Denoising Three-Dimensional Bacterial Biofilm Images Using Machine Learning

Mia Mutadich
Mentors: Dianne K. Newman, Georgia R. Squyres

Over the last two decades, the development of high-resolution, three-dimensional imaging has been instrumental in unveiling the structural and dynamical complexities of biofilms, enabling researchers to better understand their development and functions. Biofilms are structured communities of cells that given their relevance to human health, industry, and the environment, are the subject of intense research across multiple disciplines. However, direct observation of individual cell behavior within biofilms has been limited by the low signal-to-noise (SNR) of fluorescence microscopy images. We are applying deep neural networks for image restoration to three-dimensional images of biofilms to reduce noise without blurring the underlying signal – a limitation of previous techniques that we aim to overcome. To develop a neural network for this task, we started with a simpler 2D data set with a known ground truth. Because SNR in 3D biofilms varies with depth, we represented this by acquiring 2D images at various SNR levels. We trained an existing neural network architecture on these images and wrote an algorithm to compare its performance on the full range of representative SNR values. Finally, after developing a method to simulate noise in three-dimensional biofilm images, we aim to be able to retrain our optimized neural network on our semi-synthetic dataset so that it can denoise three-dimensional images.

Biofilms – communities of microbial cells and matrix components – have been ever-present for the last ~ 3.25 billion years and dominate in all habitats, accounting for 40-80% of all cells on Earth [1,3,2]. Currently, we have an incomplete understanding of biofilm formation; however, this mode of growth is integral to the survival of microorganisms in diverse environments due to the protection that the structure provides from environmental challenges like antibiotics and antimicrobial agents[1].

Over the last two decades, the development of high-resolution, three-dimensional imaging has been instrumental in unveiling the structural and dynamical complexities of biofilms, enabling researchers to better understand their development and functions[1]. Optical fluorescence microscopy is currently the technique with great potential since it allows us to observe the behavior of individual cells inside living biofilms without disturbing their structure[4]. Using this approach, we can observe phenotypic and genotypic variations within biofilms, and begin to investigate how these variations arise and interpret their functions during biofilm development[5].

A major challenge for biofilm imaging is producing clear, easily segmented three-dimensional images of living biofilms at single-cell resolution. Live imaging often requires low illumination intensities and fast imaging, which can limit signal and lead to noisy images [6]. This noise makes it difficult to see individual cells in biofilm images and causes problems for cell segmentation algorithms. In addition, biofilms are composed of densely packed bacterial cells and matrix components, which are highly scattering. Cells near the top of the biofilm, farthest away from the imaging objective, must be imaged through a thick layer of scattering material; this scattering means that the fluorescence signal is much weaker in these images. The noise becomes overwhelming, and individual cells become even more difficult to detect.

To overcome the limitations of fluorescence microscopy, we can use computational methods to denoise images of biofilms. The most commonly used mathematical method for denoising is Gaussian filtering, however, it removes noise at the expense of blurring the underlying signal. This affects the performance of image analysis because it makes the outline of the cells blurred so they are harder to identify. Machine learning has the potential to provide high-performance solutions that avoid blurring by learning from the data itself [6]. First, we give the neural network pairs of images to train on where one represents a typical input, in this case, a noisy image, and the pair image represents the desired output or ground truth, in this case, the corresponding denoised image. The network trains itself by predicting a denoised version of the noisy input image and then adjusting its logic based on how similar its predicted denoised image is to the actual ground truth. There are existing neural network architectures that have already provided highly precise denoised images of larger organisms like flatworms[6] and we intend on using and training this type of neural network on three-dimensional images of biofilms so that we can observe the behaviors of cells in the biofilm more accurately. We believe that using neural networks to denoise images of the top of biofilms will have large implications for our understanding of the dynamics inside these biological systems, and this opportunity inspires my project.

My mentor Dr. Georgia Squyres is very interested in the formation of biofilms, specifically looking at behaviors of individual cells in the structure. In order to do this, it is necessary to directly observe cells using three-dimensional imaging, but the clarity of the images is limited by their low signal-to-noise ratio. Dr. Squyres has already implemented multiple adjustments to her imaging methods to improve the signal in the images such as testing multiple colored fluorescence dyes and adjusting the microscope settings, but she identified the potential for the application of machine learning for denoising. Hence, my SURF will help provide Dr Squyres with more easily interpreted denoised images of biofilms which will allow her to observe the activity

inside a biofilm in more detail than has ever been possible and consequently put her findings at the forefront of research in her field.

