**BIOL 2500 BC**

**Final independent project**

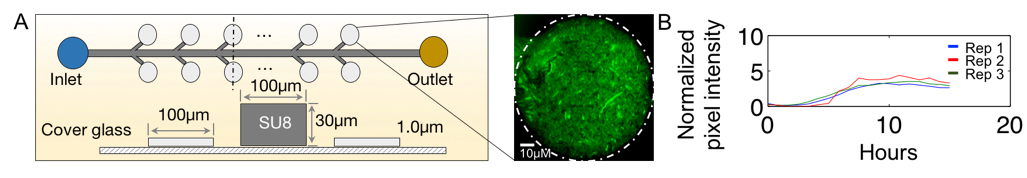
**Objective:**

The goal of this independent final project is to demonstrate your expertise of the MATLAB programming language by developing and critiquing a rigorous image analysis pipeline on real experimental data.

**Background:**

Microfluidic devices are instruments that control the flow of small volumes of liquid inside micron-sized channels. Microfluidics are used across various scientific disciplines, including microbiology, neuroscience, cell biology, and engineering, with wide-ranging applications from drug screening to point-of-care diagnostics. These platforms are useful for bacterial experiments as well: micron-sized chambers can readily trap bacterial cells in monolayers for single cell analysis; microfluidic chips can be mounted on microscope stages for fluorescence microscopy imaging; finally, since microfluidic channels allow continuous liquid flow, long-term experiments can be conducted by capturing time-lapse images.

**Experimental details:**

The data in the folder DVfiles contains time-lapse microscopy images (e.g., videos) of bacterial growth inside individual microchambers using the microfluidic device described in (Figure 1).

**Figure 1** A) The microfluidic device contains 6 channels with 24 trapping microchambers each. Cells and media are loaded through the inlet. The main channel (dark gray) is 100μm in width and 30μm in height. Branched trapping microchambers are 100μm in diameter, and ~1.3μm in height. The fluorescence image (right) shows a top-down view of a filled microchamber with GFP-expressing bacteria. B) Three growth curves from separate microchambers from one channel are plotted over hours.

Eight independent experiments were conducted. Each experiment tested a separate condition, but we don’t what they are (e.g., 8 different drugs, 2 different drugs at 4 different concentrations, etc.). Three replicates were collected per experiment using one of two cell strains G or R; G expresses green fluorescent protein (GFP), and R expresses a variant red fluorescent protein (RFP). Time-lapse images were taken at intervals of five minutes in all cases, but have been recorded for different overall durations.

|  |  |  |
| --- | --- | --- |
| **Experiment #** | **File number extensions** | **Strain** |
| 1 | 1-3 (ex: im1.dv, im1.avi) | G |
| 2 | 4-6 | G |
| 3 | 7-9 | R |
| 4 | 10-12 | R |
| 5 | 13-15 | R |
| 6 | 16-18 | R |
| 7 | 19-21 (eg 4-6) | G |
| 8 | 22-24 (eg 1-3) | G |

**Overall project description:**

The goal of this project is to determine whether you can confidently accept or reject the null hypothesis that bacterial growth is not statistically different between two experiments that will be randomly assigned to you.

Part 1 (HW#8): Answer a series of questions designed to help you develop a scaffold code for Part 2. HW#8 is submitted normally and is not included in the final report.

Part 2 (Final project and report): Develop a rigorous image analysis pipeline that reliably calculates growth rate from each video file you are assigned, despite the variability that may occur between experiments and replicates:

1. You will each be randomly assigned two experimental conditions, consisting of 3 videos each (6 in total). This will correspond to 6 video files.
2. Construct a rigorous image analysis pipeline that does the following:
   1. Plots cell growth over time for each video.
   2. Defines a metric for growth rate, and calculates this growth rate for each growth curve. This metric can be defined in many ways, several of which we learned in class; you are free to choose whichever metric you want, but must provide rationale as to why you chose it in the final report.
   3. Determines statistical difference between the average growth rates from each condition.
3. Assemble a report that meets the criteria described below.

**Provided data files:**

There are two folders of data to use for this final project.

