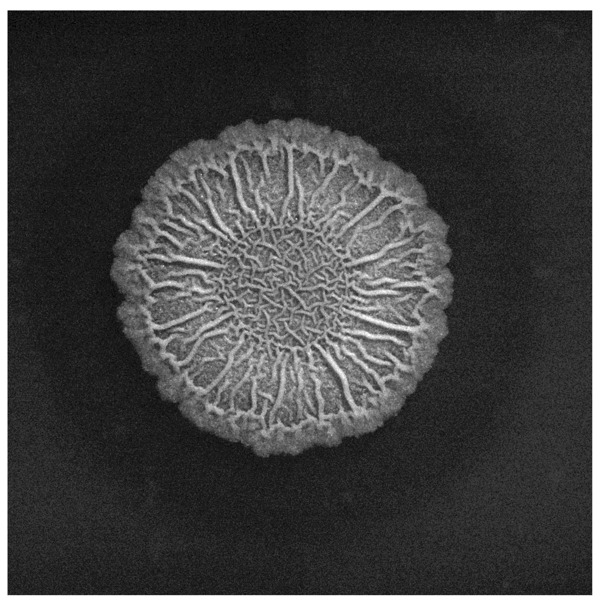
MACHINE LEARNING ANALYSIS OF BIOFILM WRINKLE MORPHOLOGY

BRITT DEWING, MOE VALI AND DIANA FUSCO

UNIVERSITY OF CAMBRIDGE

LAIDLAW FOUNDATION

Background: biofilms - bacterial communities with remarkable qualities



- Biofilms are bacterial communities characterised by the production of a protective extracellular matrix
- They are of interest because they cause chronic diseases in humans and are resistant to antimicrobials, including antibiotics.
- Their phenotypic heterogeneity as well as their complex wrinkle morphologies provide them with a fitness advantage

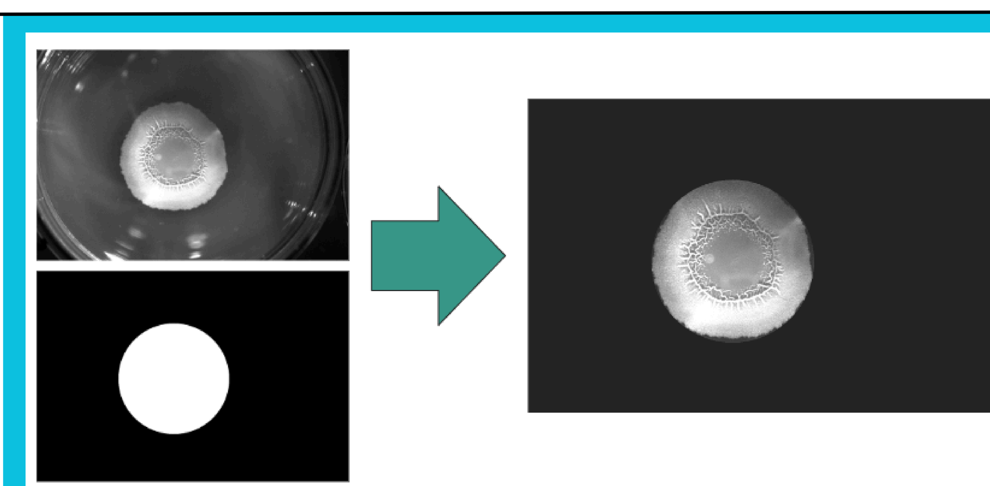
Fig 1. Brightfield image of radial biofilm colony

- In the biofilms formed by species *Bacillus Subtilis* cells are one of two mutually exclusive phenotypes: 'matrix' or 'motile' (Fig 2)

