nuclitrack Documentation

Release

Sam Cooper

Feb 01, 2018

Contents

NucliTrack is designed as an easy to use Application for tracking cell nuclei. We're glad you have chosen to try our software out, please follow the following steps to track your first video.

Alternativley check out the online video guides:

Tutorial 1

Requires HL60 example data, distributed with permission from David Svoboda, Example Data 1

Tutorial 2

Requires RPE1 dataset generated by our lab in early experiments on p21 Example Data 2

CHAPTER 1

Installation

To get started with NucliTrack we have prebuilt versions available for OSX, Linux and Windows:

Linux	macOS	Windows

No installation is required for the Application though on Windows you will need to enter the NucliTrack directory and run the .exe file, if you have any issues getting NucliTrack up and Running please see the see the *Installation Issues* section:

CHAPTER 2

Loading Time Series Images

To get started you need to specify file names for the .hdf5 files that will store your tracking data and parameters. You can either enter the location for these files directly into the text boxes and **hit enter**. Alternatively navigate to the desired folder, enter names for the files at the end of the directory path, and then **hit enter**.

Secondly you will need a video of cells that need tracking. Importantly, the video must be stored as a sequence of individual single plane image files fully supported by the pillow python module . There are then three options for loading the image time series into NucliTrack.

2.1 Auto load

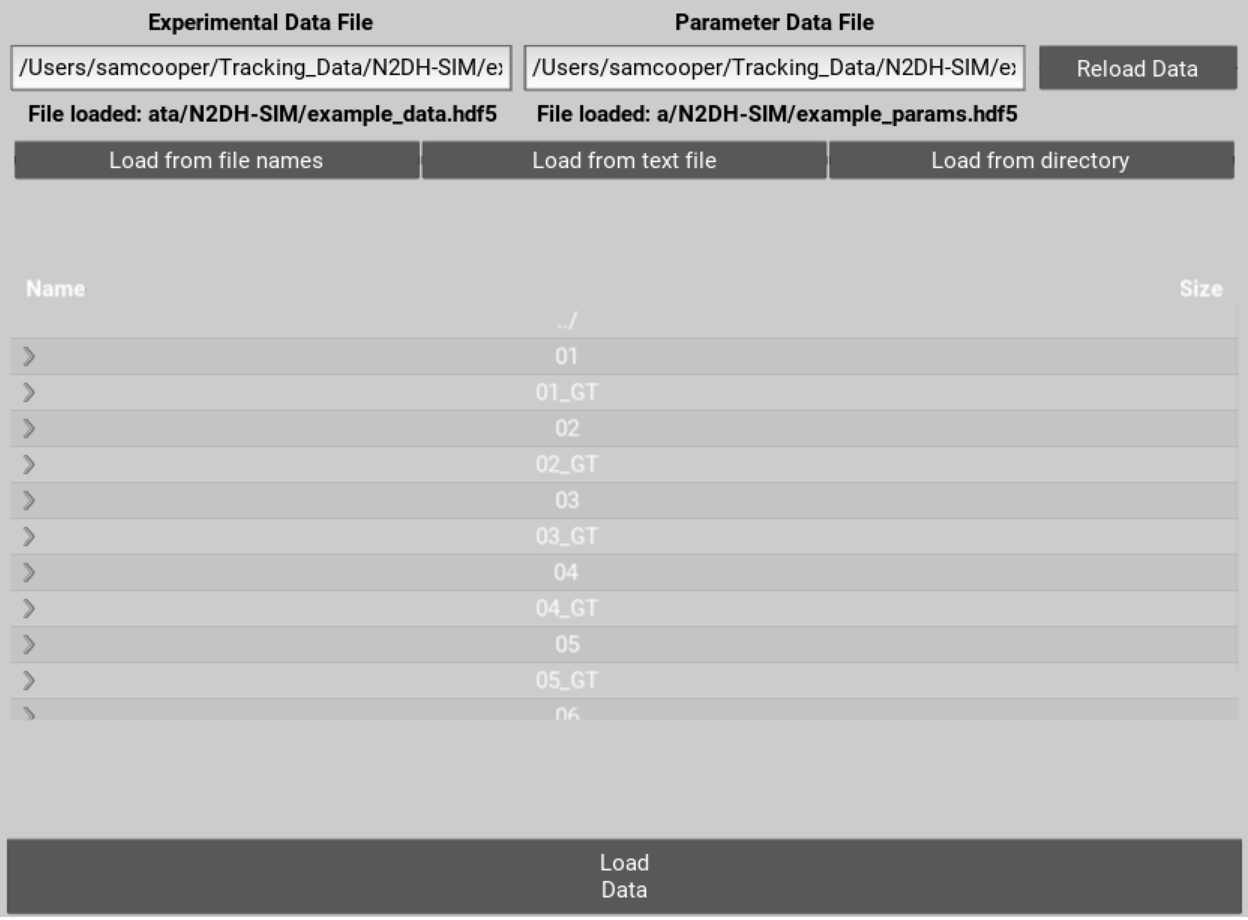
The autoloading interface allows you to load an image series based on the first and last filename of each channel. Whilst it doesn't matter how the rest of the file is named, the image series must contain a continuous series of zero padded time points e.g:

```
somename_ch1_01.tif
somename_ch1_02.tif
somename_ch1_03.tif

...

somename_ch1_99.tif
```

To load your videos navigate to the folder containing them, select the *first image* file toggle button and **double click** the first image filename. Click the *last image* toggle button, and then either modify the filename in the text input bar and hit enter or select the last image file.

