

## IdentiCyte Manual

IdentiCyte has been developed to analyse red blood cells in images taken in bright field image microscopy. It uses a user defined library of cells to identify cells in the image and allows for two different ways of interpreting results. The first, way that the results can be viewed is in the Excel document that is output after the analysis is run. The second method is by inspecting individual cell classification. Both methods will be addressed in detail in this document. We strongly recommend reading this entire document through before attempting to analyse any images.

### Requirements

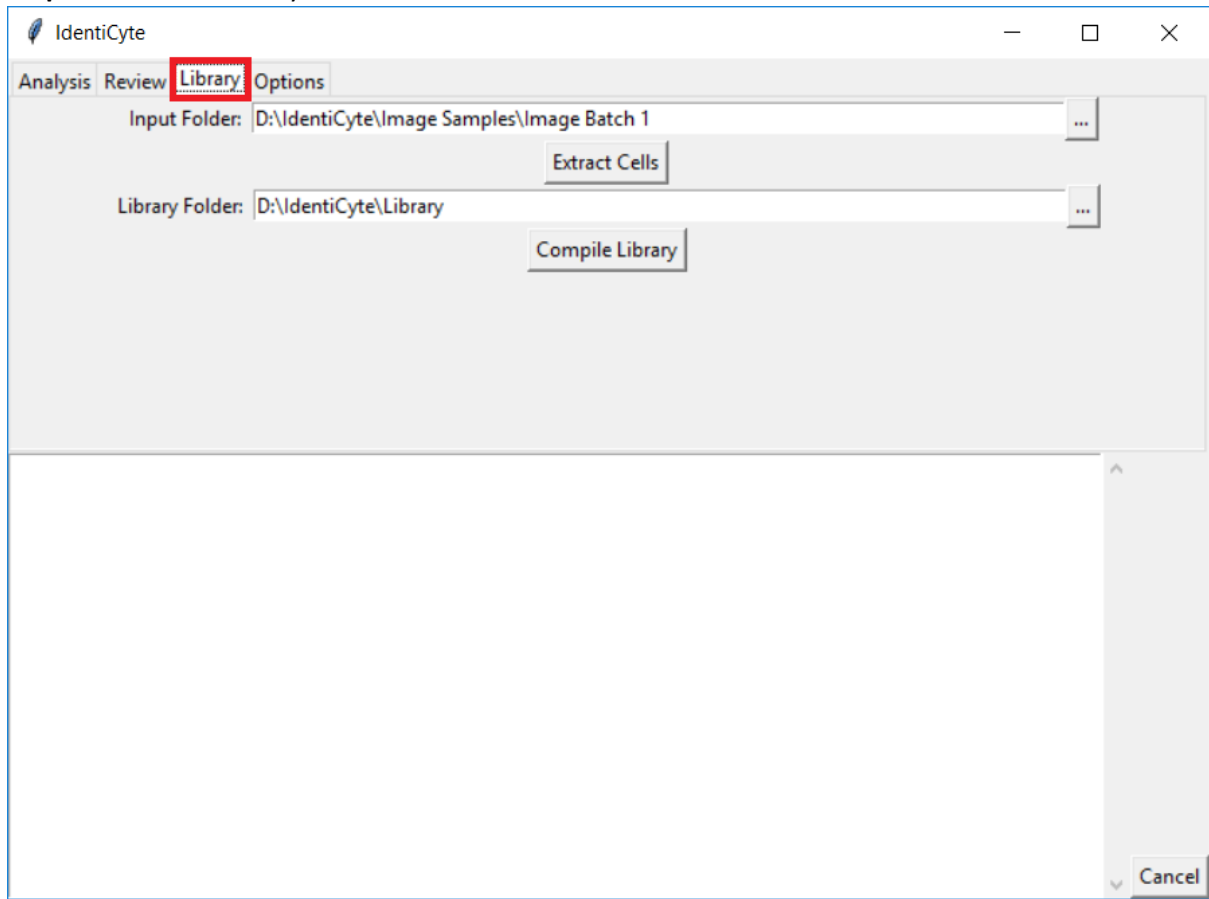
We recommend x60 or x100 objective. Best results can be obtained when cells in an image overlap as little as possible and are well spread out. All images should be taken with the same microscope settings and illumination. The images can either be TIFF, PNG or JPEG formats. TIFF stacks are not supported and each image should be a separate file. In order to use the results from this software effectively, Microsoft Excel or a suitable replacement is needed. If Excel is not available, we recommend Libreoffice Calc. When running IdentiCyte, it may take a few seconds for image analysis window will appear.


### Constructing the Library

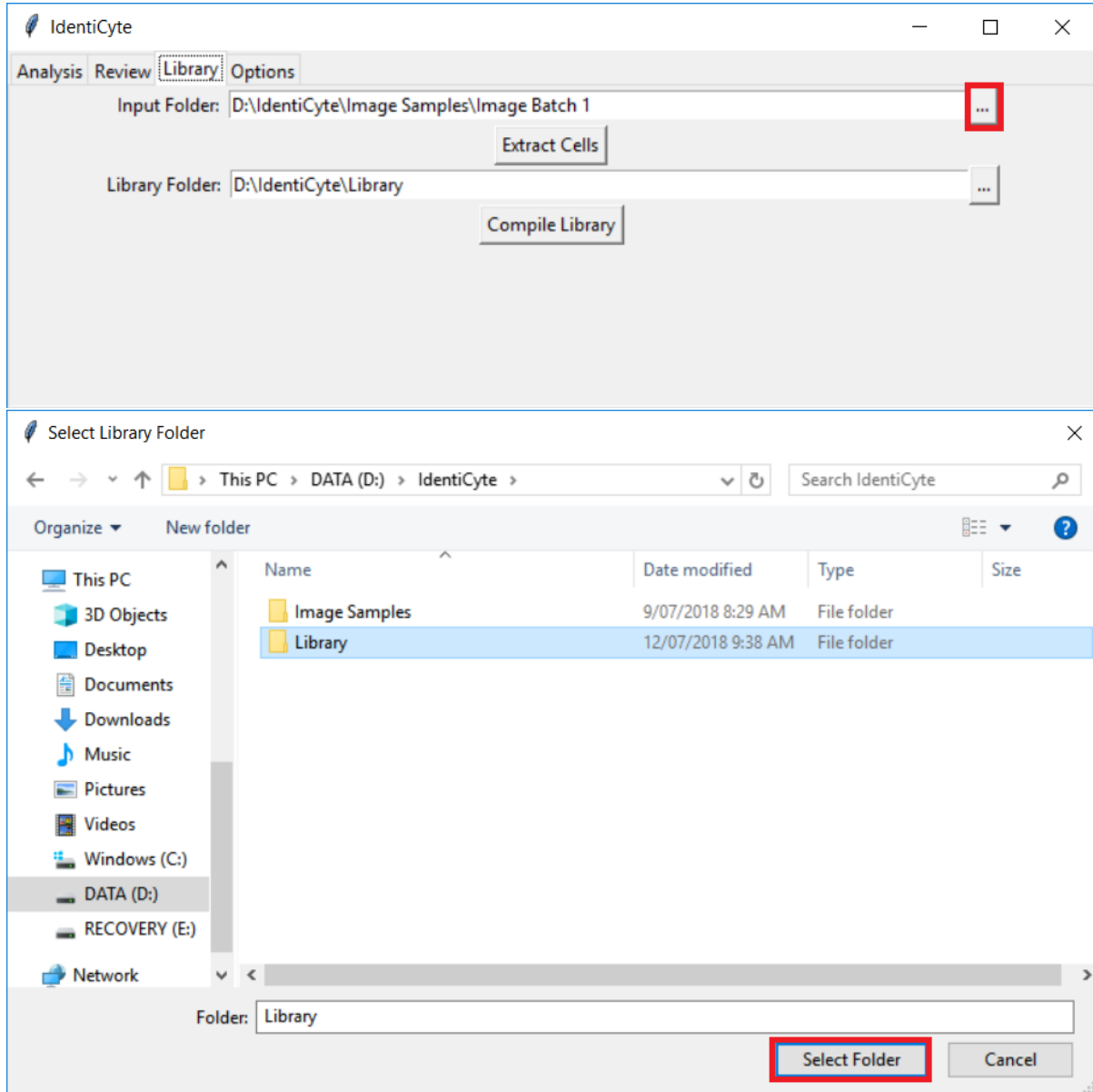
Before any analysis can begin, a library of cell examples must be constructed. A default, preconstructed library has been included with IdentiCyte, but users may wish to create their own. A new library should be construct whenever images taken with different microscope settings are to be analysed, when analysing image with different resolutions or when different categories of cells are desired. It is important that all images in the library are the same dimensions and must also all have the same number of colour channels. When creating new categories the first eight letters of each category should be distinct from the others, as while it will not affect program's execution, only the first eight letters of a category are shown when viewing the classifications overlayed onto the cells. There are two stages to constructing the library: extracting the cells and compiling the library. When creating or adding to a library, we strongly recommend using the Extract Cells method described below.

