IMageEXtractor Manual

Contents

1 Introduction

IMageEXtractor (IMEX) is an App developed for extracting data from scatter plots, enabling automatic extraction, and manual extraction and correction. IMEX is written in MATLAB R2018b, with add-on 'Computer Vison Toolbox v8.2' installed.

2 Overview
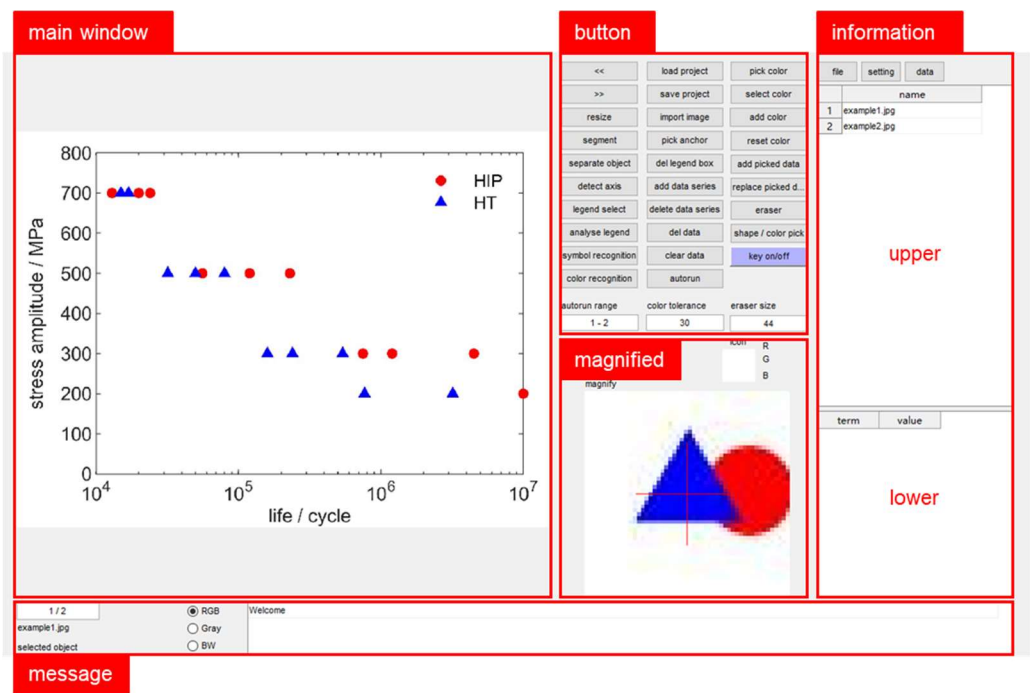
Figure 1. Major panels of IMEX.

The major panels of IMEX are shown in Figure 1. The main window shows the image to be processed. The button panel includes functional buttons and related input boxes.

The information panel displays the information on files, settings and data (Figure 2). Use the button at the top to switch between files, settings and, data information. Selecting an item in the information upper panel will trigger the lower panel to show
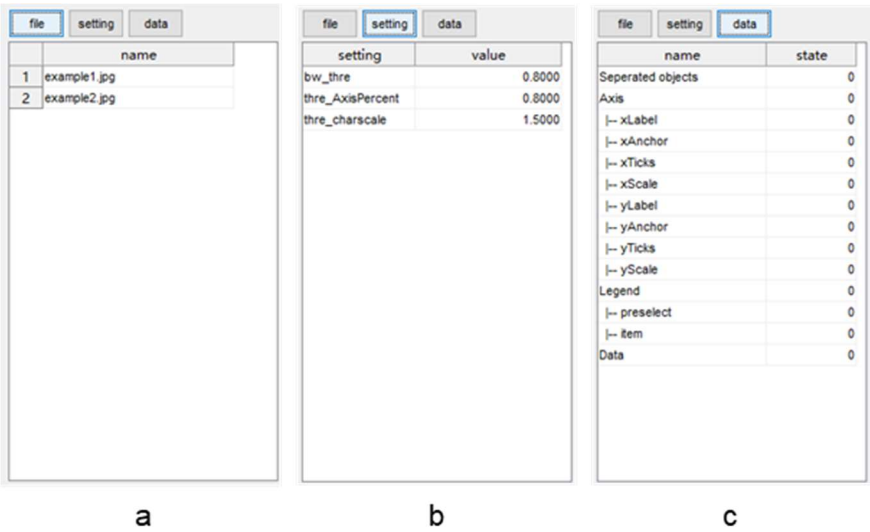
the corresponding detail data.

Figure 2. Information upper panel. (a) file, (b) setting and (c) data information.

The magnified panel displays the magnified view at the cursor position, marked by a red cross. In the upper left, there is also a small window to display the symbol of a data series and the RGB value of a pixel.

The message panel displays the message during app execution (Figure 3). In the first column, the first box shows the ID of the current image and the total number of images in the current project. It is an editable text box and entering a valid ID will trigger the main window to display the corresponding image. Below the text box, the image name and ID of selected object is shown. The second column includes 3 radio buttons to switch the color mode in the main window. The third column shows the real-time message during execution.

