DETECTING STRUCTURAL HETEROGENEITY IN SINGLE-MOLECULE LOCALIZATION MICROSCOPY DATA

Teun A.P.M. Huijben*[1], Hamidreza Heydarian*[1], Alexander Auer[2], Florian Schueder[2], Ralf Jungmann[2], Sjoerd Stallinga[1] and Bernd Rieger[1]

[1] Quantitative Imaging Group, Department of Imaging Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ, Delft, The Netherlands
[2] Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
Email: tepet@dtu.dk

Single-molecule localization microscopy (SMLM) enables imaging below the diffraction limit. The image quality can be improved further by fusing hundreds of super-resolution images of identical bio-molecular structures into a single reconstruction[1]. This approach overcomes the problem of incomplete labelling using the central assumption that all particles represent the same underlying structure. In reality, however, the sample might be heterogeneous in structure due to the biology itself, diseases, sample preparation or drug-induced variations. These potential variations between structures blur standard fusion and small subsets of structurally different particles remain undetected.

Previous work to separate classes in SMLM either uses cryo-electron microscopy-based image classification methods[2,3] or deep neural networks[4]. The first suffers from incomplete labelling, whereas the second requires the different classes to be known a-priori for training. This strong requirement for a-priori knowledge is not compatible with discovering unknown data variation and is, therefore, inapplicable to most cellular imaging applications.

We present a-priori knowledge-free unsupervised classification of structurally different particles employing pairwise registration with the Bhattacharya cost function as dissimilarity metric in combination with multidimensional scaling. We achieve 96% classification accuracy on mixtures of up to four different DNA origami structures, detect rare classes of mirrored origami occurring at a 2% rate (see **Figure 1** below), detect height variation in tetrahedron-shaped 3D DNA origami and capture variation in the ellipticity of nuclear pore complexes.

We anticipate that our training-free classification method will be instrumental in identifying rare subclasses of structurally different particles in a wide range of 2D and 3D single-molecule localization microscopy samples.

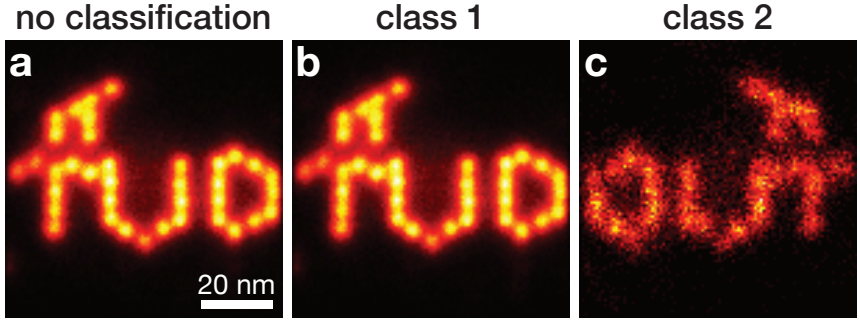

Figure 1 | Classification of 'flipped' DNA Origami. a, Particle fusion result without classification of 456 DNA origami structures of the TUD-logo. **b**-**c**, The two classes resulting from classification of **a**, containing 446 (normal) and 10 (flipped) particles per class, respectively.

[1] Heydarian, H. *et al.*, *Nature methods*, *15*(10), 781-784, 2018.
[2] Sieben, C. *et al.*, *Nature methods*, *15*(10), 777-780, 2018.
[3] Gray, R. D., *et al.*, *Scientific reports*, *6*(1), 1-8, 2016.
[4] Auer A. *et al.*, *Bioinformatics*, *36*(11), 3620-3622, 2020.