Biofilm 12 (2026) 100383

Contents lists available at ScienceDirect

Biofilm

journal homepage: www.sciencedirect.com/journal/biofilm

From annotation to analysis: a deep-learning pipeline for optical coherence tomography (OCT)-based measurements of biofilm morphology

R.I. Heroza [a,b], P. Azizinezhad [a], K.-A. Moss [c], H.M. Lewis [c], P. Pajewska [c,d], A.G. Seco de Herrera [e], A.J. Dumbrell [c], P.P. Laissue [c], N. Aldred [c,*]

[a] School of Computer Science and Electronic Engineering, University of Essex, CO4 3SQ, UK
[b] School of Computer Science, Universitas Sriwijaya, Indonesia
[c] School of Life Sciences, University of Essex, CO4 3SQ, UK
[d] School of Biological and Behavioural Sciences, Queen Mary University of London, E1 4NS, UK
[e] UNED, Spain

ARTICLE INFO

Keywords:
Biofilm
Segmentation
Optical coherence tomography
OCT
Deep learning
Morphological analysis
YOLOv8
SegFormer

ABSTRACT

Biofilms represent the predominant mode of bacterial life at solid-liquid interfaces, and understanding their composition, structure, and dynamics is critical to addressing key challenges across medical, environmental, and engineering applications. This study presents a deep learning-based framework for rapid morphological characterisation of biofilms using optical coherence tomography (OCT) imaging and an automated image processing pipeline. Images were used to train two state-of-the-art segmentation models: YOLOv8 and SegFormer. Both models delivered impressive results in delineating biofilm structures; YOLOv8 achieved 0.99 for accuracy and an intersection over union (IoU) of 0.9, while SegFormer scored 0.97 and 0.87, respectively. Model robustness was assessed across eight challenging biofilm conditions, with YOLOv8 showing superior performance in discriminating thin and non-growing biofilms, and SegFormer's superiority with stable morphologies. Additionally, we developed a framework to extract key morphological characteristics from the segmented images, including thickness, roughness and density distribution. The model-derived measurements showed strong agreement with manually generated ground truth data, confirming the reliability of the automated pipeline. Furthermore, an experiment involving four taxonomically distinct multi-species biofilms demonstrated the utility of the approach for discriminating biofilms based on their morphology. The software, containing both segmentation models, is openly available to the community and provides a foundation for future high-throughput studies examining biofilm responses to taxonomic or environmental variation.

1. Introduction

Biofilms are communities of microorganisms encapsulated within an extracellular polymeric (EPS) matrix [1–3]. They represent the predominant lifestyle of aquatic bacteria, which benefit from properties such as enhanced physical protection, increased environmental resilience, nutrient concentration, facilitation of horizontal gene transfer and distribution of metabolic processes. However, biofilms also pose significant challenges to human activity, contributing to enhanced antimicrobial resistance [4], drag, corrosion [5] and biofouling accumulation [6,7]. Beyond their central role in ecosystem functioning, biofilms therefore encroach on human activities across high-value sectors including medicine and healthcare, industrial engineering, the maritime industry and food production.

Characterising the morphology of biofilms presents challenges due to their complex, heterogeneous nature and the limitations of conventional imaging techniques. Progress over the past three decades has been closely tied to advances in microscopy and tailored analysis pipelines [8].

Confocal laser-scanning fluorescence microscopy is widely used to image biofilms in three dimensions [9]. An example is shown in Fig. 1A, where differential staining reveals a mosaic of bacterial populations in a one-day-old biofilm. This is further illustrated in Fig. 1B, where individual bacterial cells can be clearly identified. As biofilms mature, they increase in complexity; microbial communities aggregate and form structures that extend in three spatial dimensions. This can be seen in

* Corresponding author.
E-mail address: nick.aldred@essex.ac.uk (N. Aldred).

https://doi.org/10.1016/j.biofilm.2026.100383
Received 1 April 2026; Received in revised form 25 June 2026; Accepted 8 July 2026
Available online 9 July 2026
2590-2075/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

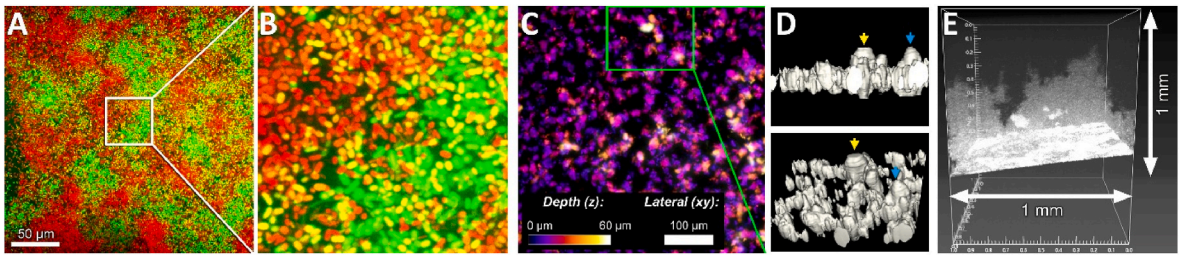


Fig. 1. A) Fluorescence micrograph at high magnification (60×) showing distinct bacterial populations in a one-day-old multi-species biofilm. **B)** Close-up of the white square in A), showing individual bacteria. **C)** Fluorescence micrograph (20x) of the same type of biofilm after two days. Bacteria form clumps and start to rise up in the third spatial dimension. False-colouring indicates biofilm depth, with a basal layer (purple/blue) of around 20 µm thickness, a less dense intermediary layer (red/orange, around 20-40 µm), and towering clumps rising up to 60 µm (yellow/white). **D)** Close-ups of the green rectangle in C). The two tallest clumps are indicated with yellow and cyan arrowheads. **E)** An OCT c-scan (3D) of the same type of biofilm after five days. The vertical component of the biofilm has further increased, although reduced contrast makes segmentation more challenging. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 1C, where depth is colour-coded to highlight variations in vertical structure. In some cases (e.g. Fig. 1C), native autofluorescence is sufficient for imaging, however, confocal microscopy typically requires fluorescent labelling of specific biofilm components [10] and relies on high-energy laser excitation [11]. Both factors are potentially toxic and may alter the biofilm properties under study [10]. Moreover, confocal microscopy is a relatively low-throughput and high-magnification technique, providing excellent resolution but only across limited length-scales [11] and to imaging depths of 200 µm.

Optical Coherence Tomography (OCT), by contrast, is a non-invasive, label-free, three-dimensional imaging modality that has proven to be highly effective for visualising biofilm structures in real time [12–14]. OCT provides micrometre-scale resolution across millimetre-scale fields of view and can penetrate hundreds of microns of depth (Fig. 1E). Unlike confocal microscopy, which offers sub-micron resolution but is limited to relatively thin biofilms, OCT can image intact, hydrated, mature biofilms growing on opaque or clinically relevant substrates without perturbing their structure. Cross-sectional B-scan images acquired through biofilms in liquid media can be assembled into three-dimensional C-scans for structural visualisation, making the technique highly suitable for studying both surface and internal morphology. By bridging the gap between cellular-scale imaging (e.g. fluorescence microscopy) and bulk morphological characterisation (e.g. visual inspection or macro-photography), OCT enables quantitative, reproducible and longitudinal assessment of biofilm structure under conditions closely resembling natural, industrial or clinical environments.

However, OCT images of biofilms can suffer from biases [15] due to low contrast, noise, and motion artifacts. These challenges complicate the extraction of accurate morphological information and have limited the widespread adoption of OCT for quantitative biofilm analysis. A key challenge remains the development of robust image processing pipelines capable of reliably quantifying biofilm structure from OCT images. Traditional methods have relied on manual or semi-automated segmentation techniques [14] which, although sometimes effective, are time-consuming, operator dependent and prone to error; especially when biofilm boundaries are ambiguous or when large volumes of data must be processed. Moreover, the lack of publicly available annotated datasets specific to biofilm OCT imaging presents a barrier to the development and benchmarking of fully automated analytical tools.

Several recent studies have applied machine-learning and deep-learning methods to segment and analyse biofilms in OCT volumes, although the field remains immature. Classical image-processing techniques and texture-based feature extraction have been combined with supervised classifiers (e.g. SVM, random forest, XGBoost) to identify and classify biofilm regions in middle-ear OCT data [16]. Other efforts have moved toward volumetric and fully automated pipelines. For example, Narciso et al. [17] introduced BISCAP, a 3D OCT image-processing framework designed to automate the characterisation and segmentation of biofilm structure from OCT stacks. Longitudinal OCT studies that track biofilm growth over time underscore the need for robust automated segmentation capable of extracting thickness, roughness and porosity metrics across timepoints. Such studies often combine bespoke segmentation methods with temporal analysis to quantify growth and detachment dynamics [18]. Classical machine learning and convolutional neural-network approaches are promising routes for OCT biofilm segmentation, but need large datasets and standardised benchmarks to work from [12].