Figure 3. Message panel.

In the following, words in red color refer to panels of IMEX, and words in purple color refer to buttons.

3 Usage

3.1 Main window

The left button of the mouse can be used to add data points or select objects in the main window, such as data points, axis anchors, selected regions.

Holding the left button and moving the mouse will box select objects or regions.

Other functions of the left button will be activated for some buttons, which will be introduced in the following sections.

The middle scroll is used to zoom in or zoom out the image.

Holding the right button and moving the mouse will pan the image.

3.2 Quickstart

imex.m is the main file and **imex.fig** is used for the layout of the graphic user interface (GUI). Running the **imex.m** in MATLAB will launch a blank IMEX GUI. The quickstart involves steps shown in Figure 4.
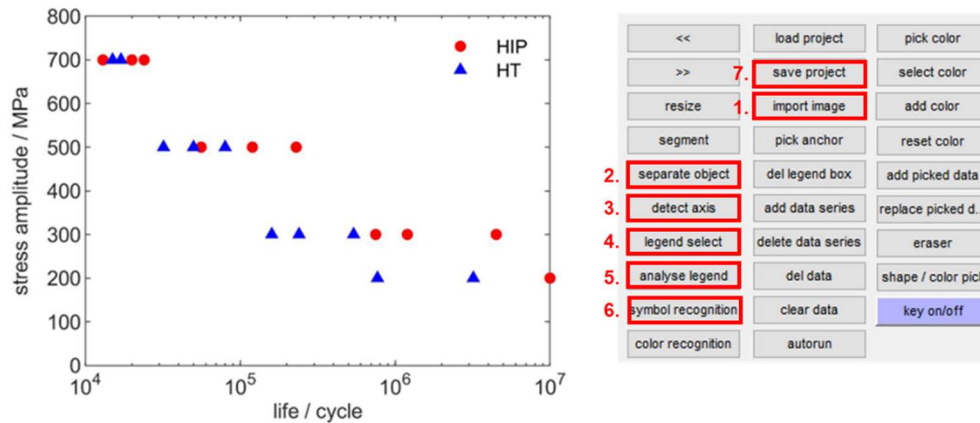
Figure 4. Steps of the quickstart.

1. Import image: click the import image button, a dialog box appears. Select the 'example' folder we provided, and IMEX will import all images under the folder.

2. Separate object: click the separate object button. A binary copy of the image is created. Connected black pixels in the binary image are identified as an object. Objects will be used for subsequent analysis.

3. Detect axis: click the detect axis button. The position of axis, axis ticks and axis labels are recognized automatically.

4. Legend selection: click the legend select button, the button is highlighted in blue. Use the left mouse button in the main window to box select the legend region (Figure 5). Click the legend select button again to finish the selection and the button will recover to gray. It is noted that color pixels irrelevant to the legend should not be included, which could influence the result of the analysis.

5. Analyse legend: click the analyse legend button, the symbols and corresponding annotations are analysed and stored as templates for the image. Click the 'item' in the information upper panel, the symbols and annotations will be wrapped by black boxes in the main window and the text of the annotation shown in the information lower panel.
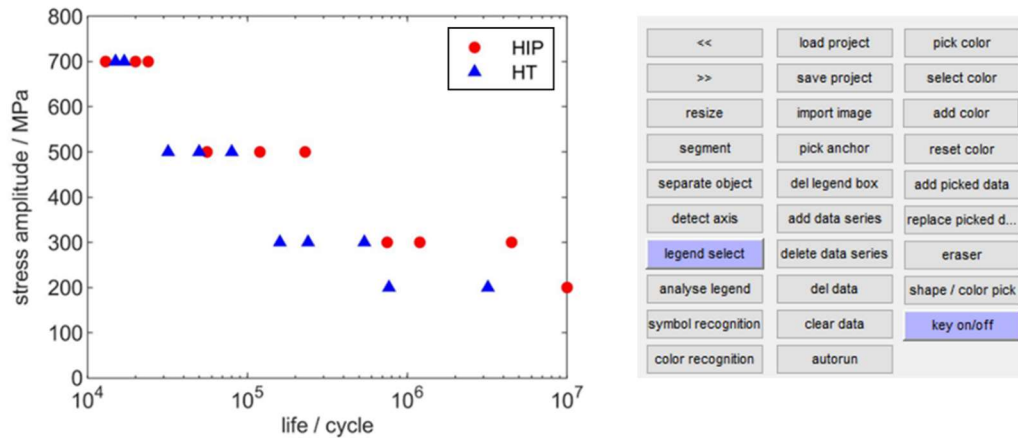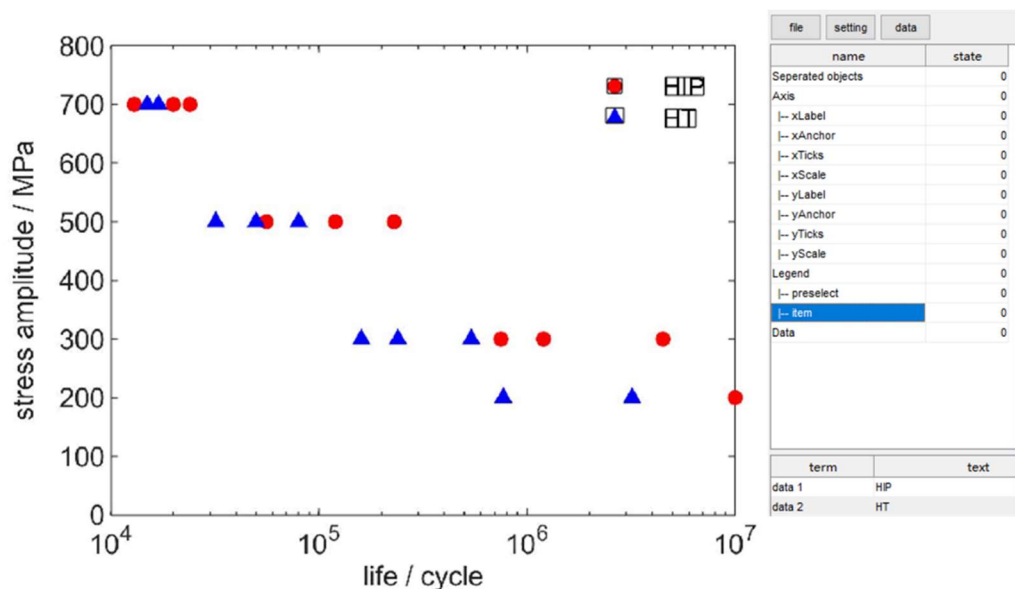
Figure 5. Box select the legend region.