**Step 1** – If IdentiCyte is not already open, double click the IdentiCyte file to run it.

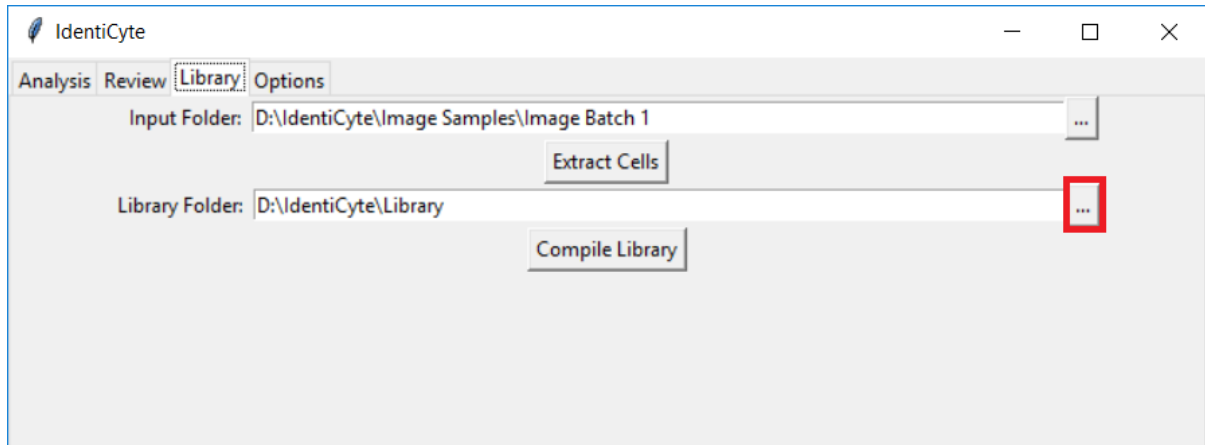
**Step 2** – Click the Library Tab.



**Step 3** – Select the Library folder to be used. If the path in the box is the desired library, this step may be skipped. This can be done by clicking the  button next to the Library Folder input box. This will bring up a window to find the Library Folder. To select the library folder navigate to it in the folder select window and click Select Folder.



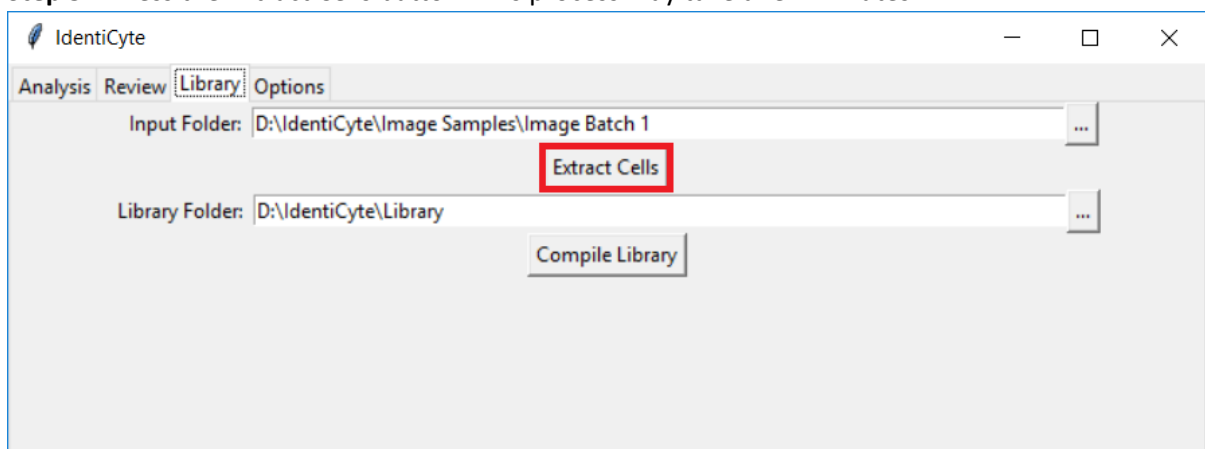
**Step 4** – Select the folder which contains the images from which cells are to be extracted. This is done in a similar way to selecting the Library, except that the folder to be selected is the one that holds the images with the cells to be extracted. This step may also be skipped if the displayed path contains the desired images. The Input Folder is only a required field for Extract Cells, the Compile Library Function does not use it.



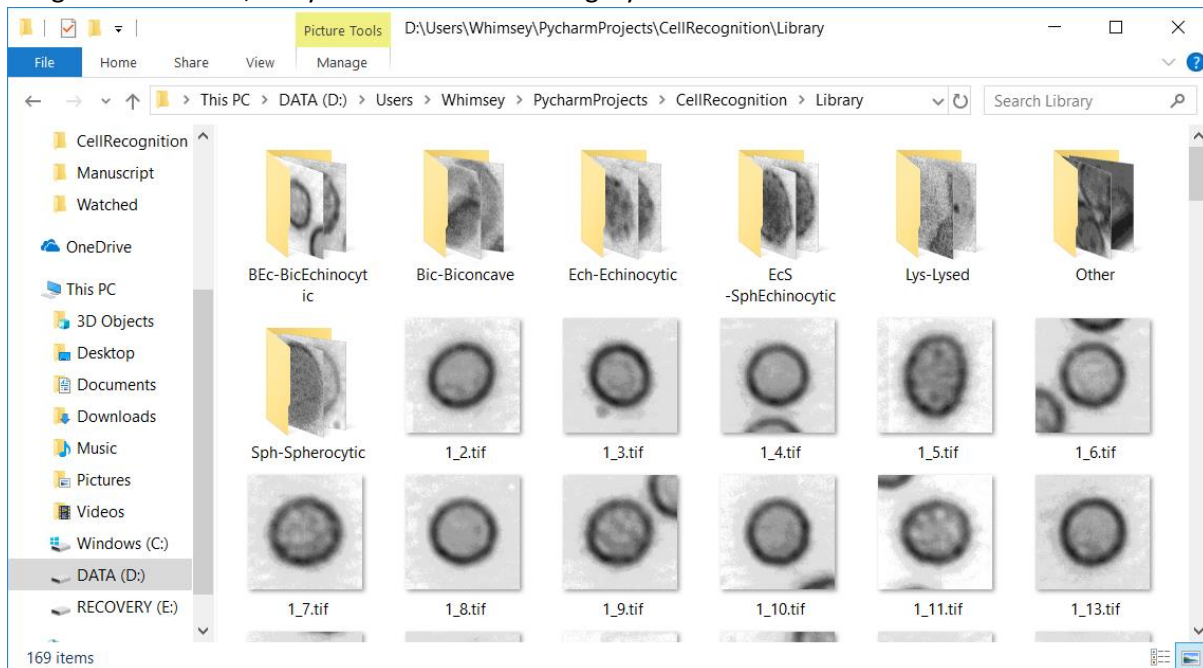
#### Extract Cells

This functionality is the recommended way of obtaining cell examples with which the library may be constructed.

**Step 5** – Press the Extract Cells button. This process may take a few minutes.



**Step 6** – Once the cell extraction has been finished, open the library folder. There should be many images in the folder, not yet sorted into the category folders.

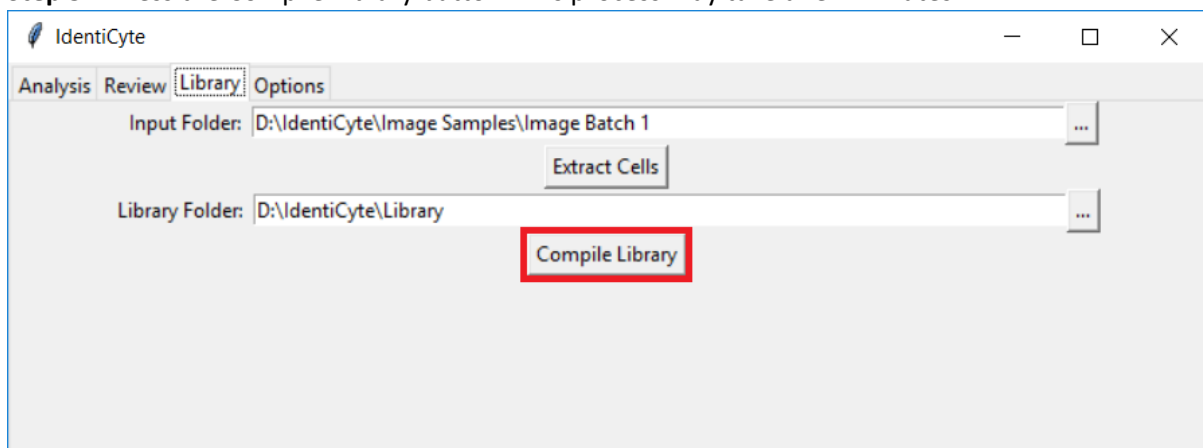


**Step 7** – Drag cell images into the folders which represent the categories to which they belong. After this, the library must be compiled before any analysis can be performed.

### Compile Library

The library must be compiled to be used for classification. This should be done any time after new cells have been added to the library or a new library has been created as described above and after every five times the analysis has been run with User Verification Enabled. To compile the library, only the Library Folder is required, the Input Folder Field is used for the cell extraction process.