Deep learning models such as YOLOv8 [19] and SAM (Segment Anything Model) [20] have demonstrated remarkable performance in object detection and instance segmentation across diverse image types. In this study, we introduce a deep learning-based framework for analysing OCT images of biofilms, combining SAM for initial segmentation and object boundary delineation with two state-of-the-art models: SegFormer, a transformer-based architecture for semantic segmentation [21], and YOLOv8. We present a training pipeline designed to overcome the limitations imposed by the scarcity of annotated biofilm datasets.

Our deep learning-based approach harnesses SAM's zero-shot segmentation capabilities, which have been trained on over one billion masks [22] and can identify objects without prior exposure to specific categories during training [23]. By using SAM-generated masks, we address the drawbacks of manual annotation, a labour-intensive and error-prone process [24], while providing a scalable and efficient solution for biofilm image analysis.

To demonstrate our analysis pipeline, we developed a custom biofilm dataset comprising OCT images categorised into eight distinct cases, representing biofilm conditions with stable, low-contrast, drifting, blurry, fragmented, and thin morphologies. This diverse dataset enables the evaluation of models under a wide range of natural biofilm appearance and behaviour, and expert validation was employed to confirm the accuracy of the segmentation results.

Specifically, we describe:

Development of a robust, automated framework: We introduce an automated pipeline for segmenting and analysing biofilm structures in OCT images, addressing challenges posed by image artifacts and diverse biofilm conditions. This framework enables more accurate and reliable biofilm characterisation, surpassing the capabilities of existing methods.

Creation of a custom biofilm dataset: We compiled a dataset of 2800 OCT images, categorised into eight distinct biofilm conditions. This dataset allows for the evaluation of deep learning models across biofilms exhibiting substantial variability in appearance and behaviour.

Assessment of deep learning models: We evaluate the performance and generalisability of deep learning models (SegFormer and YOLOv8) across a range of biofilm scenarios. Expert feedback was incorporated to further validate the models' effectiveness and robustness, particularly under complex biofilm conditions.

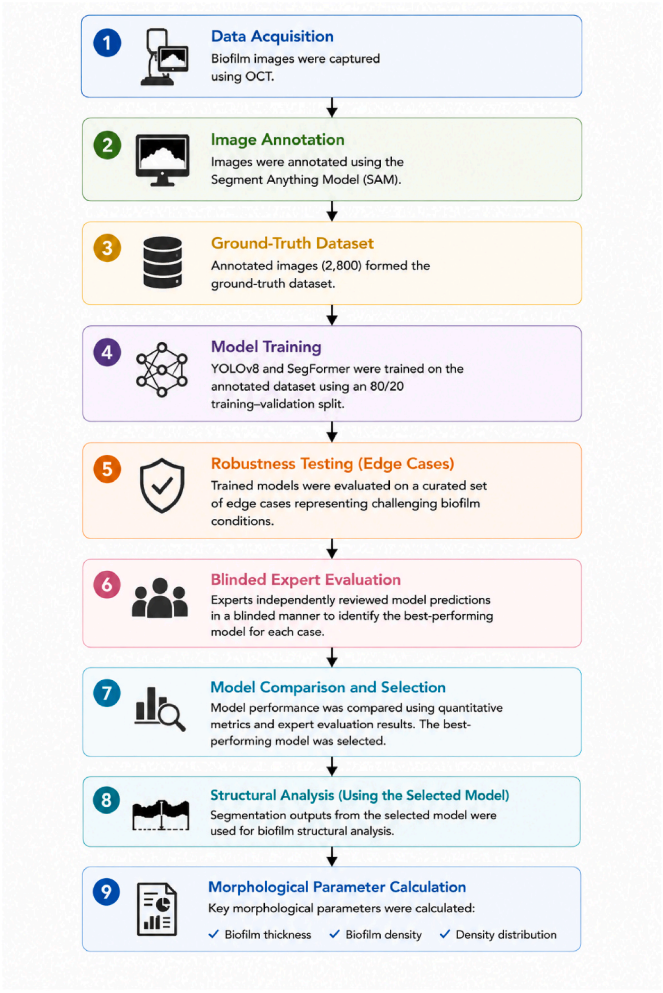


Fig. 2. Overview of the biofilm segmentation and analysis workflow.

Calculation of key morphological metrics: We compute four metrics of biofilm morphology (thickness, roughness, density, and density distribution) to facilitate the comparison of biofilm growth and structural characteristics.

2. Methods

The development process was structured into data acquisition, annotation, model training, robustness testing, model selection and structural analysis stages (Fig. 2). Biofilm images were first captured using OCT, and image annotation was performed using the Segment Anything Model (SAM). The resulting annotated data served as a ground-truth dataset for training YOLOv8 and SegFormer. The models were trained on 2800 annotated images using an 80/20 training-validation split. Following training, model robustness was evaluated on a set of edge cases representing challenging biofilm conditions. Expert evaluations were conducted in a blinded manner to identify the best-performing model predictions. Finally, the segmented outputs generated by the trained models were used to quantify key morphological parameters, including biofilm thickness, density and density distribution, demonstrating the applicability of the segmentation framework for downstream biofilm structural analysis. An overview of the complete methodology is provided in Fig. 2.

2.1. Imaging

OCT imaging was conducted using a ThorLabs spectral domain Ganymede OCT system with axial resolution of 3 μm and a LSM03-BB lens, providing lateral resolution of 8um. The working distance of approximately 25 mm allowed biofilms to be grown and imaged in polystyrene 24-well tissue-culture plates (Costar, Corning) containing 2 mL of liquid medium. All images acquired in our experiments were two-dimensional B-scans, representing a 'cross-section' through the biofilm at randomly-selected locations. Each B-scan had a resolution of 2097 × 1021 pixels, corresponding to a physical scale of 1.5 mm × 1.985 mm (X × Z). Prior to analysis, images underwent several pre-processing steps, including resizing and normalisation, to standardise them for subsequent computational processing.

Table 1
Taxonomic composition of the biofilm communities used in experiments [7,25].

Community	Isolate code	Species id
A	ATLC8b	*Simiduia agarivorans*
	H13	*Marinobacter sp.*
	H2	*Alteromonas sp.*
B	ATLC11	*Isoptericola dokdonensis*
	H13	*Marinobacter sp.*
	H2	*Alteromonas sp.*
C	ATLC3	*Rheinheimera sp.*
	H22	*Maribacter sp.*
	H13	*Marinobacter sp.*
D	ATL1	*Pseudoalteromonas sp.*
	ATL28	*Pseudoalteromonas piscicida*
	H2	*Alteromonas sp.*

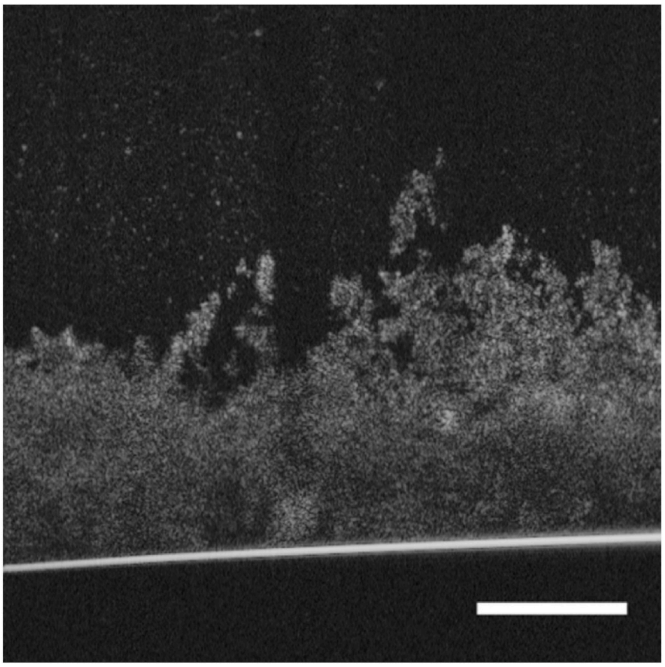


Fig. 3. An example OCT B-scan, or two-dimensional slice, through a 48h-old laboratory-grown biofilm. Scale bar = 200 μm.

2.2. Biofilm growth

Biofilms were produced using previously-isolated marine bacteria strains under controlled experimental conditions. Briefly, twelve isolates of *Alteromonas* sp, *Marinobacter* sp, *Maribacter* sp, *Simiduia* sp, *Pseudoalteromonas* sp, *Pseudomonas* sp, *Isopteriocla* sp, and *Rheinhainera* sp. were provided by the University of Copenhagen and Clemson University. Further details regarding isolation and identification are available in [7,25]. These strains were chosen for their ability to grow well individually or in multispecies communities, under laboratory conditions;

static, at room temperature in full-strength Difco Marine Broth. Every possible three-species combination was used to produce biofilms with a wide range of morphologies. Seven OCT images were taken per biofilm, resulting in a total of 2800 images used to train the deep-learning models.

The specific communities A-D listed in Table 1 were used to illustrate that the proposed framework can discriminate between taxonomically distinct biofilms. Isolates were taken from glycerol freezer stocks using sterile inoculation loops and introduced to 40 ml of Difco Marine Broth (produced according to manufacturer's instructions) and incubated on an elliptical shaker at 75 rpm for 48 h at ambient temperature. One replicate of each three-species community biofilm was produced by combining 500uL each of 3 single-species cultures. Biofilms were grown in individual wells of Corning Costar 24-well polystyrene culture plates and incubated for a 48 h period in static conditions. OCT imaging was conducted immediately thereafter. Seven 1.5 mm-long OCT B-scans were captured in each biofilm-coated well and down-stream analyses were conducted on all of the acquired images. An example biofilm B-scan image is shown in Fig. 3.

