Supporting Information

Synthetic Data-Driven Deep Learning for Label-free Autonomous Atomic Force Microscopy

Ruben Millan-Solsona,[1] Marti Checa,[1] Spenser R. Brown,[2] Amber N. Bible,[2] Bernadeta Srijanto[1], Laura Wiggins, [3] Sita Sirisha Madugula,[1] Alice L. B. Pyne,[3] Jennifer L. Morrell-Falvey,[2] Scott Retterer,[1,2] Rama K. Vasudevan,[1] Liam Collins.[1*]*

[1]Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA.

[2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA.

[3] School of Chemical, Materials and Biological Engineering, University of Sheffield, Sheffield, UK

*Correspondence to: solsonarm@ornl.gov & collinslf@ornl.gov

S1: Generation of Synthetic AFM Objects: Projection and Transformations

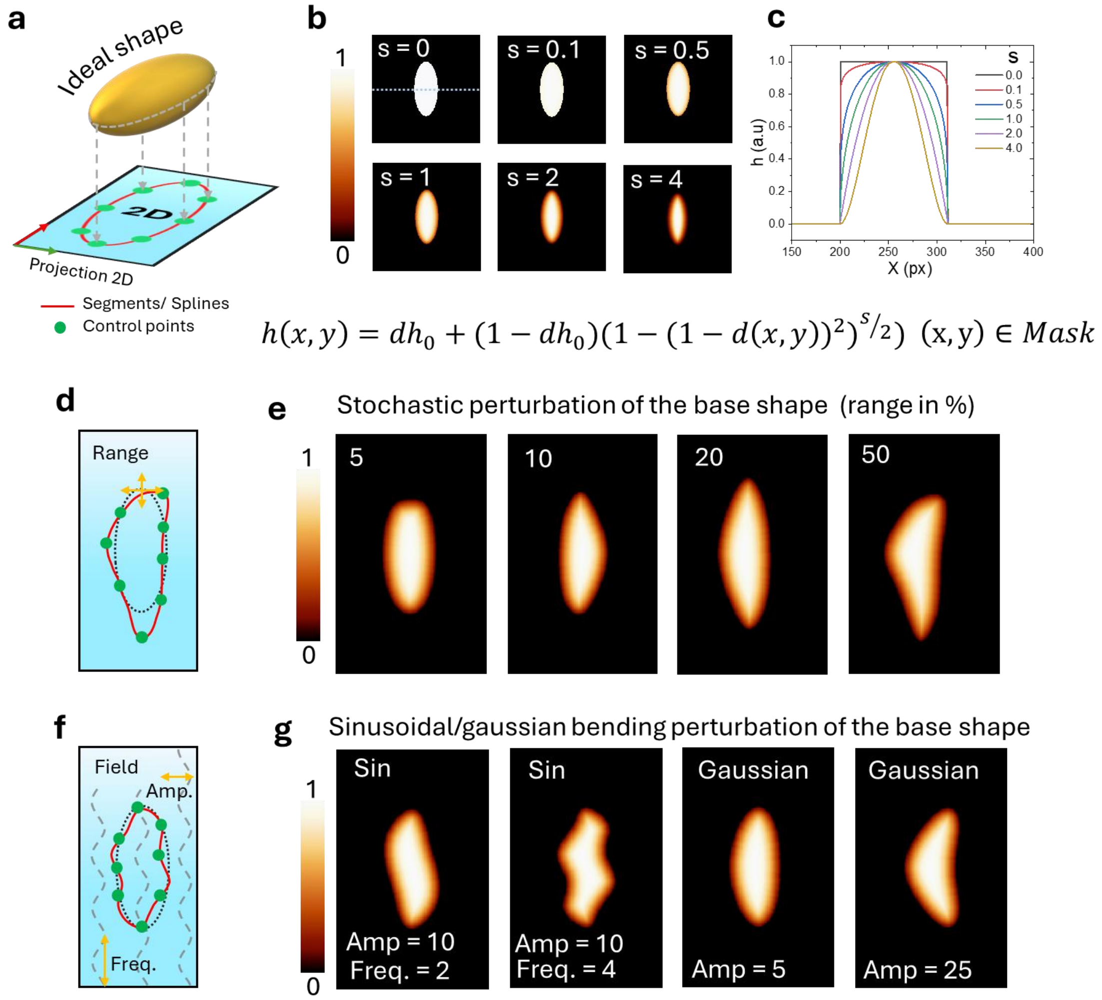

Figure S1: Synthetic AFM images are generated starting from an ideal shape, not necessarily analytical, which is projected onto the 2D plane to obtain the corresponding mask or segmentation required for data labeling. The topographic image is then constructed using a heuristic approximation *h(x,y)*, which assigns a height value based on the distance to the mask boundary *d(x,y)* and a shape parameter s that tunes the morphology. Additional transformations expand the variability of the generated objects: stochastic perturbation displaces the control points of the base shape within a defined range, producing random but controllable distortions, while bending transformations apply sinusoidal or Gaussian vector fields that deform the mask according to amplitude, frequency, and center of application. Finally, global scaling, rotation, and eccentricity adjustments can be introduced, further diversifying the synthetic dataset.

In the figure: (a) conceptual scheme of the projection of the ideal shape onto the 2D plane; (b) examples of topographic images generated for different values of the parameter s together with the defining height function; (c) longitudinal profiles corresponding to (b), illustrating morphological variability; (d) scheme of the stochastic perturbation applied to the control points of the mask;(e) examples of distorted shapes generated with different percentages of variation;

(f) scheme of the bending transformation;(g) examples of sinusoidal and Gaussian vector fields applied to the mask.