You can now either hit the load images button, or alternativley load another channel (as long as the image size and number of frames is the same). To do this choose a different channel from the *channel* drop down list, and repeat the above process. Once the first and last file names for all channels have been selected press the load images button:

```
somename_ch2_01.tif
somename_ch2_02.tif
```



Experimental Data File

/Users/samcooper/Tracking_Data/N2DH-SIM/e›

Parameter Data File

/Users/samcooper/Tracking_Data/N2DH-SIM/e›

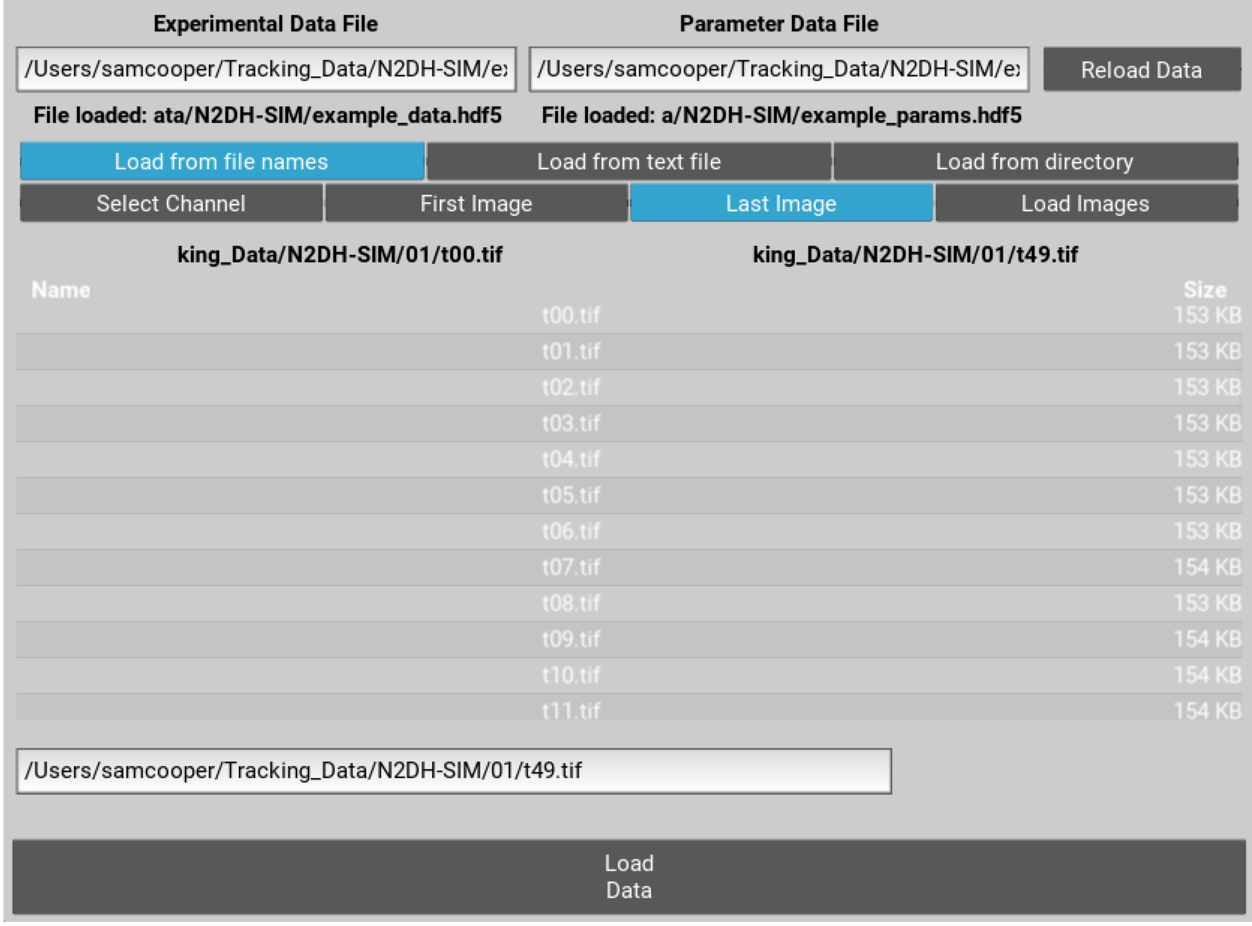
File loaded: ata/N2DH-SIM/example_data.hdf5

File loaded: a/N2DH-SIM/example_params.hdf5

Fig. 2.1: **Figure 1** Select locations to save data on the movie and parameters used.

```
somename_ch2_03.tif

...

somename_ch2_99.tif
```

Finally you can also load a series of uniquely labelled images, where the use of external segmentation software is desired. Specifically, the label series must be the same length and dimensions as the image series, and within each labelled image, each segment must be uniquely labelled. Select the *label* option from the *channel* drop down button and then the first and last image in the label series to load the images in.



/Users/samcooper/Tracking_Data/N2DH-SIM/e:

/Users/samcooper/Tracking_Data/N2DH-SIM/e:

File loaded: ata/N2DH-SIM/example_data.hdf5

File loaded: a/N2DH-SIM/example_params.hdf5

king_Data/N2DH-SIM/01/t00.tif

king_Data/N2DH-SIM/01/t49.tif

Name	Size
t00.tif	153 KB
t01.tif	153 KB
t02.tif	153 KB
t03.tif	153 KB
t04.tif	153 KB
t05.tif	153 KB
t06.tif	153 KB
t07.tif	154 KB
t08.tif	153 KB
t09.tif	154 KB
t10.tif	154 KB
t11.tif	154 KB

/Users/samcooper/Tracking_Data/N2DH-SIM/01/t49.tif

Fig. 2.2: **Figure 2** Select first and last file for each channel to autoload images.

