DeepSPT

DeepSPT is a state-of-the-art machine learning framework designed for analyzing single-particle tracking (SPT) data. It facilitates the rapid and accurate mapping of dynamic biological processes at nanometer and millisecond scales. DeepSPT's unique approach leverages diffusional behavior to reveal previously hidden biological insights, such as key timepoints in processes like rotavirus internalization and cellular localization, without requiring multi-color labeling.

Overview

DeepSPT consists of three main modules, each designed to process and analyze SPT data in a specific way:

1. **Temporal Segmentation**
2. **Diffusional Fingerprinting**
3. **Task specific Classifier**

Each module is accessible through its respective tab in the main application window. The modules are typically run sequentially, with the output of one module serving as the input for the next. This workflow allows for flexible experimentation, as users can save and reuse outputs, avoiding the need to rerun modules unnecessarily.

M1: Temporal Segmentation

Purpose

The Temporal Segmentation module is the first step in the DeepSPT analysis pipeline. It segments the particle trajectories based on their motion characteristics, which are then used in subsequent analyses.

Input Requirements

The Temporal Segmentation module accepts input files strictly in CSV format. The CSV file must contain the following columns:

- **Particle ID**: A unique identifier for each tracked particle.
- **Timepoint (Frame)**: The specific timepoint or frame number corresponding to each observation.
- **Coordinates**:
    - For 2D experiments: X and Y columns.
    - For 3D experiments: X, Y, and Z columns.

If multiple CSV files/experiments are loaded, it is important that they all have the same format.

Additional columns in the CSV file are allowed but will not impact the analysis, as you will be prompted to select the relevant columns during the setup process.

Running Temporal Segmentation

Follow these steps to run the Temporal Segmentation module:

1. **Select the Dimension**: Choose whether your experiment is 2D or 3D.
2. **Upload Data**: Browse and select one or more CSV files that contain the tracked experiment data in the required format.
3. **Specify Output Folder**: Choose a location on your computer where all the results from the modules will be saved.
4. **Map Columns**: Identify the relevant columns in your CSV file:
   ○ Particle ID
   ○ Timepoint (Frame)
   ○ X, Y, Z (if 3D)
5. **Adjust Columns**: Click the "Adjust Columns" button to confirm your selections.
6. **Set Optional Parameters**:
   ○ **Frame rate**: Default is 1 second.
   ○ **Pixel size**: Default is 0.1 μm.
   ○ **Pixel size (Z)**: Only for 3D experiments.
7. **Run the Segmentation**: Click the "Run Temporal Segmentation" button.

A progress bar will appear, showing the status of the segmentation process. Upon completion, a confirmation message will appear, and the results will be saved in the designated output folder. By default, the results will be saved in .pkl format. If needed, the user can also save the segmentation results as a CSV file by clicking the "Save in CSV" button.

Output Format

The output CSV file will contain the following columns:

- **Particle ID**
- **Timepoint (Frame)**
- **Coordinates (X, Y, Z for 3D)**
- **Ensemble Prediction**: Indicates the type of motion detected at each timepoint:
  ○ 0 → Normal
  ○ 1 → Directed
  ○ 2 → Confined
  ○ 3 → Subdiffusive
- **Ensemble Scores (0-3)**: Confidence scores for each motion type, summing to 1. The motion type with the highest score is the one predicted by the model.

M2: Diffusional Fingerprinting

Purpose

The Diffusional Fingerprinting module further analyzes the segmented data by extracting features that characterize the particle motion. These features can be used for in-depth analysis and comparison of different experimental conditions.

Starting Diffusional Fingerprinting

There are two ways to initiate the Diffusional Fingerprinting module:

1. **Proceed Directly**: Continue from the Temporal Segmentation module immediately after it completes. The application retains the segmentation data needed for fingerprinting.
2. **Load Saved Data**: If you have previously saved the output from the Temporal Segmentation module as .pkl files, you can load these files directly. Additionally, you need to reload the original CSV files containing the tracked particle data and repeat the column mapping process.

Selecting Fingerprinting Features

In the main fingerprinting screen, you will see a list of 43 feature options. By default, all features are selected, but you can customize the selection based on your analysis needs.

Types of Fingerprints

You can choose between two types of fingerprints:

- **Global Fingerprint**: Generates one set of features per particle track.
- **Temporal Fingerprint**: Generates one set of features per timepoint within each track.

Saving and Plotting Results

After the fingerprinting process is complete, save the fingerprint files by clicking the "Save Results" button. To visualize the results, click "Plotting Options." This feature allows you to:

- Navigate through particles or experiments.
- Visualize different types of motion in corresponding colors.
- Explore fingerprint features.
- Generate various 2D and 3D plots.
- Compare experiments or files.

All plots, files, and images you choose to save will be stored in the output folder specified during the initial setup.

M3: Task specific Classifier

Purpose

The Task specific Classifier module classifies the particle tracks based on the features extracted during the Diffusional Fingerprinting process. This classification can be done globally or temporally, depending on the nature of the fingerprint data.

Running the Task specific Classifier

You can run the Task specific Classifier in several ways:

1. **Sequential Run**: Use the fingerprints generated by the Diffusional Fingerprinting module and input labels to train a classification model.
2. **Manual Load**: Load previously saved fingerprints and labels, and then train the model.
3. **Load Pre-trained Model**: Use a pre-trained model to predict new labels on fresh fingerprinted data.

Classification Modes

- **Global Mode**: After generating global fingerprints, provide labels for each particle track. The model trains on these labels and can predict labels for new tracks.
- **Temporal Mode**: After generating temporal fingerprints, provide a CSV file containing one label per timepoint for each track. The model trains on these labels and can predict labels for new timepoints.

Important: The number of labels must match the number of tracks (Global) or timepoints (Temporal) in the fingerprint data.

Labeling

If you wish to assign string names to your labels, press the "Name Labels" button. Labels should be provided in a CSV file with a column named Labels. Labels must be integer values (e.g., 0, 1, 2).

Conclusion

DeepSPT is a powerful tool for analyzing single-particle tracking data, offering deep insights into dynamic biological processes. By following this guide, you can effectively use each of the three modules—Temporal Segmentation, Diffusional Fingerprinting, and Task specific Classifier—to extract meaningful information from your experiments.

For further assistance, please refer to the specific module instructions within the application or consult the documentation provided.

Contact:

Jacob Kæstel-hansen

 PhD fellow

Department of Chemistry

 jkh@chem.ku.dk

 Nikos Hatzakis, Professor

 Department of Chemistry

 hatzakis@chem.ku.dk

Konstantinos Tsolakidis

Software Engineer, Hatzakis Lab

Mail: kt@chem@ku.dk

tel: +45 50276408

University of Copenhagen

Department of Chemistry