The first aim of my project is to create and train a neural network to denoise a sample two-dimensional data set; this trained neural network will later be retrained on a three-dimensional dataset to be able to denoise three-dimensional images. My mentor Dr. Georgia Squyres provided me with a data set that she collected of two-dimensional images of bacteria but at different exposure times to simulate the different signal intensities that images at different depths of the biofilm would have when imaged – this data has been used to train my neural network. The way that we quantitively characterize the level of noise in an image is by calculating the ratio of the number of signal pixels, where the signal is the bacteria in the image, and the number of noise pixels, where the noise is the background - this is called the signal to noise ratio (SNR). The method Dr. Squyres used has allowed us to store our images in order of SNR where the 17th channel image in a series has the lowest SNR and the 2nd channel image has the highest - the 1st channel image is a phase contrast image which is not used for training. We will need to independently train multiple copies of the same neural network architecture on images of different SNR, which correspond their respective image channels, so that we can see how well the model performs as the SNR varies. Then we need to repeat this process for any neural network architectures that we think have the potential to effectively denoise our images and we compare their performances at different SNRs to decide which one is optimal on the full range of representative SNR values. For that reason, I began by producing a graph that displays the signal-to-noise ratio of the images from channels 2 to 17 in a series so that for example when we train a model on all the 6th channel images in each series, we know what SNR this corresponds to.

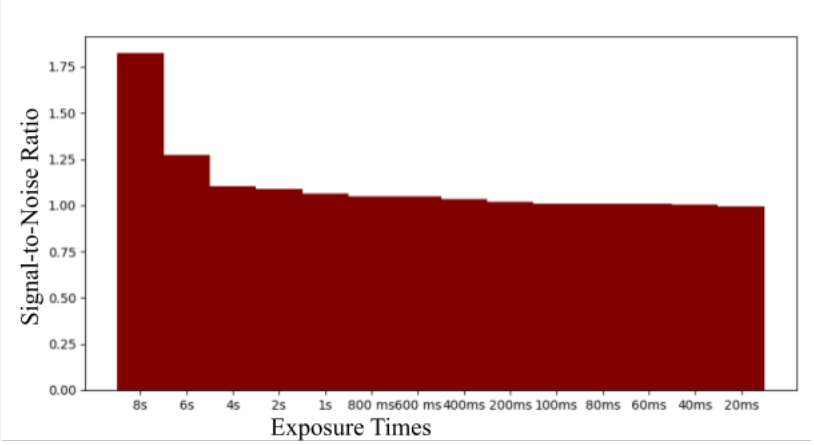


Figure 1: The signal-to-noise ratios (SNR's) of the 15 images we will use to train our neural network. *Using Otsu's method on the phase contrast image of the series to calculate a threshold, we applied the threshold mask to each of the images in the rest of the series to acquire the indexes of the signal and noise pixels in each image. Using advanced indexing in Python to access the intensity values of the signal and noise pixels in each image, we could then calculate the average signal intensity and noise intensity of each image to produce a value for the SNR. This graph will allow us to compare the performance of contending neural networks at different signal-to-noise ratios which in turn allows us to more precisely select a neural network that optimally denoises between the specific range of signal-to-noise ratios that images of biofilms fall between.*

The first deep learning architecture that looked promising, because it was mentioned in multiple papers[8], was a content-aware image restoration (CARE) network[6]. This solution has already produced consistent high-level restorations when trained on images of the flatworm *Schmidtea mediterrana*, which is especially sensitive to even moderate amounts of laser light so they must be photographed at low exposure times producing images of low SNR like ours. We trained the neural network on pairs of images where one has a high signal-to-noise ratio (SNR), known as the ground truth, and the other image has a low SNR. We always use the 2nd channel image in the series as our ground truth since it has the best SNR and for our initial training, we used the 10th channel image in the series as our low SNR image.