2.3. Dataset splitting and pre-processing

The dataset of 2800 images was split into 80% for training (2240 images) and 20% for test (560 images) to evaluate model performance on unseen data. The images underwent standard pre-processing steps to ensure compatibility with the SAM. During training, random cropping was applied to generate 1024 x 1024 image patches while preserving the local geometric characteristics of the biofilm structures. For validation and testing, images and their corresponding segmentation masks were resized to 1024 x 1024 pixels to meet the fixed input size requirements of the model. Although this transformation altered the original aspect ratio, it was applied identically to both images and masks, ensuring consistent spatial correspondence throughout model evaluation. To align with the model's pretraining, the images were normalised using ImageNet statistics (mean: [123.675, 116.28, 103.53]; standard deviation: [58.395, 57.12, 57.375]). Since SAM's backbone was trained on ImageNet, this normalisation standardised input intensities and enhanced performance across diverse datasets. Data augmentation techniques, including random cropping and flipping, were applied to improve model robustness across different biofilm structures and conditions. All preprocessing steps were implemented automatically using a standardised pipeline and did not involve manual image editing, segmentation or user-dependent parameter selection, ensuring reproducibility across the entire dataset.

2.4. Ground truth generation using SAM

Due to the lack of publicly available annotated datasets for biofilm OCT images, we employed SAM to generate initial biofilm masks. These masks were produced by manually defining the biofilm boundaries and were subsequently checked to ensure accuracy. The SAM-generated masks then served as the ground truth data for training our deep learning models.

SAM uses box-prompts to indicate regions of interest and, in our implementation, the generation of these prompts was automated. This automation streamlined the segmentation process by removing the need for manual input, thereby improving both efficiency and consistency, particularly when processing complex images.

The segmentation process began with the automated generation of box prompts, a key step in localising biofilm regions within OCT images. The first step was to identify the substrate-biofilm interface, determined by detecting the highest intensity pixel values along each vertical line in the image. These pixels typically correspond to the biofilm-substrate boundary. A regression model was then applied to these values to fit a line, providing an estimate of the substrate position and defining the lower boundary of the segmentation box. The upper boundary was

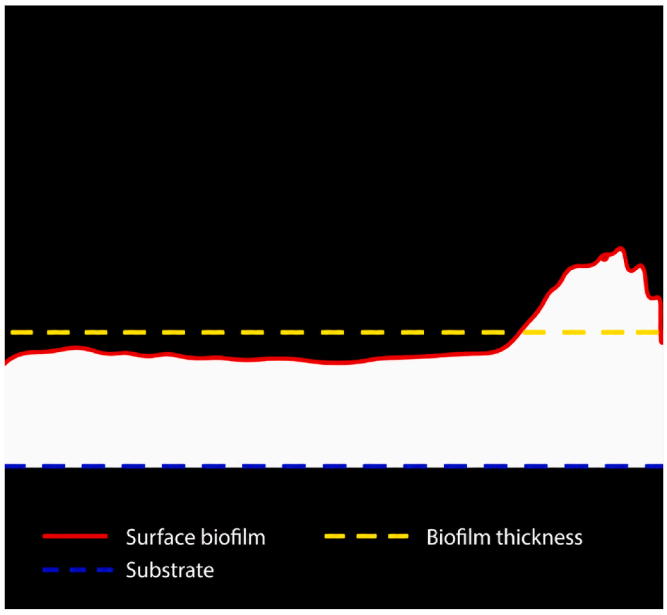


Fig. 4. Segmented OCT B-scan of a 48h-old laboratory-grown biofilm. The biofilm region is shown in white, and the background is black. The blue line indicates the lowest point from used for thickness measurements, the red line represents the highest points, and the yellow line corresponds to the thickness. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

defined by locating the nearest vertical line above the substrate in which all pixel values were at their lowest, thereby excluding floating or detached biofilm materials and isolating the region of interest. To accommodate natural variability in both substrate morphology and biofilm structure, a buffer zone was incorporated above and below the computed boundaries, ensuring that the segmentation box encompassed the full extent of relevant biofilm features.

To refine the input for the SAM model, each OCT image was thresholded using Otsu's method [26]. This enhanced the contrast between the biofilm and the background by binarising the image, thereby improving segmentation precision. The noise located beneath the detected substrate was then removed from the thresholded image. The thresholded image, together with the automatically calculated bounding box coordinates, was supplied to SAM, which generated a binary segmentation mask delineating the biofilm region. In this way, the segmentation process isolated the biofilm while distinguishing it from the substrate and surrounding background, with minimal manual intervention.

Our pipeline incorporated additional steps to validate and refine the generated bounding boxes, enhancing robustness against noise and variations in the image. To ensure comprehensive segmentation, particularly in cases where floating biofilm material was present, the algorithm recalculated the box coordinates for the region above the initially identified box. If a valid secondary box was detected in this area, it was included as an additional prompt for SAM. This enabled the segmentation model to account for detached or floating biofilm regions that were not connected to the main biofilm mass.

To create the training dataset, all 2800 images and their corresponding masks were visually reviewed to eliminate any incorrectly segmented outputs. Additionally, expert validation was conducted to ensure the accuracy and reliability of the masks. By automating the most labour-intense steps, the approach substantially reduced manual effort and enhanced reproducibility. In addition to segmentation, the masks were used to calculate key morphological characteristics of biofilms, including thickness, roughness (Ra and Rq), density, and density distribution (DD).

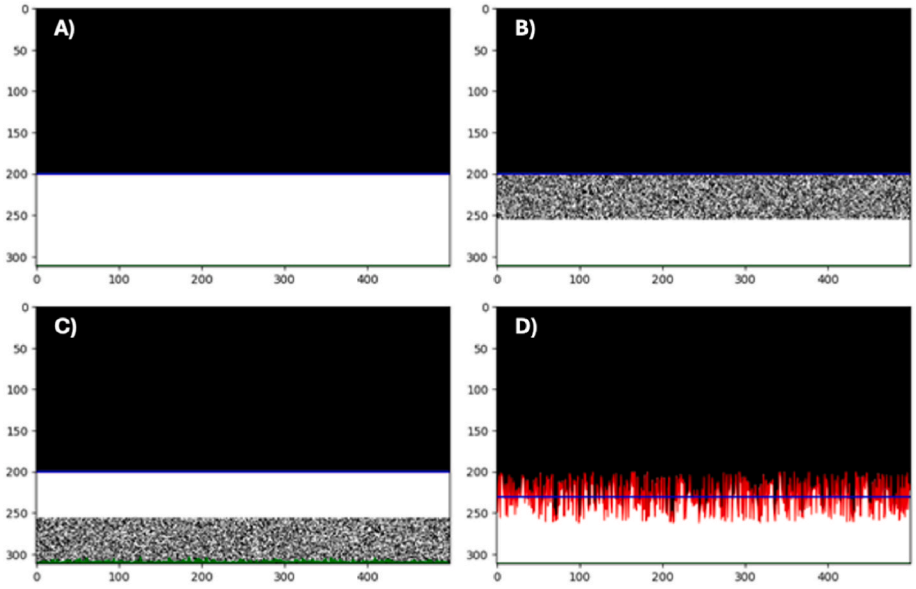


Fig. 5. Model biofilms illustrating different morphologies: A**)** A uniformly dense and smooth model biofilm. **B)** A biofilm in which the upper 50% (closer to the water interface) is less dense than the lower 50% (closer to the substrate). **C)** The reverse of B), with higher density in the upper 50%. **D)** A uniformly dense but heterogeneously rough model biofilm. The units on the X-axis and Y-axis are pixels.

2.5. Biofilm morphology

We defined the general physical characteristics of biofilms according to four traits outlined below, following the approached presented by Wagner and Horn [13].

2.5.1. Thickness (H)

Biofilm thickness is defined as the perpendicular distance between the substratum and the biofilm surface [13]. The substratum was identified as the lowest point of the biofilm mass, while the biofilm surface was defined as the highest point. Empirically, we determined the substratum as the lowest pixel and the biofilm surface as the highest pixel, as illustrated using the blue and red lines, respectively, in Fig. 4. Biofilm thickness was then computed as:

$$H = \frac{1}{n}\sum_{i=1}^{n} h_i \qquad \text{Equation 1}$$

where n represents the total number of thickness measurements, and h_i denotes the individual thickness measured at point i.

