The user manual of DEBRIS (Ver. 1.1 (2024-08-20))

Contents

System requirements

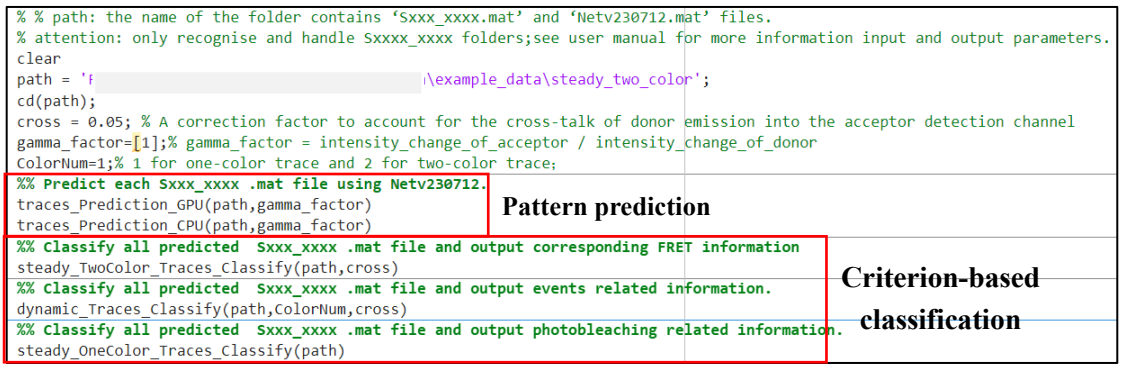
- We conducted training, validation, and testing using MATLAB (version: 9.13.0, R2022b) on a PC with an Intel Xeon W2223 CPU and an NVIDIA GeForce RTX 3090 GPU.
- Compatibility: Windows 10/11 with MATLAB (R2022b) or later version.
- Additional toolboxes required for MATLAB to run DEBRIS:
  1. Signal Processing Toolbox
  2. Deep Learning Toolbox
  3. Parallel Computing Toolbox (TM)

DEBRIS usage

DEBRIS can handle both single-color and two-color single-molecule traces of varying lengths, as well as steady fluorescence signals at equilibrium and dynamically emerging signals under equilibrium and non-equilibrium conditions.

Please download the 'DEBRIS_code' folder of latest version from https://github.com/CHENChunlai-CN/DEBRIS.git.

We have integrated DERBIS pattern prediction and criterion-based classification into the main 'DEBRIS.m' file. This allows for user-friendly automatic prediction and classification by simply adjusting the parameters and calling the corresponding functions directly.



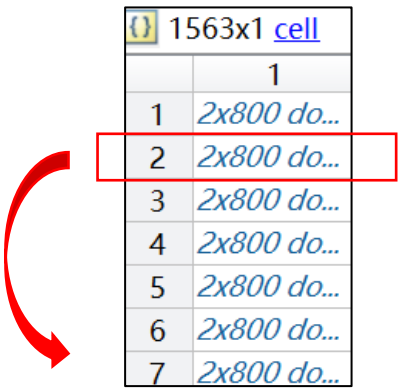
Step1: Prepare data

1. Data format:

Intensity-time data is stored as an 'Sxxxx_xxxx.mat' file.

Single-molecule traces are stored in a cell array named 'Traces' as a single raw.

Each element in the cell array contains information regarding the intensity of single-molecule traces over time, with the number of columns representing time (frame number).



Traces | Traces{1, 1}

Traces{1, 1}

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1.9794e+...	1.1622e+...	1.5218e+...	1.1532e+...	1.1704e+...	8.5409e+...	1.6496e+...	1.2195e+...	1.4765e+...	1.0549e+...	1.5722e+...	1.0657e+...	1.2999e+...
2	1.2840e+...	2.8437e+...	6.7062e+...	1.2814e+...	1.0367e+...	1.0225e+...	1.2549e+...	1.1429e+...	8.5331e+...	9.4849e+...	9.0556e+...	9.6684e+...	5.2110e+...