S2: Generation of Synthetic AFM Objects: Texture

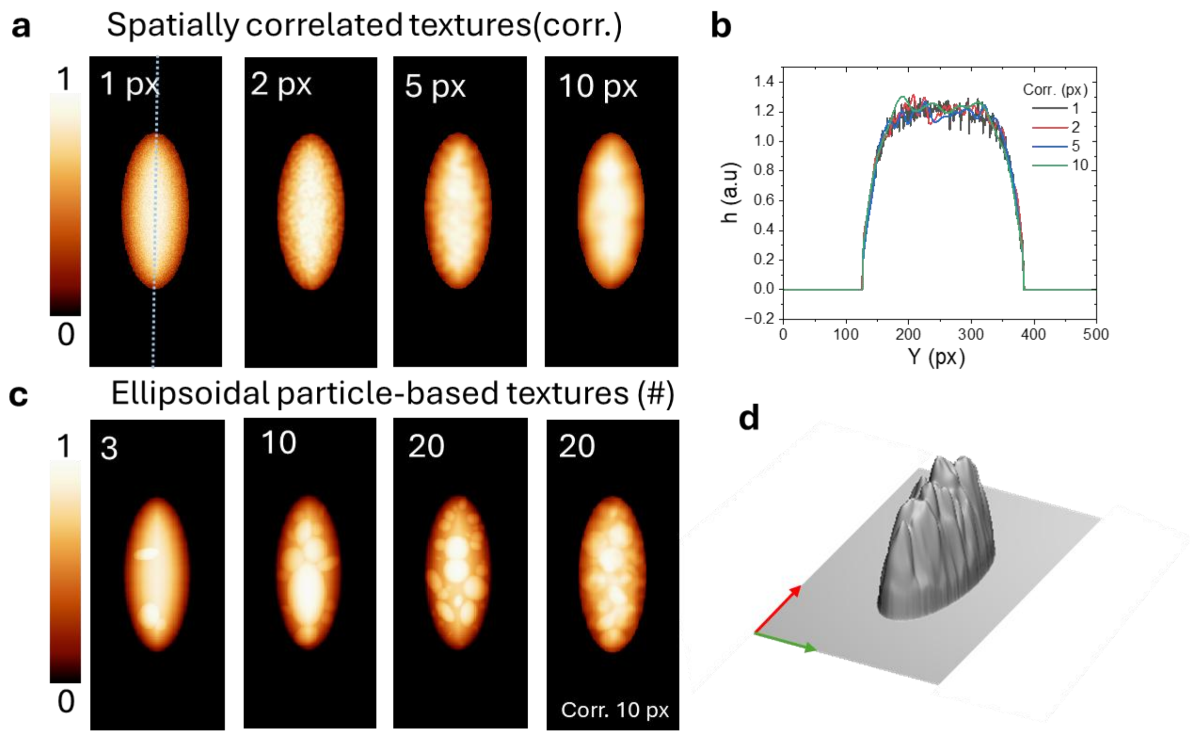

Figure S2: Once the object shape has been defined through the projection and transformation of the mask, together with its corresponding height function h(x,y), the next step is to introduce texture. Two types of textures are considered. The first corresponds to spatially correlated random noise, characterized by the correlation length (in pixels) and the amplitude, which determine the degree of spatial correlation and the height variation, respectively. In (a), topographic images are shown for correlation lengths of 1, 2, 5, and 10 pixels, all with an amplitude of 0.4. The corresponding topographic profiles for these cases are presented in (b). The second type of texture is based on ellipsoidal particles embedded within the objects. This approach provides an effective way to generate internal bumps or substructures, reproducing more realistic scenarios depending on the sample under consideration. The parameters controlling this texture include ellipse size, eccentricity, amplitude, number of ellipses, among others. These parameters provide a wide statistical variability, enhancing the richness of the synthetic data. Examples of such textures are displayed in (c). Notably, the rightmost image combines both a correlated texture (10-pixel correlation) and the ellipsoidal-particle texture. The corresponding 3D topography is shown in (d), where the richness of the statistical variability becomes evident. All these textures can be combined at different levels to further enrich the synthetic datasets.

S3: Generation of Synthetic substrates AFM

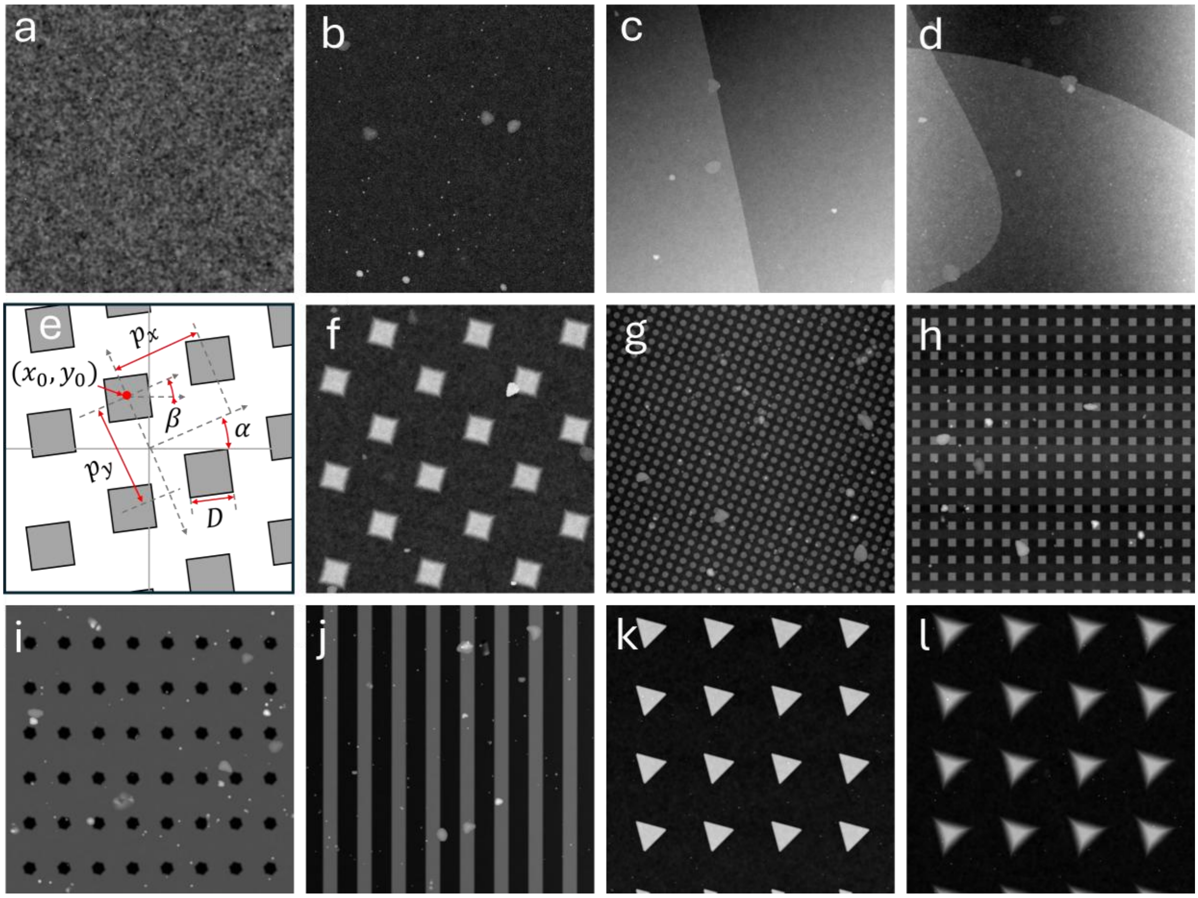

Figure S3: Examples of different substrates generated with SimuScan: (a) Flat substrate with correlated Perlin texture (correlation length of 5 pixels). (b) Substrate with texture and surface contamination. (c,d) Substrates with step features and linear and quadratic tilt, respectively. (e) Parametric schematic of substrate generation with periodic patterns. (f) Oriented square pillars with β = 30° and α = 45°. (g) Circular pillar pattern. (h) Denser square pillar pattern.
(i) Hexagon-shaped holes pattern. (j) Grating pattern. (k) Triangular pillars with zero edge smoothing and (l) pillars with 40% edge smoothing.

S4: Technical Description of Synthetic AFM Artifacts

We start from the ground truth topography *h(x,y)*(see Figure S2), representing the pristine height map of the synthetic objects projected onto the substrate. A sequence of artifact modules is then applied to reproduce realistic AFM conditions.