2.5.2. Roughness (Ra and Rq)

Biofilm roughness refers to the unevenness or texture of its surface. Empirically, roughness can be calculated as the average deviation of the biofilm surface (red line in Fig. 4) from the thickness (yellow line in Fig. 4) [27]. We quantified roughness using two standard surface roughness parameters: the arithmetic average roughness (Ra) and the root mean square roughness (Rq). Ra provides a straightforward measure of surface irregularities by averaging the absolute deviations from the thickness, whereas, Rq incorporates the square, making it a more sensitive metric that emphasises peaks and valleys in the biofilm structure. Biofilm roughnesses were computed as:

$$R_a = \frac{1}{n}\sum_{i=1}^{n} |Z_i - Z|$$
$$R_q = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Z_i - z)^2} \qquad \text{Equation 2}$$

where z_i represents the thickness at point i, $\bar{z}$ is the mean thickness, and n is the total number of points measured (the width of the measured section in pixels).

2.5.3. Density (D)

Conceptually, the density of a biofilm refers to the mass or number of microorganisms per unit volume or area within the biofilm matrix. OCT operates at a scale of magnification above individual bacterial cells, meaning that these cells are not visible. Consequently, we quantified biofilm density by analysing the number of pixels corresponding to biofilm biomass in high-resolution images. This approach enables differentiation between areas of high and low density within the biofilm without requiring cell counts.

Biofilm density was calculated using the following formula:

$$D = \frac{n_p}{A} \qquad \text{Equation 3}$$

where D is the biofilm density (pixels per unit area), n_p is the number of pixels representing the biofilm in the image, and A is the total area of the image.

2.5.4. Density distribution (DD)

While the overall density metric provides a basic measure of biofilm biomass in a cross-sectional image, it is also informative to understand how this mass is distributed within the biofilm structure. DD describes the special organisation of biofilm mass, providing insight into whether biofilm is concentrated near the surface (biofilm-liquid interface) or the substratum (biofilm-solid interface). This metric thus provides an understanding of the internal structure of the biofilm. Empirically, density distribution was calculated as the ratio of the density in the upper region of the biofilm to the overall biofilm density, as follows:

$$DD = \frac{D_{upper}}{D_{total}}$$

where D_{upper} is the density of the biofilm in the upper region (top half or top layer) and D_{total} is the overall density of the entire biofilm.

A DD value less than 0.5 suggests that the biofilm is more densely distributed near the substratum, with a larger proportion of biomass was concentrated closer to the solid interface. Conversely, a DD value greater than 0.5 indicates higher density near the liquid interface, suggesting

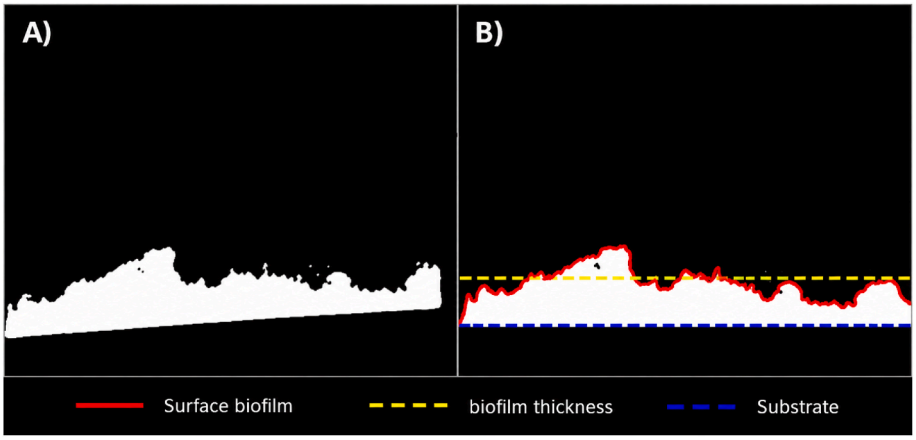


Fig. 6. A) An un-manipulated threshold mask of a biofilm in which the original image was not level at the point of capture. **B)** The adjusted mask with the biofilm aligned horizontally. The blue line indicates the lowest point from used thickness measurements, the red line represents the highest points, and the yellow line corresponds to the thickness. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

that more biomass is concentrated in the top layers of the biofilm.

2.6. Comparative analysis of biofilm metrics

Several simulations were conducted using artificial sample images to demonstrate how different metrics can characterise varying biofilm phenotypes. In these simulations, the biofilm area was represented by white pixels, while black pixels denoted voids or background. Although the samples in Fig. 5a–d have similar overall densities, their different morphologies could be distinguished using the roughness parameter, with Ra showing a value of 0 and Rq displaying a value of 0.19. This illustrated how the roughness metrics capture variations in surface texture. However, roughness alone is insufficient to describe internal morphology or differences in biofilm height. Thickness, measured as the vertical extent of the biofilm from the substratum to the water interface, is particularly useful for assessing the overall growth and the physical size of the biofilm.

In cases where internal structure is more relevant, the density metric provides valuable insight. For example, the density values of the samples in Fig. 5a and b are 1 and 0.75, respectively, highlighting differences in internal structure despite their similar overall appearances. However, density alone does not fully capture variations in biofilm structure. For instance, the samples in Fig. 5b and c have comparable density values (0.75 and 0.76), yet the spatial distribution of biomass differs markedly between them. In such cases, DD is a particularly informative metric: the sample in Fig. 5b has a DD of 0.34, indicating higher density nearer the substratum, whereas the sample in Fig. 5c, with a DD of 0.76, is denser towards the water interface. By combining these four metrics - density, roughness, thickness and DD - we obtain a comprehensive, comparative understanding of biofilm morphology, enabling rapid characterisation across different experimental scenarios.

2.7. Structural measurements

The input for the structural analysis consisted the binary mask obtained from the segmentation stage. This mask delineated biofilm pixels from the background, enabling a focused and precise quantification of biofilm structure.

Noise removal and smoothing was conducted first. Although the segmentation procedure already isolated the biofilm region, OCT images contain noise as well as light scattering and refraction artifacts. To improve the quality of the binary mask, small, isolated regions identified as artifacts or noise were eliminated; specifically, any connected component comprising fewer than 30 pixels was discarded. Additionally, outlier data points in substrate height measurements were corrected by replacing them with the nearest valid measurements. These steps minimised noise and ensured a more reliable representation of the substrate profile.

Once the mask was cleaned and smoothed, the substrate height was calculated. The substrate served as the foundational reference point for measuring the thickness of the biofilm. This was determined by analysing the binary mask to identify the lowest non-zero pixel in each column, corresponding to the substrate level. This initial estimate was then refined by correcting noisy values to ensure that the resulting height profile accurately represented the true substrate interface.

After determining the substrate height, linear regression was applied to smooth the substrate profile. This step captured the general trend of substrate elevation while reducing random fluctuations caused by noise. The resulting slope provided an estimate of the substrate's orientation. If the regression indicated that the substrate was not level, the mask was rotated to correct for this inclination, ensuring that subsequent analyses accurately reflected the true biofilm's geometry (Fig. 6).

Finally, the preprocessing phase concluded with the recalibration of surface and substrate heights based on the adjusted mask. The height of the biofilm surface was derived by locating the highest pixel values within the mask, while the substrate height was refined using the predictions from the linear regression model. This final adjustment ensured that all subsequent calculations, such as those for thickness, roughness and DD, were based on a representative reconstruction of the biofilm structure.

2.8. Deep learning models

Two deep learning models, SegFormer [28] and YOLOv8 [29], were employed for biofilm segmentation. SegFormer features a transformer-based architecture that captures complex patterns and long-range dependencies, making it particularly well-suited for analysing noisy or low-contrast images. In contrast, YOLOv8 is designed for real-time object detection, enabling rapid identification and localisation of biofilm instances within an image.

The dataset consisted of the 2800 images and their corresponding segmentation masks generated and validated using the SAM-assisted annotation pipeline. During training, random 1024 × 1024 pixel crops were extracted from the images to preserve local biofilm features while providing a consistent input size. For validation and testing, images and their corresponding segmentation masks were handled as above, including resizing, normalisation (ImageNet) and division into training and testing subsets (80:20). Data augmentation techniques including random cropping and flipping were applied to improve model generalisation and robustness against variations in imaging conditions.