Figure 6. The result of legend analysis.

6. Symbol recognition: click the symbol recognition button and the data points will be recognized according to the shape and color of the templates. Select a data 'series' in the information upper panel, recognized data points will be marked by yellow circles in the main window and corresponding numeric data shown in the information lower panel. If no 'series-n' appear in the information upper panel, click the 'data' button at the top to refresh the page.
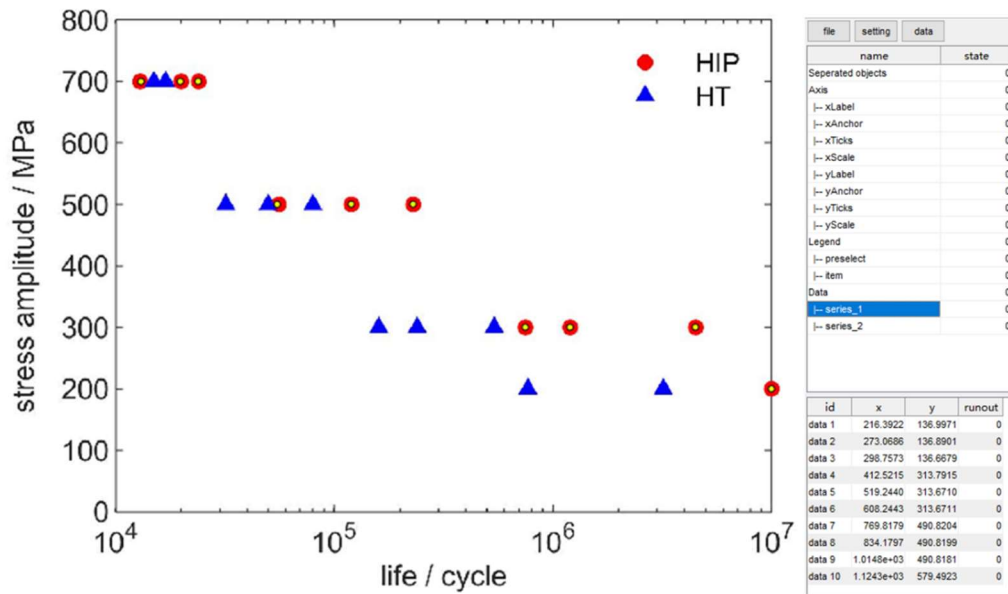
Figure 7. The result of symbol recognition.