The two-color trace displays the fluorescence intensity of the donor in the first row and the fluorescence intensity of the acceptor in the second row.

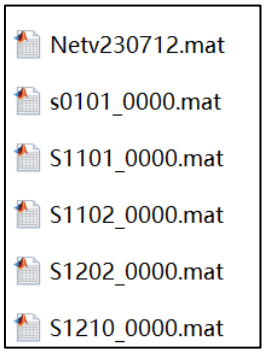
Traces | Traces{1, 1}

Traces{1, 1}

1	2	3	4	5	6	7	8	9	10	11	12	13
2.6679e+...	3.7605e+...	3.6649e+...	3.3399e+...	3.8290e+...	3.8378e+...	4.2075e+...	4.2303e+...	3.4547e+...	2.7697e+...	3.4595e+...	3.1376e+...	3.1431e+...
2.6679e+...	3.7605e+...	3.6649e+...	3.3399e+...	3.8290e+...	3.8378e+...	4.2075e+...	4.2303e+...	3.4547e+...	2.7697e+...	3.4595e+...	3.1376e+...	3.1431e+...

For a single-color trace, please copy the first row of data to the second row. Both rows should be identical. Please refer to Figure S9 of the current manuscript.

2. **File organization:** Place all ‘Sxxx_xxx.mat’ files into a designated folder. Additionally, copy the provided network file ‘Netv230712.mat’ into the same folder.



Step2: Predict pattern by DEBRIS

1. Parameters input

1) Path: the name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.
2) gamma_factor: the gamma factor, γ, accounts for the differences in quantum yield and detection efficiency between the acceptor and the donor[1]. The default value is 1. Users can approximate the input gamma factor by the ratio of the intensity changes of the acceptor and donor around the acceptor photobleaching time point or the FRET conversion time point. Furthermore, DEBRIS can tolerance to the inaccuracy of the input gamma factor value to certain degree (See Fig. S17 of the supplementary). Multiple values of Gamma_factor can be inputted as a single matrix for parallel calculation. Each input gamma factor will output a corresponding prediction results named after 'CγNet230712Sxxxx_xxxx.mat' and related classification results.

2. Function call

DEBRIS runs on both CPU and GPU. It is recommended to use GPU. The time for processing 2000 traces of 800 frames on our PC using an NVIDIA GeForce RTX 3090 GPU is ~ 7 minutes. On the other hand, using an Intel Xeon W2223 CPU, the same processing is estimated to take 27 minutes.

1) `traces_Prediction_CPU (path, gamma_factor); % use CPU to process.`
2) `traces_Prediction_GPU (path, gamma_factor); % use GPU to process.`

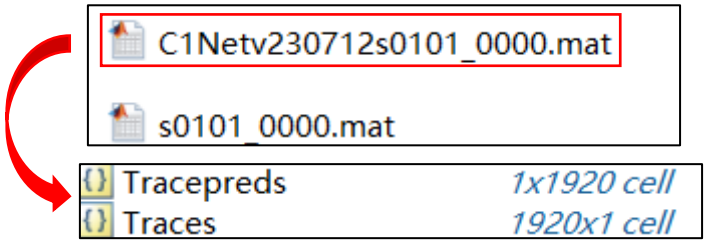
3. Output

1) Tracepreds: This cell array contains pattern-time traces generated by the DEBRIS. The contents of the cell array are matrices, and each matrix contains three columns.

   Column #1: Intensity of donor

   Column #2: Intensity of acceptor

   Column #3: Pattern-time trace of local features



2) CxNetv230712Sxxx_xxxx.mat: The file contains “Traces” and “Tracepreds”.