2.2 Load from text

Loading from text is perhaps the easiest way of loading multichannel images, and also is used by batch mode. Load from text requires you to create a text/csv file in which each image filename is specified relative to the text file, in order. To specify multiple channels indicate the channel first then the file name seperated by a comma; set the channel to 0 to load in labels. For example, where images are in a subdirectory 'ims' the text file would need to be formatted as follows:

```
1, ims/somename_ch1_01.tif
1, ims/somename_ch1_02.tif

...

1, ims/somename_ch1_98.tif
1, ims/somename_ch2_99.tif
2, ims/somename_ch2_01.tif
2, ims/somename_ch2_02.tif

...

2, ims/somename_ch2_98.tif
2, ims/somename_ch2_99.tif
0, ims/labels_01.tif
0, ims/labels_02.tif

...

0, ims/labels_98.tif
0, ims/labels_99.tif
```

Once you have created the text file, navigate to it in the file browser and **double click** to load.

2.3 Load from directory

This is probably the simplest way of loading an image series but only works where a single channel is used. Here, navigate to the folder conainging the single image series, and double click the first image in the series, this will load all other images in the directory in alphabetical/numerical order.

CHAPTER 3

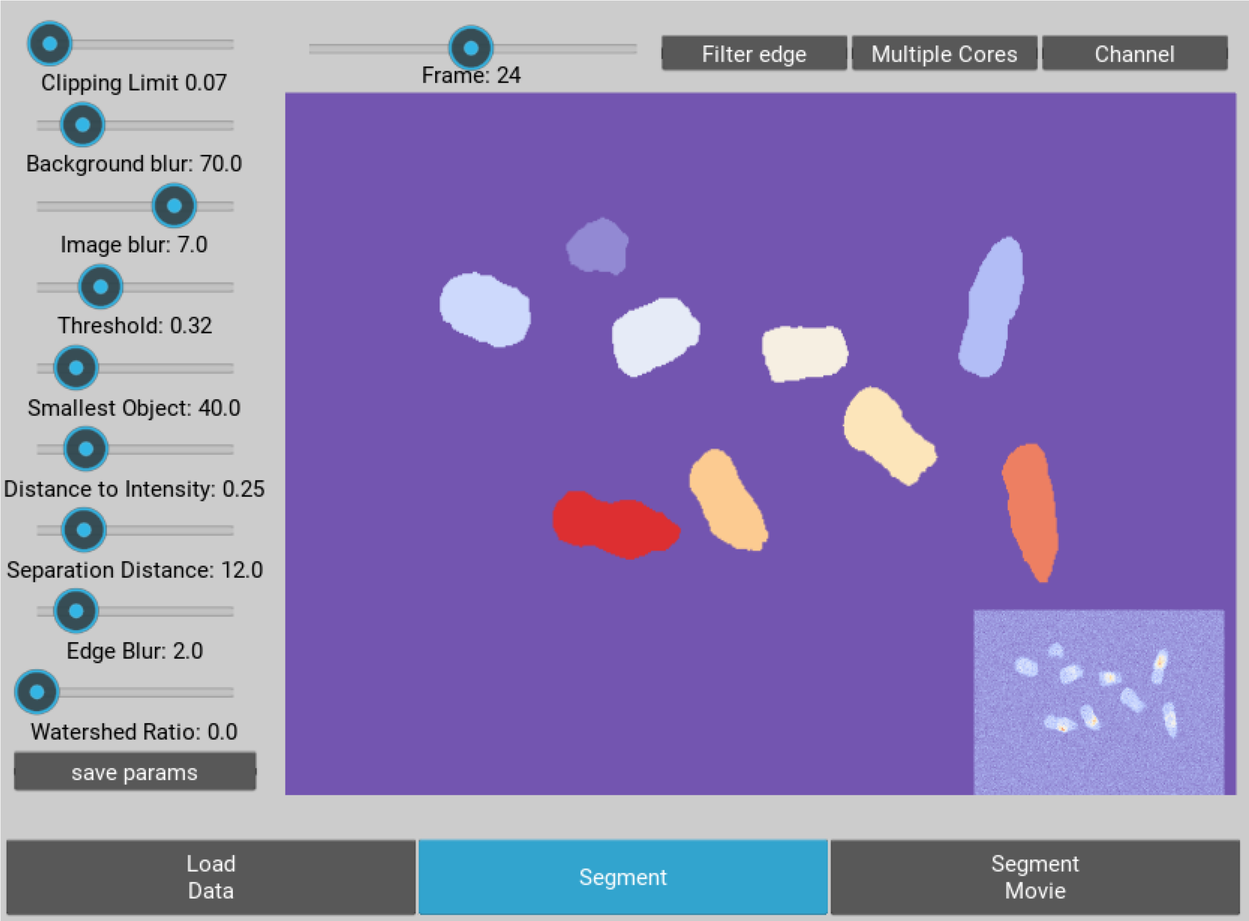
Segmentation

Now you're ready to segment your imaging series. Segmentation is based on a watershed procedure. All the steps except the last two are designed to identify the center of individual nuclei, in the last two steps the aim is to accurately mask every nuclei.