1. **Tip-sample convolution**
   The pristine surface is broadened by convolution with the probe geometry $t(x, y)$:

$$h_{topo}(x,y) = \min_{u,v}[h(u,v) + t(x-u, y-v)]$$

where $t(x,y)$ is typically modeled as a cone with spherical apex of radius R. This produces lateral resolution loss and rounding of sharp features.

2. **Feedback dynamics**

Feedback is usually simulated using a discrete proportional–integral (PI) controller:

$$z_{n+1} = z_n + K_p e_n + K_i \sum_{k=0}^{n} e_k$$

With the error term defined as

$$e_n = h_{tip}(x,y) - z_n$$

where $h_{tip}(x,y)$ is the tip-sample height and z_n the actuator position. In general, limited response time of the feedback loop produces overshoot, phase lag, and tracking errors, particularly at steep topographic edges. In our implementation we simplify the model by employing optimized values of K_p and K_i which minimize these effects. Under such conditions, adhesion between tip and sample becomes the dominant source of artifact, as it increases the effective noise amplitude and introduces stick–slip–like distortions that outweigh the residual feedback errors.

To further reproduce realistic AFM measurements, feedback-related artifacts are also modeled by incorporating line displacements (small offsets accumulated during fast-scan motion) and loss-of-contact events, both of which are commonly observed in experimental AFM datasets. Together, these effects provide a more faithful representation of the distortions introduced by feedback dynamics in real instruments.

a. **Adhesion modeling**

In real AFM experiments, tip–sample adhesion often dominates the quality of the image. To reproduce this effect in the synthetic datasets, we introduce a spatial adhesion field $A(x,y)$ that locally increases the apparent noise and produces stick–slip–like distortions. To reproduce adhesion effects, we consider three different scenarios. In some cases, adhesion is localized in random spots, representing regions where contamination or debris leads to stronger tip–sample interaction. In other cases, entire scan lines are affected, producing streaks across the image similar to those commonly observed in experimental AFM data. Finally, adhesion may occur only in a portion of a line, where short segments display increased sticking, mimicking irregular and transient events during scanning.

The adhesion effect is introduced mathematically as an increase in the standard deviation of Gaussian noise. Normally, the image contains electronic noise with a standard deviation σ_0 with adhesion, the effective noise becomes:

$$\sigma_{eff}(x,y) = \sqrt{\sigma_0^2 + \left(\beta A(x,y)\right)^2}$$

so that the random fluctuations are stronger in regions where $A(x,y)$ is high. In practice, this makes the topography appear more unstable and "rough" in sticky zones.

Additionally, adhesion is represented as small pixel shifts along the fast-scan direction, reproducing stick–slip events. In practice, this means that the recorded height values are displaced by a few pixels in regions of high adhesion, creating the characteristic irregularities observed in experimental AFM lines. Emulating stick–slip:

$$z_{adh}(x,y) = z_{topo}(x + \Delta x, y)$$

where Δx is a displacement that appears in areas with high adhesion.

b. Line-Loss Artifacts

In experimental AFM, the tip may temporarily lose contact with the sample due to feedback instabilities or sudden adhesion events. During these episodes, the scanner continues moving, but the acquired height data no longer follows the true surface. Instead, the signal appears as a straight segment bridging across the affected region. To reproduce this effect, line-loss artifacts are introduced by replacing either an entire scan line or only a portion of it with a straight interpolation between the boundary points. This substitution is applied randomly with a fixed probability, so that occasional line losses occur across the image, closely resembling the flat lines or line segments often observed in real AFM datasets.

3. Electronic and periodic noise

AFM images are also affected by different types of noise originating from the scanner electronics and the surrounding environment. To reproduce these effects in the synthetic datasets, two contributions are added to the topography.
The first one is electronic noise, which is modeled as a combination of Gaussian white noise and low-frequency $1/f$ noise:

$$\eta_{el}(x,y) = \eta_G(x,y) + \eta_{1/f}(x,y)$$

Where $\eta_G \sim \mathcal{N}(0,\sigma^2)$represents random uncorrelated fluctuations, and $\eta_{1/f}$ introduces slow drifts and correlated variations similar to those measured in real instruments.
The second contribution is periodic noise, which accounts for spurious oscillations from electrical pickup or mechanical resonances coupled into the scanning system. This is represented as a sinusoidal perturbation along the fast-scan direction:

$$\eta_{per}(x,y) = A_p \sin\big(2\pi f_p(x\cos\theta + y\sin\theta) + \phi_p\big)$$

where A_p is the amplitude, f_p the frequency in cycles per scan line, θ is the orientation angle of the pattern and ϕ_p a random phase.
The final noisy signal added to the feedback-controlled topography is then:

$$z_{moisy}(x,y) = z_{topo}(x,y) + \eta_{el}(x,y) + \eta_{peer}(x,y)$$

This simple formulation captures both the stochastic fluctuations and the structured oscillations that typically degrade AFM images.

4. Post-Processing Artifacts

One of the most common post-processing steps in AFM image analysis is flattening, where a background slope or drift is subtracted from the data. While this correction is necessary to remove scanner tilt and drift, it can also introduce characteristic artifacts if not applied carefully.
In our simulations, we reproduce this effect by applying a line-by-line flattening procedure. For each scan line (y_l), a polynomial background (typically first order, i.e. linear) is fitted to the raw data. To emulate the bias introduced by the presence of large

objects, we exclude from the fit the pixels belonging to the upper ~40% of height values, using a binary mask $M(x, y_l)$:

$$M(x, y_l) = \begin{cases} 1, & z(x, y_l) \le P_{60}(y_l) \\ 0, & z(x, y_l) > P_{60}(y_l) \end{cases}$$

where $P_{60}(y_l)$ is the 60th percentile of the heights in that line.
The flattening for each line is then obtained as:

$$z_{flat}(x, y_l) = z(x, y_l) - \hat{p}_l(x)$$

with $\hat{p}_l(x)$ the polynomial background fitted only to the pixels with $M(x, y_l) = 1$.
This approach reproduces the typical flattening artifact: regions containing tall features (e.g., large aggregates or contaminants) are excluded from the baseline, resulting in an artificial suppression of low-frequency variations and enhanced contrast of smaller features.