**Step 5** – Press the Compile Library button. This process may take a few minutes.



## Analysing images

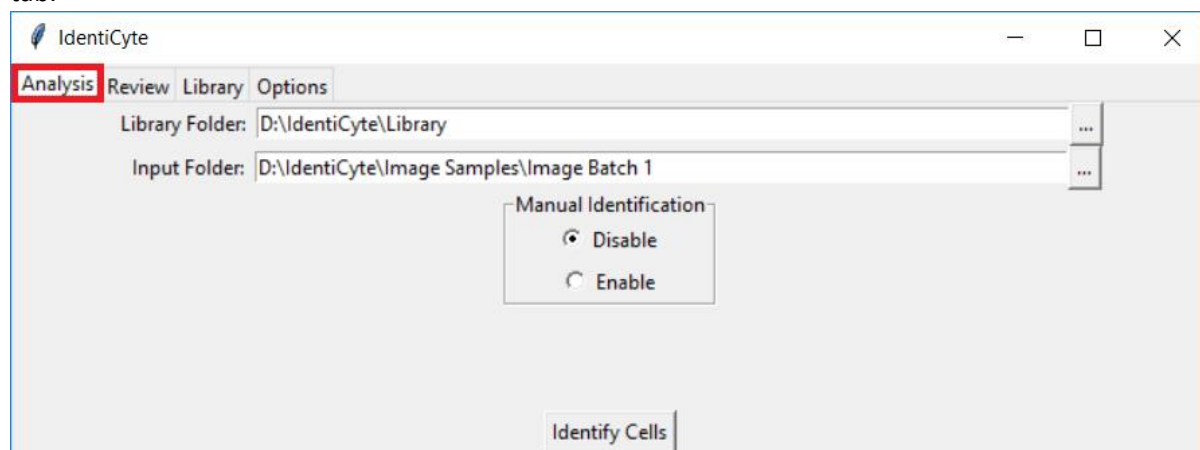
This section describes the main function of IdentiCyte, the identification of cells in a microscope image. All cells are classified with a confidence percentage which states how certain IdentiCyte is that the cell is the type it has been identified as instead of any other.


Cells can be identified with or without User Verification. If User Verification is enabled, then when a cell's confidence is too low IdentiCyte will bring up a window asking the user to identify the cell for it. Any cells identified in this way will be added to the library and will be considered in future identifications once the library is compiled again.

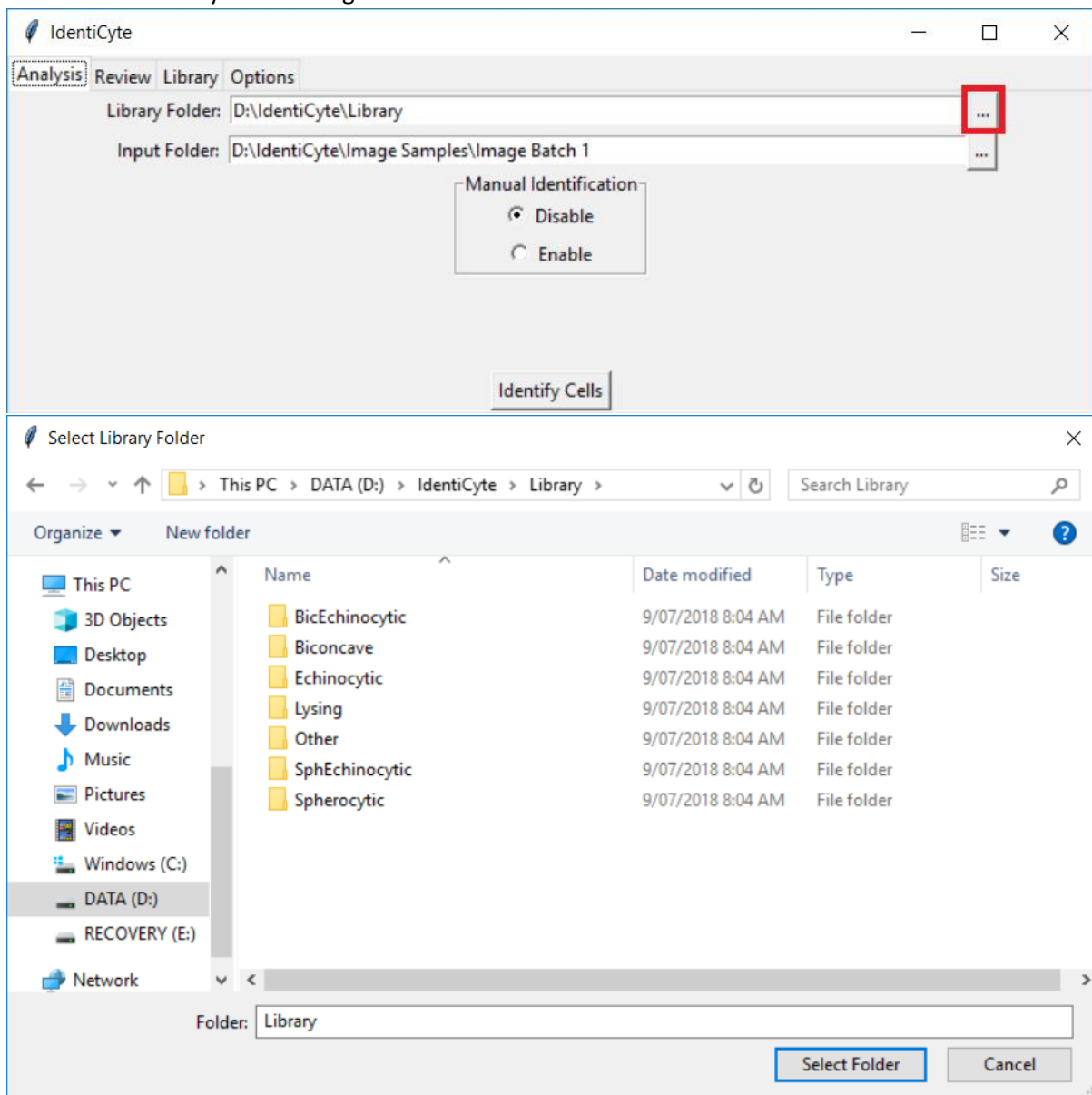
Once the analysis has been completed, two files are created in the folder containing the images that were just analysed. The first is an Excel spreadsheet with the same name as the analysed folder that gives an overview of the cells analysed and the second is a file called "Identified Cell Info" which contains the location of each cell in each image analysed along with its classification and the confidence of that classification.

The Excel spreadsheet is comprised of two sheets. The first sheet is a summary sheet, which gives the total number of cells in each category, and the percentage of Identified cells that a category represents. Cells that are too close to the edge to be analysed, or that are classified as 'Other' will not be included in this percentage. The second sheet gives a breakdown of cell classifications per image. The percentage of cells in that image of each type is given as well as the mean confidence of all cells in that image. Cells on the edge and 'Other' cells are not excluded from this statistic.

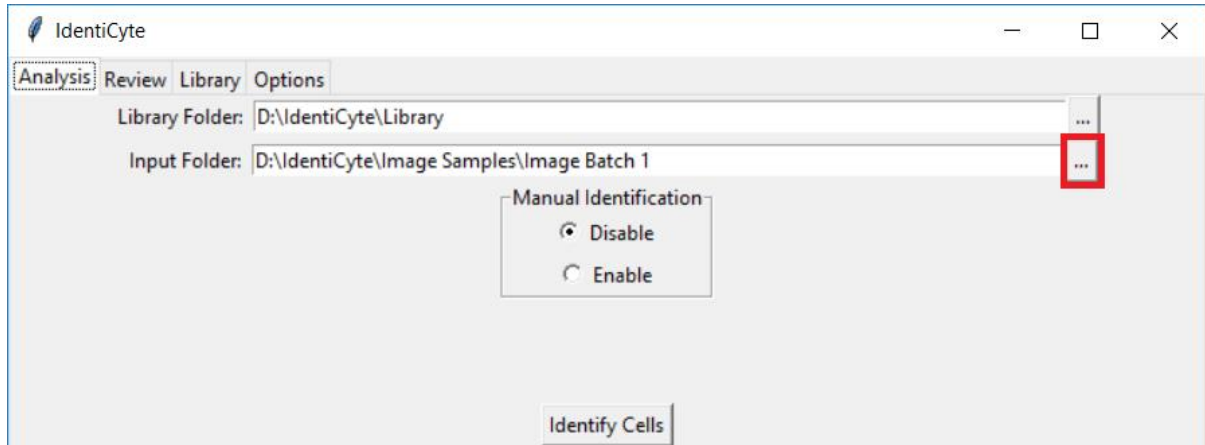
**Step 1** – If IdentiCyte is not already open, double click the IdentiCyte file to run it. Select the Analysis tab.