4. Normalization

The normalization is done via the codes 'traces_Prediction_CPU.m' and 'traces_Prediction_GPU.m' before prediction.

```
minI=2e4; % normalization: always 2e4, recommend value is 50%~70% × (ID+IA)
frag(end+1,1) = {(temptrace2(j: j+9, :)'/max(minI,mean(maxk(sum(temptrace2(j:j+9,:),2),3))))};
```

The fluorescence intensity of each 10-frame segment is normalized by dividing it by the average of the first three maxima of total intensities (donor + acceptor) or minI, whichever is larger. The default value of minI is 20,000. The recommended value of minI is 50%~70% of total intensities (donor + acceptor) of typical single-molecule traces.

Step3: Identify traces using user-defined criteria

1. Analysis of two-color traces of steady fluorescence signals

1) Parameters input

a) Path: The name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.

b) Cross: A correction factor to account for the crosstalk of donor emission into the acceptor detection channel. For our instrument, 0.05 is used for Cy3 and Cy5 FRET pair, and 0.01 is used for AF488 and Cy5 FRET pair. The default setting is 0 and no crosstalk correction is used.

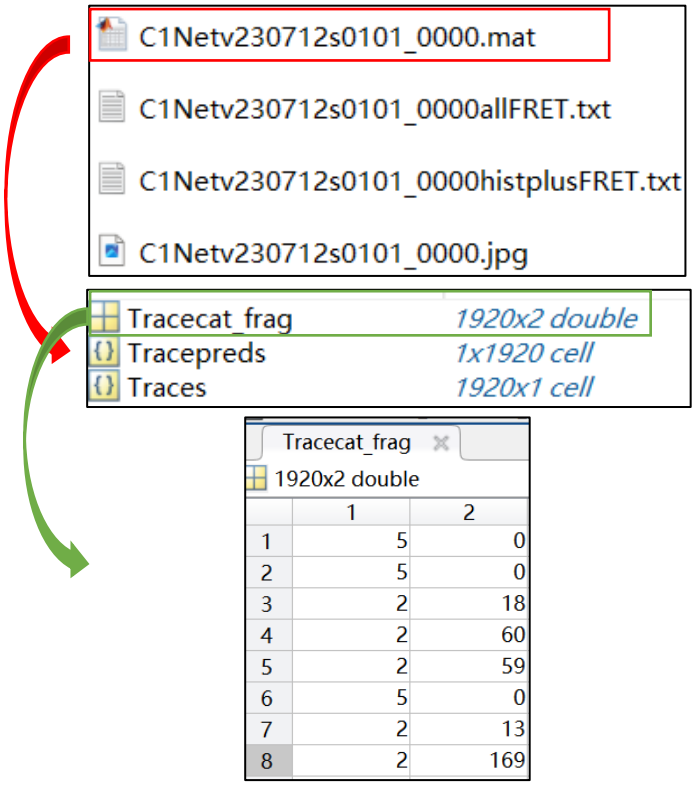
2) Function call

```
Steady_TwoColor_Traces_Classify(path,cross)
```

The time for processing 2000 traces of 800 frames on our PC using an Intel Xeon W2223 CPU is ~ 35 seconds.

3) Output

a) Tracecat_frag: Contains the category and photobleaching timepoint of each trace. The first column represents the classification of trace: 1 for 'Donor bleached', 2 for 'Acceptor bleached', and 5 for 'Discarded' (see examples in Figure 2d of the current manuscript). The second column represents the photobleaching timepoint of the current trace.



	1	2
1	5	0
2	5	0
3	2	18
4	2	60
5	2	59
6	5	0
7	2	13
8	2	169

b) C1Netv230712Sxxx_xxxxallFRET.txt: Records the FRET of each frame of all selected traces before photobleaching.

c) C1Netv230712Sxxx_xxxxhistFRETplus.txt: Histogram of FRET from allFRET.txt files, range from -0.21 to 1.21, bin size is 0.01.

d) C1Netv230712Sxxx_xxxx.jpg: FRET histogram in jpg.