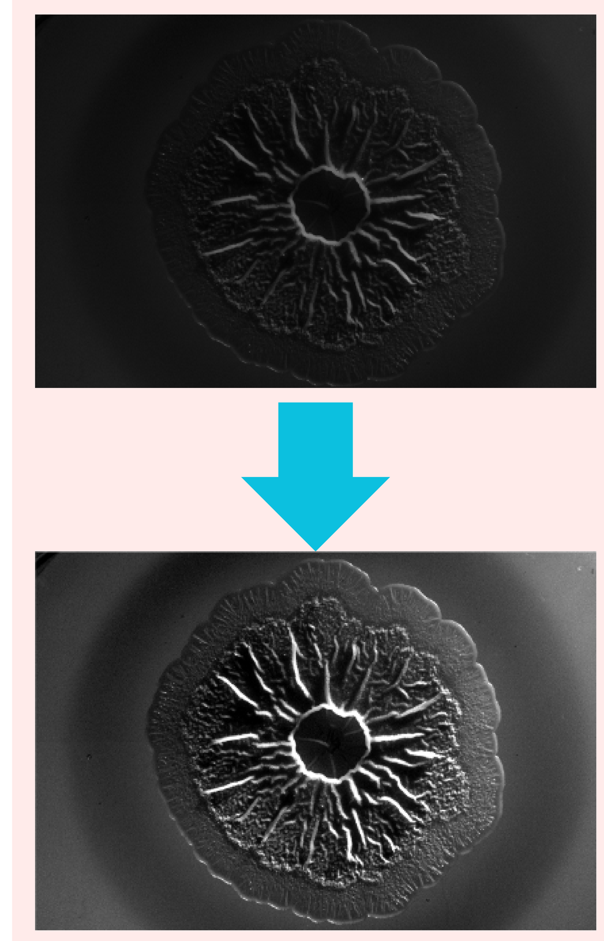
Fig 2. Dual reporter images of radial biofilm colony depicting the distribution of matrix and motile phenotypes

- It is believed that phenotype differentiation and biofilm wrinkle morphology are intrinsically related
- In order to understand how phenotypic variation and wrinkle morphology together may provide a fitness advantage to the biofilm, it is useful to isolate the wrinkles for analysis

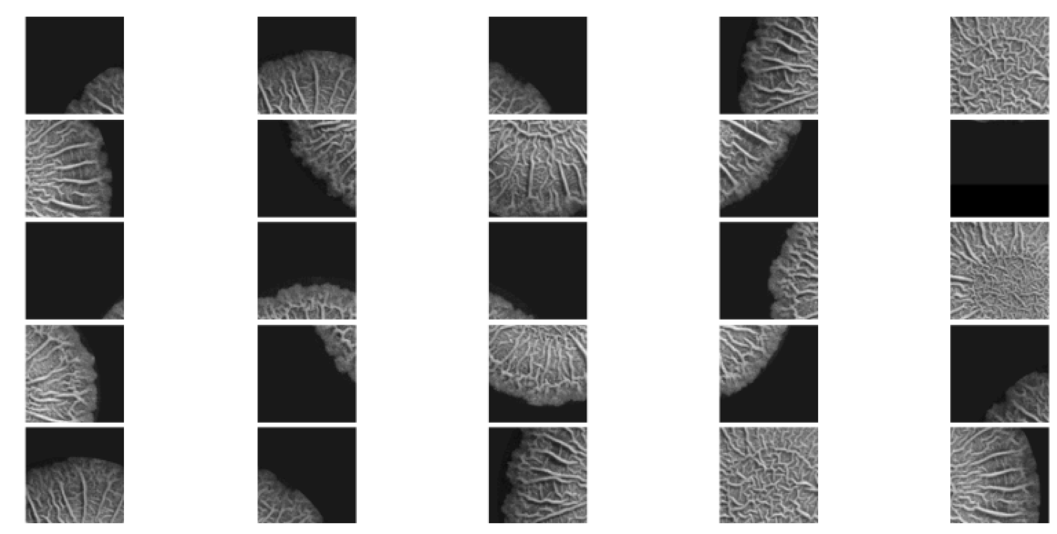
Method: data preparation



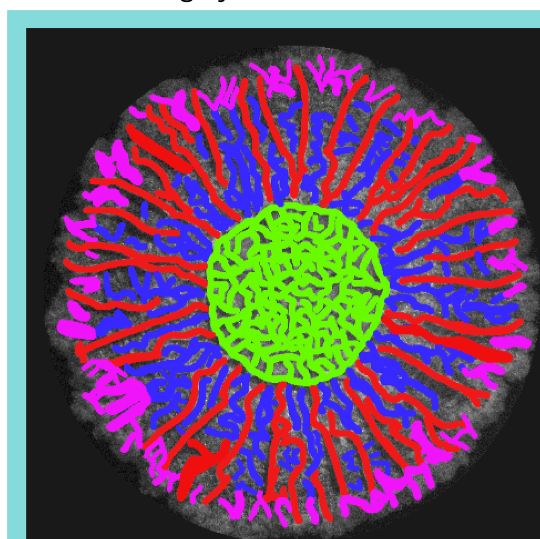
1) Background removal - Python or GIMP software