On the left hand side of the interface you should see a panel of sliders allowing you to segment the image. The best way to find out what they do is to play around, though a more detailed description for each control is below:

- Clipping threshold: this sets the maximum intensity of the image, anything above this intensity is clipped to this intensity
- Background Correction: background intensity variation is corrected by subtracting a really blurred version of the image from the original, this slider controls how blurred this image is.
- Blurring: prior to thresholding pixel noise must be reduced this slider adjusts how blurred the image used for thresholding is.
- Threshold level: the slider here controls what level of intensity should be used to define what are nuclei and what is background. Don't worry if nuclei are touching each other at this point.
- Filter small objects: if there are lots of small objects that have been thresholded but aren't nuclei you can filter them out using this slider
- Intensity or Distance: the next step is to identify local maximum within the thresholded regions, ideally local maximum will exist in the center of each nuclei, the maximum intensity often doesn't reflect the center of the nuclei or even worse may lead to multiple local maxima appearing in each cell. This function allows you to choose to what degree local maxima are defined by the intensity or distance from edge.
- Peak local maxima: behind this slider is a function which identifies local intensity peaks in the image, the slider sets the minimum distance allowed between peaks.
- Edge detection: this slider controls how blurry the image is prior to edges being marked, more blurry means edges are smoother though may match the shape less.
- Watershed: behind this slider is the watershed function which calculates where the local basins of attraction are in the image, like marking valleys in a mountainous terrain. The selected peak local maxima represent the bottom of the valleys, by adjusting the slider the lowest area is made larger or smaller, this like mining into the side of mountains to force the valley to be larger.

Use the slider in the top left hand corner to navigate around the video and try out your segmentation parameters on other frames. Also other channels can be selected for segmentaton using the *channel* dropdown list in the top right corner. Finally, cells touching the border can be filtered out by selecting the *filter edges* toggle button.



Clipping Limit 0.07

Background blur: 70.0

Image blur: 7.0

Threshold: 0.32

Smallest Object: 40.0

Distance to Intensity: 0.25

Separation Distance: 12.0

Edge Blur: 2.0

Watershed Ratio: 0.0

Frame: 24

Fig. 3.1: **Figure 3** The segmentation GUI used for identifying nuclei within the movie.

Once you have achieved good segmentation it's now time to save the parameters. Once this is done, click segment movie to apply your segmentation parameters to the entire image sequence. This may take a while depending on how fast your computer is and how large you movie is. Please don't click off the segmentation window though whilst it is running, this will result in an error.

To speed up the process you can opt to use multiprocessing. However, due to the workload on the CPUs, the GUI does not display a loading bar during this process.

When segmentation is done a button will appear allowing you to visualise the results of segmentation and decide whether you want to go back and adjust the parameters or move forward. If you are happy go ahead and click extract features.

You can also export the label files at this point as single plane tiffs. To do this create and empty directory by entering a name in the *choose directory name* text box and **hitting enter**. Re-enter the current folder and to see this directory then enter the directory and choose an output name for the label series. **Hit enter** and the series of labelled images will be saved.

Finally an optional extra pair of features giving the mean and median intensity of a peri-nuclear ring region is available. This is important where Kinase Translocation Reporter probes are being used to measure signalling dynamics. Note however, due to additional binary options being required this feature significantly slows down the extraction process.

