

CLUSTER ANALYSIS OF LOCALIZATION MICROSCOPY DATA WITH MACHINE LEARNING

David J. Williamson¹, Garth L. Burn², Sabrina Simoncelli¹, Ruby Peters¹, Juliette Griffie³, Daniel M. Davis⁴, Dylan M. Owen¹

1 - Randall Centre for Cell and Molecular Biophysics, King's College London, London UK, 2 - Max Planck Institute for Infection Biology, Berlin Germany, 3 - Laboratory of Experimental Biophysics, EPFL, Lausanne Switzerland, 4 - Manchester Collaborative Centre for Inflammation Research, University of Manchester, Manchester UK.

E-mail: david.williamson@kcl.ac.uk

KEY WORDS: localization microscopy, cluster analysis

Quantifying the clustering of points within single-molecule localization microscopy data is useful to understanding the spatial relationships of the molecules in the underlying sample. The conversion of point pattern data into a meaningful description of clustering is difficult, especially for biologically derived data, as the definitions of clustering are often subjective or simplistic. Many existing computational approaches are also limited in their ability to process large-scale data-sets or to deal effectively with inhomogeneities in clustering. Here we have developed a supervised machine-learning approach to cluster analysis which is fast and accurate. Trained on a variety of simulated clustered data, the network can then classify all points from a typical localization microscopy data-set (several million points from the entire field of view) as being either clustered or not-clustered, with the potential to include additional classifiers to describe different types of clusters. Clustered points can then be further refined into like-clusters for the measurement of cluster area, shape, and point-density. We demonstrate the performance on simulated data and experimental data of the kinase Csk and the adaptor PAG in both naive and pre-stimulated primary human T cell synapses.