Exercise

1. Materials

You will work with only 48 images. These images are located into /exercises/rf/cp/images. The plate map is a standard 96-well plate, containing 12 concentration points in columns and 4 replica rows. In addition to these images, a text file called "dose.csv" is provided, containing information about where on the 96-well plate the wells were located, and how the cells were treated. This file is located into /exercises/rf/cp/metadata.

2. Exercise: Using the CellProfiler software to identify features and obtain measurements from cellular images

a) Start CellProfiler and Open Project... name *translocation_activity.cpproj* (located into /exercises/rf/cp/).

b) By clicking on the module *Images*, you can select your images or folders containing the files.

c) Click on the module *Metadata*; this module allows you to extract information from your images or external file. Change the path next to the Sub-folder to select the dose.csv file.

d) Your pipeline is now ready to run analysis over all your 48 images.

3. Exercise: Using the CellProfiler Analyst software to visualize the data and classify cells

3.1 Starting CellProfiler Analyst and loaded the properties file

a) Start CellProfiler Analyst (Figure 1) and *Import* your *properties file* which you created along with your database in the *ExportToDatabase* module.

3.2 Visualizing the measurements in a 96-well plate layout view

In this exercise we will only focus on two of the tools: Plate Viewer and Classifier.

a) Click-on *Plate Viewer* (Figure 2) to display a 96-well plate with the feature values for each well.

b) You can visualize the data per object with *MyExpt_Per_Object* and you can display the ratio of GFP in Nuclei to GFP in Cytoplasm calculated previously by selecting *Cytoplasm_Math_IntensityRatio* under Measurement.

> Do we have homogeneous values or heterogeneous values for this measure?
> If heterogeneous what are the wells that have the lowest values?

c) In this step, you will see under "Well display" the image thumbnails.

> Do you have an idea of the wells that have mostly a positive cell phenotype?