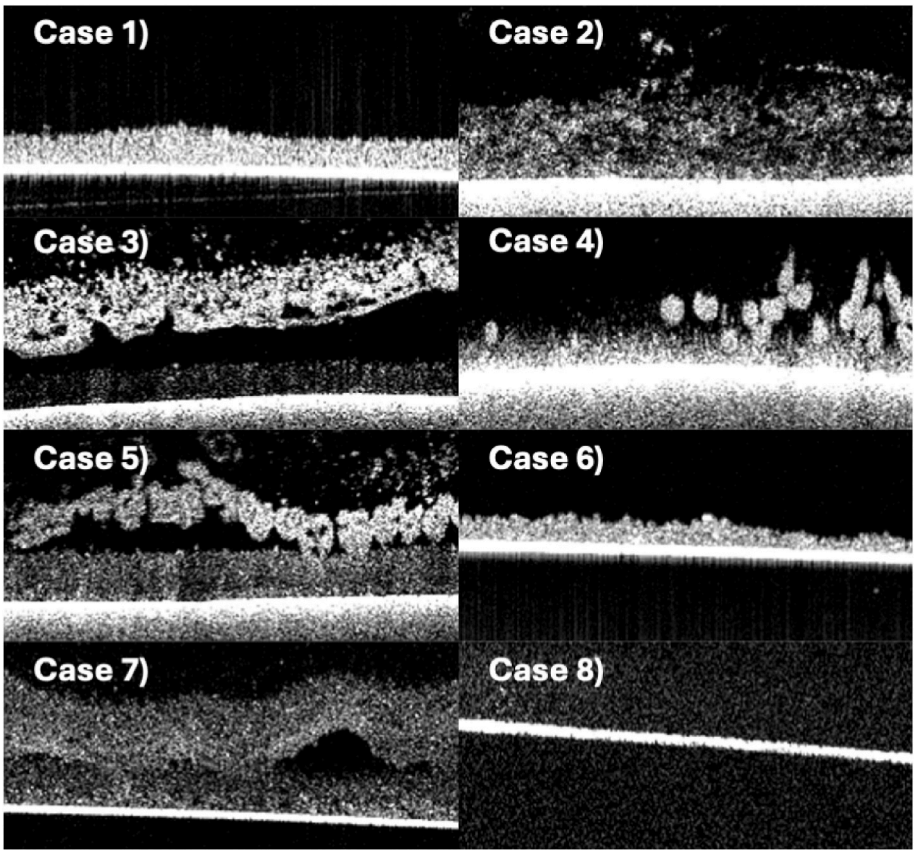


Fig. 7. Examples of eight broad categories of biofilm morphology that posed particular challenges for the automated analysis pipeline: **Case 1)** Stable biofilm, characterised by a continuous surface with uniform thickness. **Case 2)** Low contrast biofilm, exhibiting poor differentiation from the background. **Case 3)** Drifting biofilm, which may be largely detached from the substratum. **Case 4)** Blurry biofilm, with a poorly defined upper surface due to low edge sharpness or scattering effects. **Case 5)** Fragmented biofilm, presenting multiple apparently floating areas in cross section, resulting in a porous three-dimensional structure. **Case 6)** Thin biofilm, comprising only a few layers of cells on the substratum. **Case 7)** Void biofilms, relatively uniform but containing large, clearly defined voids. **Case 8)** Non-growing biofilm, in which little or no extracellular polymeric substance (EPS) is present. All Images in this figure have been contrast-enhanced to accentuate the described features.

Both models were trained for 50 epochs using the Adam optimiser with an initial learning rate of 0.0001 and a batch size of 16. Hyperparameter selection was performed experimentally by monitoring validation performance and convergence behaviour during training. The learning rate, batch size, and optimiser settings were adjusted iteratively to achieve stable convergence while minimising overfitting. A cross-entropy loss function was used to guide the training process for segmentation accuracy.

Model performance was evaluated using accuracy and intersection over union (IoU), which quantify the models' effectiveness in accurately segmenting biofilm structures. Additionally, overlays of predicted masks on original images were generated to support qualitative assessments of model performance. The complete analysis pipeline is available as a user-friendly command-line package on GitHub: https://github.com/ParastooAzizinezhad/BiofilmOCT.

2.9. Robustness testing on edge cases

To further assess the robustness of the trained models, we manually curated 302 images from the original dataset of 2800, which were categorised into eight biofilm edge cases (ambiguous or particularly challenging acquisitions or morphologies) through consensus discussion. Each case included multiple images representing particularly challenging biofilm scenarios such as low contrast, blurry images or fragmented structures (Fig. 7). This edge case dataset was used to evaluate how well the models generalised to unusual or difficult biofilm structures and did not influence the segmentation or subsequent quantitative measurements. All computational metrics were derived uniformly from the binary segmentation output, without conditioning on morphological class or structural interpretation. The categorisation of images into edge cases served only to enable qualitative and comparative assessment of model performance.

A qualitative evaluation was conducted by four biofilm researchers, who reviewed the model outputs on the edge case dataset. Each expert assessed the segmentation masks produced by the two models and selected the model they considered most biologically accurate for each biofilm type. Experts were not asked to categorise images into edge cases.

2.10. Implementation

The biofilm image processing was conducted using a custom-developed, Python-based command-line tool. This tool facilitates batch processing by accepting a directory of OCT images as input, thereby enabling the efficient handling of large datasets. Users are required to specify both the input directory, which contains the OCT images, and the output directory where the results will be stored.

The software offers flexibility in analysis through the selection of different models, including YOLOv8, SegFormer and the SAM, which can be specified via command-line arguments. Upon execution, the program processes the images and saves the outputs in a structured manner that mirrors the input folder hierarchy. The system also saves the raw OCT images in PNG format and generates corresponding mask outputs based on the selected model. It calculates all biofilm metrics and

Table 2
Performance of the two models in terms of accuracy and Intersection over Union (IoU).

Model Name	Accuracy	IoU
SegFormer	0.9695	0.8728
YOLOv8	0.9857	0.9077

records these in a CSV file for each subfolder or experiment.

The biofilm image processing tool relies on several Python libraries to facilitate deep learning-based analysis, image processing and data management. At its core, the tool uses Torch [30] for deep learning model execution, with support for CUDA acceleration when available, significantly enhancing processing speed on compatible GPUs. Ultralytics and transformers [31] are used to implement YOLOv8 [29] and SegFormer [21] models, while Segment Anything enables the use of the SAM [22] for zero-shot segmentation. Image processing operations, such as filtering and pre-processing, are handled by scikit-image [32], OpenCV (opencv-python) [33], scikit-learn [34] and Pillow. NumPy [35] and pandas [36] provide data handling capabilities, allowing for structured storage and analysis of extracted biofilm metrics. Additionally, the system incorporates matplotlib, numba, xmltodict, and ipywidgets to enhance various aspects of the project, including data visualisation, performance optimisation, metadata parsing and interactive components.

3. Results

3.1. Model performance on the full dataset

The results of the model performance on the full 2800 image dataset are summarised in Table 2. During training, both models achieved excellent convergence, with accuracy and Intersection over Union (IoU) values exceeding 99%, indicating strong learning of biofilm features present in the dataset. For the final evaluation, the YOLOv8 model achieved a higher accuracy of 0.9857 and IoU of 0.9077 compared to the SegFormer model, which recorded an accuracy of 0.9695 and IoU of 0.8728. This indicates that while both models perform well at segmenting biofilms, YOLOv8 demonstrates slightly higher accuracy and segmentation overlap.

Both the SegFormer and YOLOv8 models successfully predicted segmentation masks for the biofilm sample. As shown in Fig. 8, the mask generated by the SegFormer model captured more intricate details compared to the one produced by the YOLOv8 model, suggesting a finer segmentation capability. The morphology metrics calculated from these masks were also comparable. Specifically, the biofilm thickness measured from the SegFormer mask in Fig. 8 was 162 pixels (approximately 316 μm), while the YOLOv8 model estimated a thickness of 168 pixels (around 326 μm). Additionally, the roughness coefficient (Ra) values derived from the segmentation masks were 0.13 and 0.15 for SegFormer and YOLOv8 respectively. Both models provided reliable segmentation outcomes, with SegFormer offering slightly more detailed masks which may contribute to more precise morphological measurements of biofilm structures.

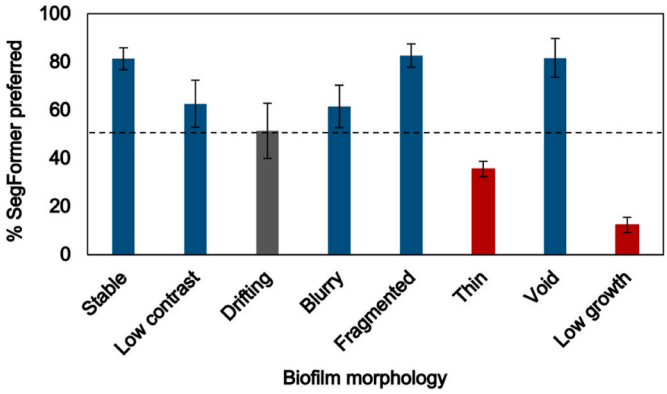


Fig. 9. Comparison of SegFormer performance across several biofilm morphologies. The bar chart shows the percentage of times SegFormer was preferred over another method for segmentation, for different biofilm morphologies. The dashed line at 50% indicates the baseline preference level. Error bars represent the standard error of the mean. Blue bars indicate a strong preference for SegFormer, the grey bar indicates neutral preference and the red bars indicate a low preferences. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