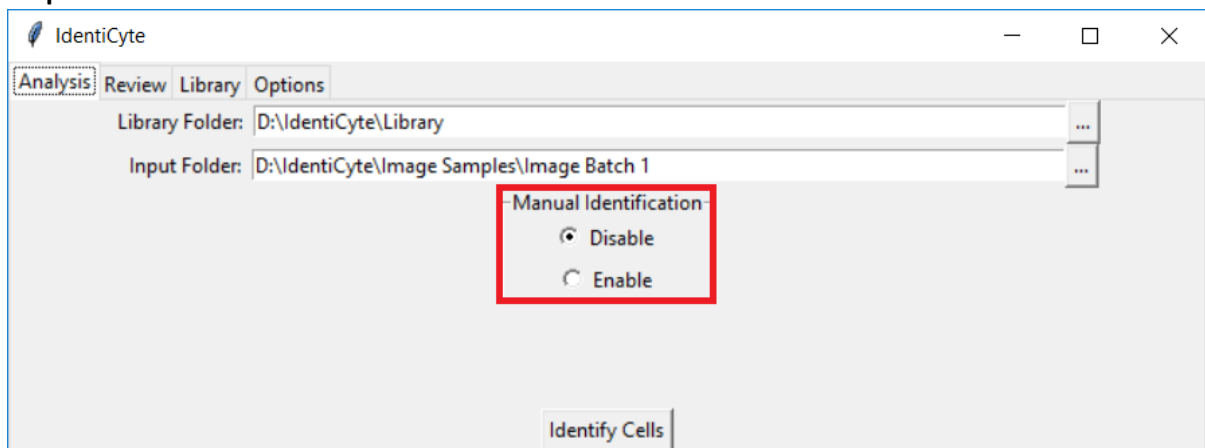
**Step 2** – Select the Library folder to be used. If the path in the box is the desired library, this step may be skipped. This can be done by clicking the  button next to the Library Folder input box. This will bring up a window to find the Library Folder. This will bring up a window to find the Library Folder. To select the library folder navigate to it in the folder select window and click Select Folder.



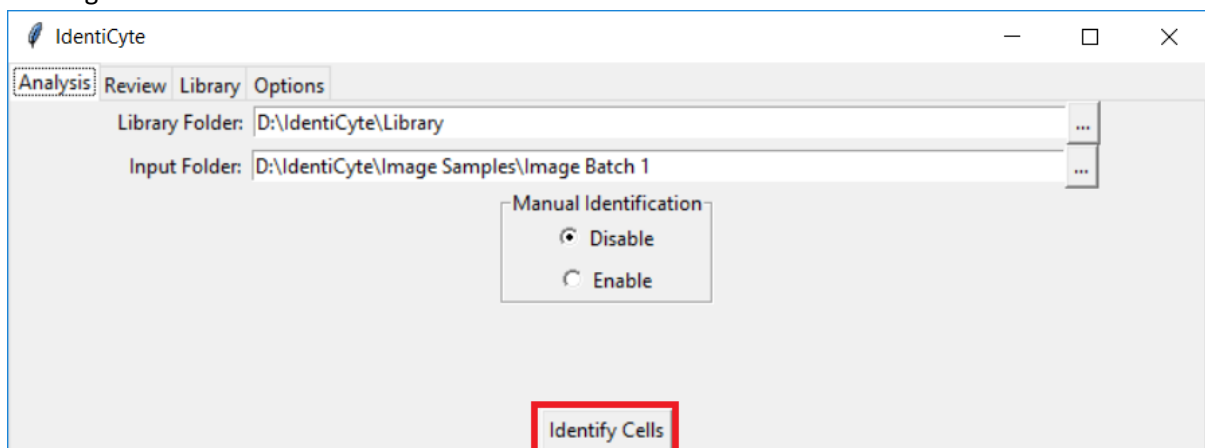
**Step 3** – Select the folder which contains the Images to be analysed.



**Step 4** – Enable or disable User Verification.



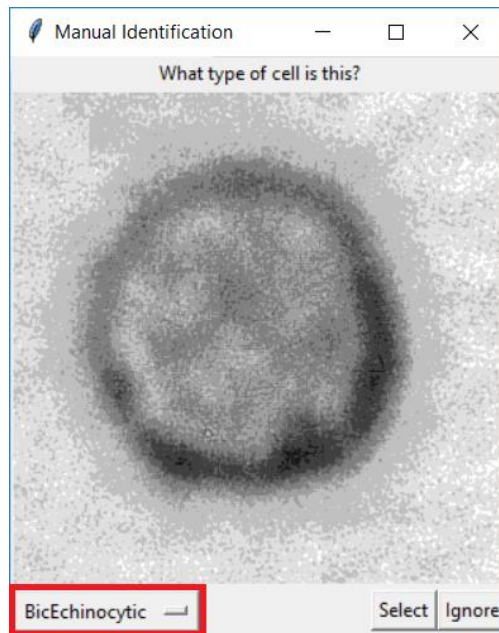
**Step 5** – Click the Classify Cells button. The program will then run, displaying its current operation in the large text box in the bottom centre of the window



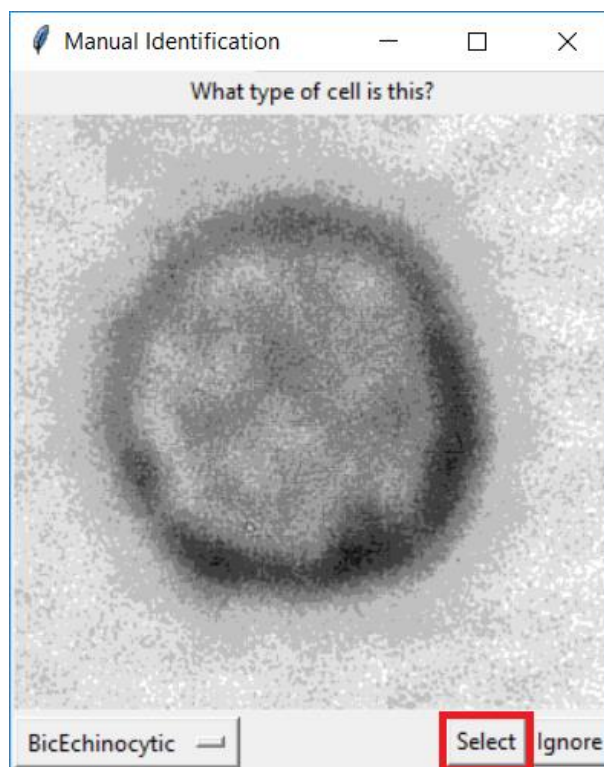
**Step 5a** – If user verification has been enabled then a window (shown below) will pop up and ask the user to identify cells that the program is unsure about.

**Step 5b** – The user verification window pops up when the confidence for a cell is below the user defined threshold in the options tab. The type can be manually identified with the dropdown menu in the bottom left corner of the window.