2. Analysis of two-color traces of dynamically emerging signals

1) Parameters input

a) Path: The name of the folder contains ‘Sxxx_xxxx.mat’ and ‘Netv230712.mat’ files.

b) Cross: A correction factor to account for the crosstalk of donor emission into the acceptor detection channel. For our instrument, 0.05 is used for Cy3 and Cy5 FRET pair, and 0.01 is used for AF488 and Cy5 FRET pair. The default setting is 0 and no crosstalk correction is used.

c) ColorNum: 1 for one-color trace and 2 for two-color trace. To analyze two-color traces of dynamically emerging signals, please set ColorNum=2.

2) Function call

```
Dynamic_Traces_Classify(path,ColorNum,cross)
```

The time for processing 2000 traces of 800 frames on our PC using an Intel Xeon W2223 CPU is ~ 50 seconds.

3) Output

a) EventsInfoS: Record 'aSxxx_xxxxEventsInfoS.txt' information of selected events.

Column #1: Number of selected traces.

Column #2: Category of the event. 1 for ‘Selected’ and 5 for ‘Discarded’ (see examples in Figure 4b of the current manuscript).

Column #3: The frame number of the event appears.

Column #4: The frame number of the FRET disappears.

Column #5: The frame number of the event disappears.

Column #6: The frame number of the last event ends, set to 0 if this is the first event of this trace.

Column #7: Number of frames between two adjacent events.

Column #8: Dwell frames of the event.

Column #9: FRET of the first frame where the event appears.

Column #10: FRET of the last frame where the event disappears.

Column #11: Average FRET of the first three frames where the event appears.

Column #12: Average FRET of the last three frames where the event disappears.

Column #13: Average FRET of the whole event.

b) PairMatrix: The file is a cell array, recording the following information of all events, including discarded ones.

Column #1: The frame number of the event appears in the pattern-time trace.

Column #2: The frame number of the event disappears in the pattern-time trace.

Column #3: Category of the event, 1 for ‘Selected’ and 5 for ‘Discarded’.

Column #4: The frame number of the FRET ends in the intensity-time trace.

Column #5: The frame number of the event ends in the intensity-time trace.

c) SyncM1: Time-dependent FRET was synchronized at the appearance of events, save as 'aSxxx_xxxxSyncM1.txt'.

Column #1: The order is frame sequence numbered from -30 to 100, with 0 representing the synchronized time point at which the event appears.

Column #2: Donor intensity at the current frame number.

Column #3: Acceptor intensity at the current frame number.

Column #4: FRET at the current frame number.

Column #5: Dwell time of the event.

d) SyncM2: Time-dependent FRET was synchronized at the disappearance of events, save as 'aSxxx_xxxxSyncM2.txt'.

Column #1: The order is numbered from -100 to 30 with 0 representing the time at which the event disappears.

Column #2: Donor intensity at the current frame number.

Column #3: Acceptor intensity at the current frame number.

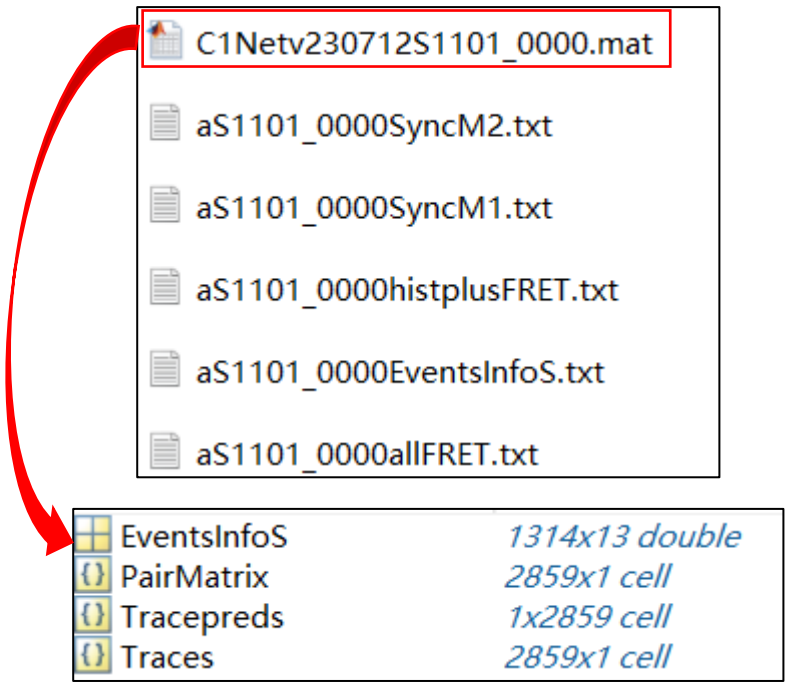
Column #4: FRET at the current frame number.

