

Interfacing CellProfiler with other software tools via files and plugins

Beth Cimini

Broad Institute of MIT and Harvard, Cambridge, MA.

Background information

General background information

This exercise is meant to extend and build upon the Beginner Segmentation exercise available from the CellProfiler tutorials page . Please consult that tutorial for general information on how to configure CellProfiler as well as information about the images. It will be generally assumed you understand the modules covered in that tutorial, including input, object creation, overlays, and saving.

What is this exercise?

There are a number of great software tools available to life scientists wishing to analyze images these days - forum.image.sc alone has more than 60 open-source tools! Sometimes, though, it helps to have a multi-tool workflow - do one step in Tool A, and then another in Tool B (and then possibly C, D, etc, but hopefully not!). In this tutorial, you'll try 3 different ways of accessing work you did in other tools.

1. Loading masks created by other segmentation tools (in this case, [Cellpose](#), but the same strategy works for many tools which use this format).
2. Accessing [ilastik](#) to use a trained pixel classification model via CellProfiler's plugins system.
3. Running Cellpose in CellProfiler via CellProfiler's plugins system and a Docker container.

Plugins

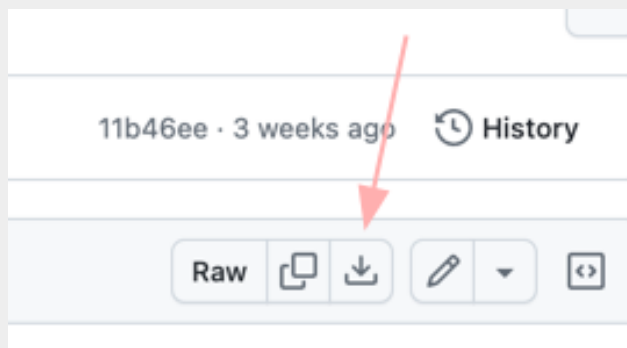
While CellProfiler doesn't have as many plugins as, say, Fiji, it does have many for you to try! You can visit [plugins.cellprofiler.org](#) to learn more. To quote from that site:

Plugins advance the capabilities of CellProfiler but are not officially supported in the same way as modules. A module may be in CellProfiler-plugins instead of CellProfiler itself because:

- it is under active development
- it has a niche audience
- it is not documented to CellProfiler's standards
- it only works with certain version of CellProfiler
- it requires extra libraries or other dependencies we are unable or unwilling to require for CellProfiler
- it has been contributed by a community member

Tip:

While that documentation has instructions for [installing plugins](#), in step 2 it suggests downloading all of the CellProfiler plugins; this isn't a bad thing to do, but you can download individual plugins from GitHub with the website button below as well or instead.



GitHub's "Download Raw Files" button

Tip:

We strongly recommend making a dedicated folder to store your CellProfiler plugins, as loading can be slow if there are a lot of other miscellaneous files around (such if they sit in the Downloads folder, for example).

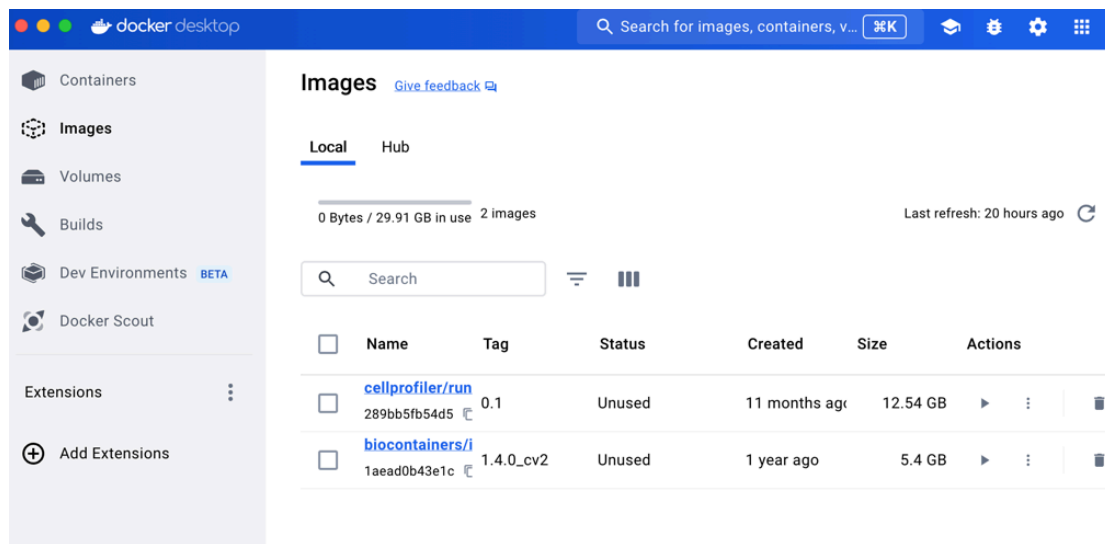
Docker Desktop

In Exercise 3 (and for Windows users, optionally Exercise 2 as well), we want to access software that we either don't want to install locally because it's painful if you're not pretty comfortable in Python (Cellpose) or that we CAN install locally but may not play nicely with other tools' multiprocessing setup (ilastik).

Computer scientists will often use software **containers** to ship tools or data that are hard to install - [here](#) is a good introduction and overview for non-computer scientists. You can think of containers as "a pre-configured operating system in a box". Because they come to you pre-configured, installation of any software happens once-and-only-once (by the creator of the container, not by you), and should stay working for long after ie the latest Mac upgrade breaks a certain older software installation. Groups like [biocontainers](#) have already containerized many of the tools you know and love. There are a number of types of software containers, but one of the most common is called a **Docker** container.