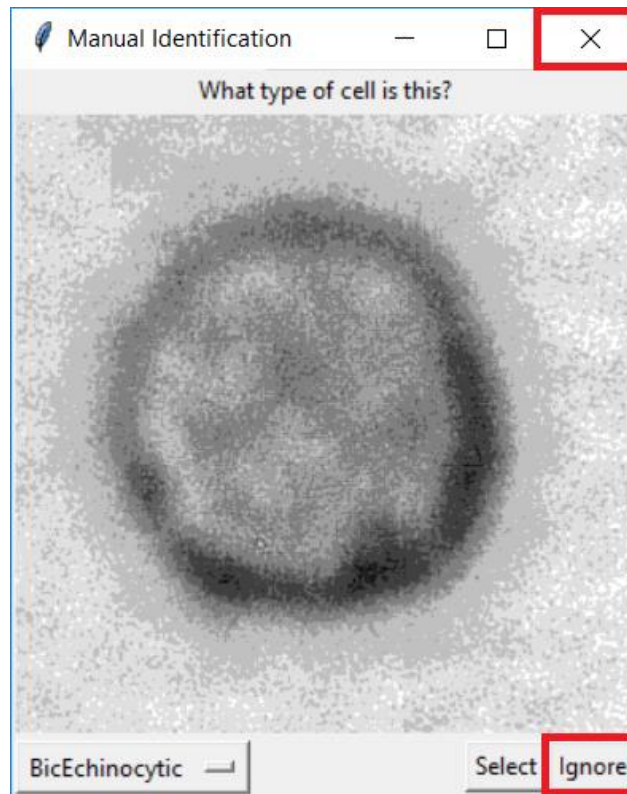




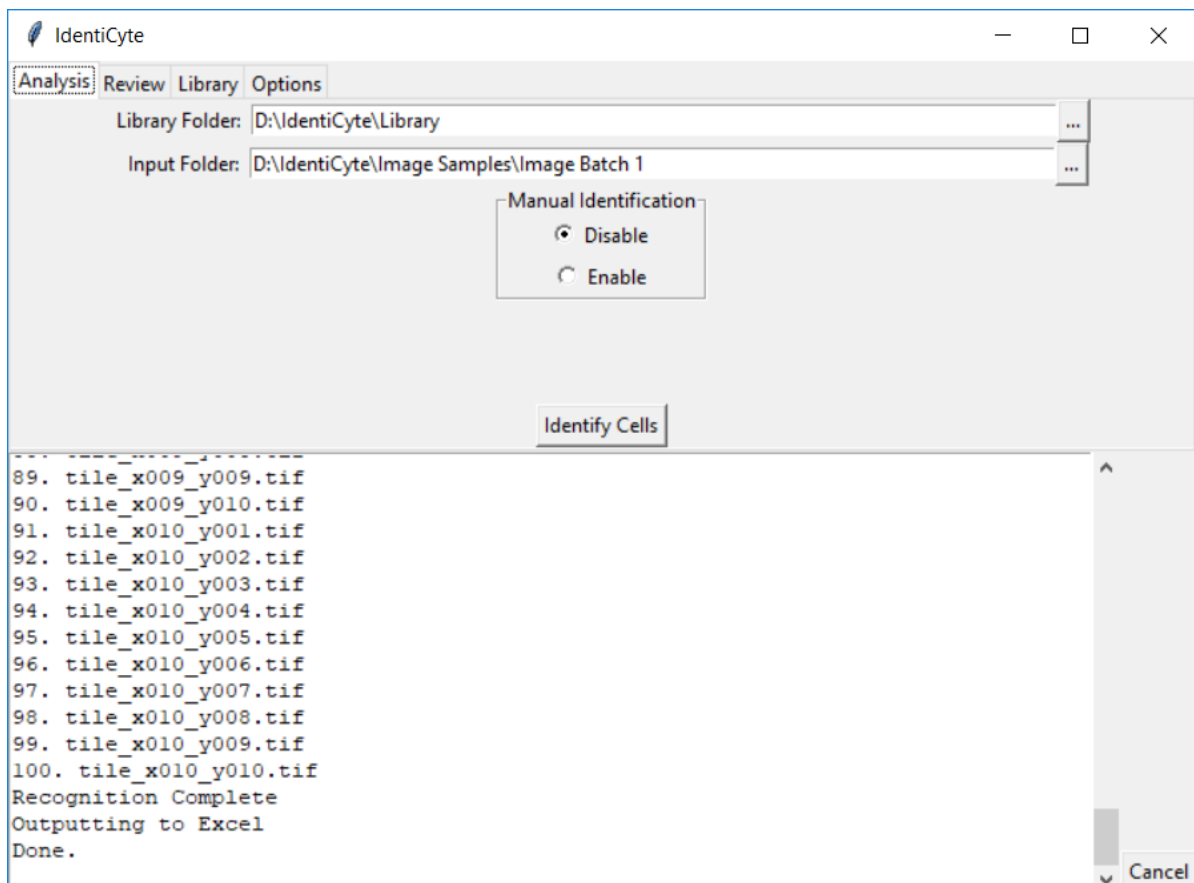
**Step 5c** – To confirm the selection, click the Select button in the bottom right corner. This will save the classified cell to the relevant folder in the library but will not add it to the compiled library.



**Step 5d** – If the cell is not to be classified, it can be ignored by clicking the ignore button in the bottom right corner or closing the window with the 'X' in the top right corner. Both options will close the User Verification window without a categorisation. An ignored cell will not be saved to the library.



**Step 6** – When the analysis is finished and the two files containing the results have been created, the program will display ‘Done.’ in the large text box at the bottom of the window. At this point the Excel document may be viewed and the classifications of individual cells may be verified in the Review Results tab.



## Reviewing Results

After a batch of images has been analysed then two files are created: an excel file in which a summary of the analysis can be found, and a file called 'IdentifiedCellInfo' which contains the locations, identifications and confidences of every cell analysed. The IdentifiedCellInfo is only used by the Review Tab.

### In the Output Excel Document

The Excel file contains two tabs: Summary and Per Image Breakdown. The summary tab lists the name of the image set in the top left corner, then gives a summary of the number of cells of each type sample and the total number of cells and the percent of all Identified cells. The 'Other' and 'Edge' categories are excluded from this number. Cells on the edge cannot be reliably classified, and the 'Other' category is provided to discard unwanted cells. This can be done by placing the cells in the Other folder in the library as with any category. Below that can be seen the mean confidence of all the cells across all the images in the sample. Finally, the user settings are shown.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Dataset Name	Image Batch 1											
2	Library Categories	BicEchinocyt	Biconcave	Echinocyt	Lysing	SphEchino	Spherocyt	Other	Edge				
3	Cells Per Category	14	5	61	0	0	327	69	77		553	Total Cells in Batch	
4	Percent of Identified Cells	3.4 %	1.2 %	15.0 %	0.0 %	0.0 %	80.3 %	-	-		407	Cells Identified	
5													
6	Mean Confidence of all Images	83.0 %											
7													
8													
9													
10	Image Location	D:\IdentiCyt\Image Samples\Image Batch 1											
11	Library Location	D:\Users\Whimsey\PycharmProjects\CellRecognition\Library - Copy											
12	User Verification	Disabled											
13	Bit Depth	3											
14	Color Channel	B											
15	Thresholding Method	Otsu											
16	Minimum Cell Size	9000											
17	Confidence Percent	50											
18	Principal Component Percent	90											

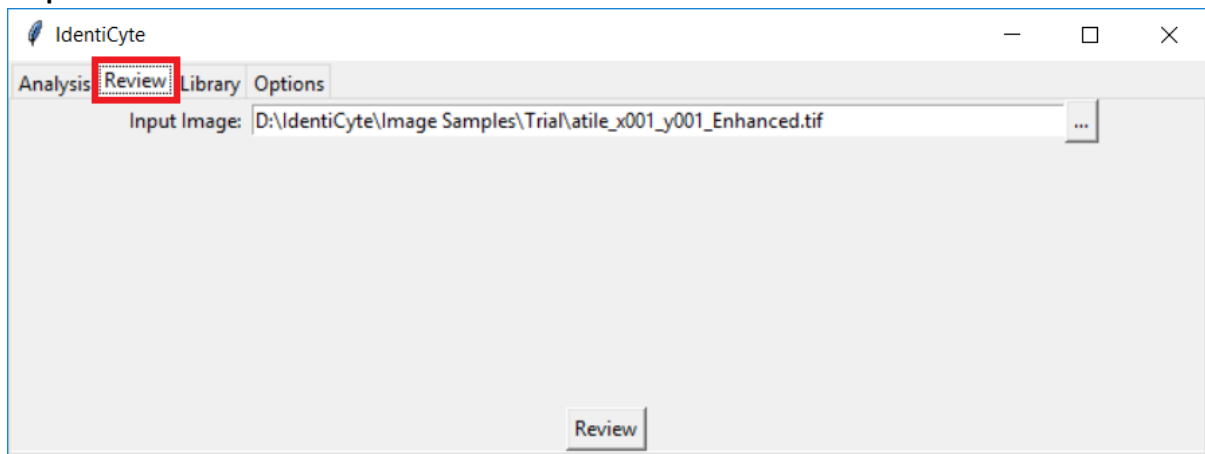
The next tab, Per Image Breakdown, shows the name of every image in the file, then the percent of the cells in that file in each category. After that, the number of cells in the image are shown and finally the average confidence of all cells in the image.


	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Image Name	BicEchinoc	Biconcave	Echinocyt	Lysing	SphEchino	Spherocyt	Other	Edge		Cells per Image	Confidence Per Image		
2	tile_x001_y001.tif	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	100.0 %	0.0 %	0.0 %		2	89.2 %		
3	tile_x001_y002.tif	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	100.0 %		1	100.0 %		
4	tile_x001_y003.tif	25.0 %	0.0 %	12.5 %	0.0 %	0.0 %	12.5 %	12.5 %	37.5 %		8	76.2 %		
5	tile_x001_y004.tif	0.0 %	0.0 %	33.3 %	0.0 %	0.0 %	16.7 %	50.0 %	0.0 %		6	63.2 %		
6	tile_x001_y005.tif	0.0 %	0.0 %	16.7 %	0.0 %	0.0 %	33.3 %	16.7 %	33.3 %		6	85.1 %		
7	tile_x001_y006.tif	25.0 %	0.0 %	0.0 %	0.0 %	0.0 %	50.0 %	25.0 %	0.0 %		4	75.1 %		
8	tile_x001_y007.tif	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	50.0 %	25.0 %	25.0 %		8	82.5 %		
9	tile_x001_y008.tif	0.0 %	0.0 %	11.1 %	0.0 %	0.0 %	66.7 %	0.0 %	22.2 %		9	89.8 %		
10	tile_x001_y009.tif	0.0 %	0.0 %	25.0 %	0.0 %	0.0 %	62.5 %	0.0 %	12.5 %		8	83.6 %		
11	tile_x001_y010.tif	0.0 %	0.0 %	50.0 %	0.0 %	0.0 %	50.0 %	0.0 %	0.0 %		2	73.9 %		

With the Review Tab

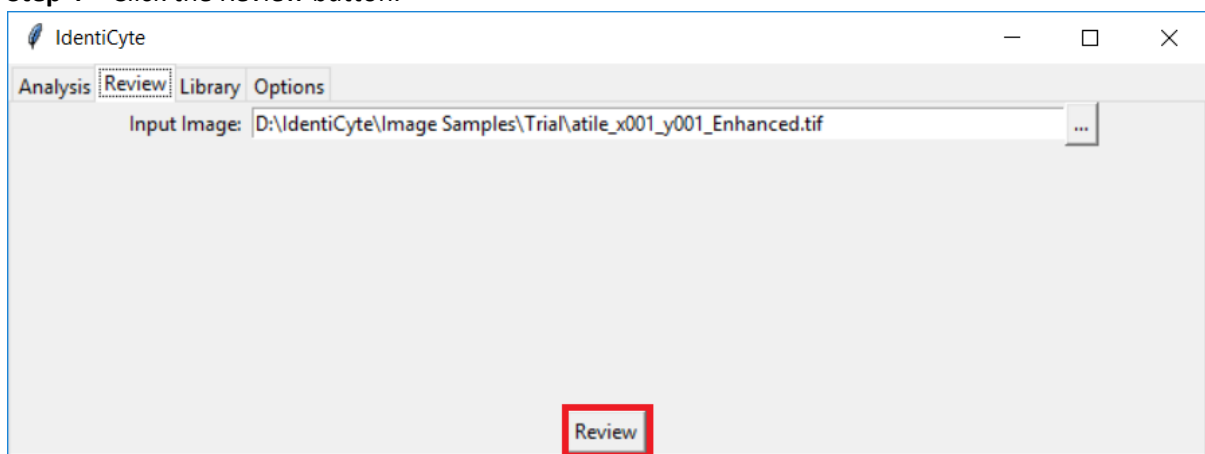
**Step 1** – If IdentiCyte is not already open, double click the IdentiCyte file to run it.