Column #5: Dwell time of the event.

e) aSxxx_xxxxallFRET.txt: Records the FRET of each frame of all selected traces before photobleaching.

f) aSxxx_xxxxhistFRETplus.txt: Histogram of FRET from allFRET.txt files, ranging

from -0.21 to 1.21, the bin size is 0.01.



3. Analysis of one-color traces of dynamically emerging signals

1) Parameters input

a) Path: The name of the folder contains ‘Sxxx_xxxx.mat’ and ‘Netv230712.mat’ files.

b) Cross: A correction factor to account for the crosstalk of donor emission into the acceptor detection channel. For our instrument, 0.05 is used for Cy3 and Cy5 FRET pair, and 0.01 is used for AF488 and Cy5 FRET pair. The default setting is 0 and no crosstalk correction is used.

c) ColorNum: 1 for one-color trace and 2 for two-color trace. To analyze one-color traces of dynamically emerging signals, please set ColorNum=1.

2) Function call

```
Dynamic_Traces_Classify(path,ColorNum,cross)
```

The time for processing 2000 traces of 800 frames on our PC using an Intel Xeon W2223 CPU is ~ 8 seconds.

3) Output

a) EventsInfoS: Record 'aSxxx_xxxxEventsInfoS.txt' information of selected events.

Column #1: Number of selected traces.

Column #2: Category of the event. 1 for ‘Selected’ and 5 for ‘Discarded’ (see examples in Figure 6b of the current manuscript).

Column #3: The frame number of the event appears.

Column #4: The frame number of the FRET disappears.

Column #5: The frame number of the event disappears.

Column #6: The frame number of the last event end, set to 0 if this is the first event of this trace

Column #7: Number of frames between two adjacent events.

Column #8: Dwell frames of the event.

Column #9: FRET of the first frame where the event appears. For one-color traces, FRET is always 0.5.

Column #10: FRET of the last frame where the event disappears. For one-color traces, FRET is always 0.5.

Column #11: Average FRET of the first three frames where the event appears. For one-color traces, FRET is always 0.5.

Column #12: Average FRET of the last three frames where the event disappears. For one-color traces, FRET is always 0.5.

Column #13: Average FRET of the whole event. For one-color traces, FRET is always 0.5.

b) PairMatrix: The file is a cell array, recording the following information of all events, including discarded ones.

Column #1: The frame number of the event appears in the pattern-time trace.

Column #2: The frame number of the event disappears in the pattern-time trace.

Column #3: Category of the event, 1 for 'Selected' and 5 for 'Discarded'.

Column #4: The frame number of the FRET ends in the intensity-time trace.

Column #5: The frame number of the event ends in the intensity-time trace.

4. Analysis of one-color traces of steady fluorescence signals

1) Parameters input

a) Path: The name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.

2) Function call

```
Steady_OneColor_Traces_Classify(path)
```

The time for processing 2000 traces of 800 frames on our PC using an Intel Xeon W2223CPU is ~ 60 seconds.

3) Output

a) Tracecat_frag: Contains the category and photobleaching time of each trace. The first column represents the classification of the trace: 1 for 'selected' and 5 for 'discarded' (see examples in Figure 5a-5b of the current manuscript). The second column represents the photobleaching timepoint (frame time) of the current trace.

b) BleachInfo: Log information about selected traces, save as 'aBleachSxxx_xxxx.txt'.