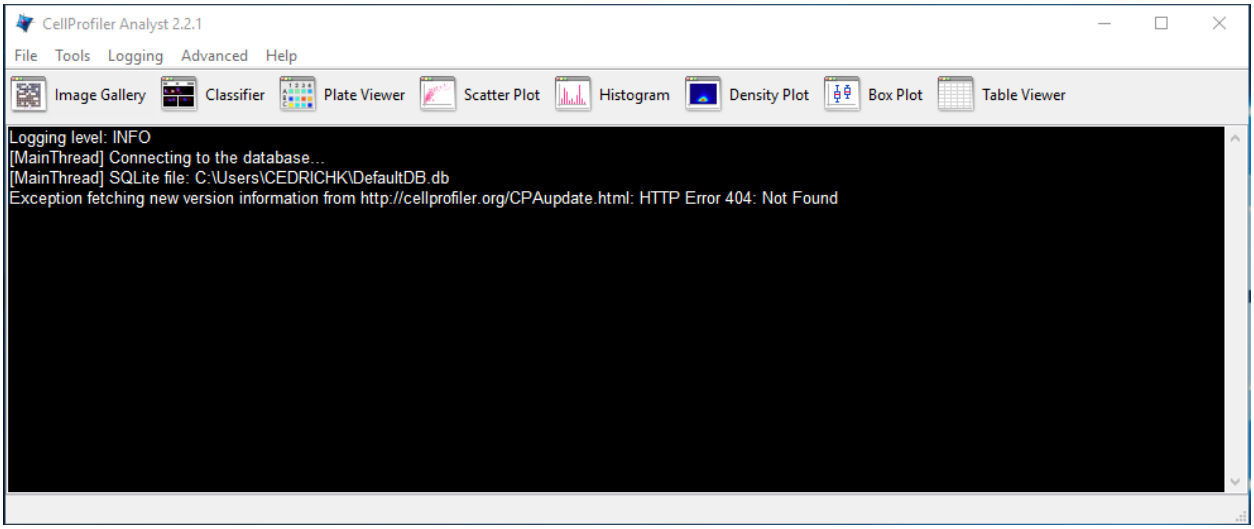

Figure 1: CellProfiler Analyst Interface

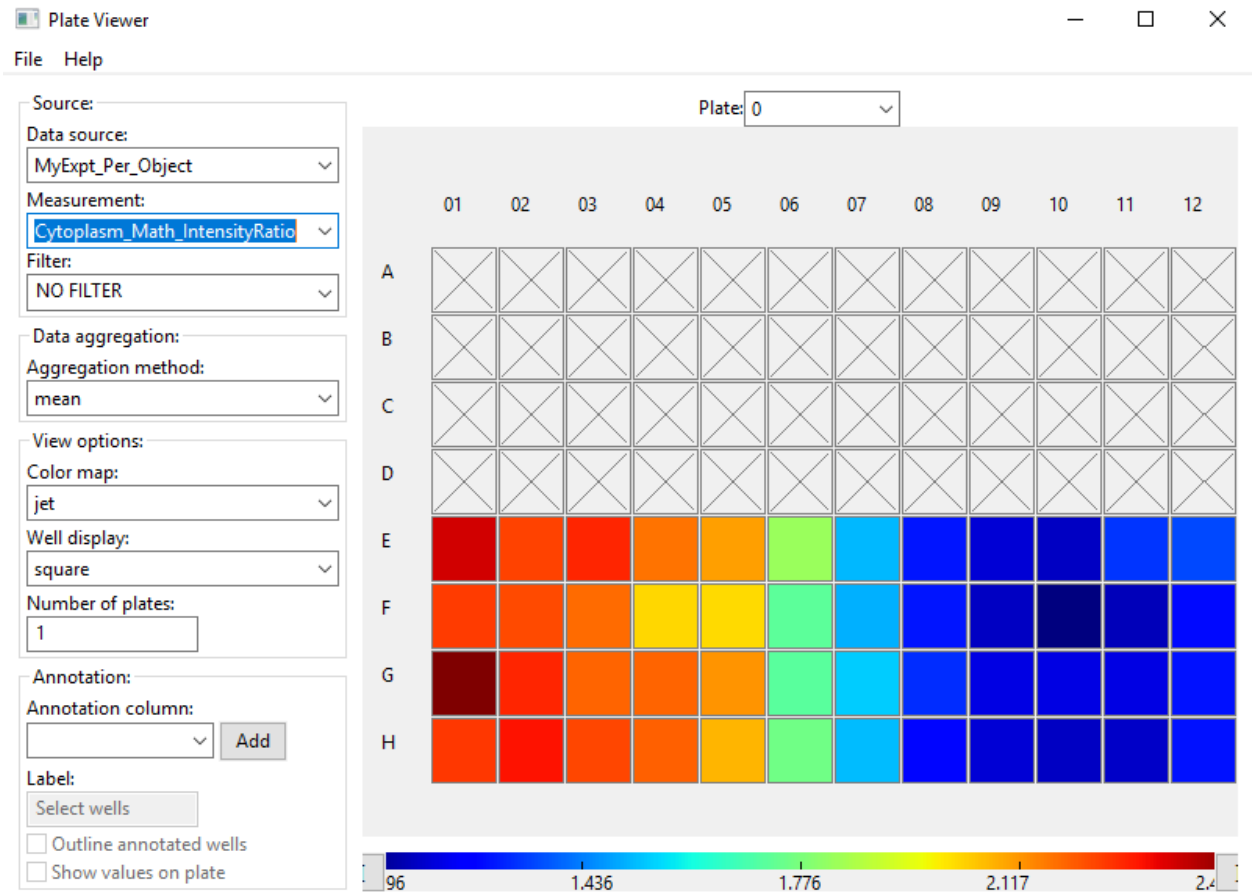

Figure 2: Plate Viewer Interface

3.3 Use the Classifier to recognize phenotypes

a) Click-on the *Classifier* (Figure 3) to classify your negative and positive cells.

b) Click on the *Fetch!* button to display randomly pictures from this experiment.

c) Manually, *drag* each thumbnail to get at least *5 cells in the positive bin* and *5 cells in the negative bin* and click on the *Train* button. A small dot is displayed in the thumbnail that show which object will be used for classification.

3.4 Reviewing the rules that determine positives and negatives cells

a) By default, it displays the top 5 of the features.

Write the top 5 features.

If you do not have interesting features, you can add images in each bin and start the training again.