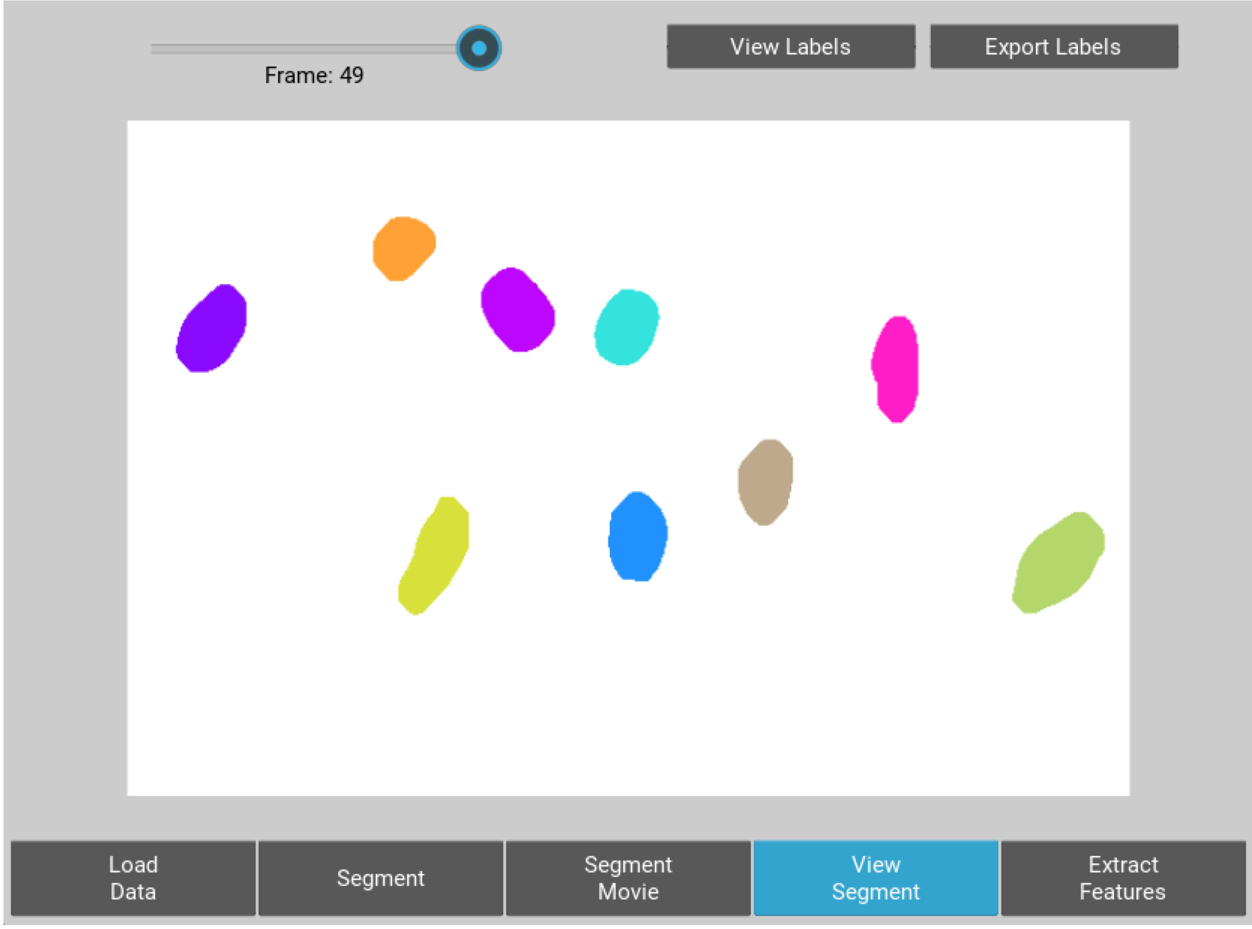
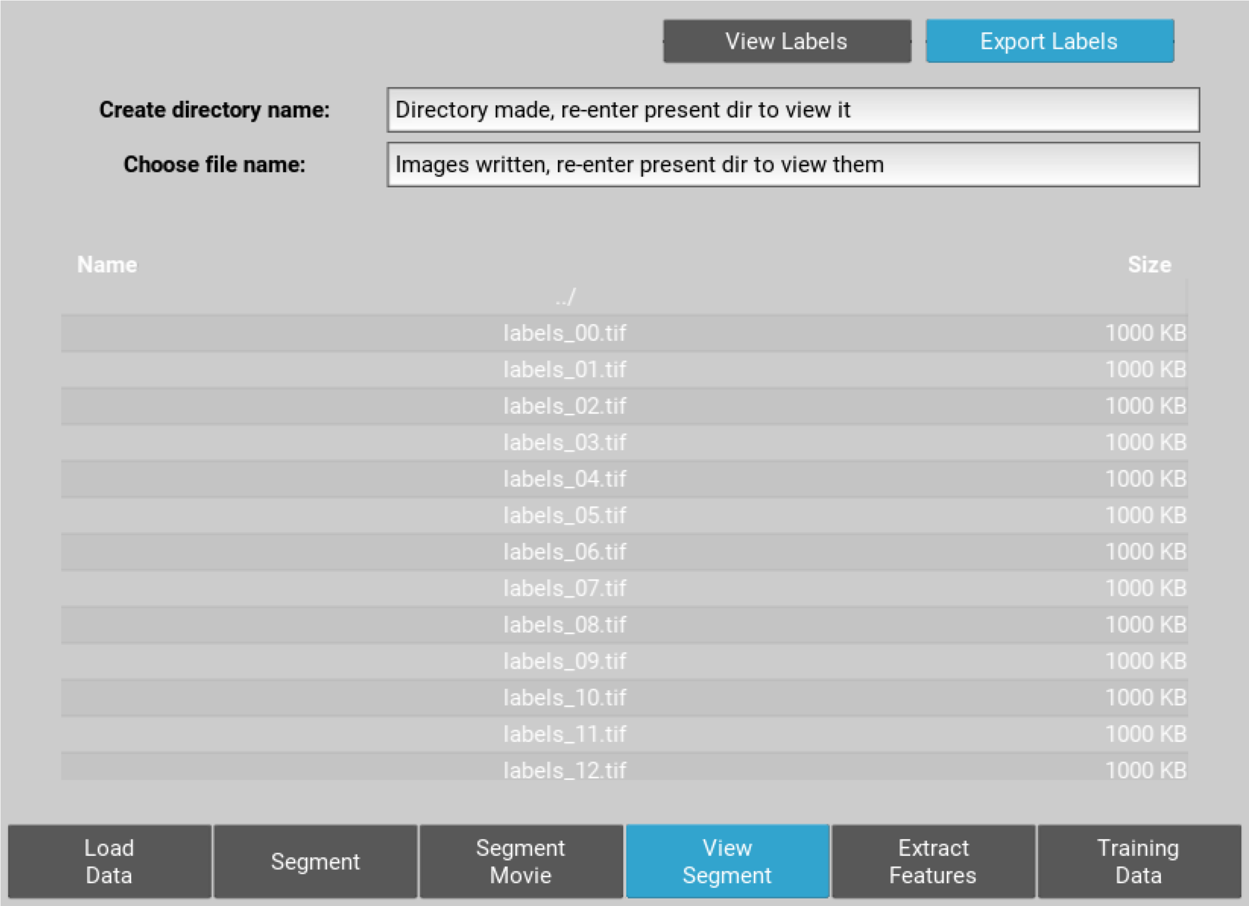


Fig. 3.2: **Figure 4** User interface for inspecting the quality of segmentation



Create directory name: Directory made, re-enter present dir to view it

Choose file name: Images written, re-enter present dir to view them

Name	Size
../	
labels_00.tif	1000 KB
labels_01.tif	1000 KB
labels_02.tif	1000 KB
labels_03.tif	1000 KB
labels_04.tif	1000 KB
labels_05.tif	1000 KB
labels_06.tif	1000 KB
labels_07.tif	1000 KB
labels_08.tif	1000 KB
labels_09.tif	1000 KB
labels_10.tif	1000 KB
labels_11.tif	1000 KB
labels_12.tif	1000 KB

Fig. 3.3: **Figure 5** Export of singeplane .tiff label images to new directory.