When a data series is selected, the left mouse button is activated to select or add data points in the main window (Figure 8). When the mouse hovers over a data point, the cursor appears as a hand. Click the left mouse button at the time will select the data point and the marker of the data point will turn to a yellow cross. Once selected, the position of the data point can be adjusted by 'w', 's', 'a' and 'd' keys and deleted by the 'delete' key on the keyboard if the button key on/off is activated in blue. When the cursor appears as an arrow, clicking the left mouse button will add a data point to the data series.
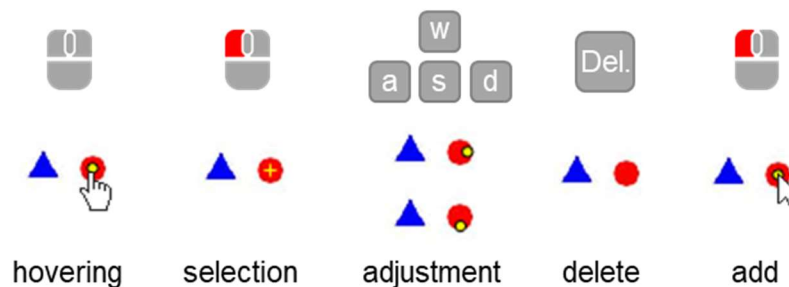

Figure 8. Interactions with the data points.

7. Save project: click the save project button and a dialog box will appear. Choose

the path and enter the filename for the project file, and save it as a MAT file (.mat). The project file records the information of axis and data in pixel coordinates for all images and can be transformed into data in physical units by post-processing. The project file can be loaded for subsequent analysis by using the load project button.
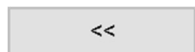
Tips: After importing images, legends can be selected at first. And then specify the autorun range and click the autorun button. The function of separate object, detect axis, analyse legend and symbol recognition will be executed sequentially for images in the autorun range.
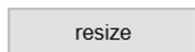
3.3 Functions of buttons

>>
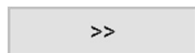
next image

Switch to the next image.

<<

previous image

Switch to the previous image.

resize

resize

Resize the image to fit the main window.
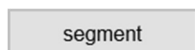
segment

segment

The button is used to segment images with multiple panels. First clicking the button will trigger the main window to enter the mode of box selection and the button turn

blue [segment] . In the main window, box select the target regions using the left mouse button. Selecting multiple regions is allowed. Clicking the segment button again completes the procedure and those regions will be saved as images to a 'segmentation' folder created in the current path.

[separate object]

separate object

A copy of a binary image is created and those connected black pixels are detected as objects. As shown in Figure 9, detected objects are wrapped in boxes. The object data will be passed to the subsequent detect axis, analyse legend, and symbol recognition functions.

Figure 9. Result of 'separate object'. Each object is wrapped by a box.

[detect axis]

detect axis

The x and y-axis are detected. According to the position of objects passed from previous separate object, objects were grouped for ticks' labels and axis labels, and then

parsed by optical character recognition (OCR). Selecting the 'xLabel', 'xTicks', 'yLabel' and 'yTicks' in the information upper panel, the corresponding object will be highlighted by a box in the main window and data shown in the information lower panel, as shown in Figure 10 and Figure 11. The scale of axis (linear/log) is determined according to the position and value of ticks.

Axis anchors are selected as the ticks at the two ends of the axes. They stand for the position and the range of axes. Selecting the 'xAnchor' and 'yAnchor' in the information upper panel, the xAnchor and yAnchor will be denoted by circle and triangle markers in the main window respectively (Figure 12). The anchors corresponding to the small and large tick value are in blue and red color, respectively. Only the x position of xAnchor and the y position of yAnchor matter.
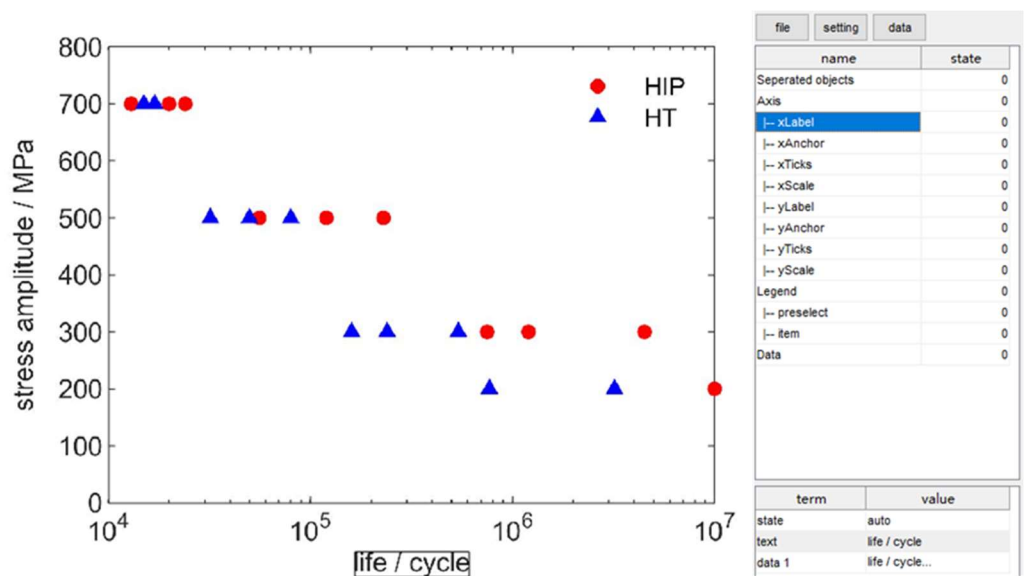