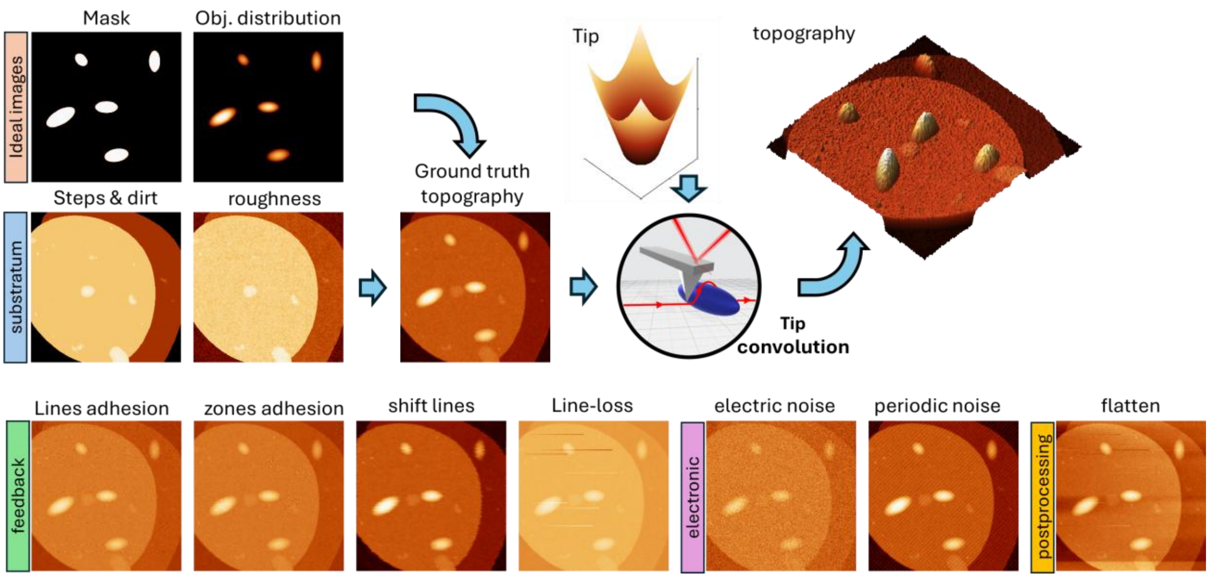

Figure S4: Scheme of synthetic image generation and artifact modeling. The process begins with a distribution of objects generated as described in the previous section, including geometric transformations and texture. In parallel, random substrates are created, composed of regions with different height levels arranged as steps, to which additional roughness and contamination are introduced. The object layer and the substrate are then combined to form the ground truth topography. Next, the image is convolved with a virtual AFM tip modeled as a cone with an opening angle between 10° and 30°, terminated by a semispherical apex of radius between 2 and 35 nm. This step produces an ideal topographic image free of artifacts. To obtain a more realistic dataset, typical distortions are subsequently added, including feedback-induced errors, electronic noise, and post-processing artifacts. The bottom row illustrates representative examples of each artifact considered, which are described in detail in Section S3.

S5: Double tip and flatten artefacts

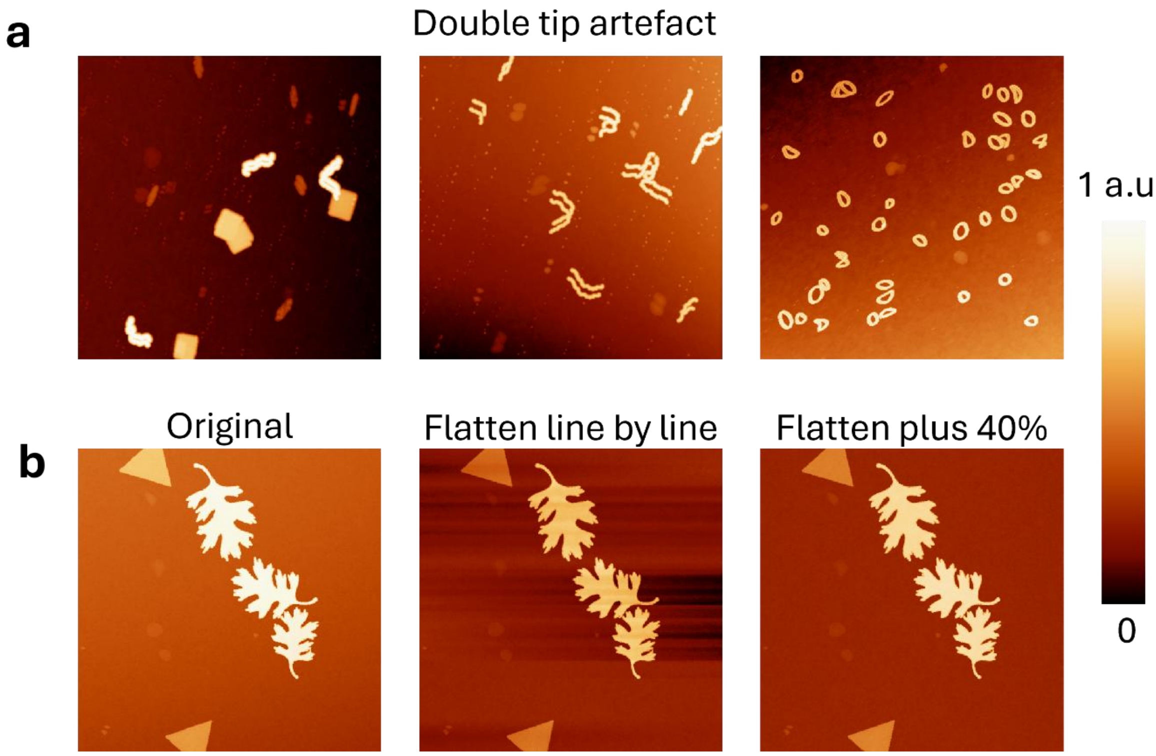

Figure S5: Double-tip and flattening artifacts. (a) Representative examples of synthetic images of linear DNA, circular DNA, and nanostructures affected by a double-tip artifact. (b) Example illustrating how flattening affects image quality and introduces artifacts that must be considered.