2) Contrast enhancement in Python. Thresholds set as *hyperparameters.*

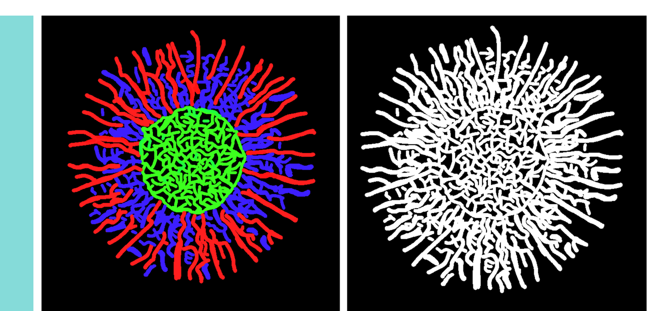


3) Patching. U-Net works optimally on 256x256 patches. Automated blank removal using Python.



1) Green - grown during nutrient abundance under isotropic stress.
2) Red - larger radial wrinkles used for nutrient transport and antibiotic escape
3) Blue - smaller wrinkles formed under anisotropic stress in nutrient depletion
4) Pink - edge wrinkles only seen in mature biofilms

4) Determining different wrinkle types and commissioning annotations (Obaid, 2024)



5) Generation of coloured and binary masks from annotations (these masks are patched in exactly the same way as their corresponding image pairs).

Aims

1) Develop a data analysis pipeline to produce binary masks predicting where the wrinkles are in radial biofilm colonies
2) Develop a data analysis pipeline to segment wrinkles in radial biofilm colonies into different types

Successful completion of these aims would allow for wrinkles and wrinkle types to be isolated from images of biofilms and quantitively analysed. This could provide a method for investigating the relationship between phenotype differentiation and biofilm wrinkle morphology.