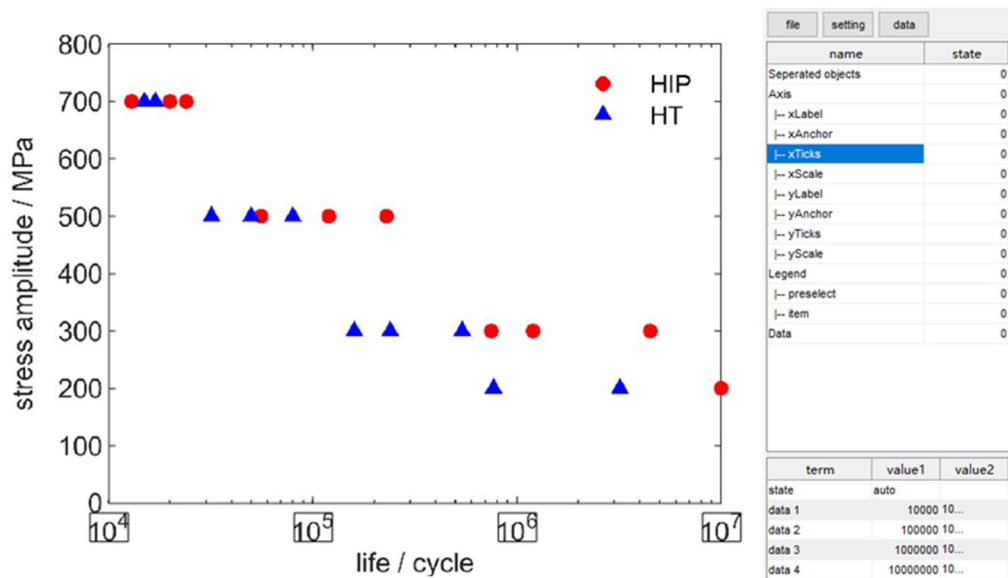
Figure 10. The result of parsing the label of x-axis.

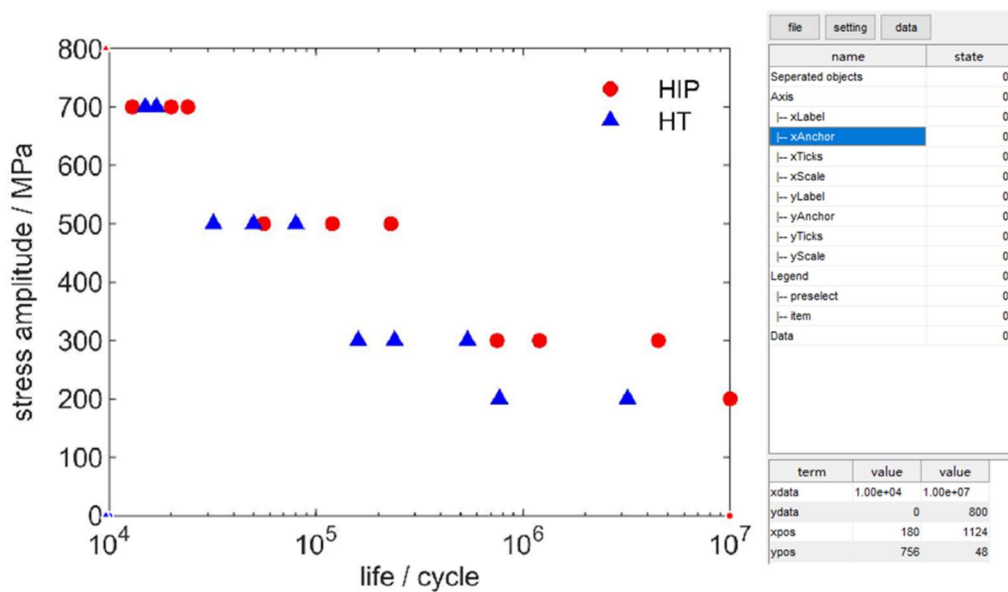Figure 11. The result of parsing the tick labels of x-axis.

Figure 12. The result of anchors of x and y-axis.

legend select

legend select

First clicking the button will trigger the main window to enter the mode of box selection and the button turn blue legend select . In the main window, box select the target regions using the left mouse button. Selecting multiple regions is allowed. Clicking the legend select button again completes the procedure and the button turns

back to gray.