**Step 2** – Click the Review Tab.

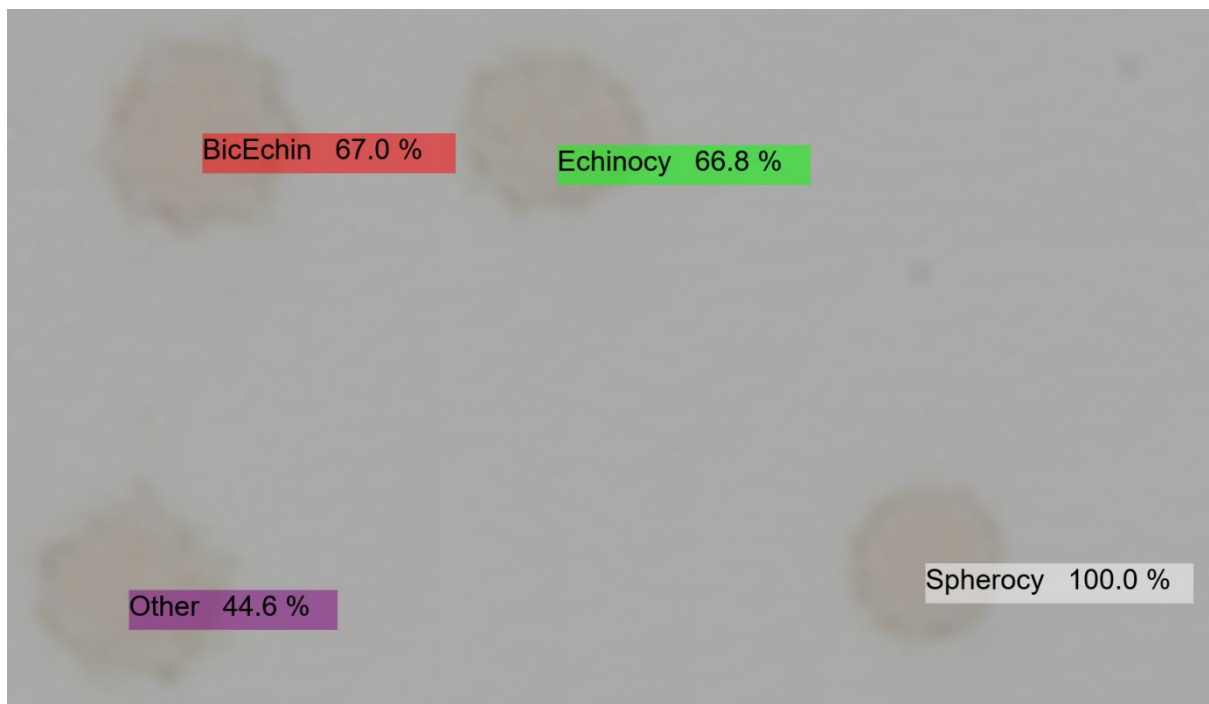
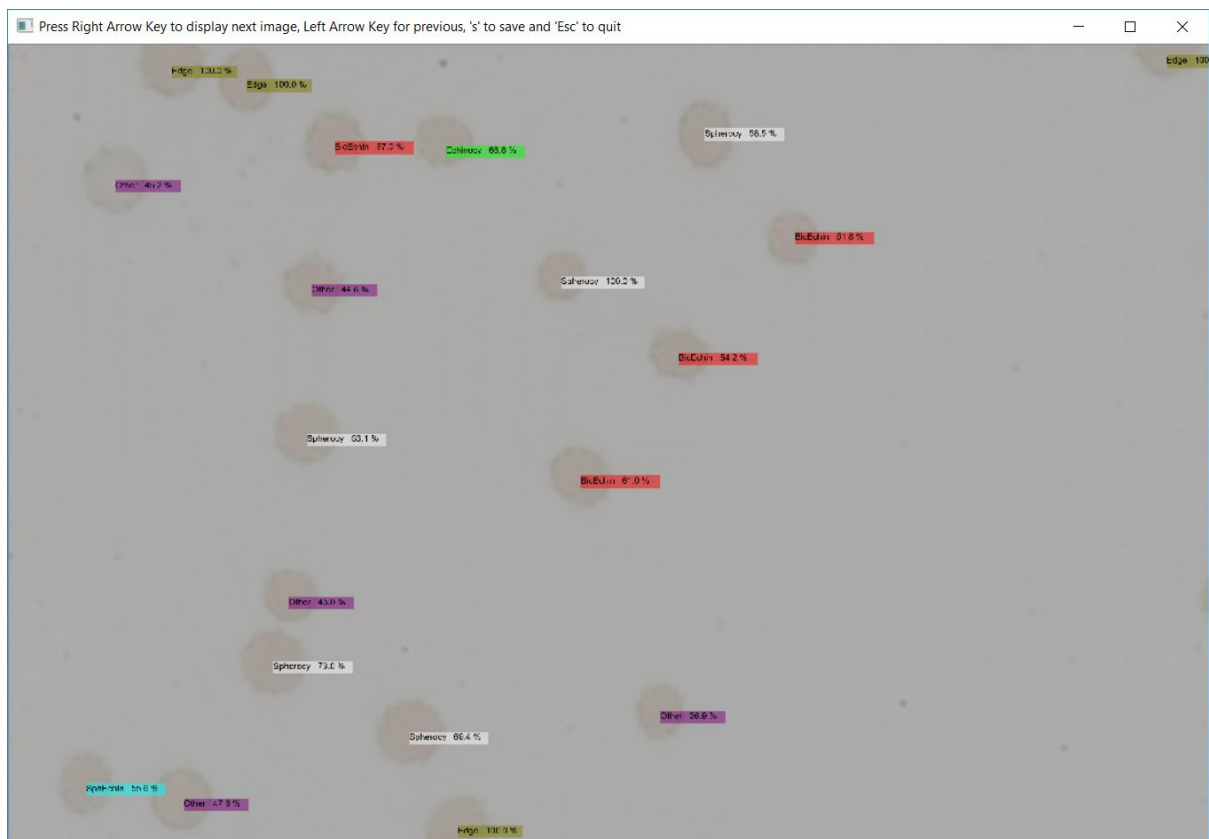


**Step 3** – Select the image to be reviewed. This can be done by clicking the  button next to the Input Image input box. This will bring up the Input Image window. To choose the desired image, navigate to its location and select it in the Input Image window.

**Step 4** – Click the Review button.

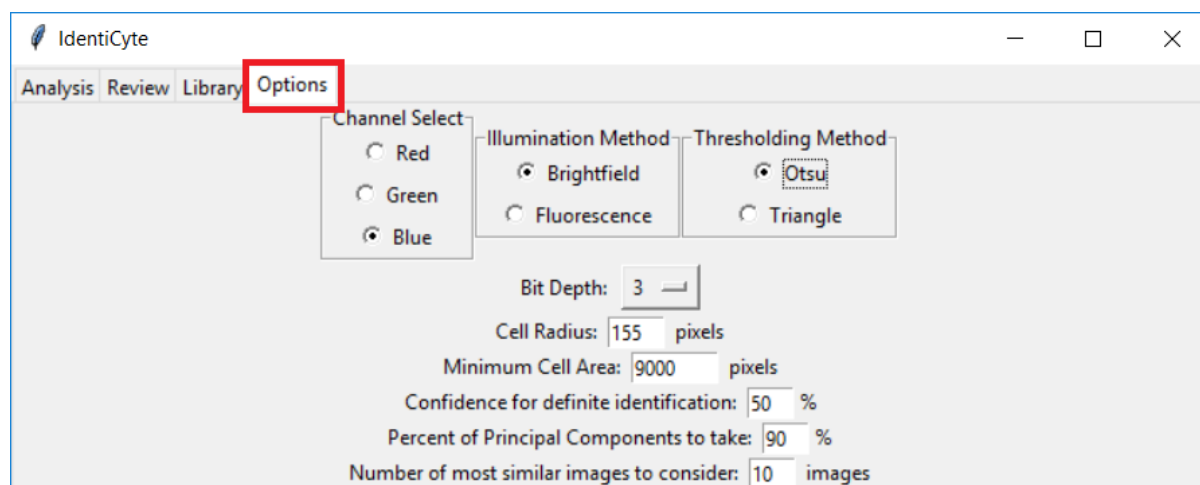


**Step 5** – The review window can be navigated with the arrow keys. The right arrow key will display the next image in the folder, the left arrow key will display the previous image. The currently displayed labelled image can be saved by pressing the s key. The saved images can be found in the 'Labelled' folder inside the folder containing the reviewed images. Finally, the review window can be closed with the red X in the top right corner, or by pressing the Esc key.