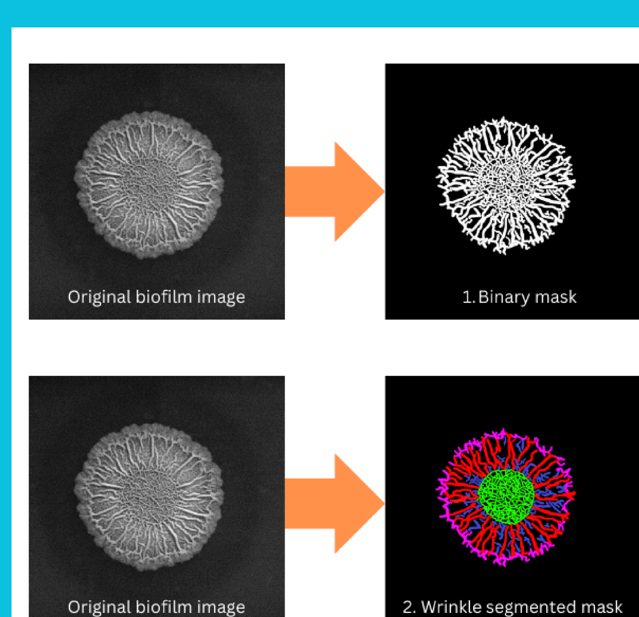


Fig 3. The aims of this project

U-Net: Training and architecture

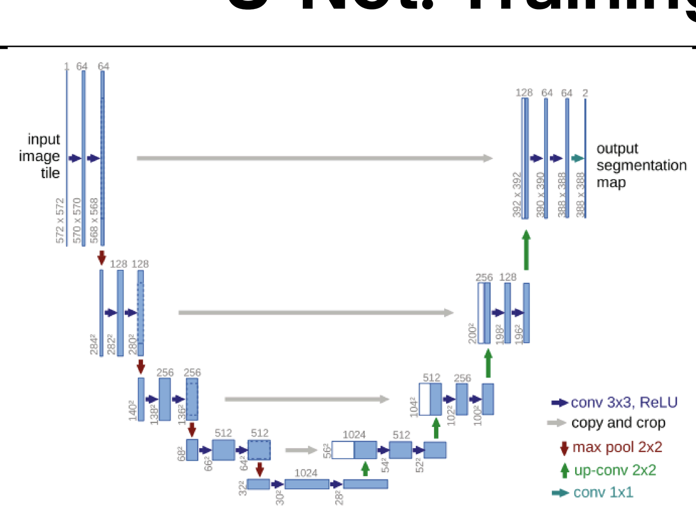


Fig 4. The U-Net architecture (Ronneberger, Fischer and Drax, 2015)

- U-Net is a convolutional neural network (CNN) architecture used for medical image segmentation (Ronneberger, Fischer and Drax, 2015).
- It is characterised by the symmetric 'encoder' / 'decoder' 'U' shape, in which features of greater complexity are extracted in each down-sampling step during training
- Feature extraction is performed using weighted convolutions. These weights are learnt during training
- The model is trained on ~100 image and mask pairs (data and 'ground truths')
- When evaluating the model takes an image as input and outputs a 'prediction'

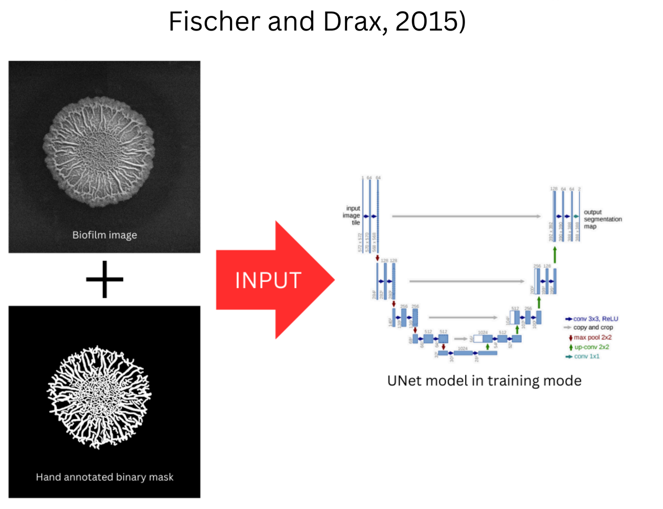


Fig 5. The training process for binary classification

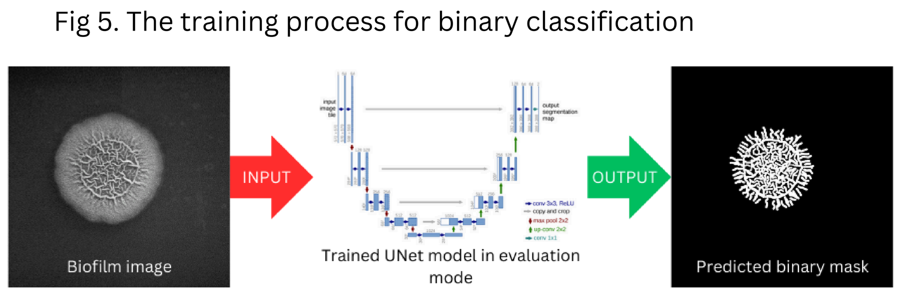


Fig 6. The evaluation process for binary classification

Alternative architectures
Architectures such as Vision Transformers can handle longer-range dependencies in spatial information. U-Net is well-trialled, but novel approaches such as ViT or hybrids like Mamba-UNet, SwinUNet and TransUNet may perform better (Wang et al, 2024)

