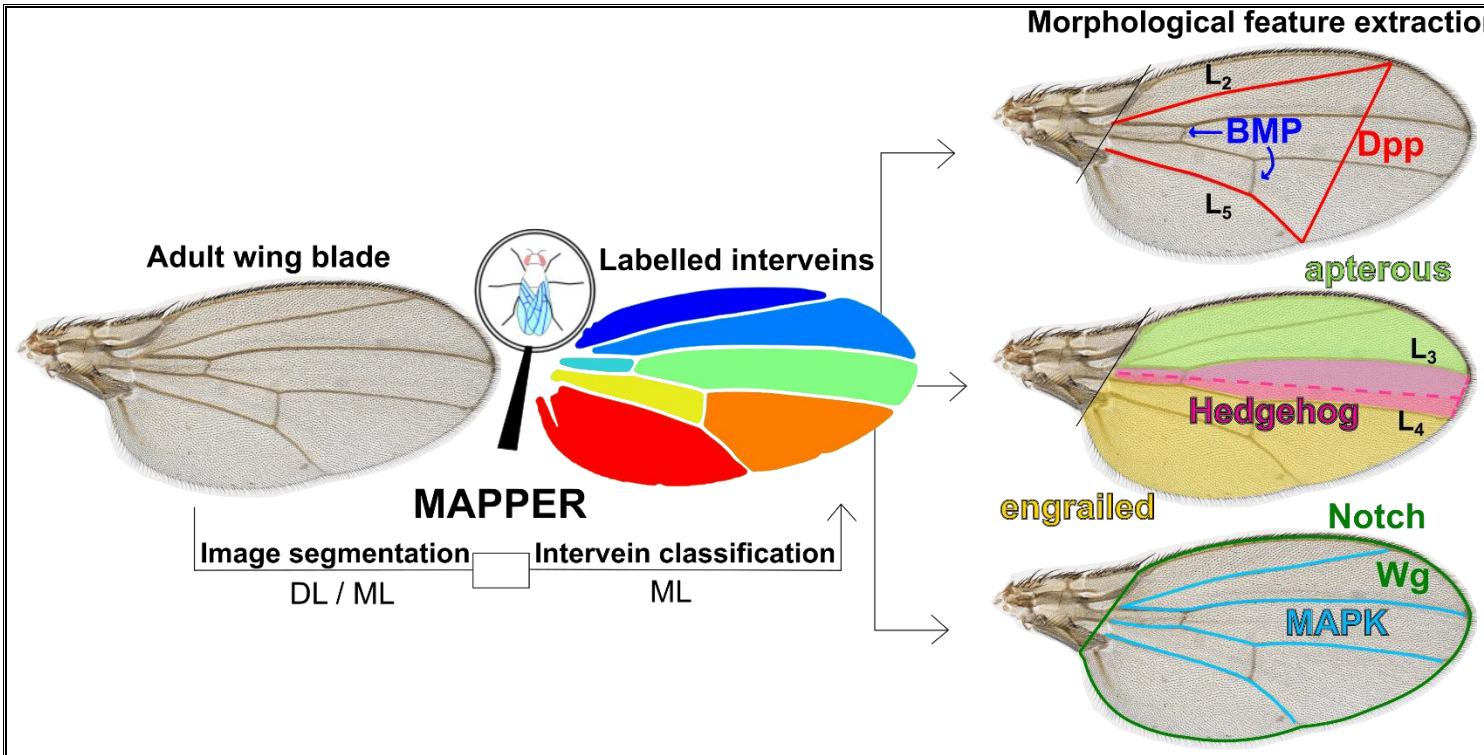


Multicellular Analysis Processing Platform for Engineering Research

User Manual and Guide

This document serves as a how-to guide for the MATLAB®-powered tool developed for the manuscript entitled: "MAPPER: A high-resolution image analysis pipeline unmasks differential regulation of *Drosophila* wing features." For more information, please see the associated publication and corresponding GitHub page.

Nilay Kumar[1], Francisco Huizar[1], Trent Robinett[1], Keity J. Farfán-Pira[2], Dharsan Soundarrajan[1], Maria Unger[1], Pavel Brodskiy[1], Marcos Nahmad[2], Jeremiah J. Zartman[1*]

[1]Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556, USA. [2]Department of Physiology, Biophysics, and Neurosciences, Center for Research and Advanced Studies of the National Polytechnical Institute (Cinvestav), Mexico City, Mexico 07360. *Correspondence: jzartman@nd.edu, (+1) 574-631-0455.

CHEMICAL AND BIOMOLECULAR ENGINEERING

MathWorks®

Table of Contents

Preface 4
Why was MAPPER created? 4
Who can use MAPPER? 4
How can I contribute to future iterations of MAPPER? 4
Troubleshooting inquiries 4
Step One: Install MATLAB and the necessary toolboxes 5
Downloading MATLAB® 5
Installing MATLAB® 5
Selecting and installing necessary toolboxes 6
(Optional) If MATLAB® version 2019b or later is already installed 6
Step Two: Download the latest version of MAPPER 7
Step Three: Place your data into MAPPER 8
Understanding MAPPER's folder structure 8
batch_output_files 8
demo_batchProcessing 8
demo_hingeRemoval_Uncropped 8
demo_individualProcessing 8
dependencies 8
GUI_Logos 8
individual_output_files 9
user_Data 9
Placing your data properly into MAPPER 9
Step Four: Wing hinge removal of raw images 10
Step Five: Image processing with ILASTIK 13
Downloading and installing ILASTIK pixel classifier 13
Selecting a classification module 14
ILASTIK module 1 14
ILASTIK module 2 14
ILASTIK module 3 14
ILASTIK module 4 15
ILASTIK module 5 15
ILASTIK module 6 15
Processing images in ILASTIK 15

Step Six: Running MAPPER 21
Understanding the input parameters 21
Size of wing filter 21
Size of intervein filter 21
Threshold distance trichome 21
Number of harmonics 21
Conversion factor 21
Labeled intervein Image output name 22
Output spreadsheet save name 22
Individual wing processing 22
Loaded raw image 22
Labeled output image 22
Biological axes display 22
User operation 22
Step Seven: Accessing and understanding your output data 22
(Optional) Train your own ILASTIK module 22

Preface

Why was MAPPER created?

The Multicellular Analysis Processing Platform for Engineering Research (MAPPER) tool was created to enable high-dimensional data analysis of wing phenotypes using *Drosophila* as a model organism. The unique capability of MAPPER to compartmentalize the *Drosophila* wing and provide diverse data for each compartment such as area, trichome counts, and landmark region measurements can allow researchers that utilize *Drosophila* to discover subtle phenotypes that are tedious to obtain by hand-measurements alone.

Who can use MAPPER?