Most life scientists aren't aware of or don't use containers, especially because typical usage involves accessing them via your terminal. But Docker doesn't have to mean terminal! There are [containers that spit out interactive websites for you to use](#), and CellProfiler has plugins that are specifically set up to call other tools that live (and are executed) inside Docker containers.

To do this, CellProfiler needs to have the infrastructure for Docker containers up and running, which you can give it by installing a free program called [Docker Desktop](#). You don't need to make an account to use it, but you probably will need to reboot your computer after installation is complete. Then, whenever you want to use a Docker-calling plugin in CellProfiler, you simply must make sure that Docker Desktop is open and running, CellProfiler will take care of everything else!



The Docker Desktop interface for Mac. Containers can be searched for in the top search bar!

Warning:

Anytime you're installing a program that can run other programs, you must of course be careful. Dockers give you access to a huge variety of tools that might otherwise be hard to install (or install together), but exercise caution in only running containers you recognize or trust!

Exercise 1: Importing masks from another segmentation tool

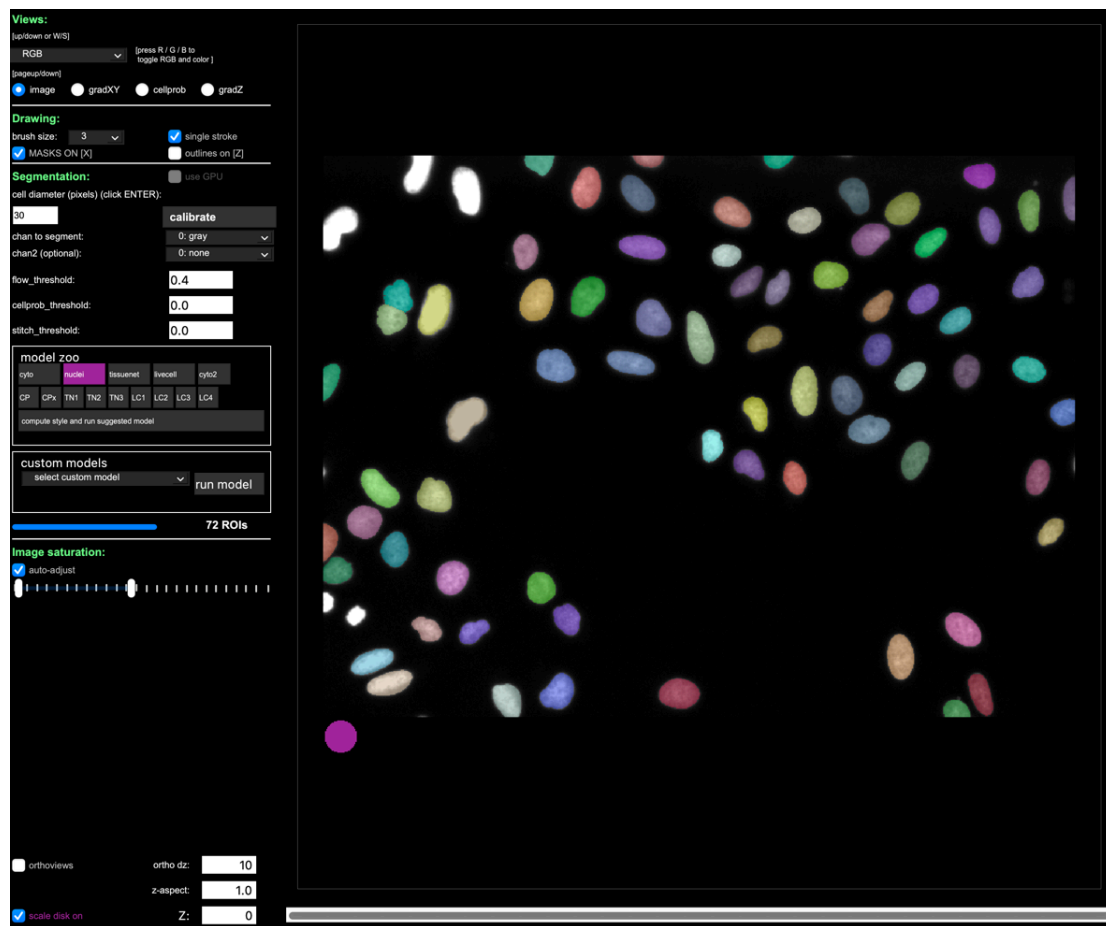
In this exercise, we will be loading in *label masks* (sometimes also called *label matrices*) made in Cellpose. Simply, a label mask is an image in which all the background pixels have a value of 0, all the pixels of object 1 have a value of 1, all the pixels of object 2 have a value of 2, etc. It is a commonly used format by many tools that do object segmentation (though unsuitable for applications where you have overlapping objects). Since many tools (including CellProfiler) can make these masks, CellProfiler has the ability to read them in and automatically detect them as objects.

We've used Cellpose 2.2.2 to generate nuclear masks from the DNA images and cell masks from the ActinGolgi images.

Only if you are curious - how did we make these label masks?

Cellpose was installed in a conda environment according to the [official instructions](#) for installation with the GUI. The software was started in that conda environment using the command `cellpose`.

Nuclear masks were made by dragging each DNA image into the GUI, selecting the 'nuclei' model with default settings, and then saving them individually as PNG using `Cmd+N` (Mac) `Ctl+N` (Windows). Since there were only 10 images and the hotkeys were available, this was not too time-consuming to do.