S6: Example outputs from models, trained exclusively on SimuScan datasets

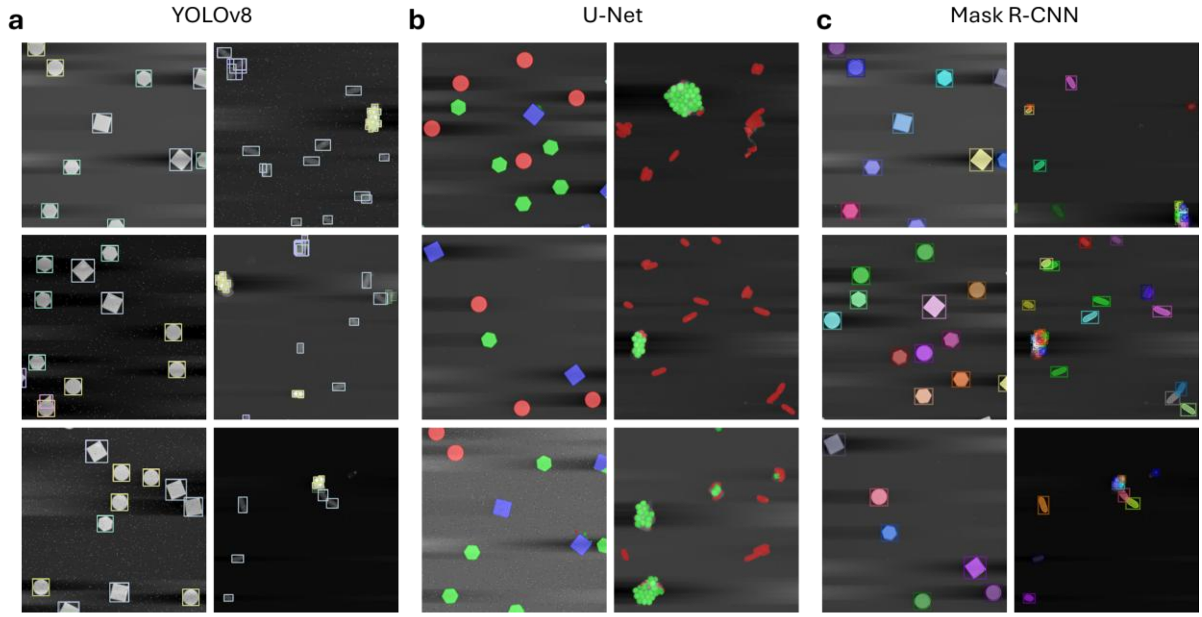

Figure S6: **Example outputs from models trained exclusively on SimuScan datasets**. (a) Representative detections obtained with the YOLOv8 model on experimental images containing both geometric shapes and biological samples of bacillus and coccus cells. (b) Segmentation results produced by a U-Net model on the same types of images. (c) Instance segmentation outputs from a Mask R-CNN model implemented with Detectron2. In all three cases, results are shown for experimental datasets including regular geometric morphologies and mixed bacterial samples, demonstrating the generalization capability of the models beyond the synthetic training data.

S7: Synthetic DNA Shapes and YOLO Predictions on Experimental AFM Images

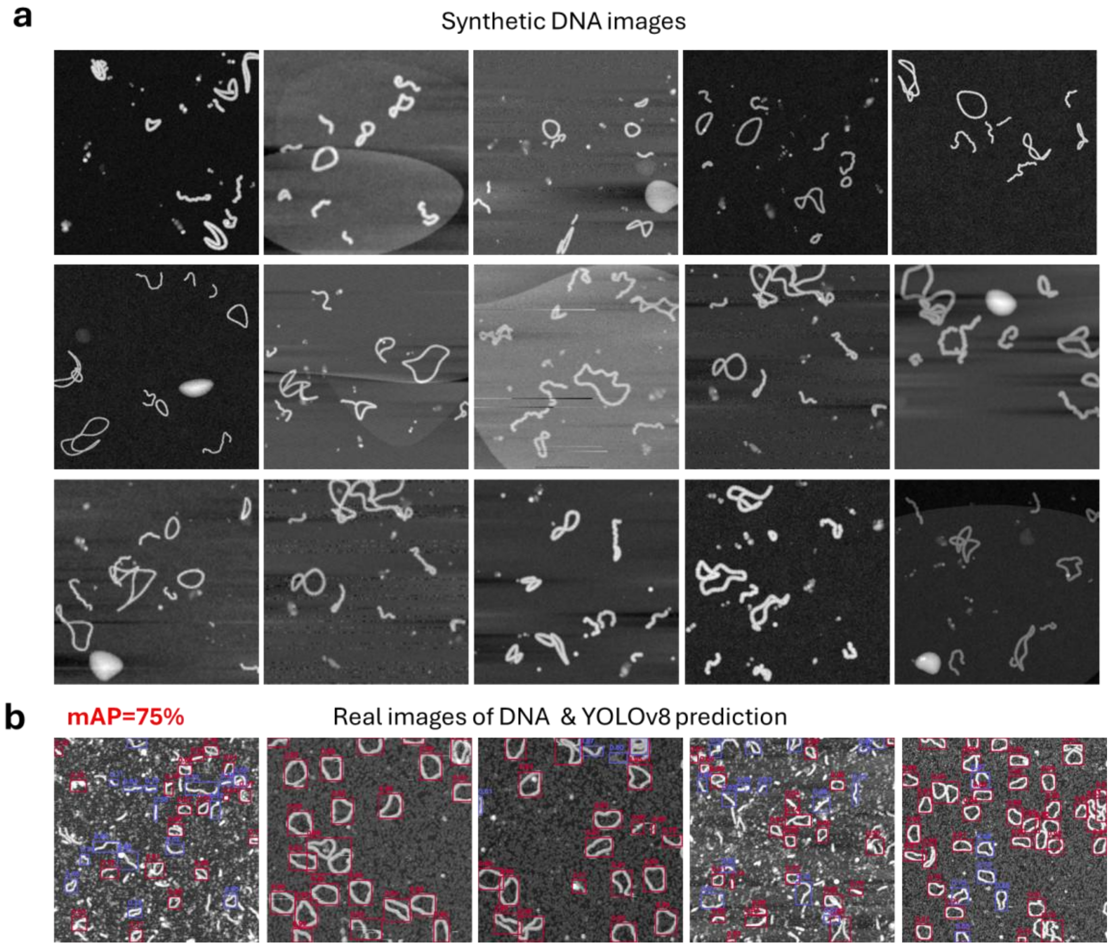

Figure S7: Synthetic DNA Shapes and YOLO Predictions on Experimental AFM Images. Examples of synthetic AFM images of circular and linear DNA are shown in (a). In (b), representative predictions from a YOLOv8 model trained exclusively on synthetic datasets are displayed on real experimental AFM images. Detections are shown with a confidence score >75%, with circular DNA highlighted in red and linear DNA in purple, illustrating the ability of the model to transfer from synthetic training data to experimental measurements.

S8: SimuScan-based training and validation on bacterial cells deposited on circular nanopillar substrates.