MAPPER was created with versatility in mind for all researchers in the *Drosophila* community. Even those with little background or experience in MATLAB or coding can implement MAPPER with the streamlined Graphical User Interface (GUI) and easy-to-access data output. Thus, the *Drosophila* community as a whole that comprises of biologists, geneticists, engineers, and bioinformaticians to rapidly obtain high-dimensional fingerprints of their *Drosophila* wing images. Modifiable parameter inputs on the MAPPER GUI enable processing of images from different imaging microscopes to further lessen restrictions on implementation of MAPPER for different *Drosophila* research groups. MAPPER's distinct image processing capabilities additionally enable any image orientation of the wing to be processed.

The latest version of MAPPER is the initial release V1.0.0 that is usable for Windows/PC, MacOS, and Linux/Ubuntu users of MATLAB. The tool was built using MATLAB 2019b. This requires users to have at least MATLAB version 2019b installed or later to run the tool.

How can I contribute to future iterations of MAPPER?

Because MAPPER utilizes deep-learning and machine-learning for image segmentation of the *Drosophila* wings, we are in constant need of new images or ILASTIK pixel-classification modules (see step four below) that can help us train our models for future iterations of MAPPER. You can view the types of wings we have already processed on our GitHub page and their corresponding ILASTIK modules. If you have a new ILASTIK module trained on a unique set of data, we would be more than happy to feature it on our GitHub page for others to have open-access to! We will highlight your institution, department, and research group alongside your ILASTIK module to ensure proper credit is given.

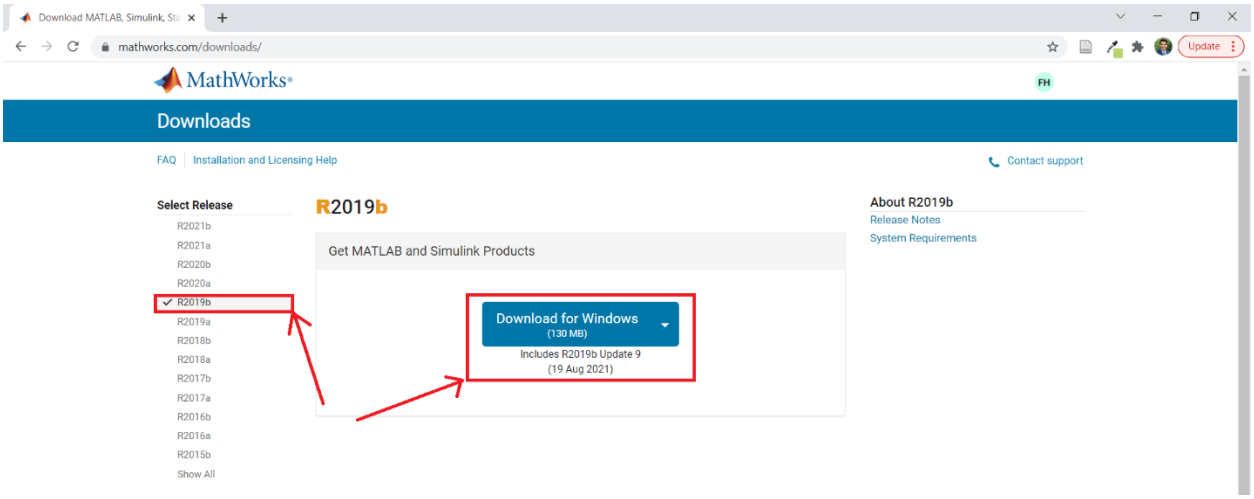
Troubleshooting inquiries

If you encounter any issues while trying to implement MAPPER, we are open to helping you troubleshoot the issue and implement fixes to ensure MAPPER's usability. All troubleshooting inquiries should be filled out using this Google Form. We will do our best to actively monitor incoming inquiries, however, if it takes longer than one week for us to respond, please reach out to the lab PI (Dr. Jeremiah Zartman) at jzartman@nd.edu. To ensure we can best address your issue, please explain in detail the issue you encountered, what steps you take that reproduce the error, or any associated MATLAB error messages that appear.

Step One: Install MATLAB and the necessary toolboxes

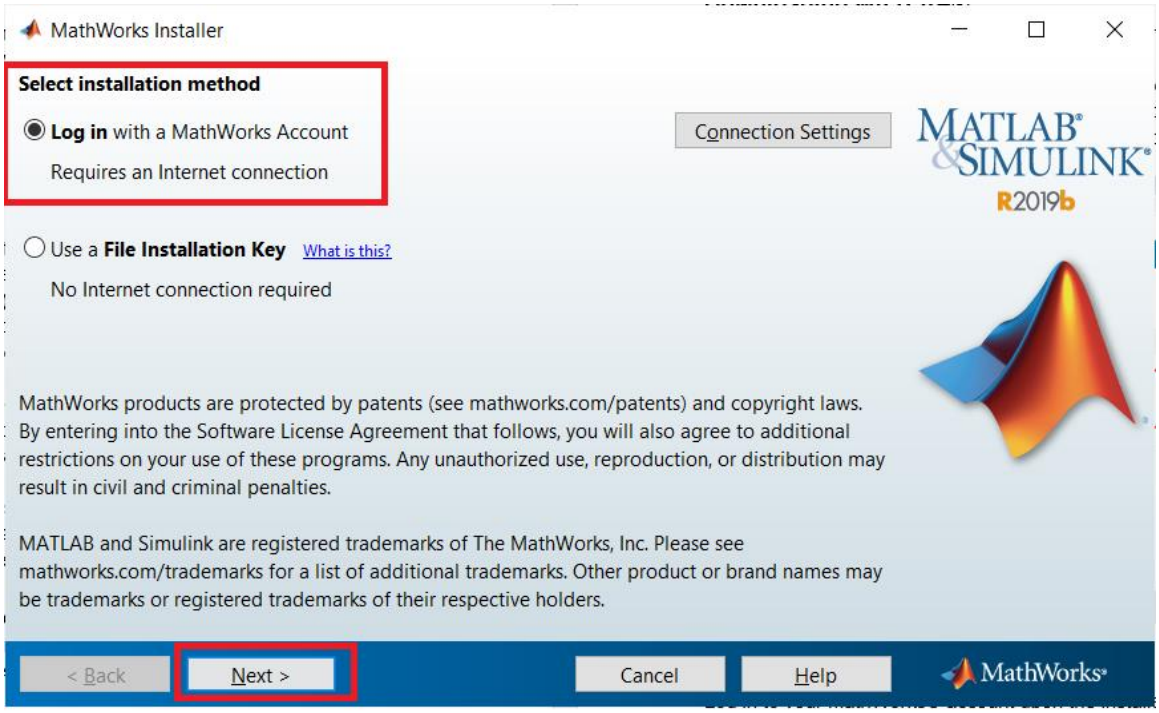
Downloading MATLAB®

In order to run MAPPER, you will first need to download and install MATLAB® from the official MathWorks website. To ensure proper functionality of MAPPER, we recommend installing MATLAB® version 2019b due to MAPPER being created with version 2019b. Later releases of MAPPER are acceptable but may encounter version-specific errors. Download the version of MATLAB® suitable for your operating system (Windows, MacOS, or Linux) and run the installer.