CHAPTER 4

Training Data

The algorithm that tracks your cells is probabilistic. Meaning, it will assign a probability to every segment giving it a measure of whether it is likely to be a single nuclei, a mitotic nuclei, on that has left mitosis recently, or an erroneous segmentation either, more than one nuclei or no nuclei.

To calculate the probabilities features must be extracted from each segment. Click extract features to do this, again this step may take a few minutes.

Now select the training data button, you will be presented with a window showing your segmented cells in gray.

The dropdown button in the top right corenr allows you to select a training data class for the probabilistic tracking algorithm. Choose a class then start clicking on the gray nuclei in the field of view. You probably want to select around 10 to 20 examples of single cells, and then 3 or 4 mitotic cells and 5 or 6 cells that have left mitosis.

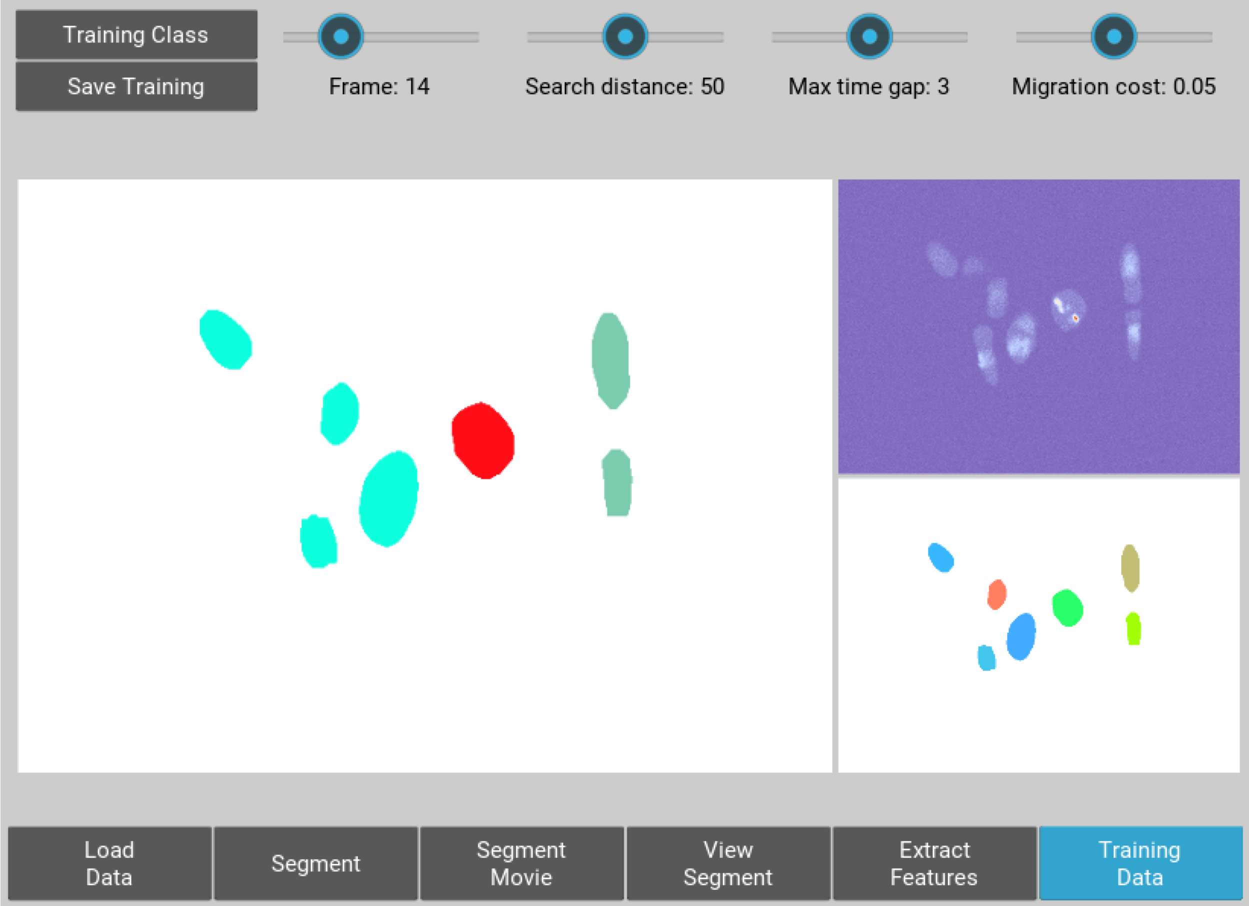
If allot of segmentation errors have been made consider selecting examples of these, but the algorithm usually works just fine without these examples.

Three new sliders also exist in the top left hand corner along with text, generally the default values will work though you may want to adjust these if tracking is particularly slow:

- The furthest left of these sliders determines how far (in pixels) the algorithm will search in the previous frame, this should roughly correspond to the largest jump that nuclei are making between frames, however larger values will also mean tracking will take longer.
- The middle of the three sliders determines the largest gap between frames that can occur in a track. Smaller values speed up tracking whilst larger values take longer.
- The final slider changes the penalty for movement distance, smaller values mean less penalty than larger values. Reduce this if cells are moving allot whilst increase this value if cells make small movements.

Once everything is selected, click save training. A button for classifying cells will now appear. Click this to assign probabilities to every cell in the image sequence.

Everything is now ready for tracking, hit this button to begin tracking.



Frame: 14

Search distance: 50

Max time gap: 3

Migration cost: 0.05

Fig. 4.1: **Figure 6** Selection of training data used for classifying nuclei into classes prior to tracking

CHAPTER 5

Correcting and Storing Tracks

With tracking complete you can now inspect the results of automated tracking by scrolling backwards and forwards through the video. For single frame movements you can also use the keyboard hotkeys **(a)** and **(d)**.