Channel 10

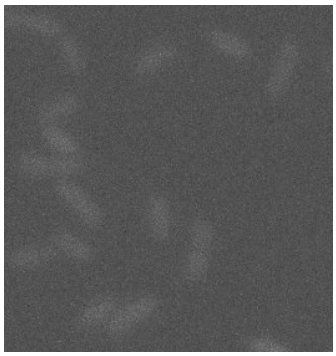


Figure 2: Low SNR image given to neural network to be denoised

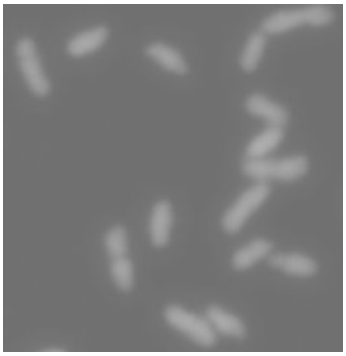


Figure 3: Predicted image denoised by neural network

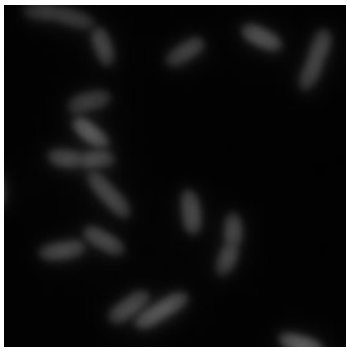
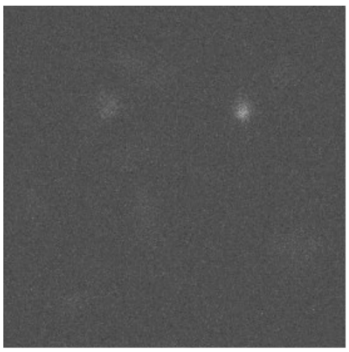
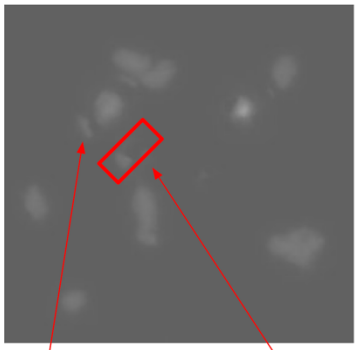


Figure 4: Ground Truth.

We know that the SNR of the 10th image in the series, which corresponds to the image with an exposure time of 200ms in my graph, is 1.15 which is very low. We can see that the neural network produced very clear denoised images which is a great sign since this SNR is already really close to that of the lowest SNR images indicating lots of potential. We were surprised that this general-purpose neural network architecture for denoising images could perform so well without optimizing it to suit our specific data, making this model particularly promising since it still has potential beyond the point at which it starts producing unsuccessful solutions.

Channel 14





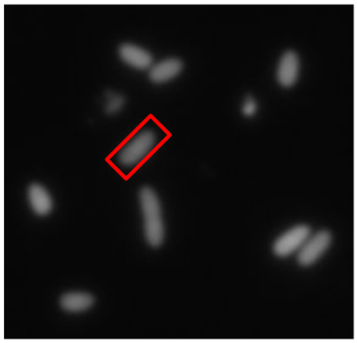


Figure 5: Low SNR image given to neural network to be denoised

Figure 2: Predicted image denoised by neural network

Figure 7: Ground Truth.

We trained the same architecture on different pairs of images where the low SNR image is different so that we can see how well the architecture performs as SNR varies. We implemented this and noticed that the model started producing unclear images after training on the 14th image of a series – this image had an SNR of 1.11. The images were unclear because although this model picked up on most bacteria in the image, it falsely identified multiple areas of the background as bacteria and in some other cases it didn't detect some bacteria at all.

Once we independently trained multiple copies of CARE on all the channels of images in our series, we needed to quantitatively rank the performance of the neural network as SNR varies. We needed to produce a value that reflects how well the model identified the bacteria in the image - which is expressed by the number of signal pixels in each image. To do this we implemented a neural network called StarDist[7] to segment our predicted images and our ground truth, returning a mask. Then we wrote an algorithm to parse through each pixel in both masks and calculate the difference between the pixels and square this value to be summed to get a final value.

The future direction of this research would be to retrain the network we selected to be able to denoise three-dimensional images. To achieve this, we would need to start by collecting an appropriate dataset. To collate the similar pairs of low SNR images and ground truth images we will need to create simulated noisy images by taking an image from the bottom of the biofilm where the noise level is low and adding a reference

noise image. Once we are satisfied with the performance of our algorithm on our artificial noisy data from aim 2, we will apply the algorithm to denoise real 3D biofilm imaging data sets. Although we won't have an independent ground truth anymore, we still need to benchmark the algorithm performance on real 3D data by checking if the output of our algorithm follows some predetermined expectations. The denoising algorithm will enable us to overcome the effects of scatter and see individual cells deep inside living biofilms.