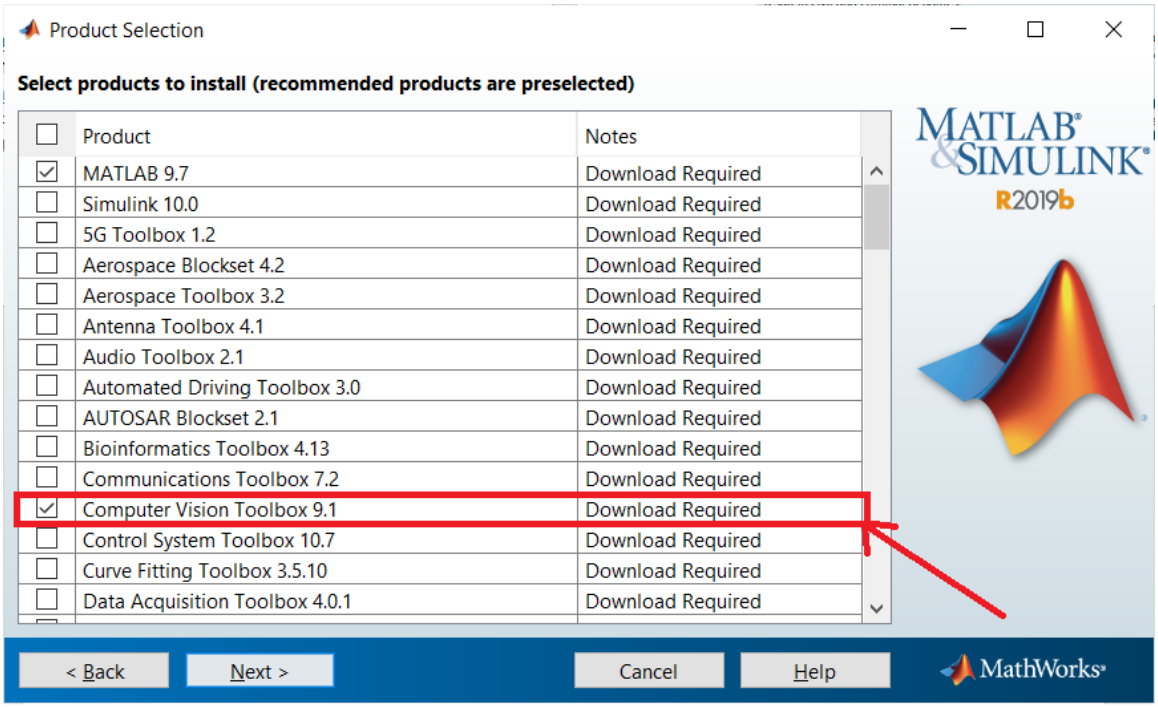
Installing MATLAB®

Log in to your MathWorks® account upon the installer's prompt and follow installation instructions.



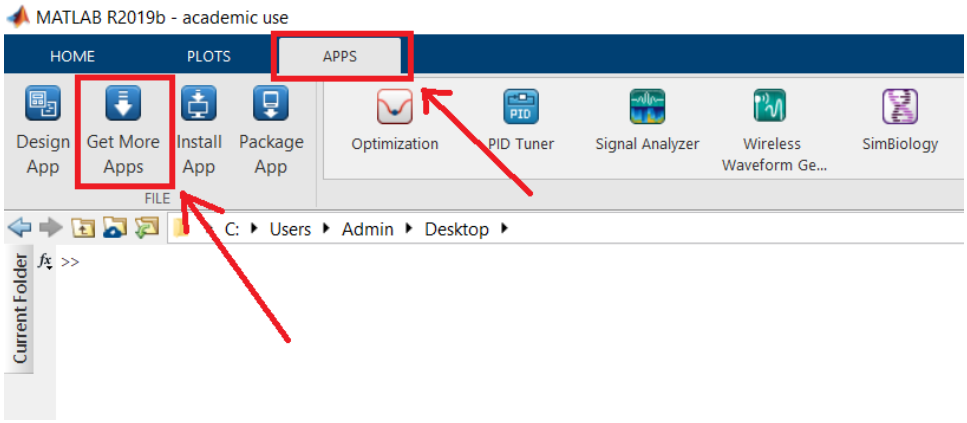
Selecting and installing necessary toolboxes

In order to run MAPPER, you will need to have installed the following three MATLAB® toolboxes: 1) Computer Vision Toolbox, 2) Image Processing Toolbox, and 3) Statistics and Machine Learning Toolbox. During the installation process of MATLAB®, you will be prompted for installation of additional product selection. To install the toolboxes, please ensure you fill out the checkbox next to the required toolbox name for all three toolboxes and continue through the installation instructions.



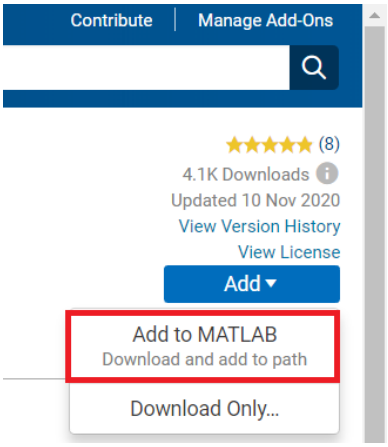
(Optional) If MATLAB® version 2019b or later is already installed

If MATLAB® is already installed, you can easily install the required toolboxes using the GUI and toolbars available in MATLAB®. Open the MATLAB® software application and navigate to the “APPS” section of the toolbar and select “Get More App” in the upper left-hand corner.



In the new “Add-On Explorer” enter the name of the required toolboxes into the upper right-hand search box and navigate to the toolbox.

After clicking on the correct toolbox, in the upper right-hand corner select “Add” > “Add to MATLAB, download and add to path.” Follow and complete the installation instructions that proceed.



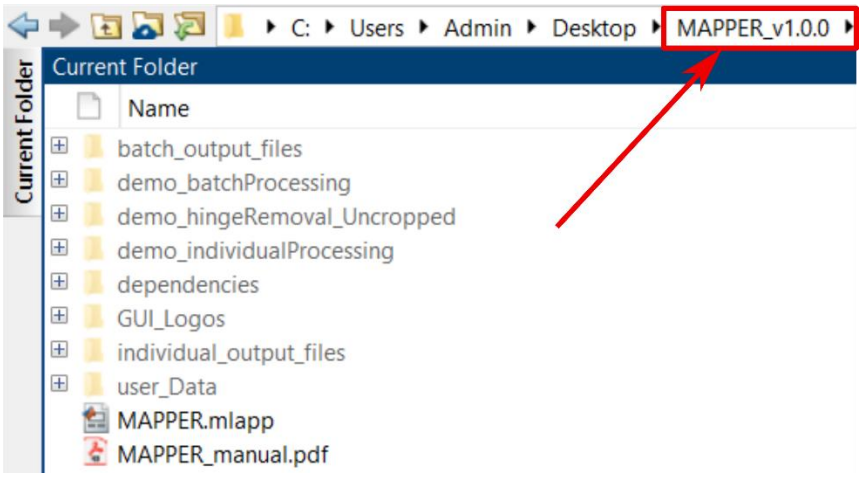
Complete this for all three toolboxes, then close and restart MATLAB.

