Project 22: Multiview image segmentation - Quick notes (working document)

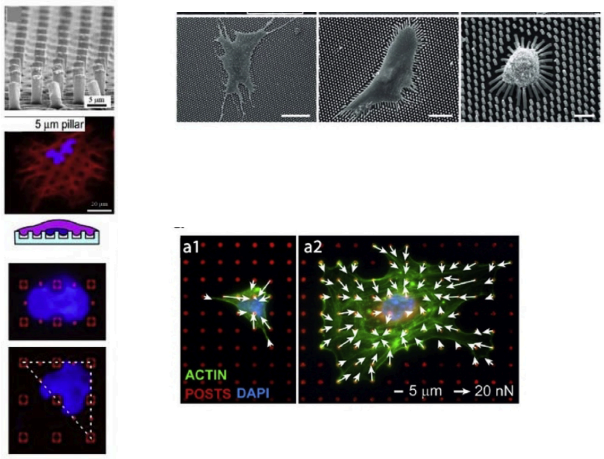
Contact the supervisor, Bjørn, via the slack channel: https://02456-2024-workspace.slack.com/archives/C07U3HUNAP6 (or a private slack channel for questions that are not relevant to everyone… emails to bjje@dtu.dk should be your last resort)

Course website:
https://docs.google.com/document/d/1jqf77FjA2w8w24tXsu-i5GuEF1WqpdFmWdqv-KKflQM/pub

Introduction:

Traction force microscopy quantifies the forces a biological cell exerts on its surroundings.

One (of several) techniques is to place biological cells on top of a grid of nanopillars that bend as the cells attach themselves to the surface (see e.g. [1,2,3]. The displacement of the pillars is (very carefully) observed using optical microscopy to identify the pillars' bending.



From [3]

This project deals with a subproblem in traction force microscopy related to a large-scale and automatic application of the techniques, specifically identifying which nanopillars cover the individual cells in an image (and potentially which parts of the cell and profiling the morphology of the cells).
The problem can be solved with relative ease if a fluorescence marker can be added to the cells (i.e., they lightly up in a certain frequency range). However, for most relevant applications, adding a fluorescence die is impossible, and it adds complexity to an otherwise simple experiment.

The end goal is to conduct analysis using brightfield (BF) microscopy, which uses standard equipment and no fluorescence markers. To this end, we need fast and robust (e.g. to noise, broken pillars, and debris in the sample) methods for imaging and identifying cells and the relevant pillars.

We have collected a dataset for training an ML model to map from a set of BF images (here 11 images at different focal points) to a mask indicating where the cell is placed (hence indicating the relevant pillars). The mask/group truth is (semi-)automatically derived using fluorescence microscopy.

Data:

The data (approx. 12-15 GB) is located on the central Gbar-cluster (https://www.gbar.dtu.dk/) at:

```
/zhome/70/5/14854/nobackup/deeplearningf24/forcebiology/data
```

The structure is as follows:

`./brightfield` Brigfield images with 11 focal point; the input to the model
`./masks` Binary masks (one image for each set of 11 input images) specifying the desired out of the model.
`./LICENSE.TXT` Please read and make sure you are OK with it!

The `./brightfield` folder contains seven subfolders indicating a well number (i.e. specific well at which the data was collected and the number of positions within the well. E.g. `Alexa488_Fibroblasts_well3_200locations` demonstrates that this is well3 and 200 positions have been imaged within the well. Additionally, `Alexa488` is the name of the fluorescent dye used (not essential for you), and `Fibroblast` indicates the cell type used in the experiments (probably not critical for you). Each folder contains the captured brightfield images taken within a given well, e.g, the file name

`Alexa488_Fibroblasts_well3_200locations_s008z05c2_ORG.tiff`

where the tag `s008z05c2` should be decoded as follows:

`s008` indicates the sample/site number in the well here 008 (out of 200)
`z05` indicates the focal point or z-position (here 05 out of 11 z-position/focal points)
`c2` indicates the channel number; c2 indicated the brightfield channel

The `./mask` folder contains the derived maks (in a flat file structure) corresponding to the cell's location (automatically estimated via the fluorescence channel). E.g.
`Alexa488_Fibroblasts_well3_200locations_s008z06c1_ORG_mask.tiff`

Here, the focus tag is always z06, which defines the optimal focus point in the fluorescence channel.
`c1` refers to the fluorescence channel.
We note that the first part of the name, i.e., `Alexa488_Fibroblasts_well3_200locations_s008z06`, should match the corresponding input of 11 images in the ./brightfield folder.