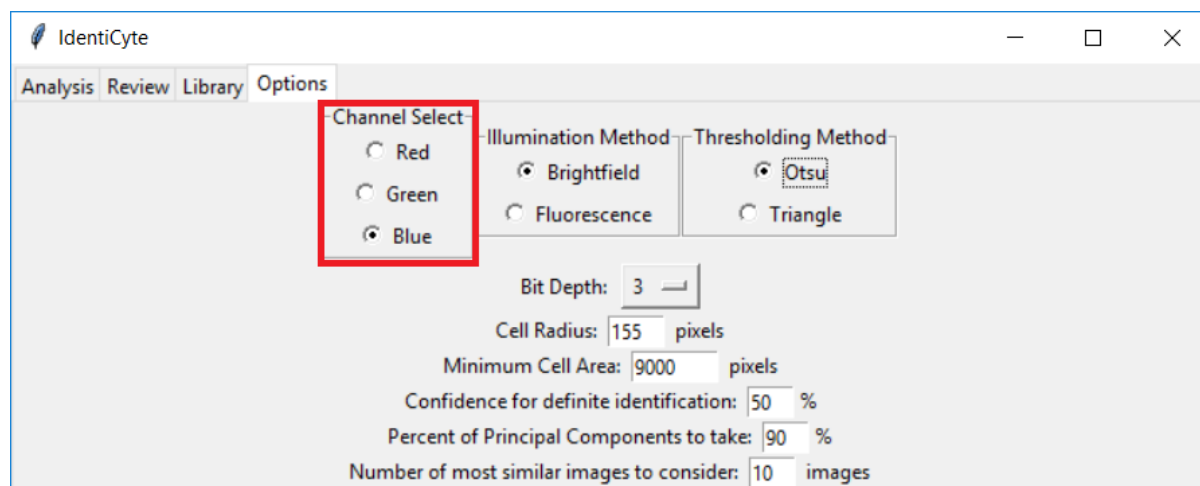


## Changing Options

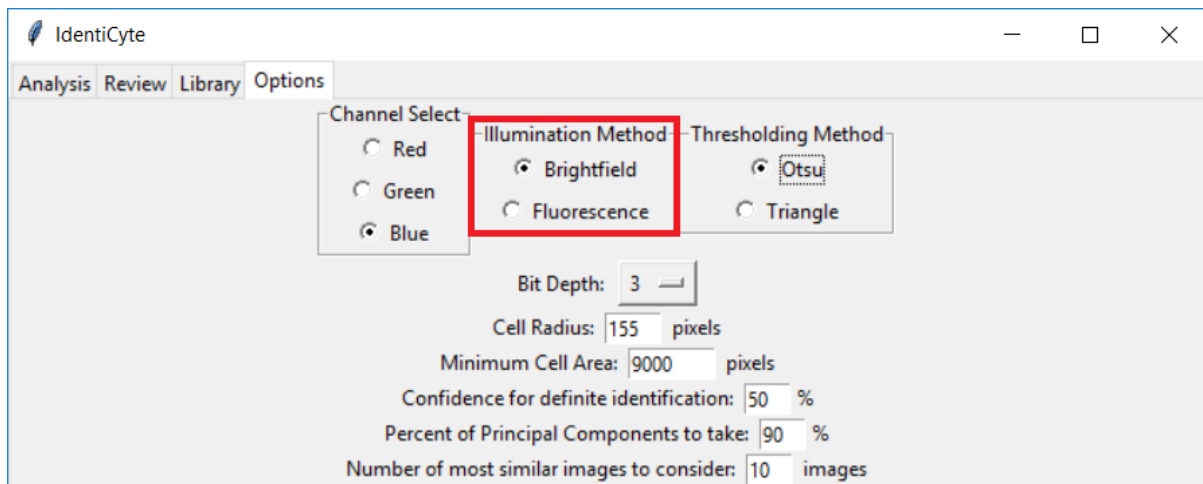
IdentiCyte has several options that can be changed to modify its behaviour. These can all be found under the options tab.



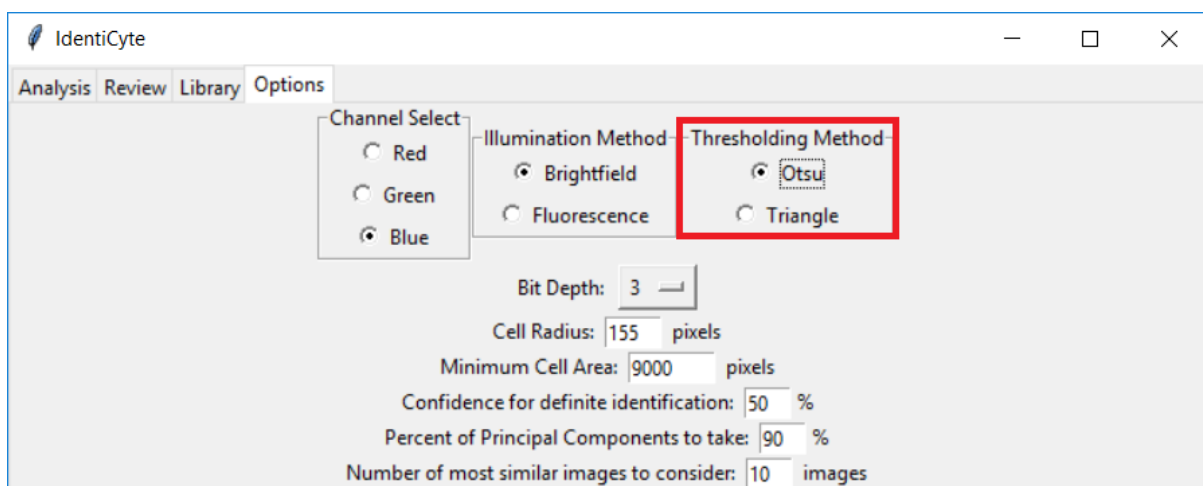
The first option is 'Channel Select'. This chooses which of the three channels is looked at when converting the cells to greyscale. The channel chosen should provide the most contrast for the objects to be identified. For example, a red cell on a white background, as is common in bright field microscopy, would have the most contrast in the **blue** channel, because both red and white have high intensities in the red channel. In the **blue** channel white will have a high intensity, but the red cell will have a low intensity. In contrast, a red cell on a black background, as might be found in dark field or fluorescence microscopy, will have high contrast in the **red** channel since the cell will have a high value in the **red** channel, but the background will have a low intensity. If the image is greyscale, this option does not matter, all channels will give the same result.



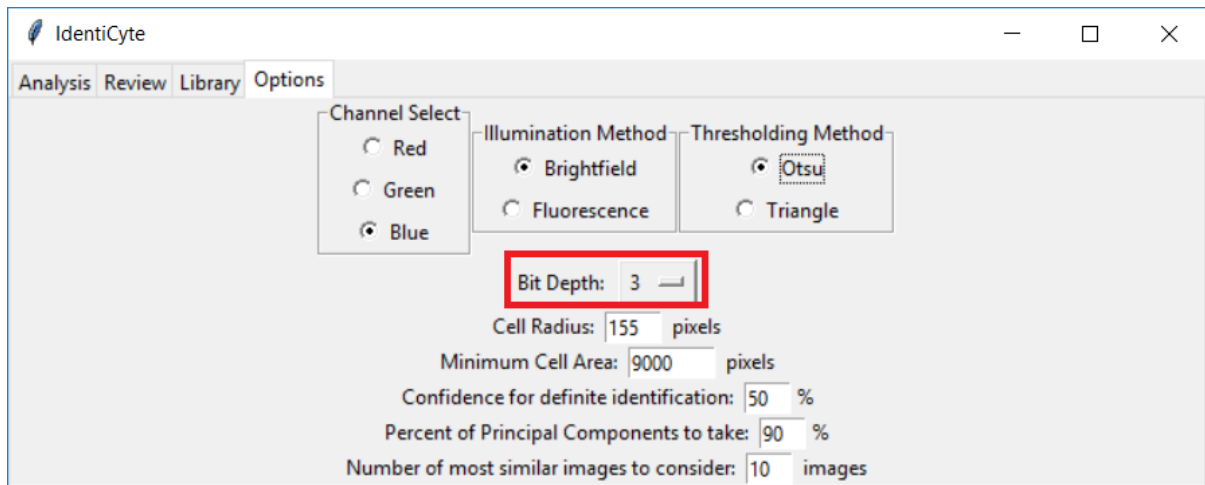
There is also an option to choose whether the samples are brightfield or fluorescent microscopy images.