Step Two: Download the latest version of MAPPER

You can download the latest version of MAPPER here. Once you have downloaded the .ZIP folder, extract the folder to an easily accessible location.

Step Three: Place your data into MAPPER

Understanding MAPPER's folder structure



MAPPER's folder and sub-folder structure was designed to contain only the application (MAPPER.mlapp) and user manual (MAPPER_manual.pdf) in the outermost folder. **DO NOT** modify contents of the sub-folders (other than the "user_Data" sub-folder) to ensure proper functionality of MAPPER. Each sub-folder contains the following:

batch_output_files

This sub-folder contains MAPPER's output after processing and extracting your data. It will contain ".csv" files and ".mat" files of your data for you to analyze at your discretion. Step seven of this manual will go more into detail of the contents of the files.

demo_batchProcessing

This sub-folder contains images and files that can be used for demonstration purposes of MAPPER's batch processing capabilities. Step six of this manual will go into more detail of how to execute batch processing.

demo_hingeRemoval_Uncropped

This sub-folder contains images that can be used for demonstration purposes of the wing hinge removal step of MAPPER. Step four of this manual will go more into detail of how to implement wing hinge removal of your raw data.

demo_individualProcessing

This sub-folder contains images and files that can be used for demonstration purposes of MAPPER's individual image processing capabilities. Step six of this manual will go into more detail of how to execute individual image processing.

dependencies

This sub-folder contains all MATLAB® functions necessary to run MAPPER. Modifying any of the contents of this sub-folder has an increased likelihood of breaking the functionality of MAPPER. **DO NOT** modify the functions within this sub-folder unless you are familiar with MATLAB® and the inner workings of MAPPER.

GUI_Logos

This sub-folder contains images that are displayed on the MAPPER GUI.

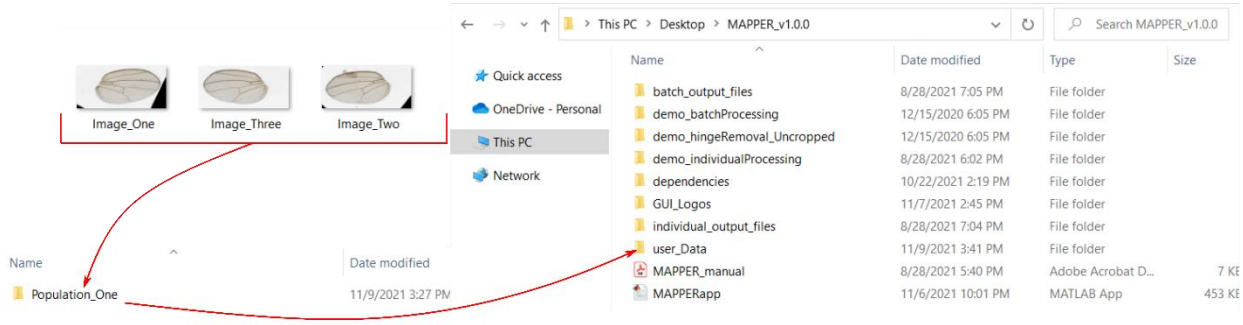
individual_output_files
This sub-folder contains a labeled segmentation of a wing processed by individual processing. The sub-folder additionally contains ".mat" files of the labeled image, the raw image, and the segmentation of the image. These files serve for troubleshooting purposes of individual wings.

user_Data
This sub-folder is empty upon the first download of MAPPER and will be where you place your raw images of *Drosophila* wings.

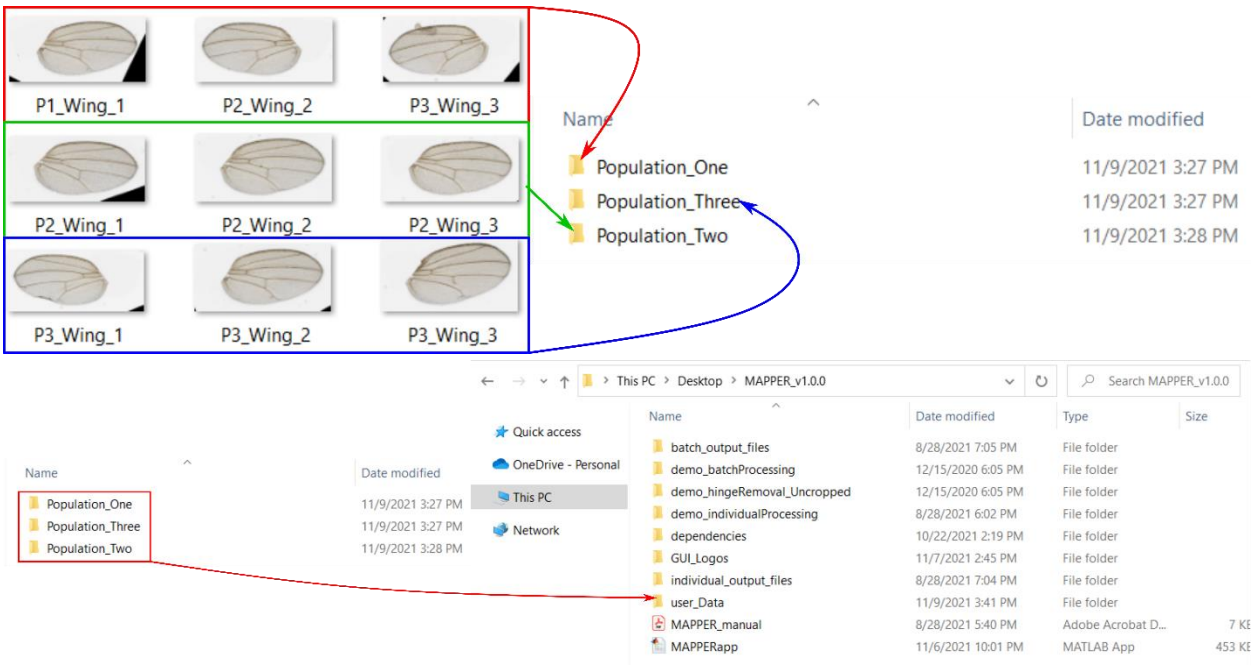
All sub-folders of the main MAPPER folder will be transparent upon starting MATLAB® because they are not yet accessible to the MATLAB® path. This is automatically taken care of upon running the MAPPER application.

Placing your data properly into MAPPER

Raw wing images (.TIF) format can be copied and pasted into the "user_Data" folder for individual wing processing (single images that need to be analyzed). However, for batch processing (analyzing a collection of images), there must be a specific folder structure for MAPPER to be able to process the images correctly. For a single set of images from one population of collected images, you can have one folder containing your raw images. This folder can be moved to the "user_Data" folder or copied/pasted to the "user_Data" folder.

