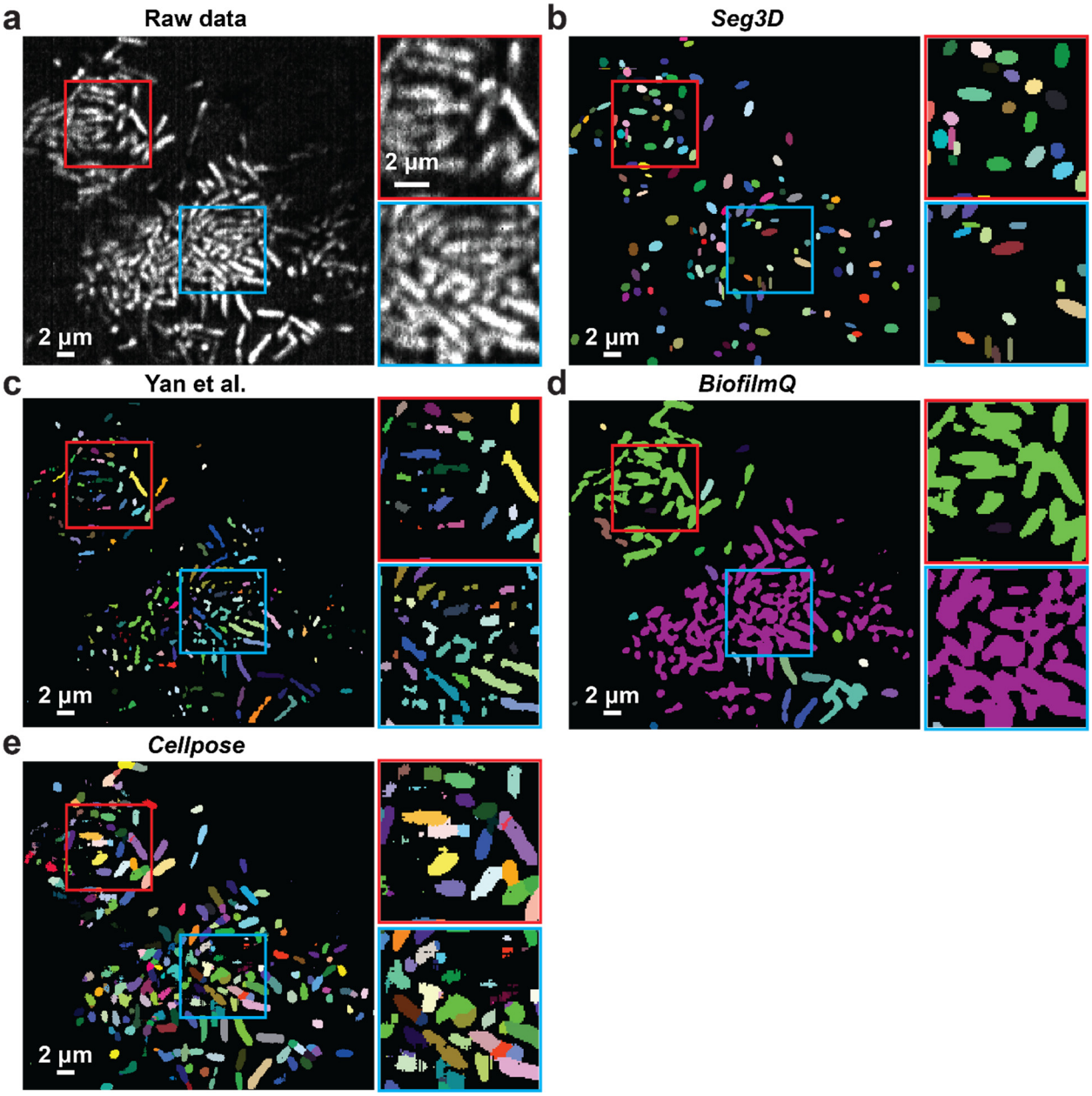


Figure S6. Visual comparison of segmentation results achieved by previous segmentation approaches that rely solely on mathematical image processing. (**a**) Experimental dataset is a *E. coli* biofilm containing GFP expressing cells 10 hours after the inoculation. (**b**) Segmentation result obtained using *Seg3D*[3]. (**c**) Segmentation result obtained using the algorithm in Yan *et al.*[4]. (**d**) Segmentation result obtained using *BiofilmQ*[5]. (**e**) Segmentation result obtained using *Cellpose*[6].