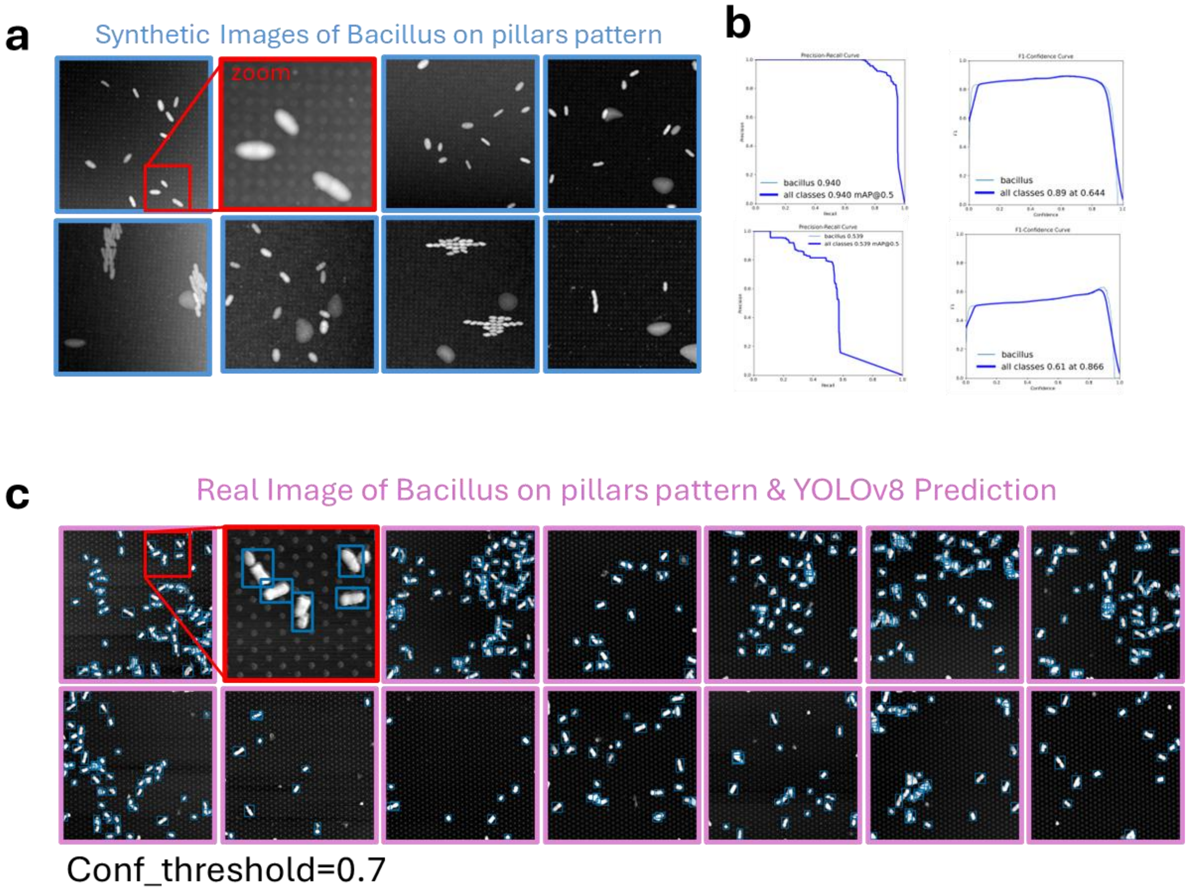

Figure S8: (a) Representative examples of synthetic AFM images generated with SimuScan, reproducing bacterial bacillus morphologies on substrates patterned with circular nanopillars. (b) Precision, recall, and F1-score curves for object detection (top) and pixel-wise segmentation masks (bottom) obtained during model training and evaluation. (c) Representative predictions on experimental AFM images of bacterial cells on circular nanopillar substrates, demonstrating accurate detection and segmentation despite the non-planar and topographically complex background.

S9: SimuScan-based training and validation on complex leaf-shaped nanostructures.

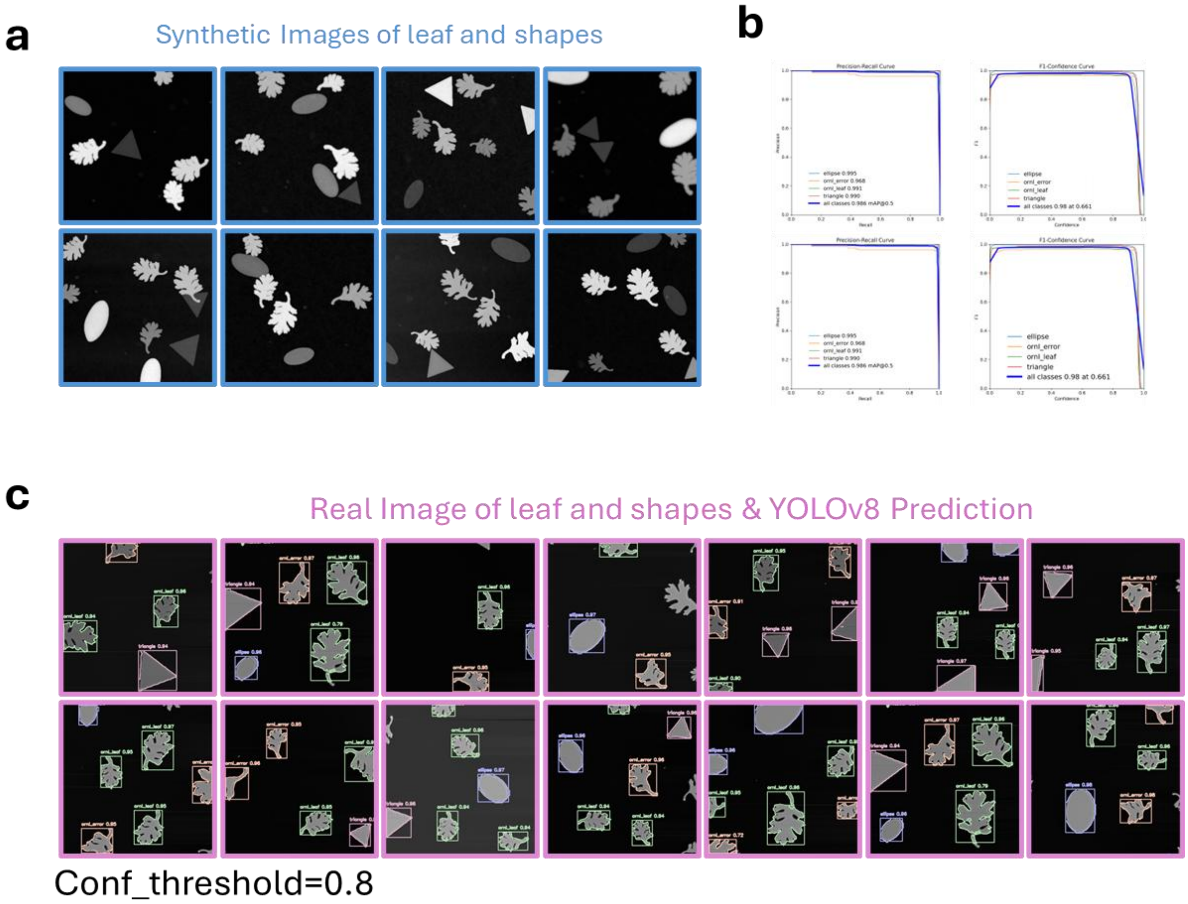

Figure S9: (a) Representative examples of synthetic AFM images generated with SimuScan for complex engineered nanostructures with leaf-like geometries. (b) Precision, recall, and F1-score curves for object detection (top) and pixel-wise segmentation masks (bottom) obtained during model training and evaluation. (c) Representative predictions on experimental AFM images of complex leaf-shaped nanostructures, illustrating robust generalization of SimuScan-trained models to experimentally acquired data with irregular and non-analytical geometries.