Methods

SNR Calculation. Images were stored in .nd2 file format with 25 series of 17 images in each series. Using Otsu's method, implemented using an OpenCV library plugin, on the phase contrast image of the series to calculate a threshold, we applied the threshold mask to each of the images in the rest of the series to acquire the indexes of the signal and noise pixels in each image. Using advanced indexing in Python to access the intensity values of the signal and noise pixels in each image, we could then calculate the average signal intensity and noise intensity of each image to produce a value for the SNR.

CARE. We implemented CARE through Keras and TensorFlow via GitHub (https://github.com/CSBDeep/CSBDeep). CARE networks were trained on an NVIDIA GPU card on the Caltech High Performance Cluster (HPC). Typically for each image volume, 8,192 patches of size 128 × 128 x 2 were randomly cropped and used to train a CARE network with a learning rate of 2×10^{-4}. From the extracted patches, 4% was used as validation data. The number of epochs for training was 100, and mean absolute error was used as loss function. Training time for a given model was 0.5–1 h, application of the model on a 2,048 × 2,044 × 17-sized image volume required ~ 15 min (further hyperparameters and training times for the datasets used in this paper are provided in Supplementary Table 4).

StarDist. We implemented StarDist through TensorFlow via GitHub (https://github.com/stardist/stardist). StarDist networks were trained on an NVIDIA GPU card on the Caltech High Performance Cluster (HPC). Masks were produced using OpenCV library to implement connected component analysis to produce labelled bacteria mask. Of these images 16% was used as validation data. The number of epochs for training was 400, they use both a binary cross-entropy loss function and a mean absolute error loss function. Training time for a given model was 10-15 h, application of the model on a 2,048 × 2,044 × 17-sized image volume required ~ 15 min (further hyperparameters and training times for the datasets used in this paper are provided in Supplementary Table 4).

Works Cited

1. Hall-Stoodley, L., Costerton, J. & Stoodley, P. Bacterial biofilms: from the Natural environment to infectious diseases. *Nat Rev Microbiol* 2, 95–108 (2004). https://doi.org/10.1038/nrmicro821
2. Flemming, H. C., & Wuertz, S. Bacteria and archaea on Earth and their abundance in biofilms. *Nature reviews. Microbiology*, *17*(4), 247–260. (2019). https://doi.org/10.1038/s41579-019-0158-9
3. Donlan R. M. Biofilms: microbial life on surfaces. *Emerging infectious diseases*, *8*(9), 881–890.. (2002) https://doi.org/10.3201/eid0809.020063
4. Sugimoto, S., Okuda, Ki., Miyakawa, R. *et al.* Imaging of bacterial multicellular behaviour in biofilms in liquid by atmospheric scanning electron microscopy. *Sci Rep* 6, 25889 (2016). https://doi.org/10.1038/srep25889
5. Hartmann, R., Jeckel, H., Jelli, E. *et al.* Quantitative image analysis of microbial communities with BiofilmQ. *Nat Microbiol* 6, 151–156 (2021). https://doi.org/10.1038/s41564-020-00817-4
6. Weigert, M. et al. Content-aware image restoration: pushing the limits of fuorescence microscopy (2018). https://www.nature.com/articles/s41592-018-0216-7
7. Weigert, M. et al. Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy (2020) https://doi.org/10.48550/arXiv.1908.03636
8. Romain F. Laine, Guillaume Jacquemet, Alexander Krull, Imaging in focus: An introduction to denoising bioimages in the era of deep learning, The International Journal of Biochemistry & Cell Biology (2021) https://doi.org/10.1016/j.biocel.2021.106077

Acknowledgements

M. Mutadich thanks the Dianne Newman lab for all their feedback and comments on this work in particular my mentor Georgia Squyres. This research was supported by the Student-Faculty Program at Caltech as well as Carolyn Ash and the Ash Family. We thank Caltech for allowing us to utilize their High Performance Cluster which significantly decreased the training time of our neural networks.