Nuclei segmented in the Cellpose GUI

Cell masks were made by navigating into the image folder and then running the command below; this requires some knowledge of the [Cellpose command line syntax](#) but is much easier for large numbers of images.

```
python -m cellpose --dir . --img_filter Ch4 --pretrained_model cyto2 --diameter 40
--save_png --verbose
```

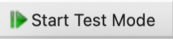


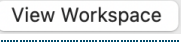
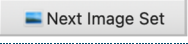
```

(bcimmini@wm4f8-761 cellpose_masks_cells % python -m cellpose --dir . --pretrained_model cyto2 --diameter 40 --save_png --verbose
2024-04-30 10:59:04,196 [INFO] WRITING LOG OUTPUT TO /Users/bcimmini/.cellpose/run.log
2024-04-30 10:59:04,196 [INFO]
cellpose version: 2.2.2
platform: darwin
python version: 3.9.16
torch version: 2.0.1
2024-04-30 10:59:04,196 [INFO] >>> using CPU
2024-04-30 10:59:04,198 [INFO] >>> running cellpose on 10 images using chan_to_seg GRAY and chan (opt) NONE
2024-04-30 10:59:04,198 [INFO] >>> using CPU
2024-04-30 10:59:04,198 [INFO] >> cyto2 << model set to be used
2024-04-30 10:59:04,198 [INFO] Downloading: "https://www.cellpose.org/models/cyto2torch_1" to /Users/bcimmini/.cellpose/models/cyto2torch_1
100%|
2024-04-30 10:59:37,787 [INFO] Downloading: "https://www.cellpose.org/models/cyto2torch_2" to /Users/bcimmini/.cellpose/models/cyto2torch_2
100%|
2024-04-30 10:59:53,084 [INFO] Downloading: "https://www.cellpose.org/models/cyto2torch_3" to /Users/bcimmini/.cellpose/models/cyto2torch_3
100%|
2024-04-30 11:00:04,675 [INFO] WARNING: MKL version on torch not working/installed - CPU version will be slightly slower.
2024-04-30 11:00:04,675 [INFO] see https://pytorch.org/docs/stable/backends.html?highlight=mkl
2024-04-30 11:00:04,877 [INFO] >>> model diam_mean = 30.000 (ROIs rescaled to this size during training)
2024-04-30 11:00:04,881 [INFO] Downloading: "https://www.cellpose.org/models/size_cyto2torch_0.npy" to /Users/bcimmini/.cellpose/models/size_c
100%|
2024-04-30 11:00:05,073 [INFO] >>> using diameter 40.000 for all images
2024-04-30 11:00:05,074 [INFO] 0%| 0/10 [00:00<?, ?it/s]
2024-04-30 11:00:05,078 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:00:13,672 [INFO] >>> TOTAL TIME 8.59 sec
2024-04-30 11:00:13,719 [INFO] 10%|# 1/10 [00:08<01:17, 8.64s/it]
2024-04-30 11:00:13,731 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:00:23,115 [INFO] >>> TOTAL TIME 9.38 sec
2024-04-30 11:00:23,152 [INFO] 20%|## 2/10 [00:18<01:12, 9.11s/it]
2024-04-30 11:00:23,156 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:00:31,916 [INFO] >>> TOTAL TIME 8.76 sec
2024-04-30 11:00:31,940 [INFO] 30%|### 3/10 [00:26<01:02, 8.96s/it]
2024-04-30 11:00:31,945 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:00:39,665 [INFO] >>> TOTAL TIME 7.72 sec
2024-04-30 11:00:39,727 [INFO] 40%|### 4/10 [00:34<00:50, 8.50s/it]
2024-04-30 11:00:39,728 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:00:48,511 [INFO] >>> TOTAL TIME 8.78 sec
2024-04-30 11:00:48,541 [INFO] 50%|#### 5/10 [00:43<00:43, 8.61s/it]
2024-04-30 11:00:48,544 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:00:57,089 [INFO] >>> TOTAL TIME 8.54 sec
2024-04-30 11:00:57,122 [INFO] 60%|##### 6/10 [00:52<00:34, 8.60s/it]
2024-04-30 11:00:57,125 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:01:05,502 [INFO] >>> TOTAL TIME 8.38 sec
2024-04-30 11:01:05,540 [INFO] 70%|##### 7/10 [01:00<00:25, 8.54s/it]
2024-04-30 11:01:05,542 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:01:13,788 [INFO] >>> TOTAL TIME 8.25 sec
2024-04-30 11:01:13,826 [INFO] 80%|##### 8/10 [01:08<00:16, 8.46s/it]
2024-04-30 11:01:13,829 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:01:22,739 [INFO] >>> TOTAL TIME 8.91 sec
2024-04-30 11:01:22,802 [INFO] 90%|##### 9/10 [01:17<00:08, 8.62s/it]
2024-04-30 11:01:22,813 [INFO] ~~~ FINDING MASKS ~~~
2024-04-30 11:01:35,845 [INFO] >>> TOTAL TIME 13.03 sec
2024-04-30 11:01:35,897 [INFO] 100%|##### 10/10 [01:30<00:00, 10.00s/it]
2024-04-30 11:01:35,897 [INFO] 100%|##### 10/10 [01:30<00:00, 9.08s/it]
2024-04-30 11:01:35,897 [INFO] >>> completed in 151.701 sec
(bcimmini@wm4f8-761 cellpose_masks_cells %