To inspect data on a track click the select track button, or use the hotkey **(z)**. Then select a labelled cell in the track segment window. Data on this track will now appear in the bottom window. You can jump from one frame to another in the movie, by clicking the jump button or pressing **(w)**, then clicking on the graph. The tracking window will then jump to the frame as selected on the x axis.

Automated tracking is likely to make errors. Buttons on the left hand side let you correct these.

- To add a segment to a track, click add segment **(c)** then select a segment on the screen.
- To remove a segment click remove segment **(v)** then select a labelled segment on the screen.
- To swap two tracks over in all proceeding frames, with a track selected, select swap tracks **(x)** then select another track on the screen.
- To add a new track, select new track **(n)** then choose an unlabelled segment on the screen.

If you want to export data on all the track click save all to CSV. Otherwise you can select individual tracks and export only these selections. To do this with a track selected click store track. It should now be marked with a black dot. Once you have selected all the tracks you want, then click export selected to csv.

Events can also be added. These are extremely useful if you want to computationally syncronise your cells to specific timepoints such as S-phase entry. There are three event options that you can choose. If you select one of these and then click on the graph window at the desired timepoint, you will see a Cyan, Yellow or Magenta line appear at this point marking the event. When you export your results to csv format, the final column will then contain the number (1, 2 or 3 depending on which event was chosen) at this time point. the hotkey for these toggle buttons are **(i)**, **(o)** and **(p)**

You can also change which features you want to visualise using the text boxes on the right. Here choose a number between one and 11, representing the features, the 'Floored Mean' feature is particularly useful for identifying PCNA foci a common fluoresecent label used to mark S-phase in cycling cells:

```
1 Area
2 Eccentricity
3 Length
```

```
Perimeter
CH1 Mean Intensity
CH1 Median Intensity
CH1 Floored Mean
CH1 StdDev Intensity
CH1 Floored Mean
CH1 Ring Region Mean Intensity (opt)
CH1 Ring Region Median Intensity (opt)
CH2 Mean Intensity
CH2 Median Intensity
CH2 Floored Mean
CH2 StdDev Intensity
CH2 Floored Mean
CH2 Ring Region Mean Intensity (opt)
CH2 Ring Region Median Intensity (opt)
CH3 Mean Intensity
CH3 Median Intensity
CH3 Floored Mean
CH3 StdDev Intensity
CH3 Floored Mean
CH3 Ring Region Mean Intensity (opt)
CH3 Ring Region Median Intensity (opt)
```

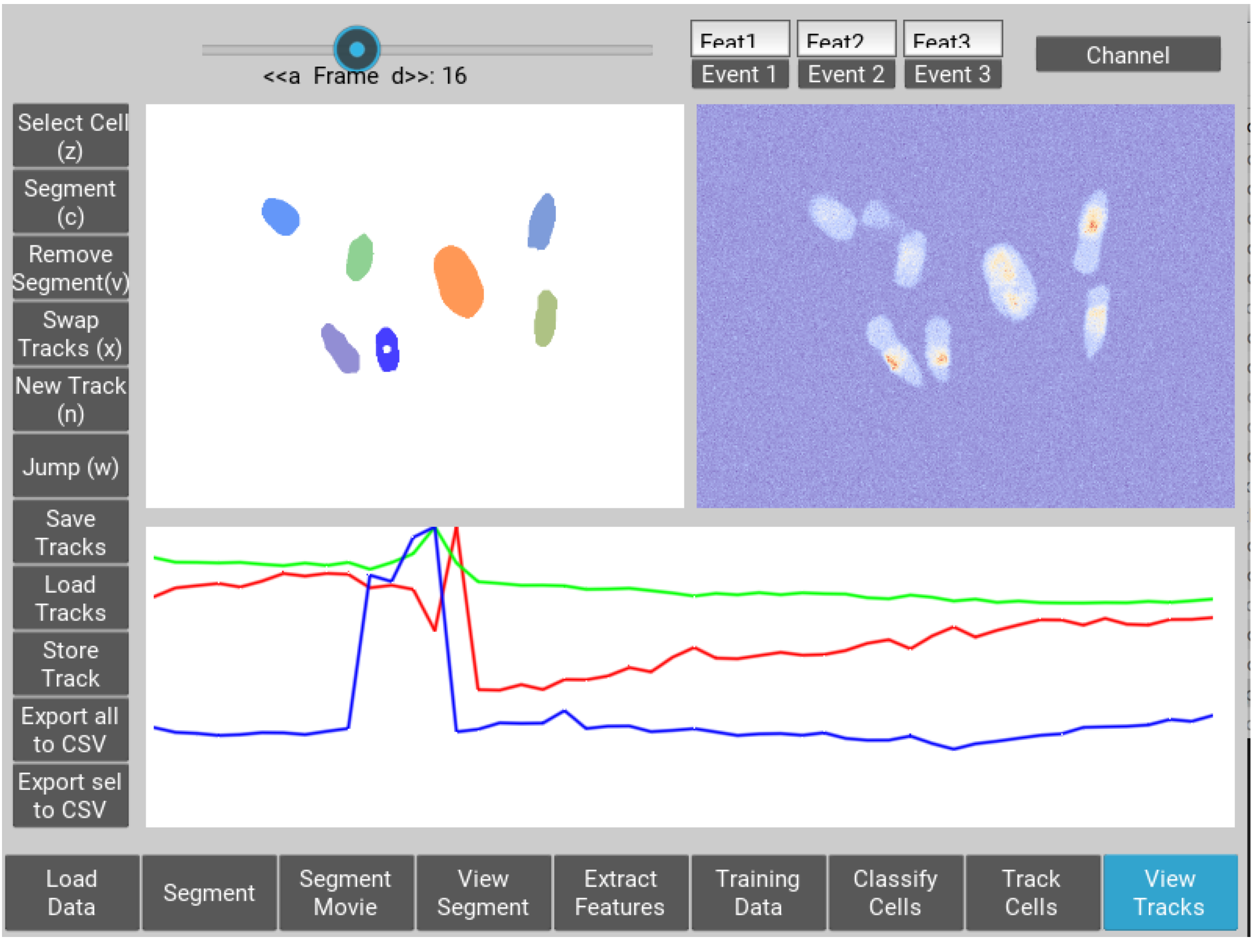


Fig. 5.1: **Figure 7** User interface for inspecting and ammending tracked nuclei