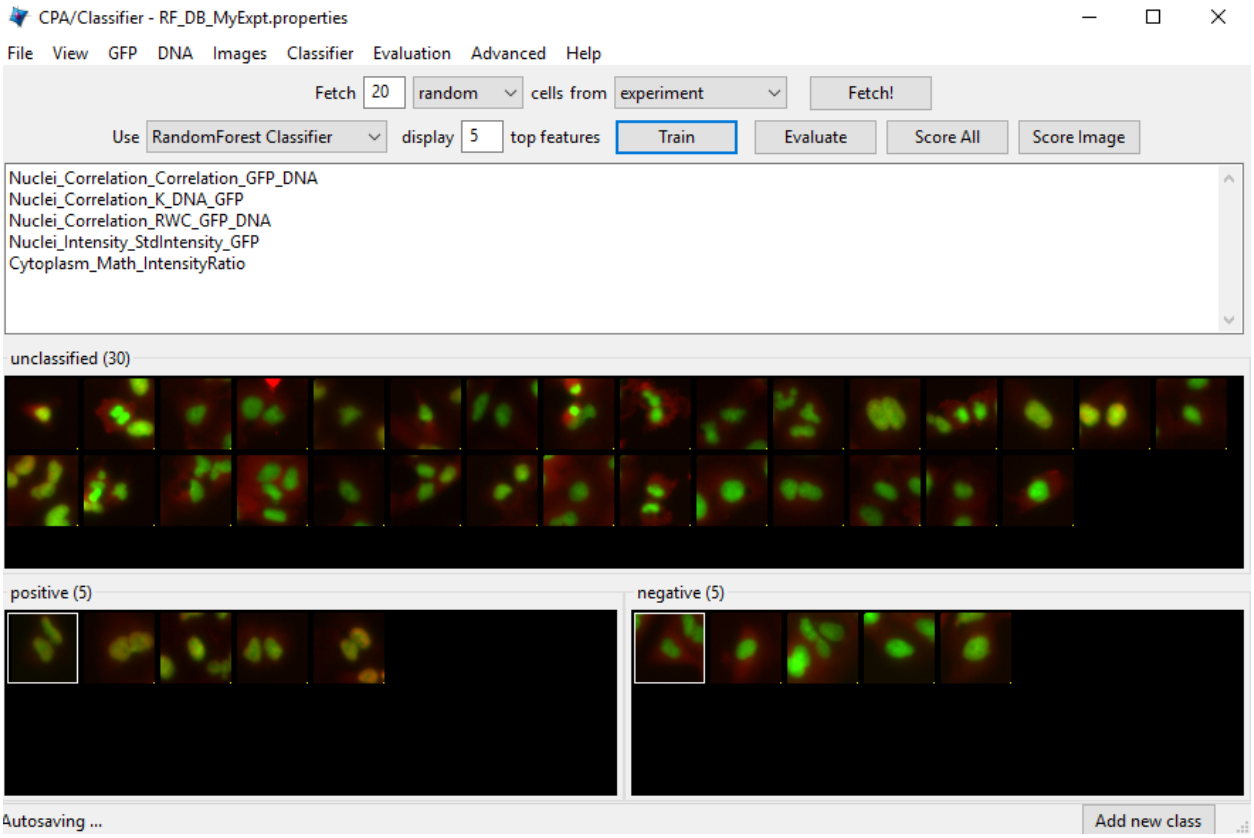

Figure 3: Classifier interface

3.5 Reviewing the accuracy of the classification with the confusion matrix

a) Press the *Evaluate* button to generate a confusion matrix for the cells you've classified so far.

How accurate is your classification after adding only a few cells to your training set?
What is the percentage of accuracy ? Note it !

3.6 Refining the training set by sorting more "unclassified" cells into the "positive" and "negative" bins

Based on the evaluation, you can choose to sort more cells or skip to the *3.7 Classifying all cells in the experiment*.

3.6.1 Refining the training set by obtaining samples from positive and negative control wells

a) Open the *Plate Viewer* and open an image from the negative controls. Click on a cell in the image that is negative for the phenotype and drag-and-drop it into the negative bin.
Repeat the above for a well containing a positive control sample, dropping the cells into the positive bin. Click the *Train* button.

b) Repeat the steps above to obtain more accuracy.

3.6.2 Refining the training set by correcting misclassified cells in an image

a) Double-click any of cell thumbnails in the positive or negative bins. From the image that opens click on *Classify Image* (Figure 4).

b) Look for up to 5 cells that are clearly misclassified. For each of these cells that you find, click on it and drag-and-drop it into the appropriate bin, then click on *Train* button again

c) Evaluate your model.

3.6.3 Refining the training set by fetching positive and negative cells

a) you can now request that the computer fetchs more examples of positive and negative cells. These new sample cells can be added to the corresponding bins, in order to improve the classifier's performance.

b) Repeat this step until you have at least 20 cells in each bin.