Column #1: Number of selected traces.

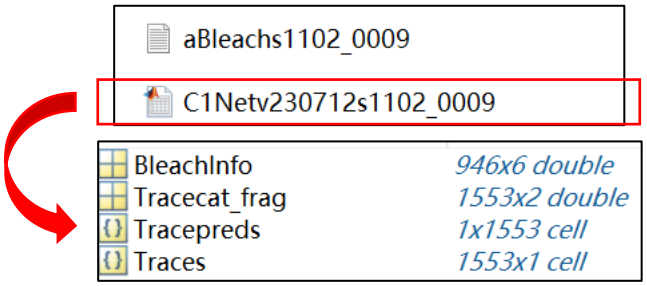
Column #2: Photobleaching frame time of the selected trace.

Column #3: Average intensity before photobleaching.

Column #4: Sum of all intensities before photobleaching.

Column #5: Standard deviation of intensity before photobleaching

Column #6: Signal-to-noise ratio before photobleaching.

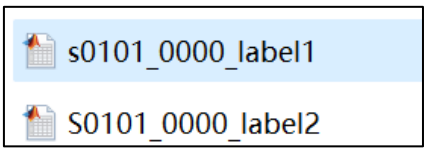


Example data demonstration

1. Analysis of two-color traces of steady fluorescence signals

```
% % path: the name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.
% attention: only recognise and handle Sxxxx_xxxx folders;see user manual for more information input and output parameters.
clear
path = '                                    \example_data\steady_two_color';
cd(path);
cross = 0.05; % A correction factor to account for the cross-talk of donor emission into the acceptor detection channel
gamma_factor=[1];% gamma_factor = intensity_change_of_acceptor / intensity_change_of_donor
ColorNum=1;% 1 for one-color trace and 2 for two-color trace;
%% Predict each Sxxx_xxxx .mat file using Netv230712.
traces_Prediction_GPU(path,gamma_factor)
% traces_Prediction_CPU(path,gamma_factor)
%% Classify all predicted  Sxxx_xxxx .mat file and output corresponding FRET information
steady_TwoColor_Traces_Classify(path,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output events related information.
% dynamic_Traces_Classify(path,ColorNum,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output photobleaching related information.
% steady_OneColor_Traces_Classify(path)
```

In the example data of 'steady_two color', we provide two files in the 'manual labelling by experts' folder. Each file contains the 'Tracescat_label' and 'Traces' variables, where 'Traces' stores the intensity-time traces. 'Tracescat_label' contains two columns. The first column represents the classification of the trace: 1 for 'donor bleached', 2 for 'acceptor bleached' and 5 for 'discarded'. The second column represents the photobleach time (frame number) of the current trace.



2. Analysis of two-color traces of dynamically emerging signals

```
% % path: the name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.
% attention: only recognise and handle Sxxxx_xxxx folders;see user manual for more information input and output parameters.
clear
path = '                                    \example_data\dynamic_two_color';
cd(path);
cross = 0.05; % A correction factor to account for the cross-talk of donor emission into the acceptor detection channel
gamma_factor=[1];% gamma_factor = intensity_change_of_acceptor / intensity_change_of_donor
ColorNum=2;% 1 for one-color trace and 2 for two-color trace;
%% Predict each Sxxx_xxxx .mat file using Netv230712.
traces_Prediction_GPU(path,gamma_factor)
% traces_Prediction_CPU(path,gamma_factor)
%% Classify all predicted  Sxxx_xxxx .mat file and output corresponding FRET information
% steady_TwoColor_Traces_Classify(path,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output events related information.
dynamic_Traces_Classify(path,ColorNum,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output photobleaching related information.
% steady_OneColor_Traces_Classify(path)
```