The ***official*** train/test split on which you should report is
Test: Well 1 (i.e. samples in the folder Alexa488_Fibroblasts_well1_50locations)
Training/validation: Well 2-7

Warning: This is real-world and uncurrated data, and there are likely noisy/problematic samples in the dataset. You are free to perform your own curation of the training sets. The test has been validated (i.e., well 1), and you should not remove samples when evaluating it.

Note: More data may arrive as we go along, but do not count on it…

Models:

- A classic image processing baseline (ideally, you should include this!): Filter everything based on insight about pillar structure. *Hint*: Try to do an FFT of the images and look at the peaks in the spectrum… Could the peaks filter the image (i.e. remove the pillars)? Small exercise: is there a connection between an FFT (the magnitude) and a convolutional neural network…?
- A simple convolutional neural network: e.g., 2-3 convolution layers, with 4-16 filters and a fully connected layer followed by a similar set of transposed convolution layers…. or a simply a fully connected layer outputting the resulting mask). *Hint*: Consider how the 11 channels are merged and if you could experiment with this.
- U-Net is a well-known segmentation model [6], and several (pre-trained) examples are available on Huggingface and elsewhere.
- Vision transformers (you likely need to use pre-trained models and then fine-tune them!)

- … your favorite model.

… expected problems: You only have ~1100 examples…. and likely way more parameters, so you must think carefully about ensuring your model does not overfit!

- Regularization: Explore if regularization can provide some assistance (you'd likely need to explore various types, e.g., dropout, L2, L1, and the strengths)
- Data augmentation (rotation, translation, cropping), applies the same augmentation to all (11) images in the input.
- Pre-trained models, e.g. find pre-trained U-net models for segmentation and extend them to the specific case and fine-tune them,
- Find other related data sets to potentially pre-train parts of your network before fine-tuning.

Experiments:

There are several aspects to be explored (potentially) and reported/visualized via empirical experiments:

- Comparing model type/architecture, size, activation functions, learning rates, and other hyperparameters.
- Learning curves (not optimization traces) are plots showing the number of training samples vs. performance. Performance could be measured on both training, validation, and test sets.
- Input (and output) resolution: Experiment with different image resolutions (subsample the images) or the number of input images per sample (e.g., reduce from 11 to 3 input images) and compare the performance.
- Explainability techniques/sanity checks, e.g., grad cam, visualization techniques (e.g. [4, 5]).
- …

Practicalities:

Meetings/supervision: On Mondays, I (Bjørn) will aim to be in room ***303A/databar 048*** at 13:00 and stay for a few hours during the allocated course timeslot (I'll let you know via Slack if changes occur). I am unavailable for meetings outside the scheduled course timeslot, but you can write on Slack, and I'll try to respond.
NOTE: ***Due to other obligations, I (Bjørn) won't be in 303A/databar 048 until 14:45ish on Monday, 4/11. Please start working on the suggested activities listed below.***

Coding/HPC assistance: Please use the TAs as usual. The HPC IT support is also much better at this than I am.

The poster presentation on December 10th will be postponed as Bjørn is at the NeurIPS conference on that day. We will agree on a later date in December or January.

Suggested activities in the first project week:

- Get access to, load, visualize, and study (a subset of) the data (i.e. show the input and output images for a few examples using Python).
    - It would be beneficial to do a spectral analysis of the images, i.e., do an FFT and visualize it.
    - … perhaps you need to familiarise yourself with the Gbar system (login, ssh, etc) to get started. https://www.gbar.dtu.dk/
- Start working on a dataloader for PyTorch.
- Start setting up a basic CNN, taking as input 11 images and outputting a single mask.

- Reflect on how you will evaluate the model outputs(e.g. intersection over union). Start implementing the evaluation pipeline/code.
- Skim some of the background papers [1-3]
- …

References:

[1] Tan JL, Tien J, Pirone DM, Gray DS, Bhadriraju K, Chen CS. Cells lying on a bed of microneedles: An approach to isolate mechanical force. Proceedings of the National Academy of Sciences 2003;100(4):1484–1489.
[2] du Roure O, Saez A, Buguin A, Austin RH, Chavrier P, Silberzan P, et al. Force mapping in epithelial cell migration. Proceedings of the National Academy of Sciences 2005;102(7):2390–2395.
[3] Ermis M, Antmen E, Hasirci V. Micro and Nanofabrication methods to control cell-substrate interactions and cell behavior: A review from the tissue engineering perspective. Bioactive Materials 2018;3:355–369.
[4] Visualizing and Understanding Convolutional Networks, Matthew D Zeiler, Rob Fergus, https://arxiv.org/abs/1311.2901
[5] Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, Dhruv Batra https://arxiv.org/abs/1610.02391
[6] U-Net: Convolutional Networks for Biomedical Image Segmentation, Olaf Ronneberger, Philipp Fischer, Thomas Brox
…