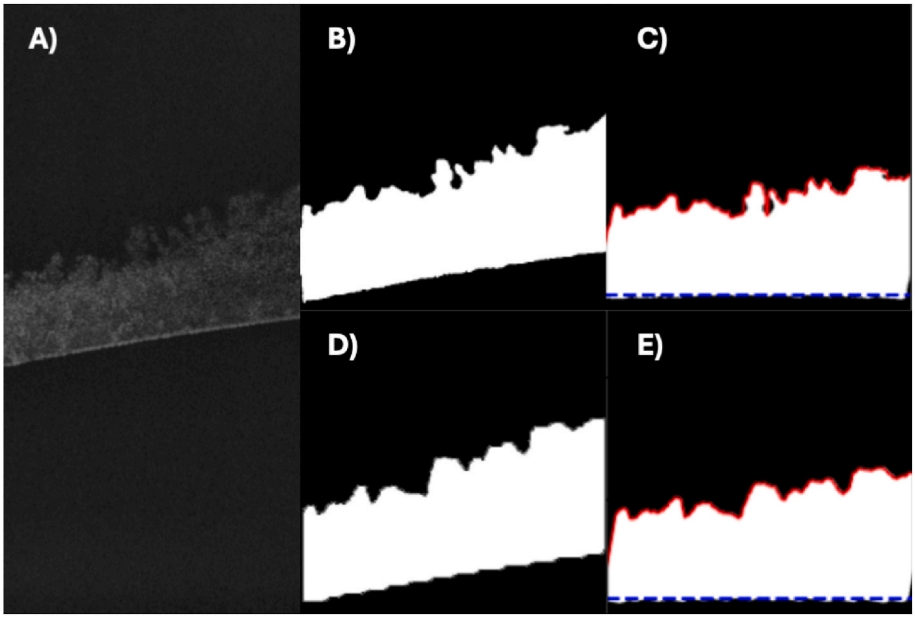


Fig. 8. A) An unmanipulated OCT B-scan of a 48h-old laboratory-grown biofilm. **B)** Initial segmentation mask and **C)** corrected mask with upper (red) and lower (blue) boundaties of the biofilm defined by SegFormer. **D)** A corrected mask and **E)** upper (red) and lower (blue) bounds of the biofilm defined by YOLOv8. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

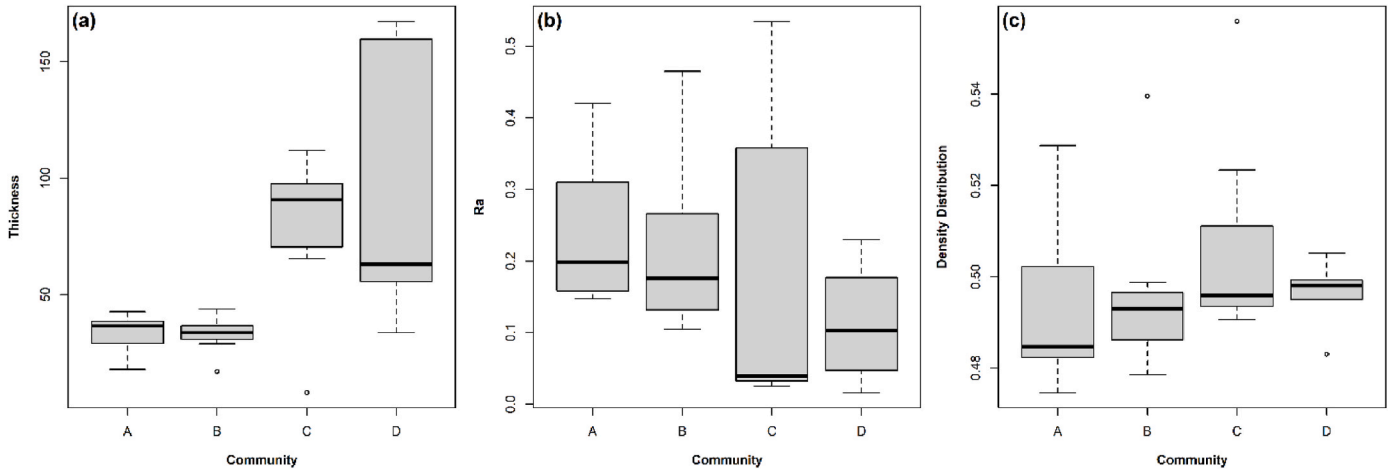


Fig. 10. Morphological and structural characteristics of four distinct biofilms communities (A, B, C, and D). The box- and-whisker plots illustrate the median (horizontal line), interquartile range (box), and range (whiskers) for four parameters: **A)** Thickness (micrometres); **B)** Surface Roughness (Ra); **C)** Density Distribution (DD). Outliers are represented by open circles. Communities C and D biofilms were thicker than those of communities A and B (Post-hoc Dunn tests: $p < 0.03$ unadjusted).

3.2. Robustness testing on edge cases

To evaluate the robustness of the trained models, we tested their performance on a curated set of eight edge cases representing challenging biofilm conditions. Four experts were asked to review the segmentation results for each of the cases and select the model that produced the most accurate biofilm mask. The experts' selections made are summarised in Fig. 9.

The evaluation of the segmentation performance demonstrated a strong dependence on the biofilm morphology. SegFormer exhibited a high preference, achieving over 60% agreement for five out of the eight categories, while YOLOv8 was preferred in two. SegFormer demonstrated particular strength in stable biofilms (81.4% agreement) and fragmented biofilms (82.8% agreement). In contrast, YOLOv8 performed better with thin biofilms (64.4% agreement) and non-growing biofilms (76.4% agreement). Both models were equally effective at segmenting drifting biofilms. Although there was variability in the agreement between the models and the expert evaluators across all cases, there was also some disagreement among individual experts. The lowest inter-expert variability was observed for stable and fragmented biofilms segmented using SegFormer, whereas highest variability occurred with drifting biofilms, where expert opinions diverged regarding the most accurate. Overall, the most consistent expert agreement was observed for stable biofilms (SegFormer: 81.4%), fragmented biofilms (SegFormer: 82.8%), and thin biofilms (YOLOv8: 64.4%). In summary, SegFormer was considered superior in most cases, while YOLOv8 maintained strength in specific biofilm morphologies.

3.3. Experimental validation

We demonstrated the utility of the analysis pipeline in an experimental context, by using it to compare the morphologies of biofilms produced by four taxonomically distinct biofilm communities (A, B, C and D; Table 1). For each community seven OCT images were taken at random points across the biofilm (one biofilm replicate per community, seven pseudo replicates), and the analysis pipeline was applied to extract the morphology metrics (Fig. 10). Statistical analysis was carried out using R [37]. The most pronounced difference between biofilm morphologies was identified by the thickness metric (Kruskal-Wallis: $\chi 2 = 12.374$; $p = 0.006$). Communities C and D produced biofilms that were significantly thicker and more voluminous than those produced by communities A and B (Dunn tests: $p < 0.03$ unadjusted; all pairwise comparisons A-C, A-D, B–C, B-D. Dunn test from FSA package [38]. Visually, community D biofilms appeared far more variable in thickness compared to those from community C, however, further replication would be required to confirm this observation (see Fig. 10).

Since multiple morphological features were measured, a Principal Components Analysis (PCA) was conducted on the data produced by the pipline (metrics: Thickness, Ra, Rq, Volume, Density Distribution, Pixel-based Density). This was to establish if one or two morphology metrics (extracted principal components) could describe the biofilm morphology. Data used were extracted from seven randomly placed images across the biofilm for communities A, B C and D as described above (see also methods and Table 1), without "community" as an identifier. The PCA indicated that the four communities (A-D) broadly clustered into two groups (communities A & B, and communities C & D, see Fig. 11) based on their combined morphology metrics. Communities A & B can be separated from communities C & D on the second principal component score (Wilcoxon Rank-Sum test: $W = 28$, $p < 0.001$). PC2 is associated with thickness and density metrics. PC1 is most highly associated with roughness metrics; PC1 score for group A & B was not significantly different to PC1 score for group C & D (Wilcoxon Rank-Sum test: $W = 136$, $p = 0.08$).

4. Discussion

Methods that allow rapid, high-throughput analysis of biofilm morphology are currently lacking, yet they are essential across many industries where biofilms pose operational challenges. Being able to predict biofilm morphologies from known taxonomic or environmental parameters would enable more accurate modelling of biofilm impacts in specific industrial context. At a fundamental level, large-scale experimental datasets are required to advance our understanding of the factors that shape biofilm morphology. However, the absence of tools capable of generating and processing biofilm morphology data rapidly and at scale continues to hinder progress in understanding community dynamics. Developing high-throughput imaging and analysis pipelines is therefore critical to meeting both applied and fundamental research needs.