> Has the accuracy of the model increased ? No, Go back to the first step and repeat, until the classifier displays the desired level of accuracy.

3.7 Classifying all cells in the experiment

a) If your are satisfied with the result, then you can press the *Score All* button. A "Hit table" window will appear containing the summarized scores for every image. The "Total Cell Count" is reported, as well as the number of positive and negative cells classified. The last column is the enrichment score.

> Which wells have a very high enrichment score? What is the minimum concentration that generate the expected effect?

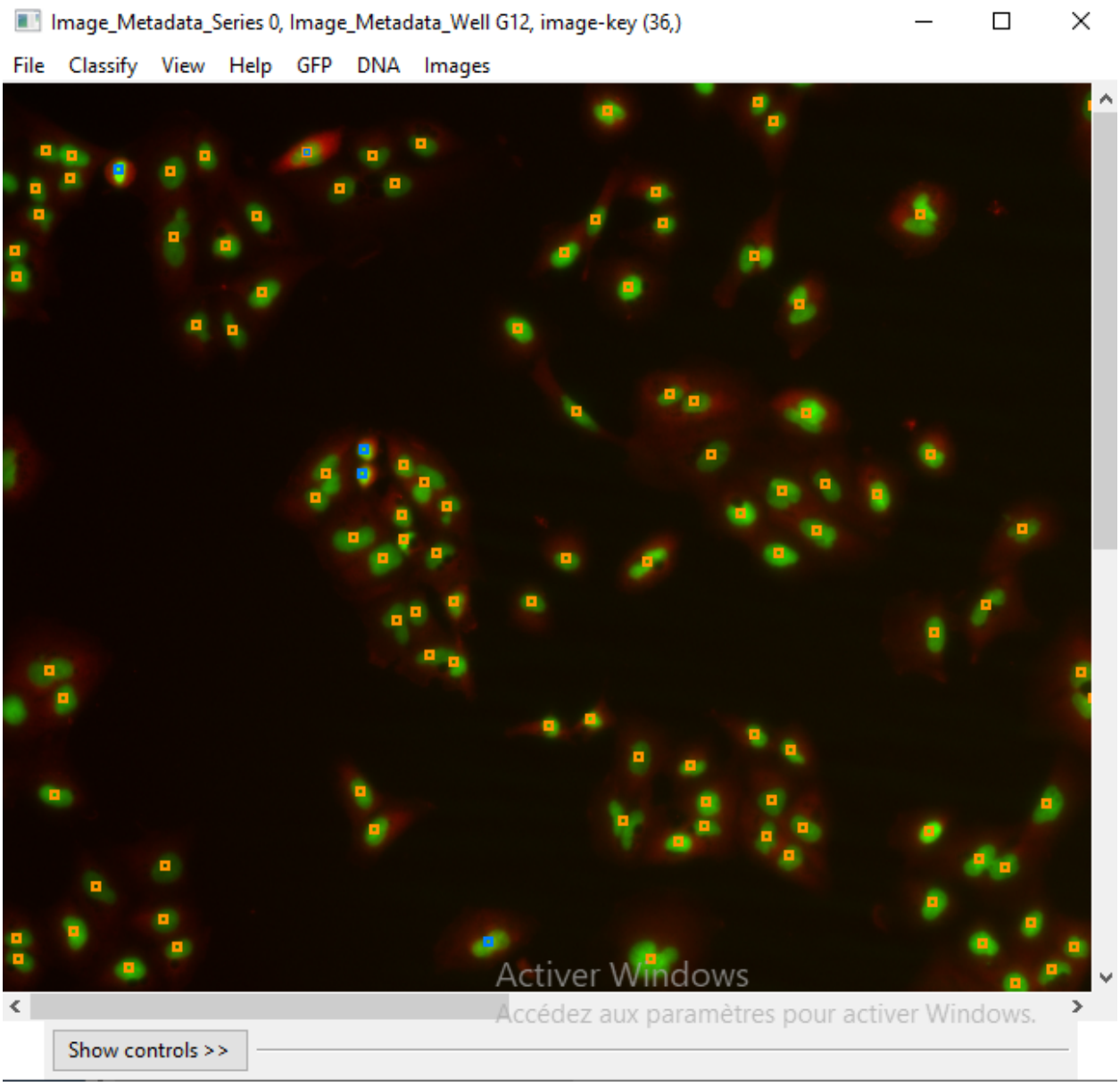

Figure 4: Classify the image