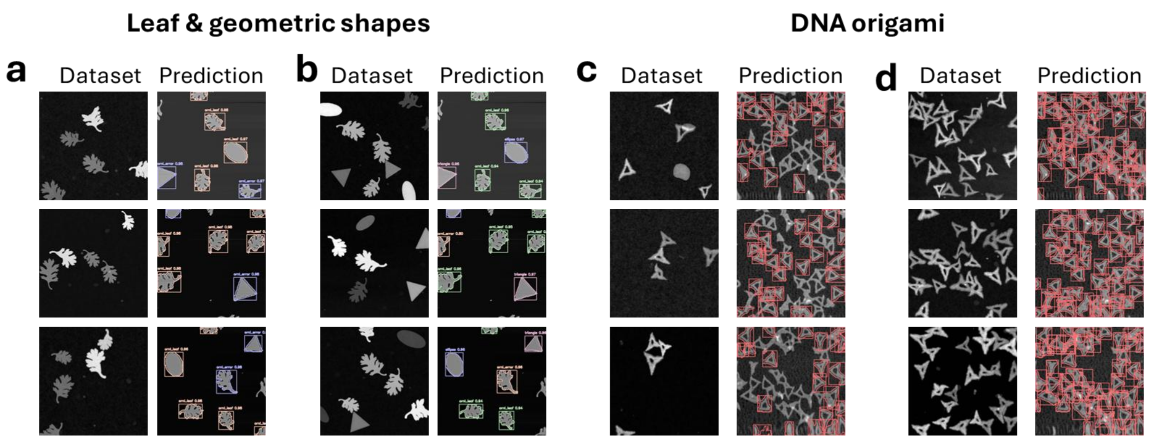

Figure S10: Impact of synthetic data composition on YOLO detection and segmentation performance. Several examples illustrate how the design of synthetic training datasets affects model predictions. (**a**) Synthetic data generated without including additional objects present in the experimental sample (e.g., ellipses and triangles). As a consequence, the YOLO-based object detection and instance segmentation incorrectly identifies some ellipsoidal objects as targets. (**b**) Improved prediction accuracy was realized when these additional object classes are incorporated into the synthetic training dataset, reducing false detections. A second example based on **triangular DNA origami** highlights the role of object density. (**c**) When the training dataset mainly contains isolated objects, the model preferentially detects well-separated origami structures, leading to reduced performance in crowded regions. (**d**) By increasing the density of triangular origami objects in the synthetic dataset to better match experimental conditions, detection and segmentation of strongly overlapping structures are significantly improved.

S11. Resolution Dependence of YOLO Detection Accuracy (25 × 25 µm)

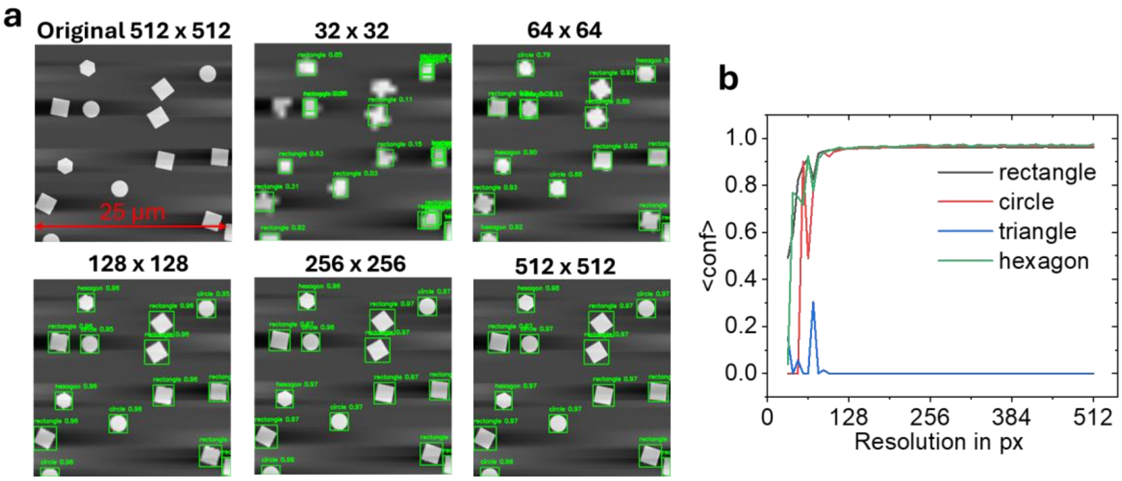

Figure S11. Dependence of YOLO detection accuracy on image resolution. (a) Example predictions on geometric shapes for input images of 512 × 512, 256 × 256, 128 × 128, 64 × 64, and 32 × 32 pixels, all corresponding to a 25 × 25 µm field of view. (b) Confidence versus image resolution for each class. Confidence remains above 90% for all classes when the resolution is at least 128 × 128 pixels, indicating the minimum sampling density required for robust detection at this field of view.

S12. Resolution Dependence of YOLO Detection Accuracy (50 × 50 µm)

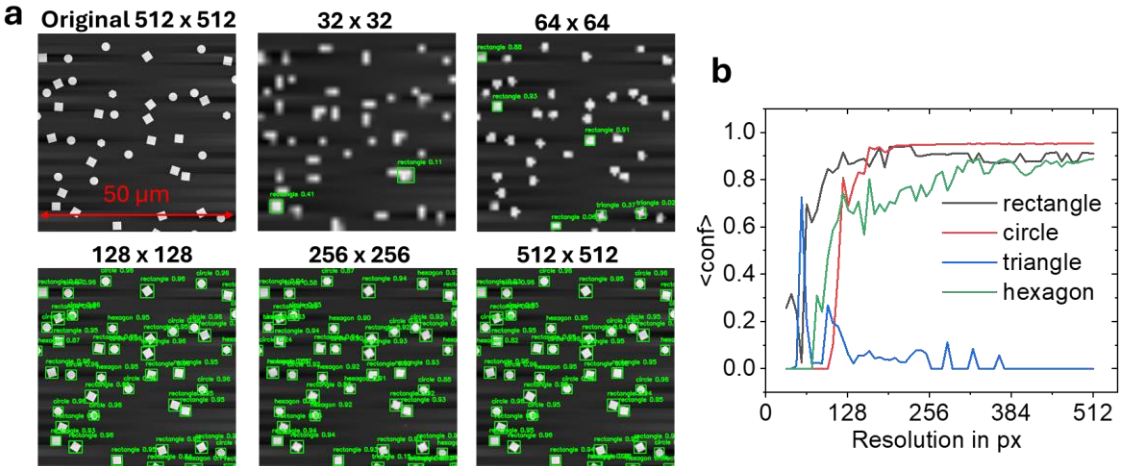

Figure S12. Dependence of YOLO detection accuracy on image resolution. (a) Example predictions on geometric shapes for input images of 512 × 512, 256 × 256, 200 × 200, 128 × 128, and 64 × 64 pixels, all corresponding to a 50 × 50 µm field of view. (b) Confidence versus image resolution for each class. In this case, confidence values exceed 90% only when the resolution is at least 200 × 200 pixels, reflecting the greater sampling requirements imposed by the larger field of view.