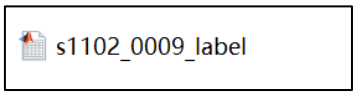
3. Analysis of one-color traces of dynamically emerging signals

```
% % path: the name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.
% attention: only recognise and handle Sxxxx_xxxx folders;see user manual for more information input and output parameters.
clear
path = '                                    \example_data\dynamic_one_color';
cd(path);
cross = 0; % A correction factor to account for the cross-talk of donor emission into the acceptor detection channel
gamma_factor=[1];% gamma_factor = intensity_change_of_acceptor / intensity_change_of_donor
ColorNum=1;% 1 for one-color trace and 2 for two-color trace;
%% Predict each Sxxx_xxxx .mat file using Netv230712.
traces_Prediction_GPU(path,gamma_factor)
% traces_Prediction_CPU(path,gamma_factor)
%% Classify all predicted  Sxxx_xxxx .mat file and output corresponding FRET information
% steady_TwoColor_Traces_Classify(path,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output events related information.
dynamic_Traces_Classify(path,ColorNum,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output photobleaching related information.
% steady_OneColor_Traces_Classify(path)
```

4. Analysis of one-color traces of steady fluorescence signals

```
% % path: the name of the folder contains 'Sxxx_xxxx.mat' and 'Netv230712.mat' files.
% attention: only recognise and handle Sxxxx_xxxx folders;see user manual for more information input and output parameters
clear
path = '                                    \example_data\steady_one_color';
cd(path);
cross = 0; % A correction factor to account for the cross-talk of donor emission into the acceptor detection channel
gamma_factor=[1];% gamma_factor = intensity_change_of_acceptor / intensity_change_of_donor
ColorNum=1;% 1 for one-color trace and 2 for two-color trace;
%% Predict each Sxxx_xxxx .mat file using Netv230712.
traces_Prediction_GPU(path,gamma_factor)
% traces_Prediction_CPU(path,gamma_factor)
%% Classify all predicted  Sxxx_xxxx .mat file and output corresponding FRET information
% steady_TwoColor_Traces_Classify(path,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output events related information.
% dynamic_Traces_Classify(path,ColorNum,cross)
%% Classify all predicted  Sxxx_xxxx .mat file and output photobleaching related information.
steady_OneColor_Traces_Classify(path)
```

In the example data of steady_one color, we provide a file in the 'manual labelling by experts' folder. Each file contains the variables 'Tracescat_label' and 'Traces', where 'Traces' stores the intensity-time traces. Tracescat_label' contains two columns. The first column represents the classification of the trace: 1 for 'selected' and 5 for 'discarded'. The second column represents the photobleach time of the current trace.



FRET calibration after steady two-color traces classification

```
%%for the calculation of the gamma factor and spectral crosstalk based on 'A photobleaaching' classification.
clear
path = '                                                    \Steady2Color\DEBRIS';
cd(path)
files = dir(fullfile(path,'*Netv230712*.mat'));
```

1. Code

FRET_correction.m

2. Parameters input

Path: location of DEBRIS prediction results storage

3. Output

```
gamma factor of s0101_0000:1.248392
cross talk of s0101_0000:0.045666
```

Using the 'acceptor photobleaching' category of the steady two-color traces, the gamma factor [1] ($\gamma = \frac{\Delta I_A}{\Delta I_D}$) is defined as the ratio of the change in acceptor intensity to the change in donor intensity around the acceptor photobleaching time point. Spectral crosstalk[1] ($cross = \frac{I_A}{I_D}$) is defined as the ratio of the acceptor channel intensity to the donor channel intensity after acceptor photobleaching.

If the proportion of 'acceptor photobleaching' traces is less than 25% of the total selected traces, the output will be "Note: too few traces of acceptor photobleaching, less reliable gamma factor".

References

1. Rahul, R., Sungchul, H. & Taekjip, H. A practical guide to single molecule FRET. *Nat Methods* **4**, 507–16 (2008).