**ARI\_IoU\_calculation readme**

1. **Installing R**

The code runs in the R programming language. To install, download R from: <http://www.r-project.org/>

We then recommend the use of R-Studio, which provides a graphical user interface for R. This is also freely available at:

<http://www.rstudio.com/>

The first time the code is run, additional R plugins are required and can be installed using the tools -> install packages menu. The plugins required are:

* mclustcomp
* dbscan
* gtools
* imager
* spatstat
* contoureR
* sp
* RColorBrewer
* grDevices

1. **Data format.**

We have included an example data set, however, should you wish to test the code with your own data it needs to be in a specific format. – the same as our example

1. Create a folder in the same folder as the two R files with the name of the experiment, in our example it is “test”.
2. Place the ground truth data into this folder.
   * It must contain 3 columns: x coordinates (in nm), y coordinates (in nm) and ground truth index for each molecule, in that order. The format must also be .csv
   * The headings must be “x”, “y”, and “index”, like in the example file (data.csv).
3. Create another folder within the test folder called “classes”, and place in here the result indexes from your clustering algorithm.
   * You can place multiple result index files in here, for example, if you used several different user setting for your algorithm.
   * The file must be a single column (heading – “result”), with the cluster result indexes for that particular clustering result, and in .csv format.
     + For example, if there are 600 points in total in the ground truth data, then there must be 600 indexes in the file, i.e., all points must be assigned an index. In our example case, for both the ground truth and the cluster results, noise/non-clustered points are assigned a zero as an index, and clustered points get non-zero indexes.
4. **Running the ARI\_IoU\_calculation R script**

Once the data are in place the ARI and IoU values can be calculated using “ARI\_IoU\_calculation.R”

1. Open RStudio
2. Open the file named “ARI\_IoU\_calculation.R”
3. Enter the file path where the R scripts and data were placed. For example:
   1. “~Desktop/ARI\_IoU\_Indexes/”
4. In line 10, type the name of the folder you want to process. For example:
   1. FolderName = “test/”
   2. IMPORTANT – remember to include the forward slash (“/”) after the FolderName.
5. In line 12, type the name of the ground truth data you want to use. For example:
   1. FileName = “data.csv”
   2. IMPORTANT – remember to include the file extension .csv
6. In line 14, enter the size of one dimension of your region of interest in nanometers.
   1. We recommend square regions no larger than 2000\*2000 nm.
7. Click “Source” to run the code, and the code will then run. Time to complete will depend on the number of files placed on the class folder.
8. Once the code has finished it will have created a new timestamped folder within your data folder. In this folder will be:
   1. A single .csv with the ARI and IoU values for each of the files within the class folder (filenames retained for easy identification after analysis)
   2. A .PDF with plots of the polygons extracted for the IoU calculations for each class (plots: ground truth point pattern, ground truth polygons, class file results polygons, overlay of ground truth and results polygons).
   3. A .RData file, which can be opened in R Studio if you wish to use the output and source data for any further analysis, processing or plotting.