```

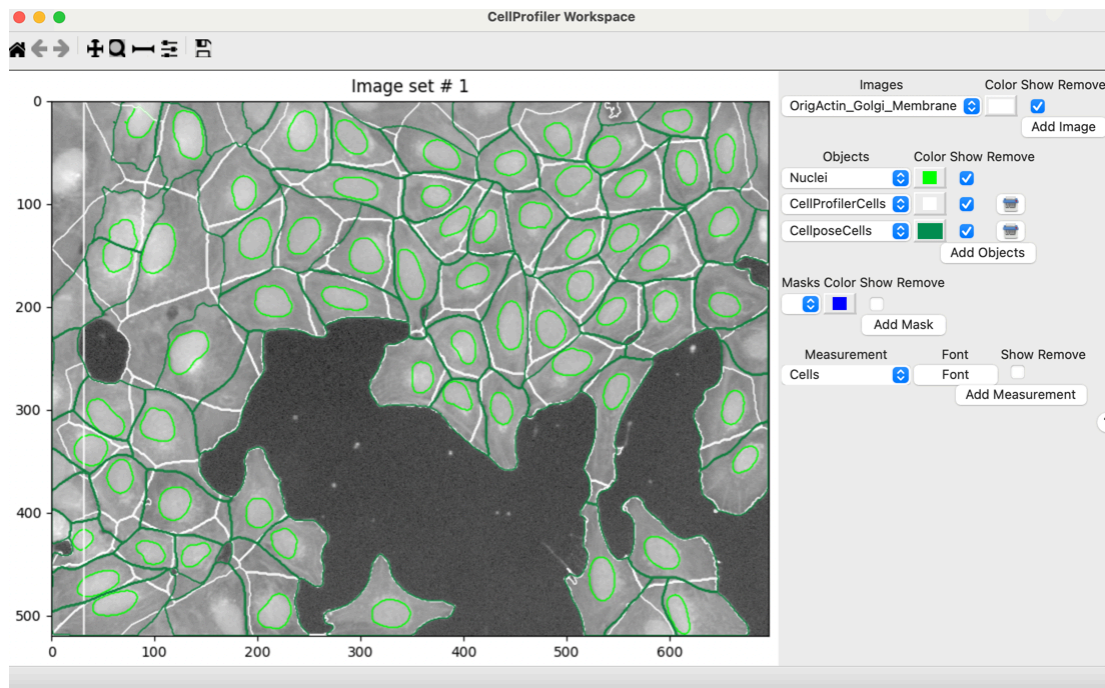
Command line execution of Cellpose

Loading the label masks into CellProfiler

1. Open up a clean copy of CellProfiler (or run File -> New Project) and drag `bonus_1_import_masks.cppipe` into the pipeline panel.
2. Drag the `images_Illum-corrected` subfolder from the main Beginner Segmentation exercise and also the `cellpose_masks_cells` subfolder from this exercise. *Do not drag in the `cellpose_masks_nuclei` folder.*
3. Put CellProfiler into TestMode , open the eye icon  next to OverlayOutlines, and then hit  3 times to create a classical segmentation and compare it with the Cellpose-generated version. You can check the settings in the OverlayOutlines module to see which outline color corresponds to which segmentation in the output.
4. Optionally, open the Workspace Viewer using  to create easily on-the-fly customizable overlays.
5. Hit  and repeat a couple of times to examine the segmentations on more images.

Question for you:

Are there cases where you think classical segmentation typically performs better? Where Cellpose performs better?



Visualization of images and objects in the Workspace Viewer

Reverse engineer how this worked

1. Open the NamesAndTypes module.
2. Scroll down in NamesAndTypes to find the entry that adds the masks - is there any setting you notice about it that is different than for the other images?
3. Find the entry in NamesAndTypes that has 2 rules the image has to pass, not just one - do you understand why that is?

Bonus-to-the-Bonus - add the provided nuclear segmentations as well

1. Drag and drop the `cellpose_masks_nuclei` folder in the Images module to add the nuclear masks to the image list.
 2. Go to the NamesAndTypes module and configure it so both kinds of masks can be loaded (hint: there's a duplicate button you can use to make a second copy of a channel!)
- You'll probably have to modify the existing settings for loading the cell masks by changing its rule or adding a second rule. Do you understand how/why?

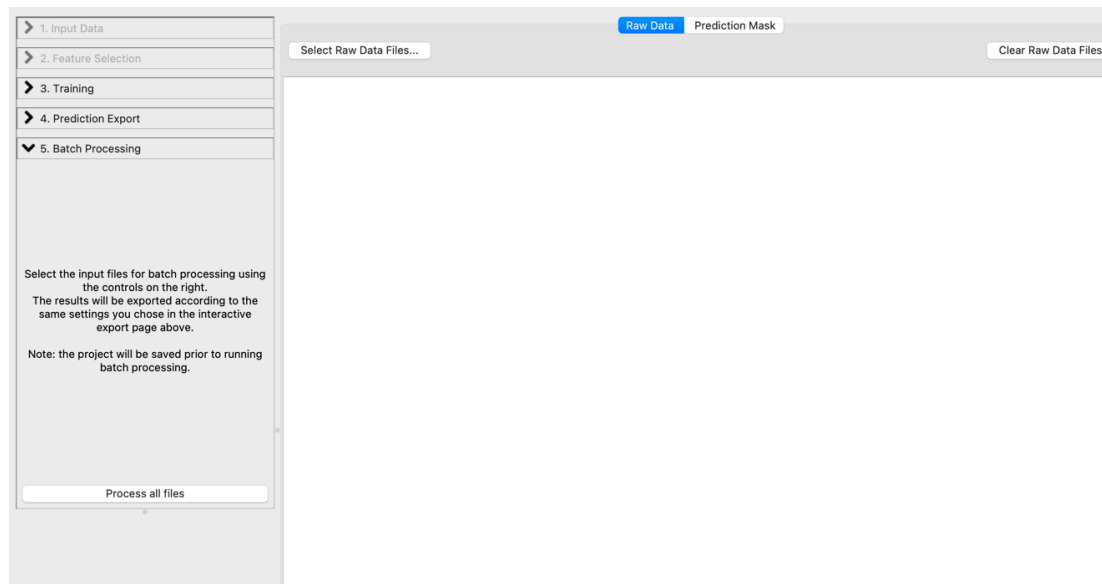
Exercise 2: Running ilastik locally or from a Docker container

Classical segmentation requires the thing you care about in an image be bright and every single other pixel dark(er). If you have a good clean fluorescent signal for the thing you care about, great! If not, you may need to resort to some tricks.