Our work presents a bottom-up approach to developing a unified software tool that enables users to rapidly extract meaningful and comparable metrics of biofilm growth from OCT images. The purpose was not to capture fine-scale morphology or to associate specific structural features with particular taxa, but rather to provide general descriptors that encompass a large proportion of the morphological variability observed across conditions, and to ground-truth these descriptors as a basis for comparing treatments. To achieve this, we

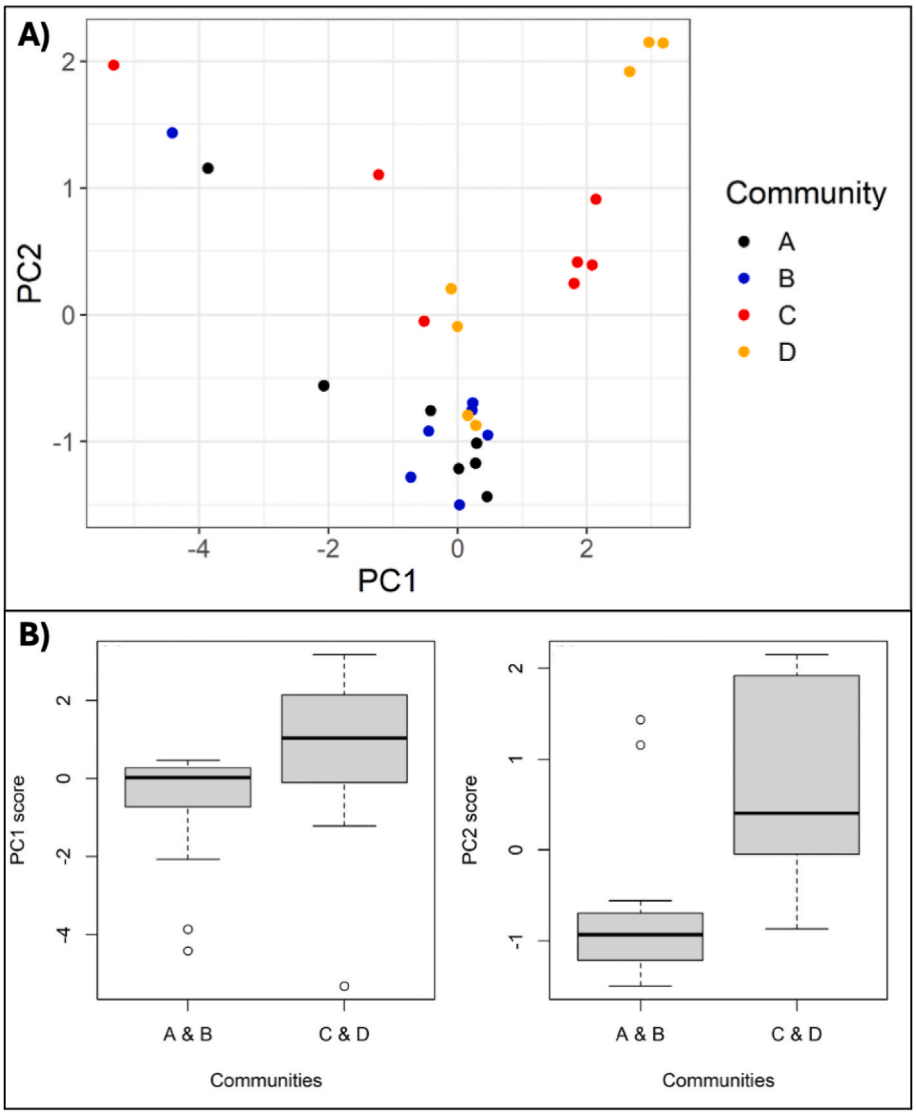


Fig. 11. A) Principal Components Analysis (PCA) ordination plot illustrating the separation of the four biofilm communities. The plot visualises the variance in the biofilm characteristics across the four communities (A, B, C, and D) in a two-dimensional space. The axes represent the first two components (PC1 and PC2), which account for the maximum variance in the input data. Each point represents a replicate, coloured according to its community designation. Communities A & B (grouped together) have significantly different PC2 scores than communities C & D (grouped together) (Wilcoxon Rank-Sum test: $W = 28$, $p < 0.001$), indicating the communities in the groups differ most in thickness and density. **B)** Box and whisker plots illustrating the difference in PC2 scores between communities A & B and C & D.

establish a development pipeline involving laboratory growth of a large number of model biofilms, a reproducible and computationally light-weight imaging protocol and the identification of four simple yet informative morphology metrics. These metrics were implemented on regions classified as biofilm by two machine learning models. The full workflow was integrated into a single software package, which is freely available to the community. A PCA (Fig. 11) indicated that four chosen bacterial communities clustered broadly into two groups based on this method: communities A and B, and communities C and D. This pattern was consistent with the high-level differences observed in Fig. 10, but also illustrates the value of integrating multiple metrics into a single multivariate analysis. Notably, communities A and B share two of their three members in (Table 1), which may contribute to their more similar biofilm morphologies and the tighter clustering observed in Fig. 11 across all combined metrics.

The results from training the machine learning models demonstrated that both SegFormer and YOLOv8 achieved high accuracy and IoU scores, with YOLOv8 slightly outperforming SegFormer in both metrics. YOLOv8 achieved an accuracy of 0.9857 and IoU of 0.9077, compared to SegFormer's accuracy of 0.9695 and IoU 0.8728. However, when examining edge cases the results presented a more nuanced pattern. In cases of particularly challenging biofilm structures, the superiority of the different models depended heavily on the specific biofilm characteristics, indicating that neither model consistently outperformed the other across all conditions. Thus, both are included as options in our published software.

4.1. *SegFormer's strength in stable and fragmented biofilms*

Expert evaluators (Fig. 9) identified SegFormer as the most accurate approach in cases involving more complex biofilm structure, such as stable, fragmented and void biofilms. Across these scenarios, SegFormer was selected on more than 70% of evaluations, indicating a consistent preference among experts. This suggests that SegFormer's pixel-wise segmentation approach is particularly well suited for scenarios where biofilm boundaries are clearly defined and morphological features are

more consistent. In edge cases involving low-contrast, drifting and blurry biofilms, SegFormer also demonstrated a slight advantage over YOLOv8. These cases typically pose challenges for segmentation due to poor visual distinction or motion artifacts. SegFormer's transformer-based architecture, which captures global contextual information, likely contributes to its robustness under visually ambiguous or noisy conditions.

4.2. YOLOv8's superior performance in complex biofilms

Conversely, YOLOv8 showed greater strength in handling ambiguous or challenging biofilm cases, particularly thin and non-growing biofilms. In these cases, YOLOv8 was preferred in more than 60% of evaluations, with particular dominance in the non-growing biofilm category. This pattern suggests that YOLOv8's object detection-based architecture may generalise more effectively to biofilms with complex, or poorly defined boundaries, enabling it to outperform SegFormer in these more dynamic and variable scenarios.

While there is growing interest in automating OCT-based biofilm analysis, previous approaches have relied on manual or semi-automated segmentation techniques that can be time-consuming, operator-dependent and difficult to apply consistently across large datasets [14]. More recent studies using machine-learning and image-processing workflows relied upon supervised classifiers and human-defined image features [16,17]. While these demonstrate the feasibility of automated biofilm characterisation, the availability of large annotated datasets for training and evaluating deep-learning models is limited. By integrating the Segment Anything Model SAM as an annotation aid we were able to generate a large ground-truth dataset for model training and remove one of the key barriers to the broader adoption of deep-learning methods in OCT biofilm analysis. Further, while previous studies have focused on segmentation or classification [39], the workflow presented here enables direct extraction of biologically relevant parameters such as biofilm thickness, density and density distribution from model-generated segmentation masks. This provides a practical route toward high-throughput and reproducible quantitative biofilm analysis using OCT imaging.

5. Conclusions

In this study, we successfully developed and validated a robust, automated deep learning pipeline for the rapid morphological characterisation of biofilms using OCT images. The pipeline reliably extracts four key morphological metrics (thickness, roughness, and density distribution) which collectively enable comprehensive and quantitative assessment of complex biofilm structures.

The capability presented here opens opportunities for large-scale combinatorial studies that explore biofilm growth across multiple conditions and very high levels of replication. Once the imaging steps, which can themselves be automated, are completed, the processing and analysis of thousands of images can be carry out within hours and without any user intervention, particularly when supported by dedicated graphical processing units or a high-performance computing cluster. The availability of datasets of this scale will, in turn, support modelling approaches that have previously been constrained by limited data. Such models will underpin a more comprehensive understanding of the factors shaping biofilm morphology, their interactions and, and their consequences for industrial processes.

Funding

The team thank the Defence Advanced Research Projects Agency (award HR0011-23-2-0008) and the Office of Naval Research (awards N00014-23-1-2177 & N00014-25-1-2529) for supporting the research described here.

CRediT authorship contribution statement

R.I. Heroza: Investigation, Methodology, Software, Writing – original draft. **P. Azizinezhad:** Formal analysis, Investigation, Methodology, Software, Writing – original draft. **K.-A. Moss:** Data curation, Investigation, Methodology. **H.M. Lewis:** Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **P. Pajewska:** Investigation, Methodology. **A.G. Seco de Herrera:** Conceptualization, Writing – original draft, Supervision, Funding Acquisition. **A.J. Dumbrell:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **P.P. Laissue:** Conceptualization, Formal analysis, Resources, Supervision, Writing – original draft. **N. Aldred:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Nick Aldred reports financial support was provided by Defense Advanced Research Projects Agency. Nick Aldred reports financial support was provided by Office of Naval Research. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The team thanks Prof. Mette Burmølle and Dr Cristina Amador of the University of Copenhagen, and Prof. David Karig of Clemson University for providing the bacterial strains used in this study.

Data availability

Data will be made available on request.