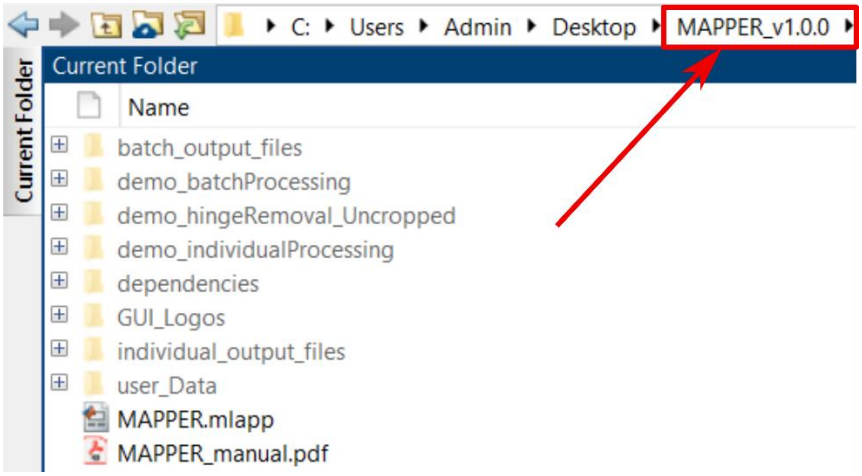


For multiple sets of images (*e.g.,* images from multiple populations of flies), label each population with a different name for you to be able to identify them. Additionally, it may be beneficial for you to label the raw images in a manner that allows them to be connected back to the correct population the images come from. Each folder should contain the raw .TIF images for the population needing to be processed. All folders should then be moved or copied/pasted to the "user_Data" folder.

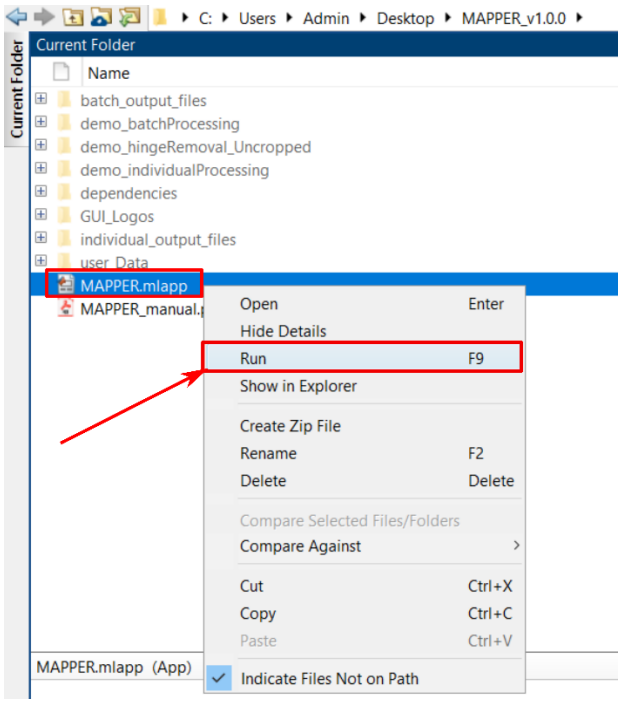


Step Four: Wing hinge removal of raw images

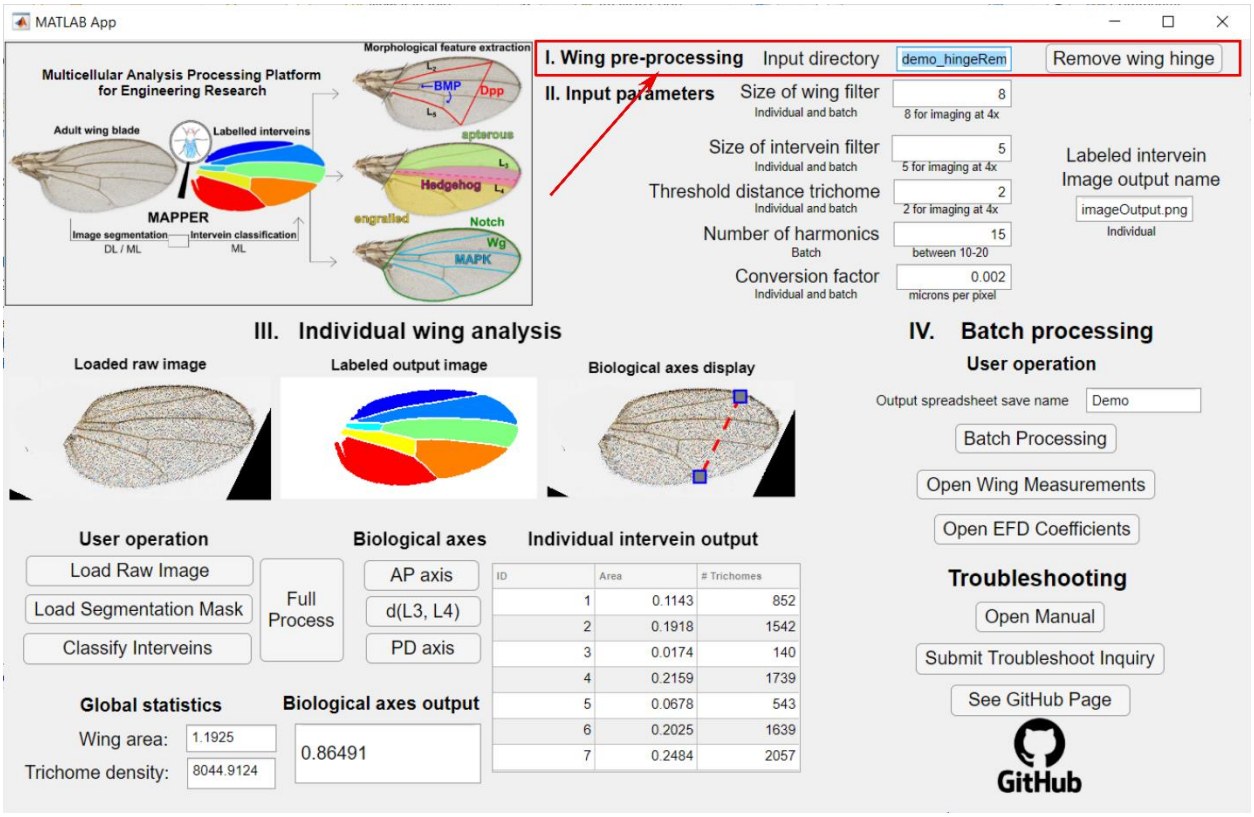
Start the MATLAB® software and navigate to the location where you have placed the main MAPPER folder.



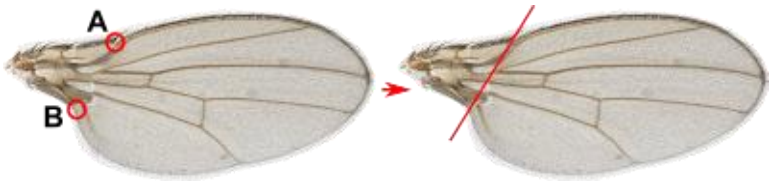
Right-click the “MAPPER.mlapp” and select “Run” to start the application and have the GUI available.