analyse legend

analyse legend

Before using the analyse legend button, legend select should be conducted in advance. Symbols and corresponding annotations are extracted in the selected legend region. It is assumed that symbols are placed in front of annotations and each line contains only one record of symbol and annotation. Therefore, for those legend with multiple columns, each column can be selected as a single legend region. After completing the analysis of legend, clicking 'item' in the information upper panel, the symbols and annotations will be highlighted by a box in the main window and annotation shown in the information lower panel (Figure 13). The symbols are stored as templates and passed to symbol recognition or color recognition function.
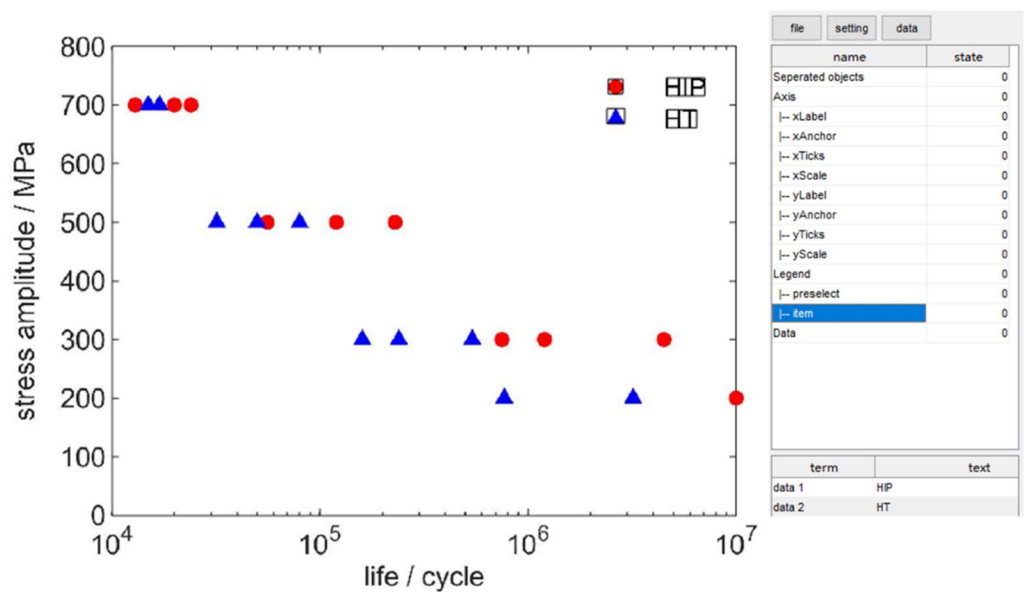

Figure 13. The result of analyse legend.

symbol recognition

symbol recognition

Using the templates of symbols from the analyse legend function, data points are extracted from the image using both shape and color. Click the data button at the top of

the information panel to refresh the page, and then click 'series-*n*' in the information upper panel. Data points will be highlighted by yellow circles in the main window (Figure 14). See **3.2 Quickstart** for the interaction with data points.

color recognition

color recognition

Using the templates of symbols from analyse legend, pixels are extracted from the image using color. The function is used where symbols are heavily overlapped and locating the symbol position is difficult, such as the fatigue crack growth rate (FCGR) figures. Clicking 'series-*n*' in the information upper panel, pixels will be highlighted by yellow in the main window (Figure 15).

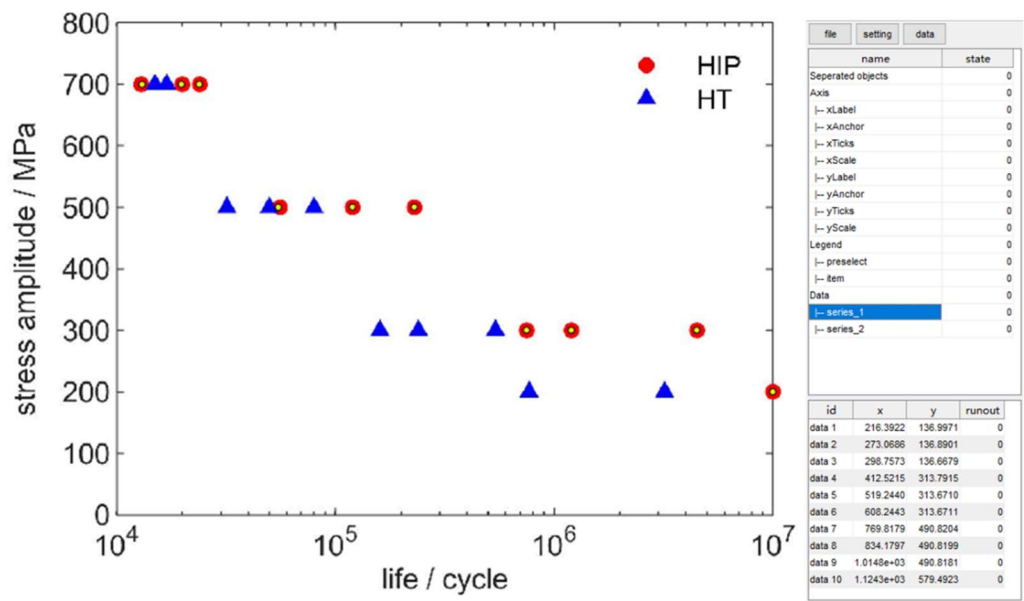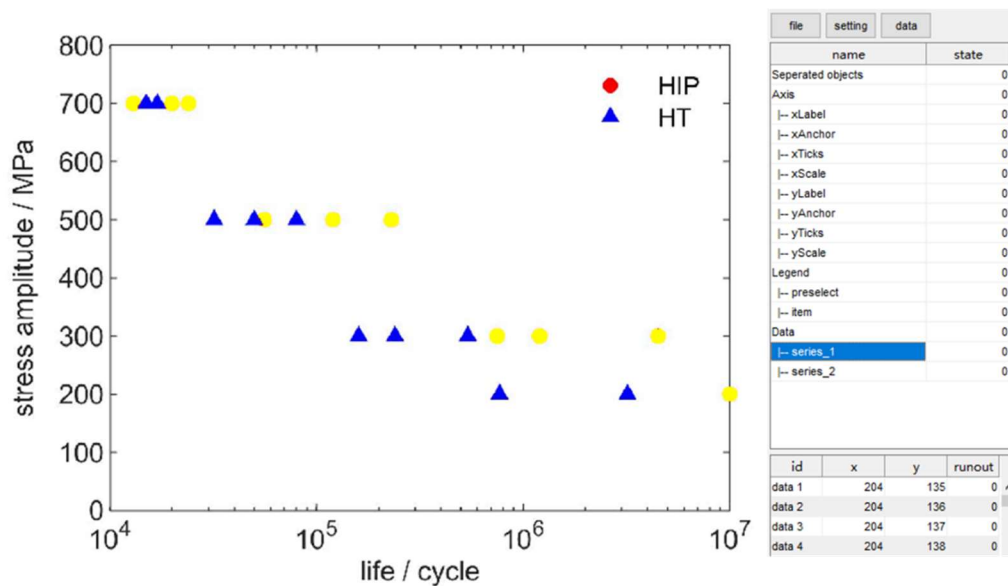Figure 14. The result of symbol recognition.

