An open-source platform for single-molecule localization

A Janelia Open Science Software Initiative proposal

Srini Turaga, Jakob Macke, Jonas Ries

Single-molecule localization microscopy (SMLM), has revolutionized biology as it extends the resolution of light microscopy by up to 50-fold. SMLM, and related techniques like single particle tracking, single molecule FISH and *in situ* sequencing depend on advanced analysis algorithms a) to determine the precise positions of single fluorophores in the camera images, b) to visualize the data in 2D and 3D and c) to extract quantitative measures from the specific data (point-clouds, coordinates of fluorophores with uncertainties) to drive biological interpretation.

We propose to develop a complete python-based user-friendly GUI-based open source platform for single-molecule localization that will become the standard software for SMLM and targets various end-user applications.

We have recently developed DECODE (Speiser, Mueller, et al Nature Methods 2021), a deep learning method for highly accurate single molecule localization in challenging data. DECODE has been publicly released as an open source (pytorch based) python project (https://github.com/TuragaLab/DECODE). It can be installed locally or run in the cloud in Google colab via a Jupyter notebook interface. DECODE has been downloaded over 600 times in the last 6 months, including for use in the Colab notebooks.

We have also been working on an extension to DECODE for the analysis of single molecule FISH datasets. While conceptually similar to SMLM based super-resolution microscopy, this technique has been used to localize and count mRNA transcripts as an assay of gene expression. In preliminary unpublished work (Speiser et al in prep), we show the effectiveness of DECODE at analyzing these datasets.

Our current implementation of DECODE suffers from two drawbacks. First, it relies on a non-interactive Jupyter notebook based user interface. Second, for key steps in the analysis pipeline including calibration, and sophisticated visualizations, DECODE currently relies on another Matlab-based package called SMAP (Ries Nature Methods 2019). SMAP is a full-fledged user-friendly Matlab-based SMLM software which provides an extensive interactive GUI.

Both versions of DECODE described above will benefit from the development of a stand-alone platform integrating analysis and visualization in a seamless manner. We propose to start the development of this open source platform for single molecule localization by building a Napari-based interface, which is an open source python based visualization & processing platform for microscopy data. SMAP will provide the design framework for this new platform. We will start by incorporating the SMAP algorithms and GUI from Matlab into the new napari+DECODE python platform.

Impact at Janelia
Several labs and projects involve the localization of single molecules. The AIC and James Liu's lab independently perform SMLM using the lattice light sheet microscope and the iPALM microscope respectively. Both James and Ulrike Boehm (AIC) have expressed interest in developing and adopting DECODE for their use. They currently rely on outdated matlab and labview based custom algorithms and are eager to modernize.

Further, several ambitious projects at Janelia use multi-round FISH to assay gene expression and classify neuronal cell type in large brain tissue volumes in mouse cortex and the zebrafish brain. Efficient and accurate algorithms for the detection and localization of mRNA transcripts will be important for these projects. We have already been collaborating with Stephan Preibisch to develop a DECODE based FISH analysis pipeline.

Broader impact and community outreach
The most popular open-source softwares for SMLM like QuickPALM, RapidSTORM or ThunderSTORM are outdated (last major commit for ThunderSTORM was 2014) and lack functionality and performance. SMAP provides the most functionality of all SMLM software, has been continuously developed for the last 10 years to include state-of-the-art algorithms, and has an increasing user-community (on average approx. 10 unique clones per week), but its dependence on MATLAB renders it not suitable for a community-driven open-source project. Napari is an open source python based visualization platform for microscopy data, recently developed by Nick Sofroniew and colleagues at CZI. Its aim to replace ImageJ as the standard image analysis platform has been boosted by a massive funding program by CZI. Thus, Napari is a natural platform for the proposed software. The Napari team has received numerous requests to integrate SMLM rendering and several SMLM groups (Leterrier, Royer, Sibarista, Ries) have started to organize a community to contribute to this goal.

Several labs in the community have already adopted DECODE in their SMLM analysis pipeline. But broader adoption will continue to be limited by the lack of a full-fledged DECODE GUI, and the clunky reliance on Matlab-based SMAP for key steps in the analysis pipeline. We have identified several key outside collaborators/power users including Christophe Leterrier (CNRS) and Wes Legant (UNC) who will help design and test the GUI.

We also hope a well-documented and modular open source platform for SMLM with broad adoption will also prove to be a useful resource for algorithm developers. Such a platform, along with datasets already available, will lower the barrier to entry for computer scientists, enabling them to easily plug in and evaluate their own algorithm innovations without needing to develop an entire analysis and visualization pipeline.

Proposed plan
We propose to develop an open-source community-driven Napari-based SMLM software as a close collaboration between the Turaga, Ries, and Macke research groups.

GUI for DECODE: We propose to integrate DECODE with napari as the interactive GUI front-end.

Point spread function (PSF) model generation in Napari+DECODE: Currently, the PSF model necessary for DECODE and standard 3D fitting is generated from experimental bead stacks using SMAP. Both the Turaga and the Ries lab are developing new algorithms to determine PSF models using inverse modeling. Thus we propose to integrate these newly developed algorithms in Napari for a fully integrated DECODE workflow.

3D SMLM rendering in Napari: We propose to extend Napari with 3D visualization routines for SMLM. This includes rendering every localization as a Gaussian with a size determined by the localization and user-defined color coding and visualization as a sum projection and flexible filtering for localizations.

Extending functionality of SMLM software platform: In the future, we will successively integrate the main functionality from SMAP to the new software platform. This includes post-processing steps such as merging of fluorophores in consecutive frames, multi-color assignment or drift-correction and routines to extract biological information from the point-cloud data, including model fitting and single-particle averaging (Wu et al, bioRxiv 2021).

Community involvement: We aim to conduct virtual training workshops and hackathons to nucleate an active developer and user community around our proposed platform. Our previous participation in the I2K workshop last year already led to several active users of the DECODE software. We see these activities as an important way to publicize our project and to train users.