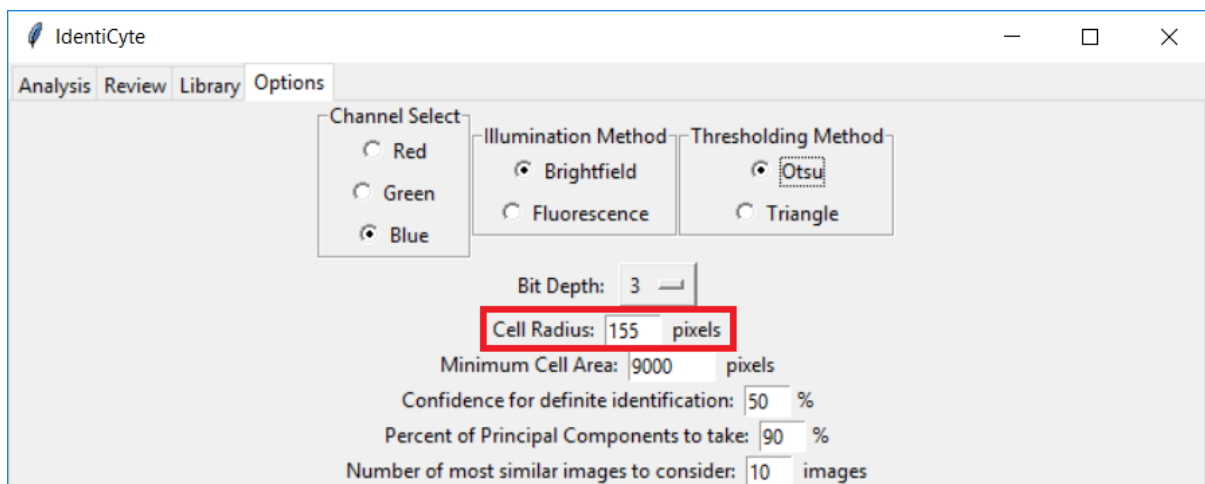
The next option is the thresholding method. This is the method by which the images are transformed from greyscale to binary, and it affects how cells are detected. Otsu generally causes fewer cells to be missed, and Triangle misses more cells but does a better job at distinguishing joined cells.



Then there is the Bit Depth option. This simply changes the number of levels of grey that are used to represent cells in the library and when analysed cells are compared to the library. The number of shades of grey are  $2^{\text{depth}}$ . We have found that 3 bits gives a good balance between retained detail and noise elimination. Values between 1 and 8 can be selected.

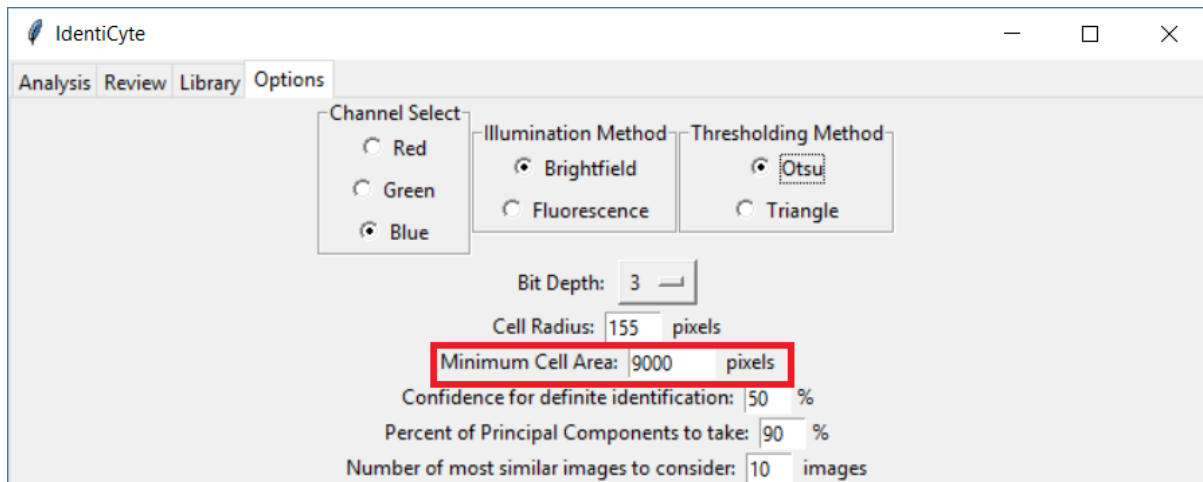


Directly below that can be seen the Cell Radius option. This affects the size of the image that is extracted with the Extract Cells function. The radius is the number of pixels that are extracted in each direction around a central pixel, so the extracted cell image will be a  $(2 \times \text{radius} + 1)$  pixel square.

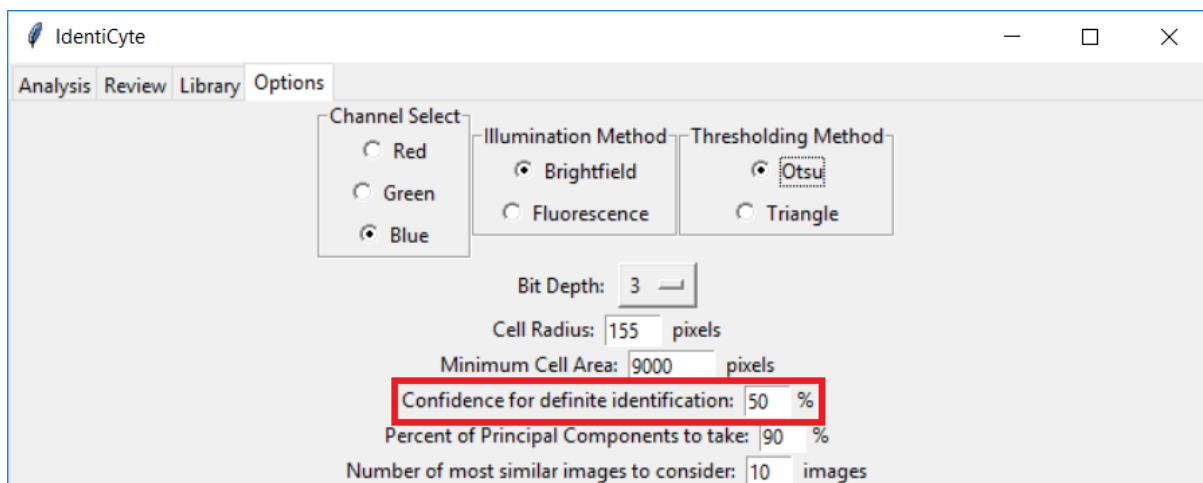




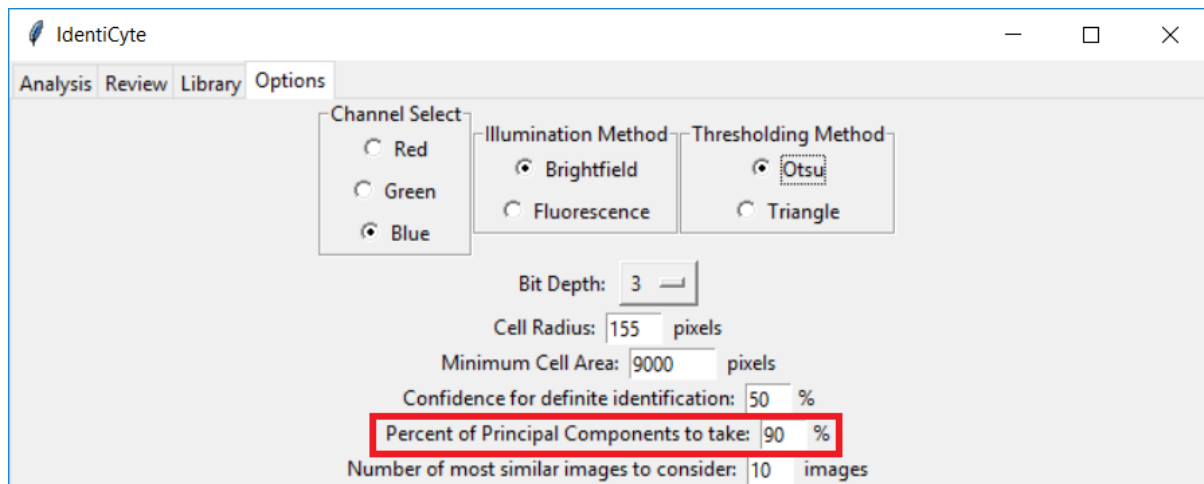
Next is the minimum cell size. This is a measure of area and dictates how many pixels a dark spot must have to be considered a cell. This option is mostly used to get rid of debris smaller than the cells.



Then we see the Confidence for user verification. Increasing this will increase the frequency with which the analysis will ask the user to confirm a cell's classification. Decreasing will make the popup appear less often. If user verification is disabled, cells with a confidence lower than this threshold will be classified as 'Other'



Now we see the percent of principal components to take. The program uses Principal Component Analysis to reduce the number of features of each cell to compare. This number changes the percent of principal components used in comparison, represented as a percentage of the total number. Accuracy raises sharply with this number to a variable point, after which it starts decreasing. The default of 90% generally gives good results. This variable can be changed to optimize results, but caution should be exercised when doing so.



Finally, we discuss the number of most similar images to consider. When the cell is compared to the library, this gives the number of images from the library that are used to classify it.