Results

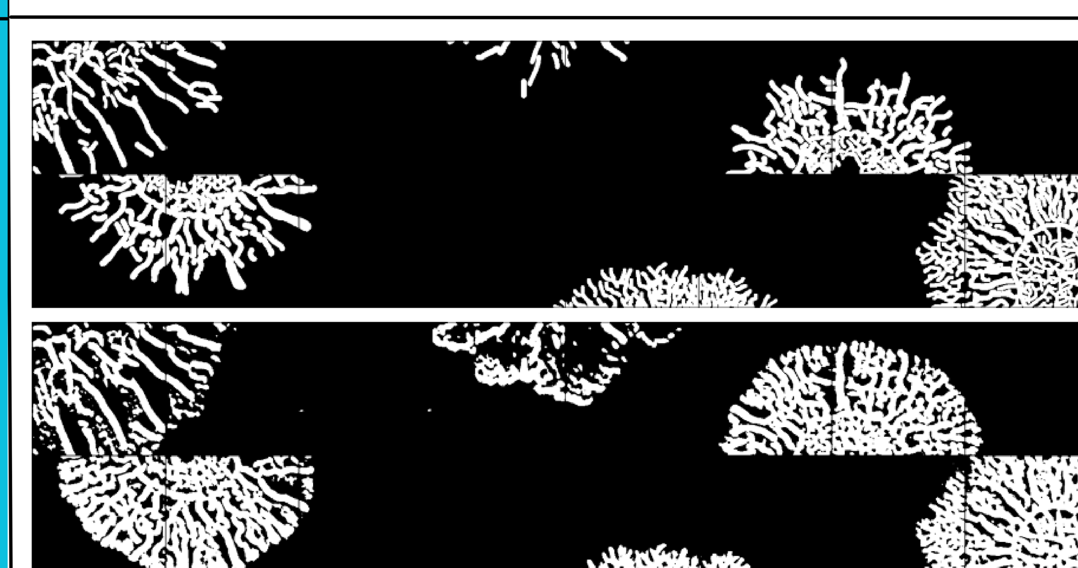


Fig 7. Top: patched ground truths of the binary classification task. Bottom: the corresponding predictions

- Results for task 1 are shown in Fig 7.
- Predictions observed to 'overfill' wrinkles, possibly due to loss of feature precision in contrast adjustment
- Appears that the model is detecting biofilm edges as wrinkles - would perhaps be better handled by a ViT.