One such trick is to train a classifier on a pixel-by-pixel basis to say “here is what I think is the likelihood that this is a pixel you want to end up in your segmentation”; this classifier then ends up creating for each pixel a *probability value* that corresponds to how likely it thinks it is you want to segment that pixel. If

your classifier is good, that will give you an image where the pixels you care about are high-probability (bright) and everywhere else is low-probability (dark). That's exactly what we want! Life scientists tend to call this “**Pixel Classification**”; computer scientists will sometimes refer to it as “**Semantic Segmentation**”.

There are a few popular Fiji plugins for doing this, including Weka Trainable Segmentation and Labkit, and you should absolutely use them if they work better for you! We tend to use ilastik, because it makes it easy to automate creating a classifier from a very small number of images and then bulk-applying it to many others in “Batch Processing” mode. You can check out a tutorial we have written for running ilastik, and *then* CellProfiler at tutorials.cellprofiler.org (look for Pixel-based Classification).



ilastik's Batch Processing mode

Why do two steps though, when you can do one instead? In this tutorial we'll take a pre-trained ilastik classifier and run it inside our CellProfiler pipeline, so we can find and measure objects all in one step.

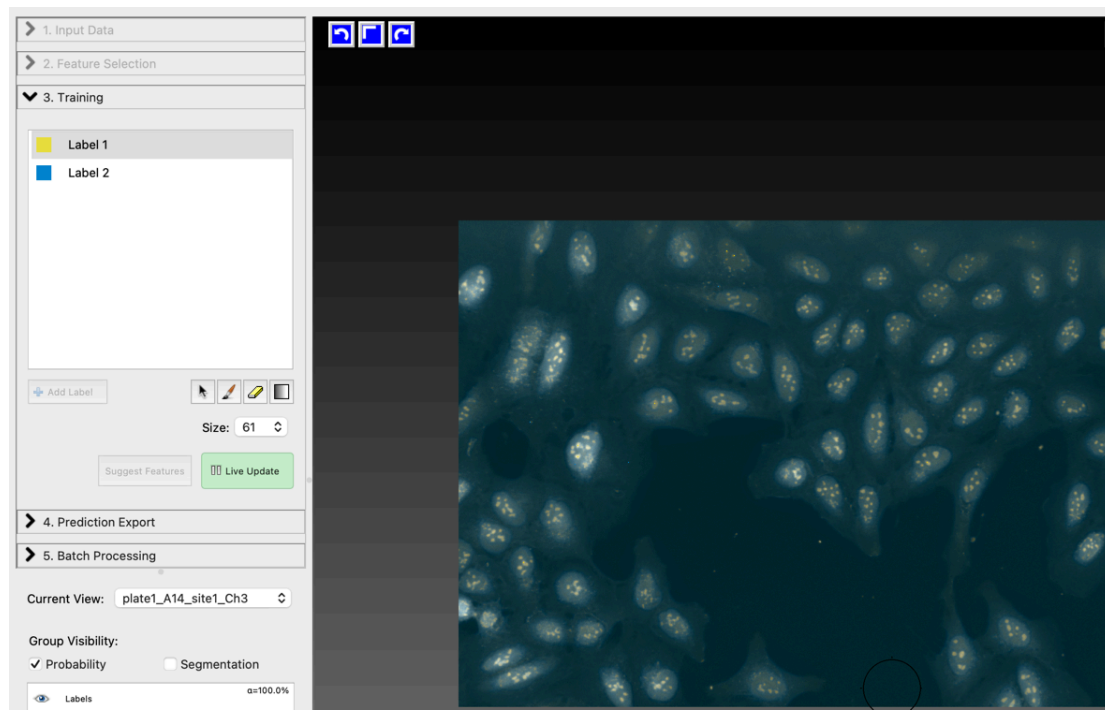
You will need either ilastik or Docker Desktop installed on your computer for this exercise. If ilastik, make sure it is CLOSED, if Docker Desktop, make sure it is OPEN. We recommend installing ilastik because it is a good and helpful tool, and will allow you to do this exercise's bonus challenge.

Warning:

If you are on Windows, execute RunIlastik in Local mode (working on an installed copy of ilastik, rather than a copy inside a Docker file), this exercise will work in TestMode, but not will not run in analysis mode - we are working with the ilastik developers to determine why that is. This is fine for the purpose of this exercise; if you have a lot of your own data you want to run later, you can still use ilastik in a two step process, and/or use Dockerized ilastik.

Only if you are curious - how did we train the classifier?

This classifier was made in ilastik 1.4.1b5 by training on 4 images (A14_site1, E18_site1, D16_site1, and C12_site1) in which we labeled just two classes: one class for nucleoli (yellow in the figure below), and one class for every other part of the image (blue in the imagebelow). One COULD have made more classes, but this, in practice, worked.



Screenshot of this ilastik classifier

When using ilastik for fluorescence microscopy, you will likely get the best performance if you **keep your annotations extremely minimal** - the classifier you're going to use was trained by using 31 total pixels of annotation across the 4 images (13 pixels of annotation inside nucleoli, and 18 pixels outside of them); no image had more than 14 pixels annotated in total. We strongly suggest you make your classifiers one pixel at a time, doing point annotations! Resist the urge to draw squiggly lines all over the image! This feels counterintuitive but we promise it's true.