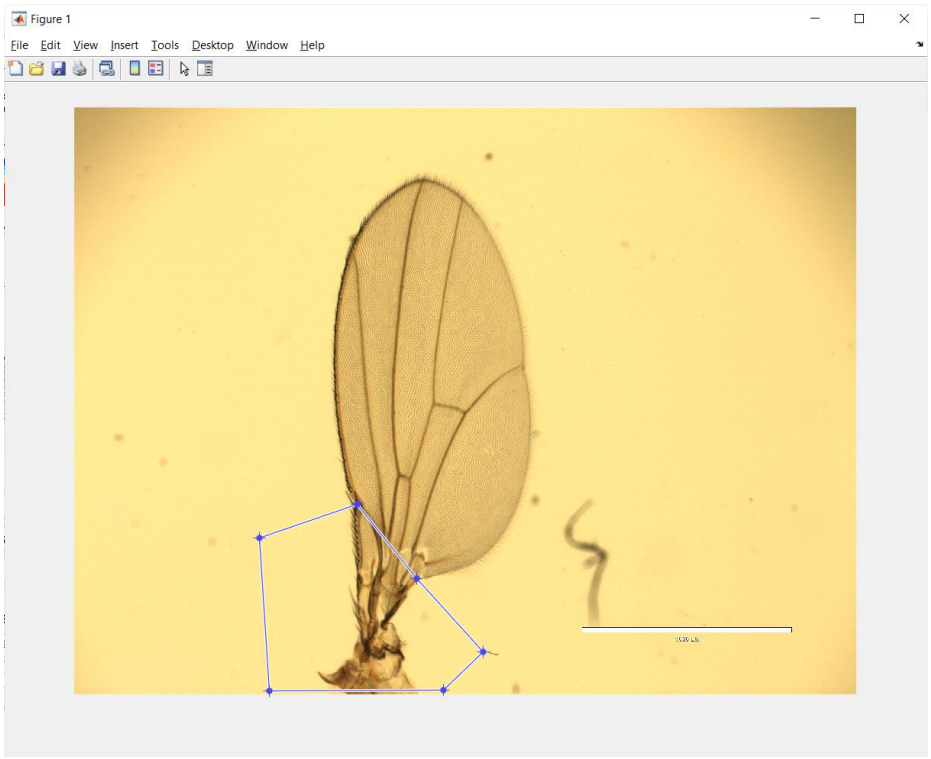
In section **I. Wing pre-processing** of the GUI in the upper-right corner, change the “Input directory” text from “demo_hingeRemoval_Uncropped/” to “user_Data/FOLDER_NAME/”. Where the “FOLDER_NAME” should be the name of the folder you placed your raw wing images in (*e.g.,* “user_Data/Population_One/” if following the manual’s naming scheme). Proceed to click the “Remove wing hinge” button to process your images for the specified folder. Alternatively, you can leave the text as is to run a demonstration of the wing hinge removal step.



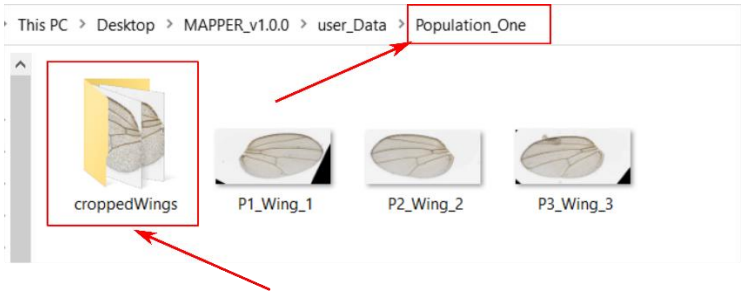
Upon clicking the “Remove wing hinge” button, a MATLAB® Figure with the first image in your folder will appear. You will need to click points that crop out the wing hinge region of the adult *Drosophila* wing. For our images, we have identified a suitable cutout region for the wing hinge by drawing an extended line through the subcostal break (point A in the below image) and through the costal break (point B in the below image).



To do this in MAPPER, each point you click on the Figure will generate a blue point on the figure. You will need to generate a polygon surrounding the wing hinge region using subsequent clicks. We have found the easiest method to do this is to have the first two points generated by clicking be the subcostal and costal breaks. After finishing creating a polygon around the wing hinge, double click the starting point to proceed with processing of the next image. Remove the wing hinge from each image in the folder using the same procedure.



After removing the wing hinge from all of your images, a new folder containing the cropped images named “croppedWings” will appear in the folder you provided in the “Input directory.”



If you are processing a single folder, you can proceed to the next step. If you are processing multiple folders/populations, you will need to perform the wing hinge removal step for each of the folders (*i.e.,* enter “user_Data/Population_Two/” into the “Input directory” section and clicking the “wing hinge removal” button).

Step Five: Image processing with ILASTIK

Downloading and installing ILASTIK pixel classifier

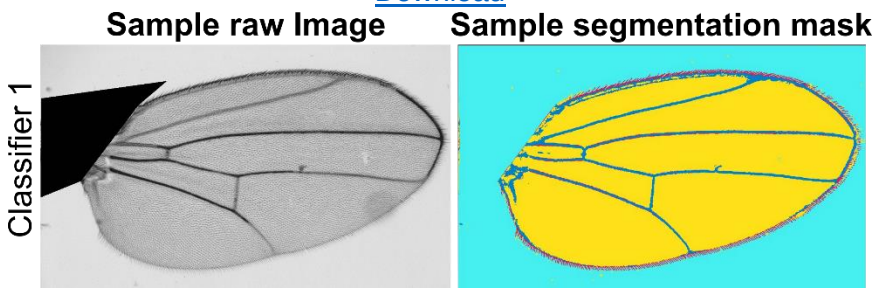
In order to properly segment the input *Drosophila* wing images, you will need to utilize the Pixel Classification workflow using ILASTIK. You must first download and install ILASTIK for your operating system.

Selecting a classification module

We have pre-trained pixel classification modules utilizing ILASTIK for several wing images that have already been processed. As we train new modules and new modules are provided to us from the *Drosophila* community, we will iteratively add to our collection of pre-trained modules available for you to use on our GitHub page. In this manual, we provide links to download six pre-trained ILASTIK modules. Below each module link, you will find a representative image of the *Drosophila* wings that were used to train the module. You should download and use the ILASTIK module that has the closest resemblance in lighting, background, brightness, contrast, and saturation to the images you would like to process. If none of the available ILASTIK modules closely resemble the images you would like to process, there are detailed instructions in a later step on how to train your own ILASTIK module. **NOTE:** The number of channels of your images must match the number of channels in the training data for the ILASTIK module you choose (*i.e.*, RGB channel images must have an ILASTIK module trained on RGB images).