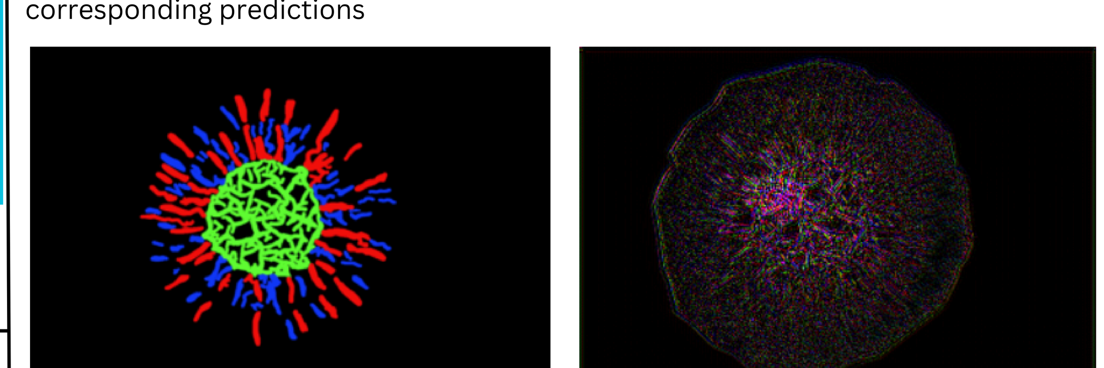


Fig 8. Ground truth (left) and prediction for image-to-image translation method

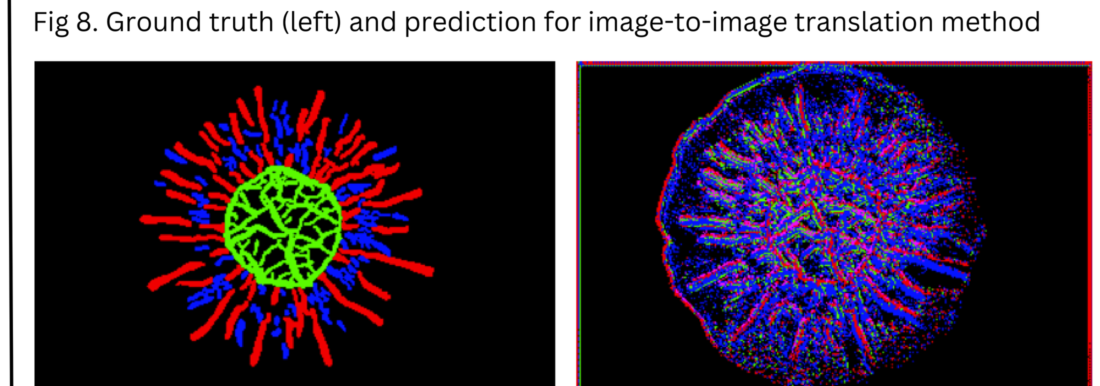


Fig 9. Ground truth (left) and prediction for multiclass classification

Performance was evaluated using the following metrics: binary classification - dice score; image-to-image translation - Structural Similarity Index Measure; multi-class classification - mean Intersection over Union.

- Results for task 2 in 2 different approaches
- Image-to-image translation: output pixels have the RGB channels and can hence be any colour
- Multi-class classification: output pixels restricted to one of five colours (4 wrinkle types and background)
- Multi-class classification shows promise when predicting wrinkle types one at a time

Further analysis

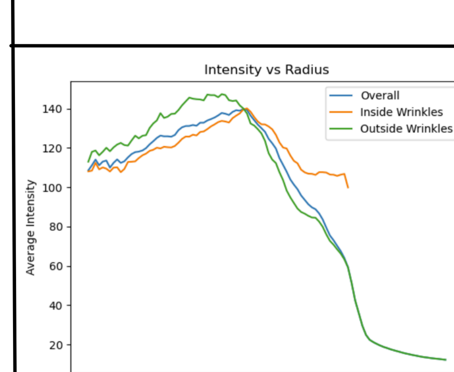


Fig 11. Using the binary masks, the areas inside and outside the wrinkles can be isolated for analysis of the intensity of different phenotypes across the radii. This could provide insights into the relationship between phenotype and wrinkle morphology

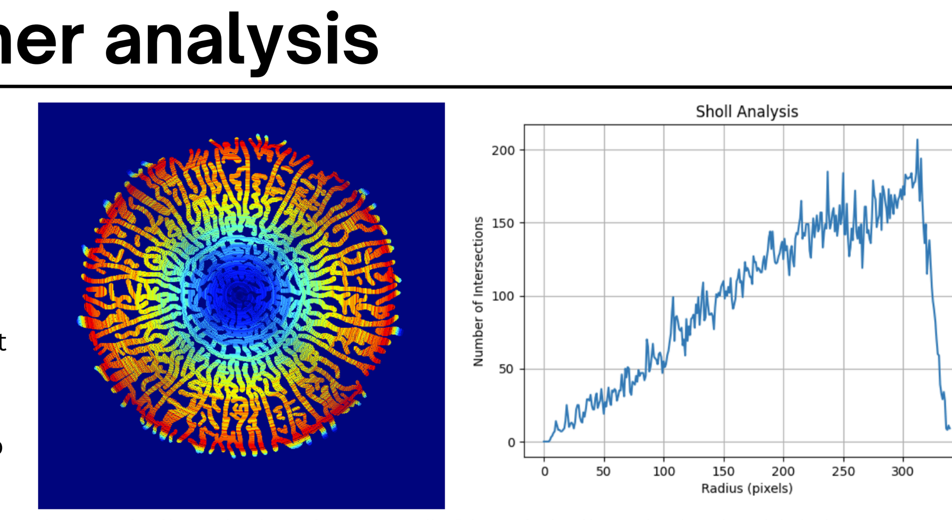


Fig 10. Sholl analysis colourmap (left) and sholl analysis plot (right)

Sholl analysis is a method of determining wrinkle density that can be applied to the produced binary masks. The density of the wrinkles at different radii can be shown using a colourmap.

References
Biofilm images: Krishnan (2022) and other members from Fusco Lab
Annotations: Obaid (2024)
Krishnan et al. (2024) [Unpublished review]
Ronneberger, Fischer and Drax (2015) *MICCAI 2015*
Wang et al. (2024) *arXiv preprint: 2402.12345*

Email: bd452@cam.ac. uk
LinkedIn: Britt Dewing