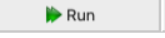
Grab the Runilastik plugin

1. Download [the plugin](#) into a folder on your local computer. As stated above, we strongly suggest a folder that contains ONLY plugins.
2. In CellProfiler's File -> Preferences menu (CellProfiler -> Preferences in Mac), set the **CellProfiler plugins** directory to the folder containing the plugin.
3. Close and reopen CellProfiler to load the plugin.

Load the classifier and evaluate it

1. Drag **bonus_2_ilastik.cppipe** into the pipeline panel.
2. Drag the **images_Illum-corrected** subfolder from the main exercise into the Images panel.
3. Open the Runilastik module, set the path to your project file (**NucleoliDetection.ilp**) and then also:

- ilastik-installed (recommended): set the path to your local installation of ilastik.
- Docker-installed: change the top setting from Local to Docker.

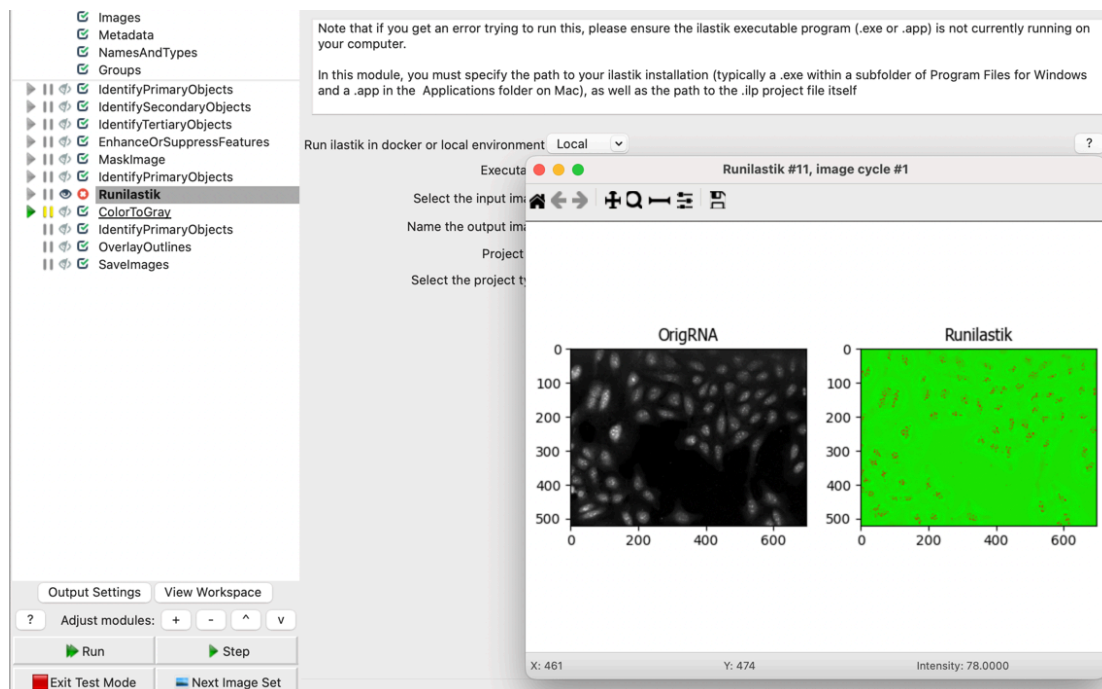
3. Put CellProfiler into TestMode , open the eye icons  next to Runilastik and OverlayOutlines, and then hit . You can check the settings in the OverlayOutlines module to see which outline color corresponds to which segmentation in the output.

- You may wish to put a pause  next to SaveImages, or uncheck it , to keep it from saving images, but that's up to you.

Note:

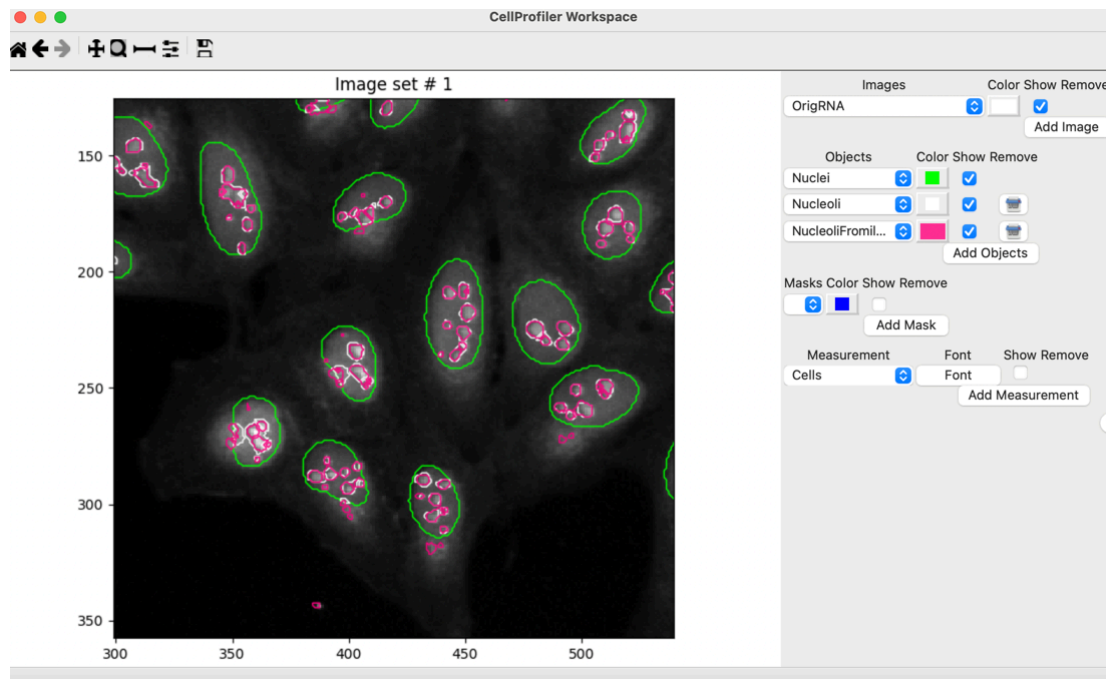
If using Docker, the very first time you execute the Runilastik module, it will need to download a ~5GB file, which may be slow depending on your internet connection. You only need to do this step once, however!