Figure S7

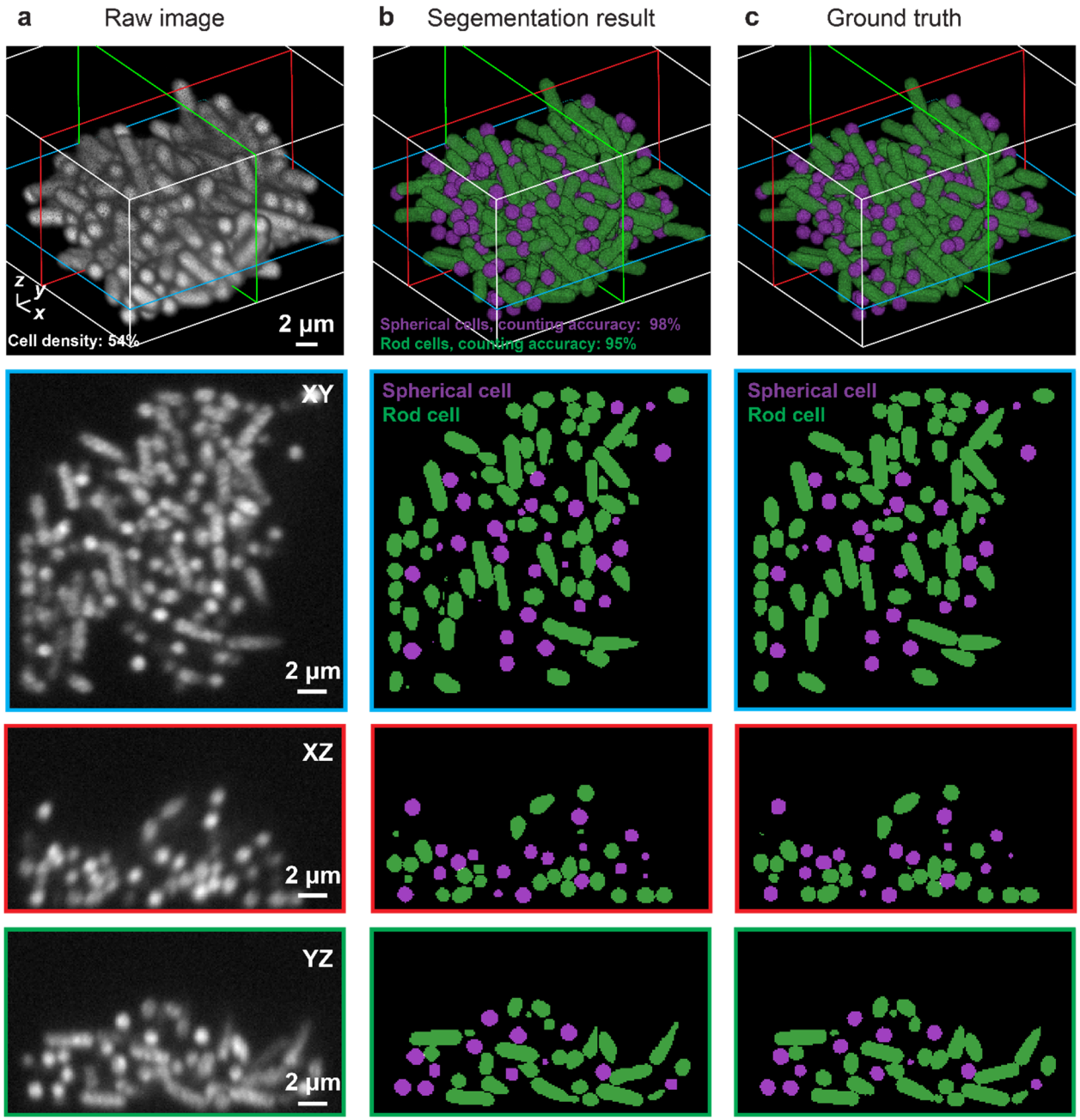


Figure S7. Segmentation of mixed population biofilms containing spherical cells and rod-shaped cells. (**a**) Simulated fluorescence image of a mixture of spherical cells and rod-shaped cells. The mixing ratio is 50:50. (**b**) Segmentation result obtained by *BCM3D* and (**c**) ground truth. Three orthogonal planes are shown below each 3D image.