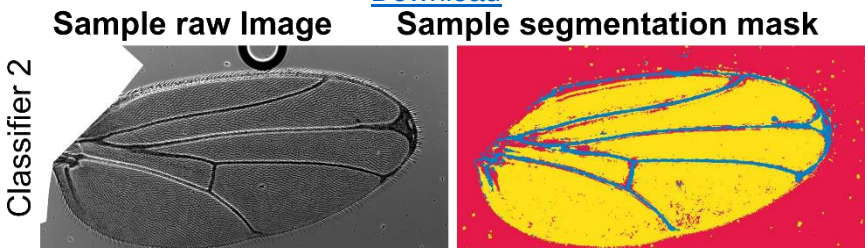
ILASTIK module 1

Download



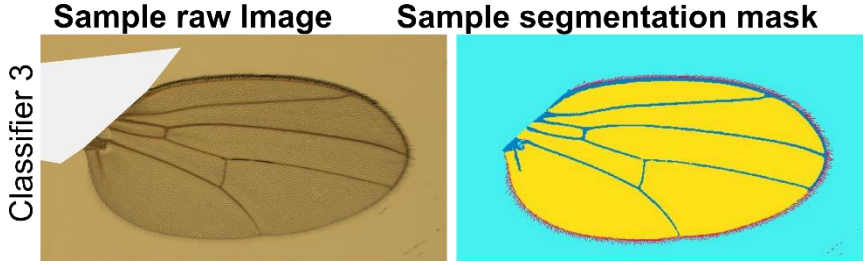
ILASTIK module 2

Download



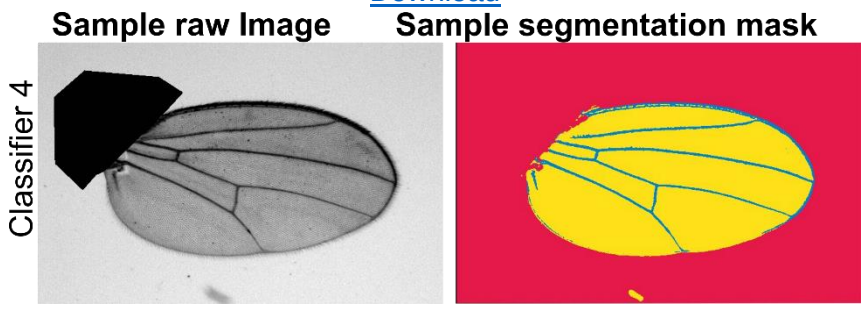
ILASTIK module 3

Download



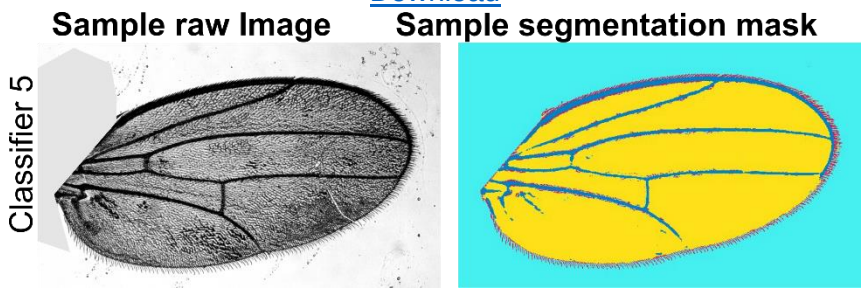
ILASTIK module 4

Download



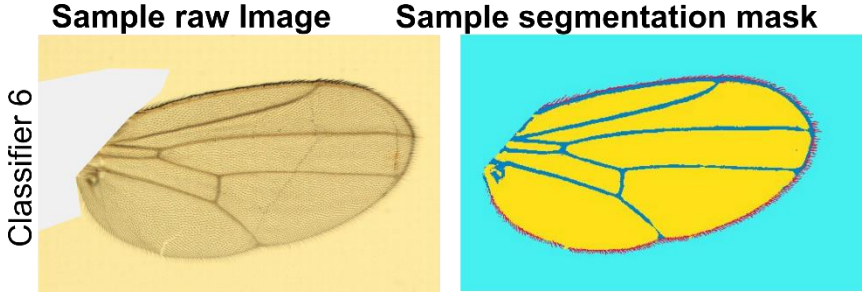
ILASTIK module 5

Download



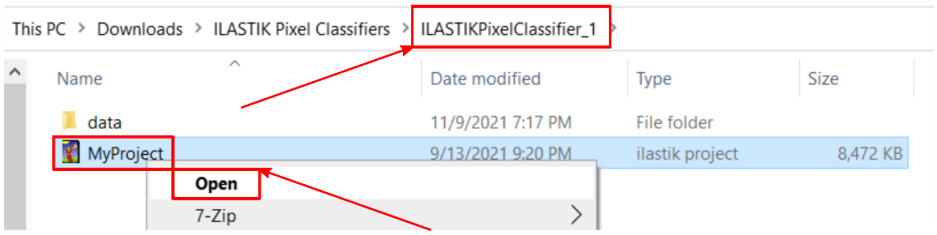
ILASTIK module 6

Download

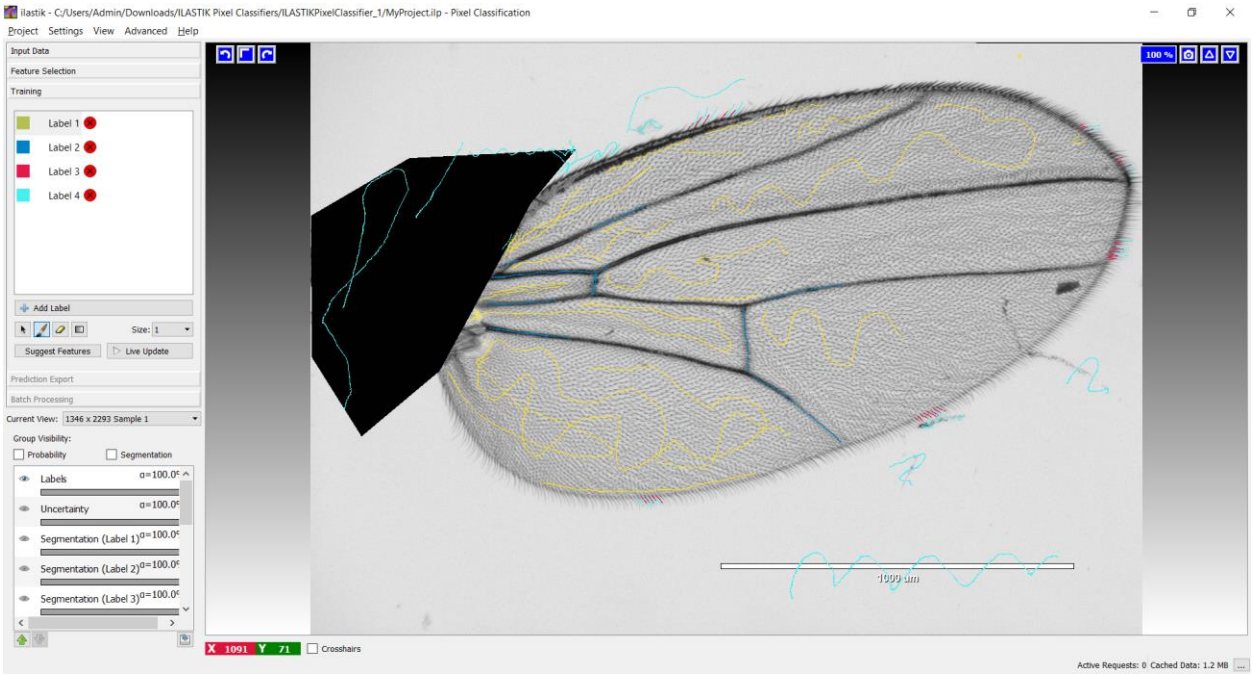


Processing images in ILASTIK

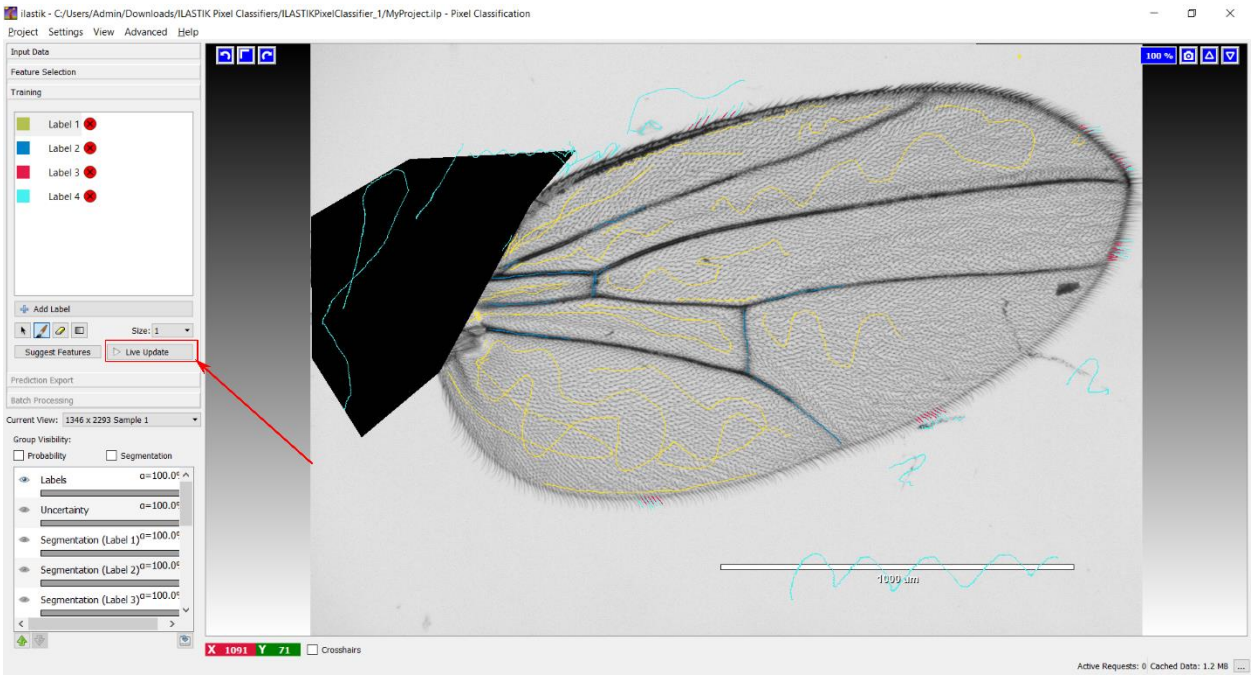
After selecting and downloading your module of choice, extract the .ZIP folder to an easily accessible location. The folder will contain 