Figure 15. The result of color recognition.

load project

Load a project file (MAT file) whose format is consistent to IMEX.

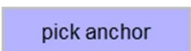## save project

Save the data of current project to a MAT file.

import image

Import images under a folder.

pick anchor

Use to correct the position of axis anchor or place axis anchor manually. After clicking, the pick anchor button turns blue [pick anchor], and the main window enter the mode of anchor selection. Select 'xAnchor' or 'yAnchor' in the information upper

panel to display axis anchors. A specific anchor can be selected directly from the main window or clicking the corresponding table cells that displaying the value of 'xpos' and 'ypos' in the information lower panel. When an anchor is selected, clicking the left button in the main window will move the anchor to the position of the cursor. 'w', 's', 'a' and 'd' keys can be used to adjust the position of anchors when the key on/off button is activated. The axis tick value of the anchor can be modified in the table cells of 'xdata' and 'ydata' in the information lower panel.

del legend box

delete legend box

After legend selection, clicking the 'preselect' in the information upper panel will display all the selected regions by boxes. Select a region by left clicking it and delete the region by del legend box button.

add data series

add data series

After clicking the add data series button, it will turn blue add data series and the main window enters the mode of symbol selection. Box select a symbol for the data series in the main window and click the add data series button again to complete the procedure. A new data series will display in the information upper panel and selecting it triggers the mode of manual data points addition or correction, whose operation is the same as **3.2 Quickstart**.
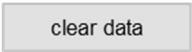
delete data series

delete data series

Select a data series in the information upper panel and use the delete data series button to delete it.

del data

delete data

Select a data point or box select a group of data points and use the del data button to delete them. When the key on/off button is activated, the 'Delete' or 'Del.' keys have the same function.
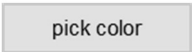
clear data

clear data

Data for all data series are deleted.

autorun

autorun

After selecting the legend region for images and specifying the range of autorun in the autorun text box (format: '*n* - *m*'), the autorun button executes the functions of separate object, detect axis, analyse legend and symbol recognition sequentially for images in the range.

pick color

pick color

Click the pick color button will turn it to blue | pick color | and trigger the mode of picking a color in the main window. Clicking the left mouse button in the main window and the color of the pixel under the cursor is displayed in the magnified panel. Clicking the pick color button again will turn off the color picking mode and the button turn back to gray.

Figure 16. The result of pick color.

select color

select color

When a color is picked (the pick color button should be in blue), those pixels with similar color can be selected by the select color button. The selected pixels are highlighted in yellow. The threshold of RGB to define similar pixels can be specified in the color tolerance text box.