Figure S8

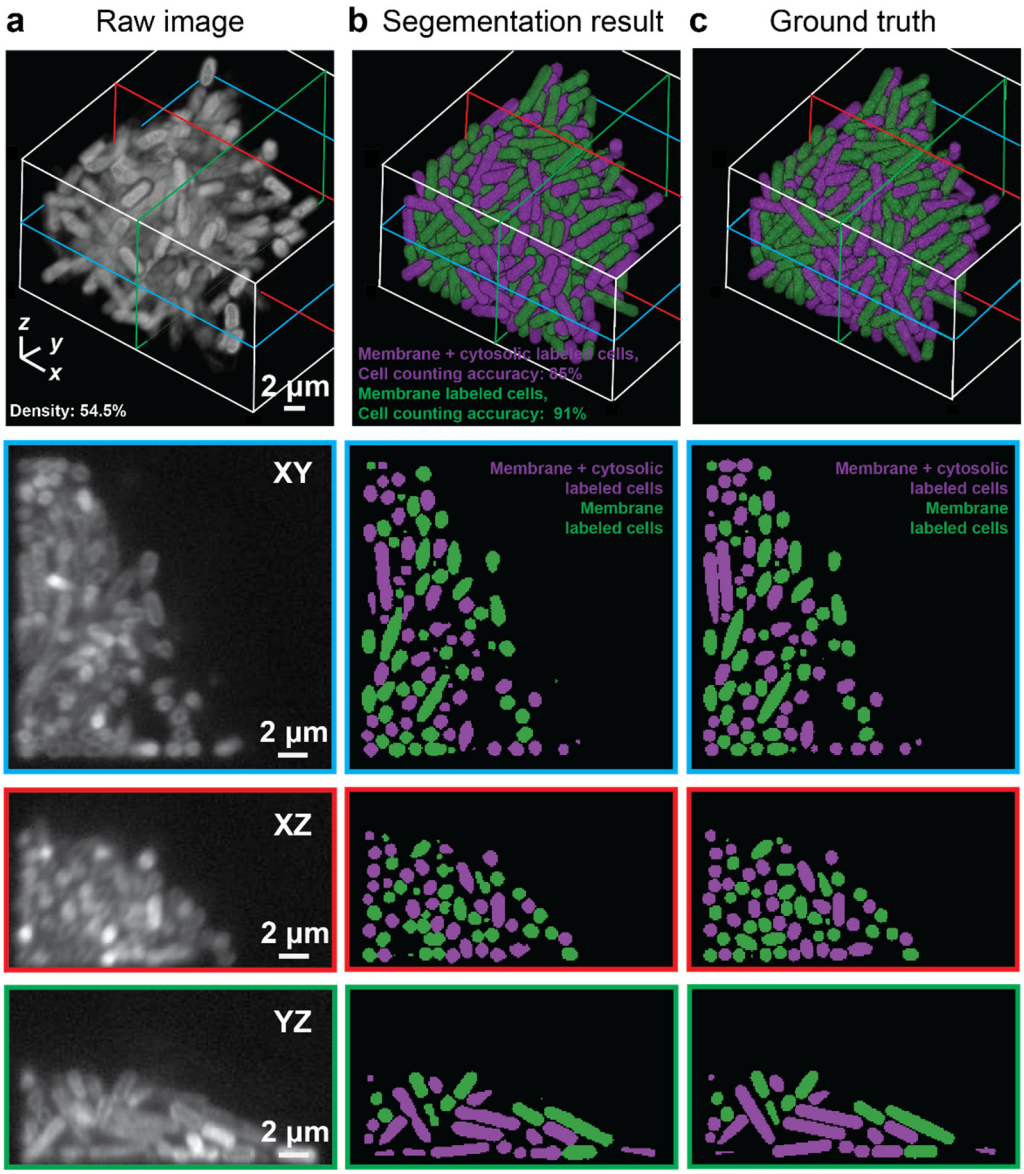


Figure S8. Segmentation of mixed population biofilms containing membrane-stained cells and membrane-stained cells that additionally express an intracellular fluorescent protein. (**a**) Simulated fluorescence image of a mixture of membrane-stained cells and membrane-stained cells that additionally express 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 S9

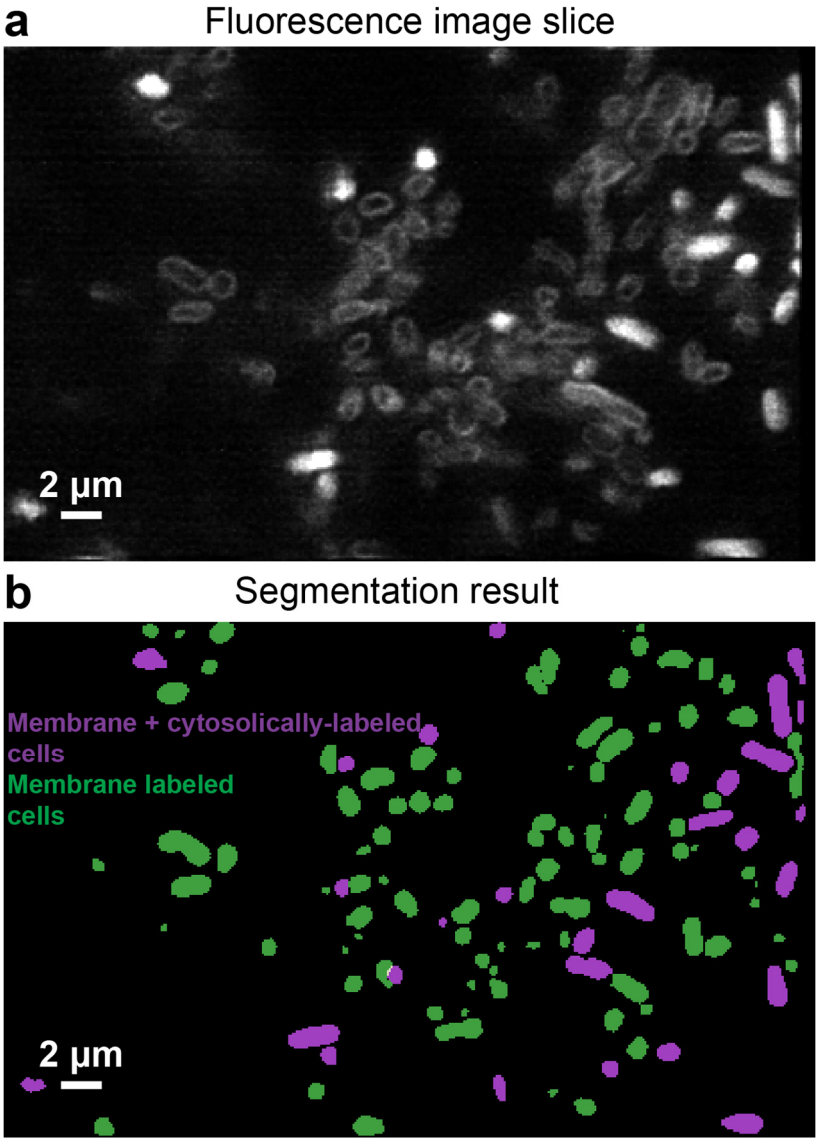


Figure S9. (**a**) Experimental 2D slice of a mixed *E. coli* population containing membrane-stained cells and membrane-stained cells that additionally express an intracellular fluorescent protein. The mixing ratio at the time of inoculation was 50:50. All cells were labeled by the FM4-64 membrane-intercalating 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).

6. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *bioRxiv*, 2020.2002.2002.931238 (2020).