- Evaluate your prediction in Runilastik across a few image sets - how well does it perform? Does it perform worse on images it wasn't trained on?



Screenshot of the Runilastik module

- Evaluate your segmentations across a few image sets using OverlayOutlines and/or the WorkspaceViewer - are there cases where you think the ilastik model-based segmentation performs better? The filtering-and-masking segmentation?



Evaluating segmentations in the Workspace Viewer

Bonus-to-the-Bonus - train your own ilastik model

Based on your evaluations above, can you identify some places where additional training might help fix some issues in the ilastik model?

1. Open `NucleoliDetection.ilp` in ilastik.
2. Navigate to the **Training** tab.
3. Turn **LiveUpdate** on.
4. Pick an image you think needed some help and add (a very small number of single-pixel) annotations. Did things get better or worse?
5. Add some very large annotations (a long, squiggly line, for example) to one image, then switch images in order to prove to yourself large annotations can actually harm in some cases rather than help (if you want to save your classifier, go back and use the eraser tool to delete these!).
6. If you think you've sufficiently improved the prediction, save this model and return to CellProfiler. Did it help in the cases where you thought it would?

Exercise 3: Running Cellpose from a Docker container

RunCellpose is by far our most popular plugin, simply because a) Cellpose is awesome! and b) `conda` installing software when you aren't very computationally comfortable isn't. You can use the plugin in either of two modes - using a local `conda` or `python` installation that contains both CellProfiler and Cellpose, OR using Docker. The run time with Docker is substantially slower (about a minute more per image, in our testing), but if installation would take you a long time and be frustrating, in this sense you can "trade" your personal hands-on frustration time for time where CellProfiler is running on your computer (but you aren't there). For many people, this is a good trade!

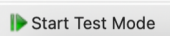


Start Docker Desktop

1. If you have not already installed Docker Desktop from the link above, please do so! This may involve rebooting your computer.
2. Start Docker Desktop.
3. Optional but strongly recommended - once Docker Desktop opens, use the search functionality to search **Cellpose** and look for the **cellprofiler/runcellpose_with_pretrained** model and pull it. If for whatever reason this isn't working, move on, but it will save you some time later.

Grab the RunCellpose plugin

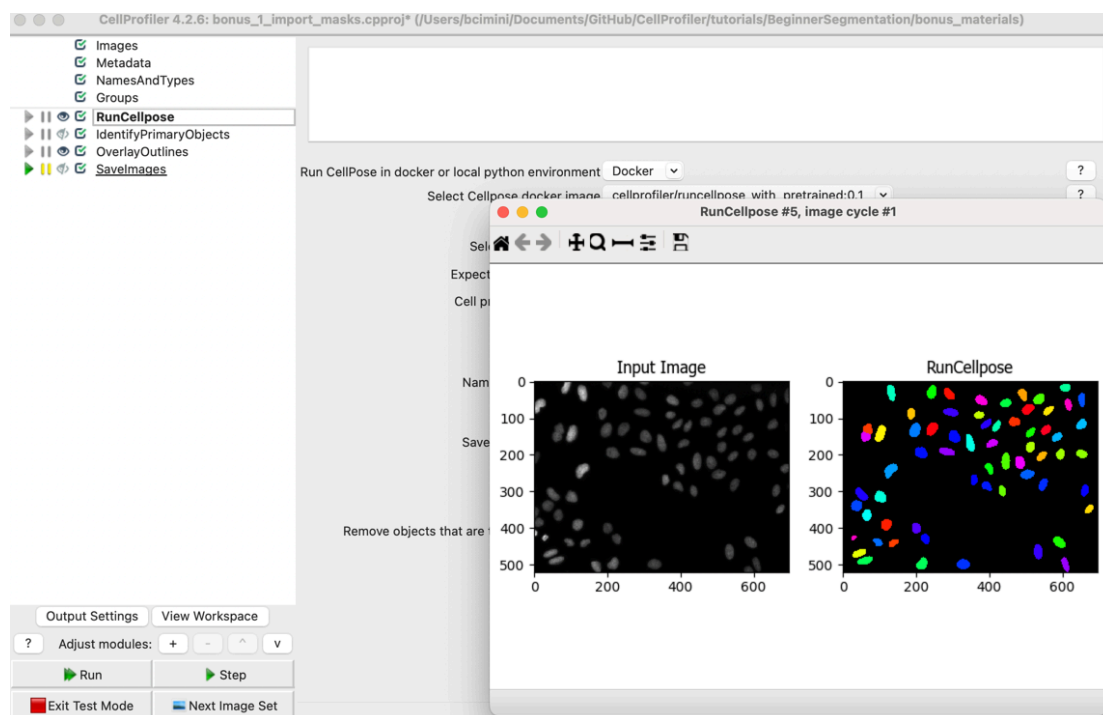
1. Download [the plugin](#) into a folder on your local computer. As stated above, we strongly suggest a folder that contains ONLY plugins.
2. In CellProfiler's File -> Preferences menu, set the **CellProfiler plugins directory** to the folder containing the plugin.
3. Close and reopen CellProfiler to load the plugin.