S13. High-resolution AFM image of a Bacillus cell

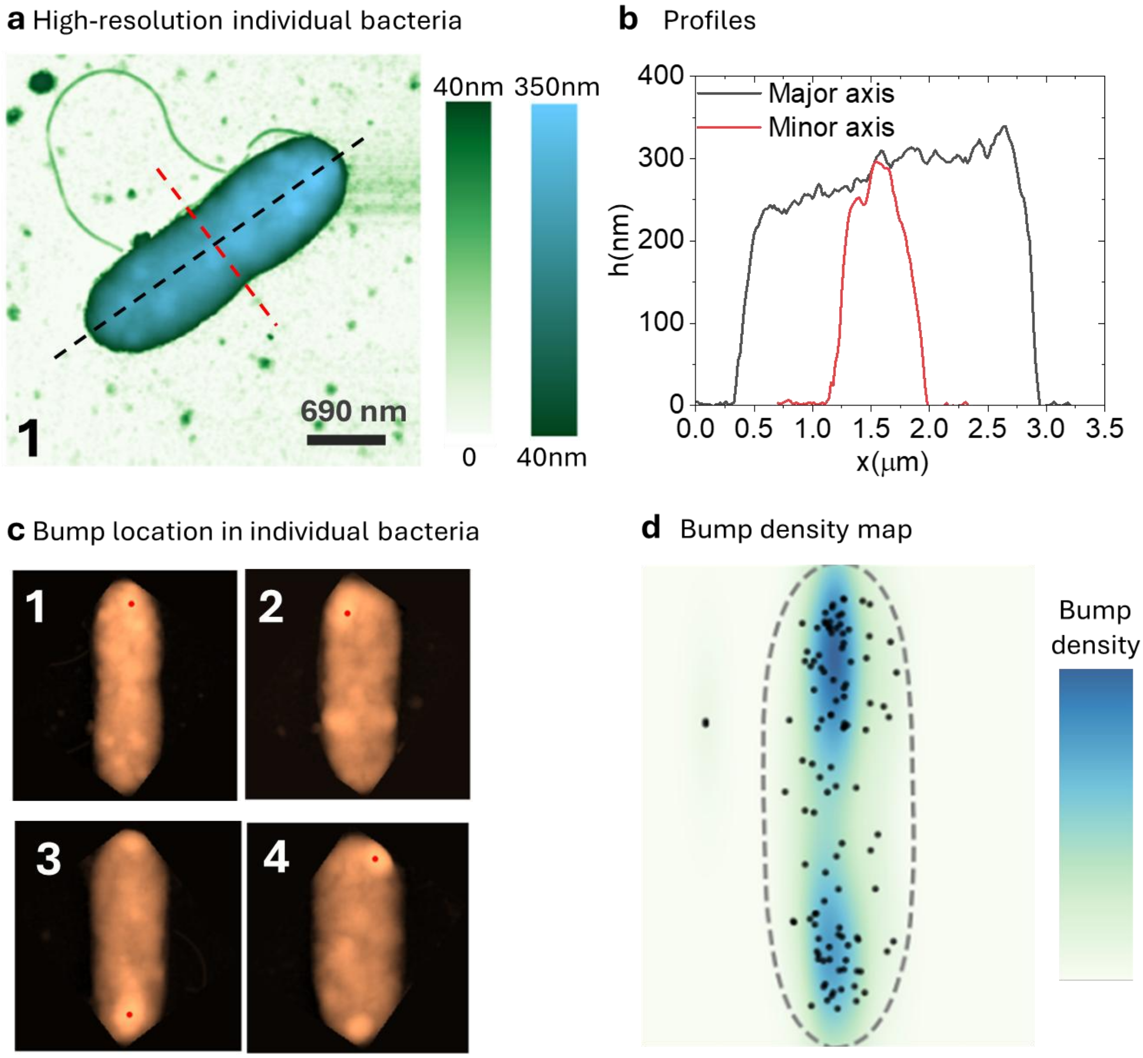

Figure S13. (a) High-resolution AFM image of a bacillus cell shown at two different scales to highlight the structural richness captured by the autonomous workflow. (b) Topographic profiles along the major and minor axes of the bacterium shown in (a). (c) Representative examples of multiple *P. aeruginosa* cells with localized membrane protrusions, where red points mark the detected maxima. (d) Bump density map compiled from 100 bacillus cells, normalized to the major axis (scaled to 1) to improve visualization of the recurrent polar distribution of protrusions.

S14. Zoomed-in views corresponding to the large-area analysis shown in Figure 7

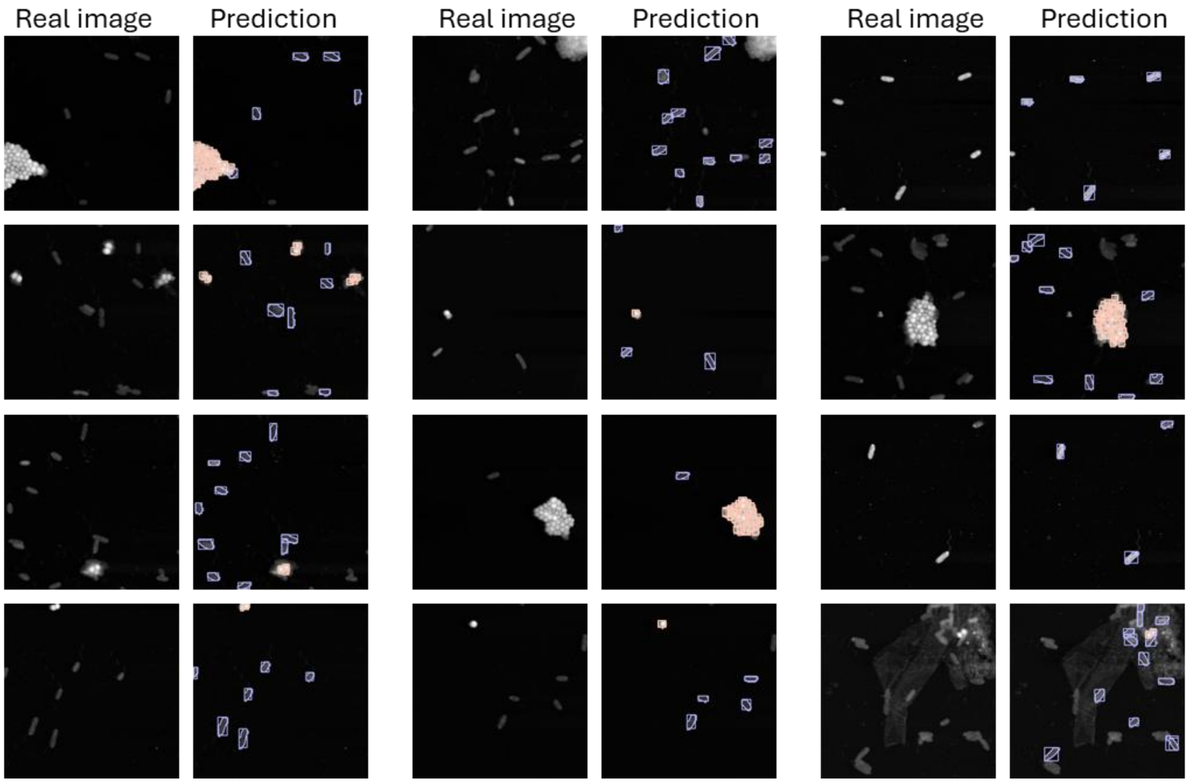

Figure S14 High-magnification views of representative regions from the experimental AFM dataset shown in Figure 7, displayed alongside the corresponding model predictions. Spherical coccus-type bacteria are highlighted in orange in the prediction maps, while rod-shaped bacillus-type bacteria are shown in purple. Consistent with the experimental topography, coccus cells generally exhibit a greater apparent height than bacillus cells.