References

[1] Flemming HC, Wingender J. The biofilm matrix. Nat Rev Microbiol 2010;8:623–33. https://doi.org/10.1038/nrmicro2415.
[2] Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: an emergent form of bacterial life. Nat Rev Microbiol 2016;14:563–75. https://doi.org/10.1038/nrmicro.2016.94.
[3] Steinberg N, Kolodkin-Gal I. The Matrix reloaded: how sensing the extracellular Matrix synchronizes bacterial communities. J Bacteriol 2015;197:2092. https://doi.org/10.1128/JB.02516-14.
[4] Burmølle M, Webb JS, Rao D, Hansen LH, Sørensen SJ, Kjelleberg S. Enhanced biofilm formation and increased resistance to antimicrobial agents and bacterial invasion are caused by synergistic interactions in multispecies biofilms. Appl Environ Microbiol 2006;72:3916–23. https://doi.org/10.1128/AEM.03022-05.
[5] Puentes-Cala E, Tapia-Perdomo V, Espinosa-Valbuena D, Reyes-Reyes M, Quintero-Santander D, Vasquez-Dallos S, Salazar H, Santamaría-Galvis P, Silva-Rodríguez R, Castillo-Villamizar G. Microbiologically influenced corrosion: the gap in the field. Front Environ Sci 2022;10:924842. https://doi.org/10.3389/FENVS.2022.924842.
[6] Dang H, Lovell CR. Microbial surface colonization and biofilm development in marine environments. Microbiol Mol Biol Rev MMBR 2015;80:91–138. https://doi.org/10.1128/MMBR.00037-15.
[7] Amador CI, Fatima N, Jakobsen ASS, Macario L, Pichon P, Aldred N, Burmølle M. Directed assembly of biofilm communities for marine biofouling prevention. Appl Environ Microbiol 2025;91:e01392. https://doi.org/10.1128/aem.01392-25. 25.
[8] Li M, Jiang Y, Zhang Y, Zhu H. Medical image analysis using deep learning algorithms. Front Public Health 2023;11:1273253. https://doi.org/10.3389/FPUBH.2023.1273253.
[9] Hoogstraten SWG van, Kuik C, Arts JJC, Cillero-Pastor B. Molecular imaging of bacterial biofilms—a systematic review. Crit Rev Microbiol 2023;50:971. https://doi.org/10.1080/1040841X.2023.2223704.
[10] Barbosa A, Miranda S, Azevedo NF, Cerqueira L, Azevedo AS. Imaging biofilms using fluorescence in situ hybridization: seeing is believing. Front Cell Infect Microbiol 2023;13:1195803. https://doi.org/10.3389/FCIMB.2023.1195803.
[11] Mhade S, Kaushik KS. Tools of the trade: image analysis programs for confocal laser-scanning microscopy studies of biofilms and considerations for their use by experimental researchers. ACS Omega 2023;8:20163–77. https://doi.org/10.1021/acsomega.2c07255.

[12] Wagner M, Taherzadeh D, Haisch C, Horn H. Investigation of the mesoscale structure and volumetric features of biofilms using optical coherence tomography. Biotechnol Bioeng 2010;107:844–53. https://doi.org/10.1002/BIT.22864.
[13] Wagner M, Horn H. Optical coherence tomography in biofilm research: a comprehensive review. Biotechnol Bioeng 2017;114:1386–402. https://doi.org/10.1002/BIT.26283.
[14] Fabbri S, Dennington SP, Price C, Stoodley P, Longyear J. A marine biofilm flow cell for in situ screening marine fouling control coatings using optical coherence tomography. Ocean Eng 2018;170:321–8. https://doi.org/10.1016/J.OCEANENG.2018.10.030.
[15] Hartmann R, Jeckel H, Jelli E, Singh PK, Vaidya S, Bayer M, Rode DKH, Vidakovic L, Díaz-Pascual F, Fong JCN, Dragoš A, Lamprecht O, Thöming JG, Netter N, Häussler S, Nadell CD, Sourjik V, Kovács ÁT, Yildiz FH, Drescher K. Quantitative image analysis of microbial communities with BiofilmQ. Nat Microbiol 2021;6(2):151–6. https://doi.org/10.1038/s41564-020-00817-4. 2021; 6.
[16] Zaki FR, Monroy GL, Shi J, Sudhir K, Boppart SA. Texture-based speciation of otitis media-related bacterial biofilms from optical coherence tomography images using supervised classification. J Biophot 2024;17:e202400075. https://doi.org/10.1002/jbio.202400075.
[17] Narciso DAC, Pereira A, Dias NO, Monteiro M, Melo LF, Martins FG. 3D Optical Coherence Tomography image processing in BISCAP: characterization of biofilm structure and properties. Bioinformatics 2024;40. https://doi.org/10.1093/bioinformatics/btae041. btae041.
[18] Won J, Hong W, Khampang P, Spillman DR, Marshall S, Yan K, Porter RG, Novak MA, Kerschner JE, Boppart SA. Longitudinal optical coherence tomography to visualize the in vivo response of middle ear biofilms to antibiotic therapy. Sci Rep 2021;11:5176. https://doi.org/10.1038/s41598-021-84543-9.
[19] Yue X, Qi K, Na X, Zhang Y, Liu Y, Liu C. Improved YOLOv8-Seg network for instance segmentation of healthy and diseased tomato plants in the growth stage. Agric For 2023;13:1643. https://doi.org/10.3390/AGRICULTURE13081643. 2023 Vol 13 Page 1643.
[20] Kirillov A, Mintun E, Ravi N, Mao H, Rolland C, Gustafson L, Xiao T, Whitehead S, Berg AC, Lo WY, Dollár P, Girshick R. Segment anything. Proc IEEE Int Conf Comput Vis 2023:3992–4003. https://doi.org/10.1109/ICCV51070.2023.00371.
[21] Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P. SegFormer: simple and efficient design for semantic segmentation with transformers. Adv Neural Inf Process Syst 2021;15:12077–90. https://doi.org/10.48550/arXiv.2105.15203.
[22] Yang M, Lu J, Kim H-S. SAM-guided Pseudo label enhancement for multi-modal 3D semantic segmentation. 2025.
[23] Osco LP, Wu Q, Lemos EL de, Gonçalves WN, Ramos APM, Li J, Marcato J. The Segment Anything Model (SAM) for remote sensing applications: from zero to one shot. Int J Appl Earth Obs Geoinformation 2023;124:103540. https://doi.org/10.1016/J.JAG.2023.103540.
[24] Wang J, Tabassum N, Toma TT, Wang Y, Gahlmann A, Acton ST. 3D GAN image synthesis and dataset quality assessment for bacterial biofilm. Bioinformatics 2022; 38:4598–604. https://doi.org/10.1093/BIOINFORMATICS/BTAC529.
[25] Chaulagain D, Shamabadi NS, Leslie SA, Karig DK. From natural microbe screening to sustained chitinase activity in exogenous hosts. ACS Synth Biol 2024;13: 1165–76. https://doi.org/10.1021/acssynbio.3c00637.
[26] Otsu N. Threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–6. https://doi.org/10.1109/TSMC.1979.4310076. SMC.
[27] Adams T, Grant C, Watson H. A simple algorithm to relate measured surface roughness to equivalent sand-grain roughness. Int J Mech Eng Mechatron 2012. https://doi.org/10.11159/IJMEM.2012.008.
[28] Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P. SegFormer: simple and efficient design for semantic segmentation with transformers. https://doi.org/10.48550/ARXIV.2105.15203; 2021.
[29] Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. 2023.
[30] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, Killeen T, Lin Z, Gimelshein N, Antiga L, Desmaison A, Köpf A, Yang E, DeVito Z, Raison M, Tejani A, Chilamkurthy S, Steiner B, Fang L, Bai J, Chintala S. PyTorch: an imperative style, high-performance deep learning Library. https://doi.org/10.48550/ARXIV.1912.01703; 2019.
[31] Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, Cistac P, Rault T, Louf R, Funtowicz M, Davison J, Shleifer S, Von Platen P, Ma C, Jernite Y, Plu J, Xu C, Le Scao T, Gugger S, Drame M, Lhoest Q, Rush A. Transformers: State-of-the-Art natural Language processing. Proc. 2020 conf. Empir. Methods nat. Lang. Process. Syst. Demonstr. Association for Computational Linguistics; 2020. p. 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6. Online.
[32] Van Der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T. scikit-image: image processing in Python. PeerJ 2014;2:e453. https://doi.org/10.7717/peerj.453.
[33] Bradski G. The OpenCV Library. Dr Dobb's J Softw Tools 2000.
[34] Pedregosa F, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2011;12:2825–30.
[35] Harris CR, Millman KJ, Van Der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, Van Kerkwijk MH, Brett M, Haldane A, Del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE. Array programming with NumPy. Nature 2020;585:357–62. https://doi.org/10.1038/s41586-020-2649-2.
[36] McKinney W. Data structures for statistical computing in python. 2010. p. 56–61. https://doi.org/10.25080/Majora-92bf1922-00a. Austin, Texas.
[37] R Core Team. R: a Language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2023. https://www.R-project.org/.
[38] Ogle DH, Doll JC, Wheeler AP, Dinno A. FSA: Simple Fisheries Stock Assessment Methods (R package version 0.10.1). https://CRAN.R-project.org/package=FSA; 2026.
[39] Yerly J, Hu Y, Jones SM, Martinuzzi RJ. A two-step procedure for automatic and accurate segmentation of volumetric CLSM biofilm images. J Microbiol Methods 2007;70(3):424–33. https://doi.org/10.1016/j.mimet.2007.05.022.