Load the pipeline and evaluate segmentation

1. Drag **bonus_3_cellpose.cpipe** into the pipeline panel.
2. Drag the **images_Illum-corrected** subfolder from the main exercise into the Images module
3. Put CellProfiler into TestMode , open the eye icons  next to RunCellpose and OverlayOutlines, and then hit 
 - You may wish to put a pause next to SaveImages, or uncheck it, to keep it from saving images, but that's up to you

Note:

If you didn't already pull the container in the Docker Desktop section above, the very first time you hit the RunCellpose module, it will need to download a ~13GB file, which may be slow depending on your connection. You only need to do this step once however!

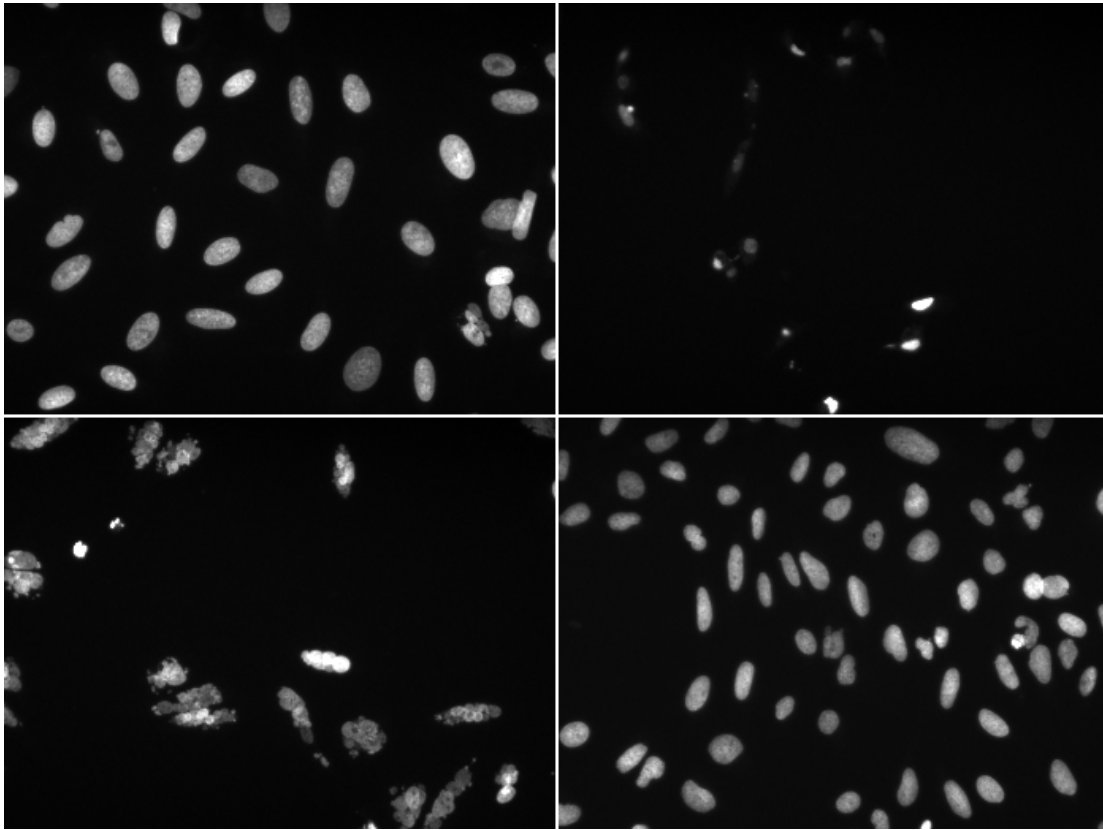


The output of the RunCellpose module

4. As before, using OverlayOutlines and/or the WorkspaceViewer, evaluate segmentation on a few images. Where is CellProfiler doing better, and where is Cellpose doing better?
5. Use the [?](#) button to learn more about the different parameters you can pass to Cellpose (we don't offer all of them, but many!) - how does tweaking these affect your output? How about changing the model you're using, and/or the image you're segmenting?

Bonus-to-the-Bonus 1 - test on a more varied data set

The data we gave you in this exercise was all very clean, and all from the negative control wells of this experiment. The [folder linked here](#) contains images with a broader variety of phenotypes - try replacing the images in the **Images** module with this folder instead and see how well the two kinds of segmentation work in different conditions!



A broader range of phenotypes can be harder to consistently account for in conducting segmentation.

Tip:

You're encouraged to explore the performance in some random images on your own (you can do this from the **Test** menu with the **Random Image Set** option), but if you find yourself consistently ending up with images that look similar you can try examining images from the following list of wells- A08, A12, B12, B18, C7, D6, D19, D22, E3. You can pick these with **Test -> Choose Image Set**

Bonus-to-the-Bonus 2 - improve Cellpose segmentation

Based on your evaluations above, can you identify some places where additional training might help fix some issues in the Cellpose model? If you happen to have Cellpose on your computer (or want to try installing it - see Exercise 1 for a link to the official install instructions), open Cellpose with some of the Ch1 images and see if you can train a model that will perform better.

Bonus-to-the-Bonus 3 - test your classical segmentation skills

The classical segmentation parameters provided here are based on expert tuning from a very experienced CellProfiler user - how close can you get? Add another IdentifyPrimaryObjects module to the pipeline, tune it yourself, and then look in the workspace viewer to see how it performs, especially in the more diverse image set from Bonus-to-the-Bonus 1. Compare your parameters to the expert ones - how do they differ? Do you understand why some of these settings may have been chosen?

What next? Want to know more about CellProfiler plugins and modules?

1. Read the [CellProfiler-plugins paper](#)
2. Watch this [video](#) to learn how to write a module