CHAPTER 6

Python Package Installation

If your looking to develop your own features, tracks hundreds of videos in batch, or contribute to the nuclitrack project you will need, Python3 . We recommend using the Anaconda distribution for Windows users.

Once you've got Python3 up and running you will need packages called Cython and Numpy (Anaconda already has them installed). These are easily installed using the pip installer, from the terminal type:

```
> pip install cython
> pip install numpy
```

Now you're ready to install NucliTrack:

```
> pip install nuclitrack
```

Notes:

- On windows you will likely need to install kivy prior to nuclitrack, a guide for this can be found on the kivy website .

From experience after installing the dependencies download the latest wheel e.g. "python3.6, 64bit" from the kivy website then run the following commands:

```
> python -m pip install C:\Kivy-1.9.1.dev-cp27-none-win_amd64.whl
```

Here substitute in the location of the downloaded .whl file.

- If Python2 is installed on your system, you will probably need to specify the pip installer to use e.g. pip3.6, to ensure everything works properly.

On more basic systems some dependencies may cause issues, on RHEL/Fedora linux I have noted a few:

- RHEL/Fedora linux: Kivy may require the openGL dev library to install properly > yum install mesa-libGL
- RHEL/Fedora linux: Kivy may require xclip/xsel to import properly, install EPEL, then > yum install xclip,xsel
- RHEL/Fedora linux: Kivy may require bzip to to import properly > yum install bzip2-devel then configure and make python3.6 again

CHAPTER 7

Python Package Functions

The python NucliTrack package contains two functions that allow it to be run either as a GUI or in batch mode (discussed later). In both cases, asserting that the script is being called as main is vital for multithreading to work properly.

To run nuclitrack using the GUI create a new python script, cut and paste the following code into it, and then run the script:

```
import nuclitrack
if __name__=='__main__':
        nuclitrack.main.run()
```

To run nuclitrack in batch mode which doesn't require the kivy library, create a new python script and paste the following into it.

```
import nuclitrack
if __name__=='__main__':
        nuclitrack.batchanalyse.batchanalyse('myfile.txt','myparams.hdf5','myoutput')
```

Here, 'myfile.txt' represents a text file for loading images, in the format described in the *Load from text* section. The 'myparams.hdf5' file must be created by using the GUI on a reference movie, and contains the parameters selected for segmentation and tracking, as well as training data chosen in the training data GUI. Finally 'myoutput' is the name that both the 'output.hdf5' and 'output.csv' file will be saved as. The 'output.hdf5' file can then be loaded into the GUI and track correction and inspection carried out. Alternatively results can be directly analysed from the 'output.csv' file.

To process multiple movies, the batchanalyse function can be called inside a loop where multiple text files are used to index different image series, with the output file name varied accordingly.

CHAPTER 8

Installation Issues

Windows: Due to multiprocessing and loading issues, we have opted not to compile Nuclitrack into a single file executable. Please unzip the file following download, enter it and then double click NucliTrack.exe from within the file, this will launch a console which will then load NucliTrack.

Mac OSX: error 67062, this is to do with gatekeeper interfering with the terminal launch as the App is not trusted. To solve this error, launch terminal and enter the following command:

```
> sudo spctl –master-disable
```

Enter your password on request, then launch the NucliTrack Application which should now work. The App will now be trusted so you can re-enable gatekeeper with the following command, and it will continue to work.

```
> sudo spctl –master-enable
```

Nuclitrack tries to open closes with no warning, potential issue with codsigning solve as below, if this doesn't work email me / raise bug:

```
> sudo codesign -f -s - Path_to_app_folder/NucliTrack.app/Contents/MacOS/SDL2

> sudo codesign -f -s - Path_to_app_folder/NucliTrack.app/Contents/MacOS/SDL2_ttf

> sudo codesign -f -s - Path_to_app_folder/NucliTrack.app/Contents/MacOS/SDL2_mixer

> sudo codesign -f -s - Path_to_app_folder/NucliTrack.app/Contents/MacOS/SDL2_image
```

Linux: We have had issues with graphics drivers, either update these otherwise please install and locally build the Python package. We also built the package on EL7, we're working on releasing linux versions built on other distributions, for the moment please use the Python package if the pre-built version does not work.

If you can launch the application from terminal please let us know the error message when it fails to load, this way we'll be able to resolve the issue for future releases.

OpenGL version is less than 2.0: Update your graphics drivers.

NucliTrack is also availale as a Python 3 package for development, and batch processing of image, see the *Python Package Installation* section for details of how to install Nuclitrack as a Python package.

CHAPTER 9

Contribute

- Source Code: https://github.com/samocooper/nuclitrack

CHAPTER 10

Support

If you are having issues, please let me know my email is sam^socooper.com sub in the @ for ^. Also check out my website socooper.com

CHAPTER 11

License

The project is licensed under the MIT license.