add color

add color

Pixels with different colors can be selected by sequentially pick color and add color. The difference between select color and add color is that select color clears previous selected pixels whereas add color preserves them.
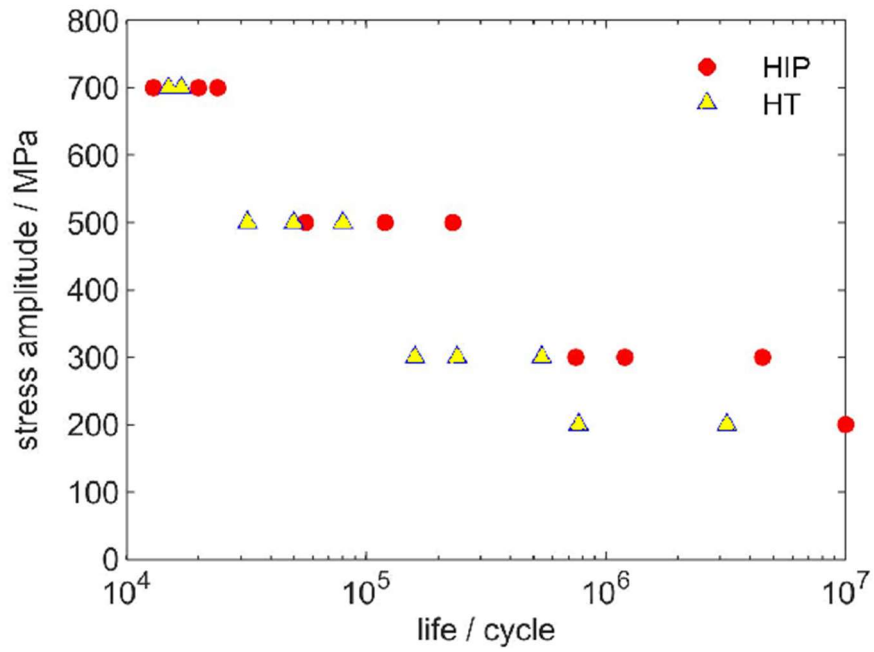
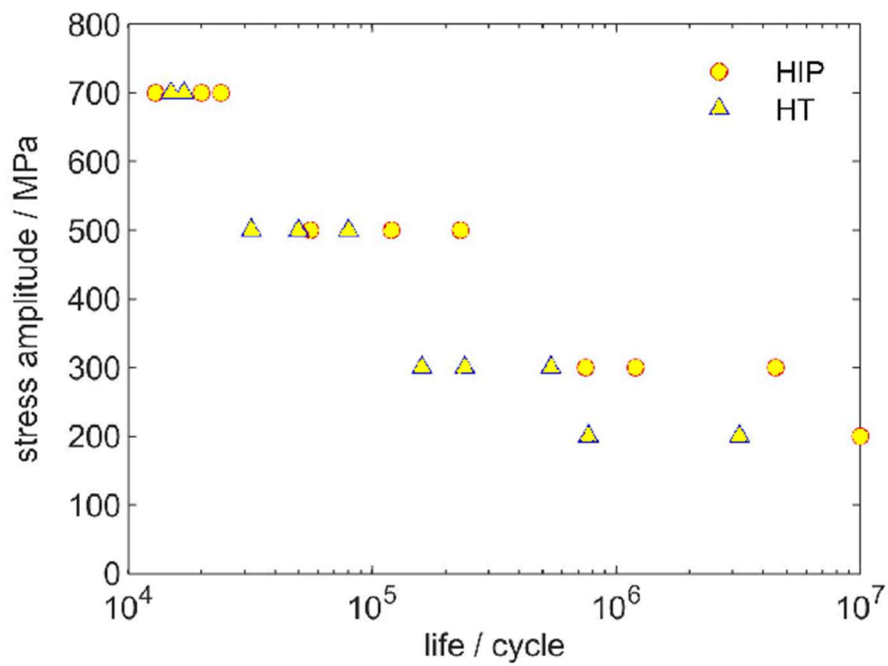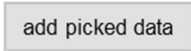Figure 17. The result of select color. Blue pixels are selected.

Figure 18. The result of add color. Red pixels are added to the selection in addition to blue pixels.

reset color

reset color

The pixels highlighted in yellow are restored to their original color by the reset

color button.

add picked data

add picked data

When pixels are selected by pick color, select color and add color, choosing a data series in the information upper panel and using the add picked data button can add pixels data to the data series.

replace picked data
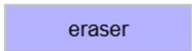
replace picked data

When pixels are selected by pick color, select color and add color, choosing a data series in the information upper panel and using the replace picked data button can replace the original data by the newly picked pixel data.

eraser

eraser

Clicking the eraser button turns it blue eraser and triggers the eraser mode in the main window. The eraser is in the shape of a rhombus and clicking the left mouse button erases the data within the rhombus. The size of eraser can be specified in the eraser size text box, in the unit of pixel. Clicking the eraser button again turns off the eraser mode and the button turns gray.

Figure 19. The result of eraser.

shape / color pick

shape / color pick

The button is used to switch the type of data extraction, using shape information and locating the position of symbols, or simply extracting pixels by color. It affects the how the data is displayed when selecting data series in the information upper panel. For those data recognized with shape information, each data point is represented by a yellow circle marker. For those pixels recognized only by color, the color of the pixels is directly turned yellow in the main window.

key on/off

key on / off

When the key on / off button is activated ![key on/off], the keyboard takes effect. Clicking the button again will turn off the function of the keyboard.