1. DVfiles – contains movie files in the format .dv; these are your **raw dataset** **and should be used for all MATLAB analysis.**
2. AVfiles – contains movie files in the format .avi for visualization purposes only; files are compatible with the implay function in MATLAB and standard movie plays on macbooks and PCs.

**Assignment:**

1. Written report (100 pt):

**Main text:**

Abstract High-level overview of pipeline, conclusions, and limitations

Methods Description of the computational and statistical methods:

* Describe which datasets were used in the analysis
* Describe each step in your final image analysis pipeline, including references to file names where the analysis steps were performed
* Describe your metric for growth rate and how it was calculated, including any relevant equations and references to relevant file names
* Describe the statistical test(s) that was used, including references to relevant file names

Results Main results should be reported in the format of a scientific article:

* Walk the reader through every step of building a rigorous image analysis pipeline from your initial version to the final algorithm. Include as many figures with captions as you need to describe the data and processing steps
* Report figures of growth over time from each experiment
* Provide rationale for the growth rate metric you chose to use, and report the calculated growth rate from each experiment using figures and/or tables
* Clearly state the outcome of the statistical test

Discussion Summarize the strengths and weaknesses of your image analysis pipeline:

* Discuss whether you believe this conclusion is reliable based on your knowledge of how your image analysis pipeline works. In other words, explain the potential caveats in your conclusions (e.g., what are the largest weaknesses in your pipeline?).
* Discuss how your algorithm could be improved upon, along with your overall takeaways of image analysis

Supplemental table: A table with each script/function name, brief description, and where the data it generates can be found (e.g., a figure number from the main text).

**MATLAB Files**: Hand in all associated .m files that correspond to the file names from your supplemental table. As with all HW’s, include a main file to run everything in order. This should generate every figure from your main text.

1. Oral presentation (16 pt):

Each student will have 8 minutes to present their findings to the class, followed by 2-3 minutes for questions. The presentation will be graded on rubric below:

|  |  |  |
| --- | --- | --- |
| **Action** | **Description** | **Points** |
| Introduction | Provide high-level data overview | 2 |
| Hypothesis | Clearly state null and alternative hypothesis | 2 |
| Methods | Provide detailed overview of analysis pipeline, growth rate metric, and statistical test used | 4 |
| Results | Share the outcome of your statistical test | 4 |
| Discussion | Do you feel confident in your conclusion? Why or why not. If no, elaborate on one or more aspects that you found particularly challenging, and how the code could be improved in the future. If yes, explain how your code handles variability across conditions. | 3 |
| Overall professionalism | No less than 7 or more than 9 minutes; comes prepared to present; slides are professional looking | 1 |

**Formatting:** Follow all formatting guidelines above along with those outlined as good MATLAB practices. All code should be professionally commented and documented throughout. Use file names that are descriptive. Include a master pipeline script that runs all your files in the order they were meant. Suppress all intermediate outputs and use any combination of scripts/functions you prefer.

**Honor code:** Every experiment is different. No two algorithms will work on every pair. Feel free to discuss concepts amongst each other, but all code should be written individually. If I suspect cheating, students will be asked to complete a separate exam to pass the course.

OVERALL PROJECT TIMELINE AND POINTS DISTRIBUTION

|  |  |  |  |
| --- | --- | --- | --- |
| **Action** | **Points** | **Due date** | **Where to hand it in?** |
| **Written report:**  main text and MATLAB files | **100** | **Dec 12th, 12:59p** | **Canvas** |
| **HW #8** | **16** | **Nov 29th, 12:59p** | **Canvas** |
| **Oral presentation** | **16** | **Dec 6th, 12:59p** | **Canvas** |

|  |  |
| --- | --- |
| **Number** | **Experiments** |
| 1 | 1 & 3 |
| 2 | 2 & 4 |
| 3 | 3 & 6 |
| 4 | 7 & 8 |
| 5 | 1 & 3 |
| 6 | 2 & 4 |
| 7 | 2 & 8 |
| 8 | 6 & 3 |
| 9 | 1 & 4 |
| 10 | 5 & 8 |
| 11 | 1 & 7 |
| 12 | 2 & 6 |
| 13 | 5 & 6 |
| 14 | 2 & 7 |
| 15 | 3 & 5 |
| 16 | 4 & 8 |