1) *Drosophila* wing images (with the wing hinge removed) that were used as training data to build the ILASTIK classifier and 2) an ".ILP" file that serves as the ILASTIK pixel classifier that has been pre-trained. Right-click the ILASTIK project file and select "Open" to start up ILASTIK. **NOTE:** ILASTIK start-up time may take several minutes.



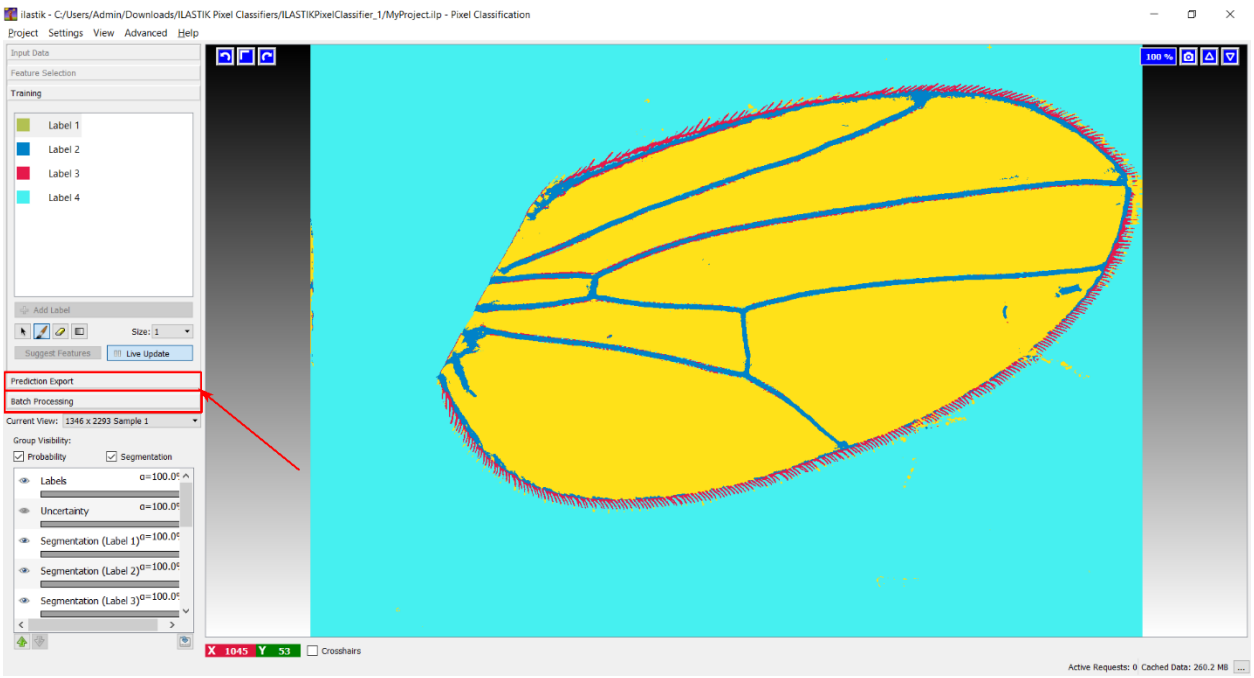
You will know ILASTIK start-up has completed when you can see a sample wing image marked with four different colors corresponding to different components of the image. The meanings of these colors are expanded upon in a later section for training your own ILASTIK module.



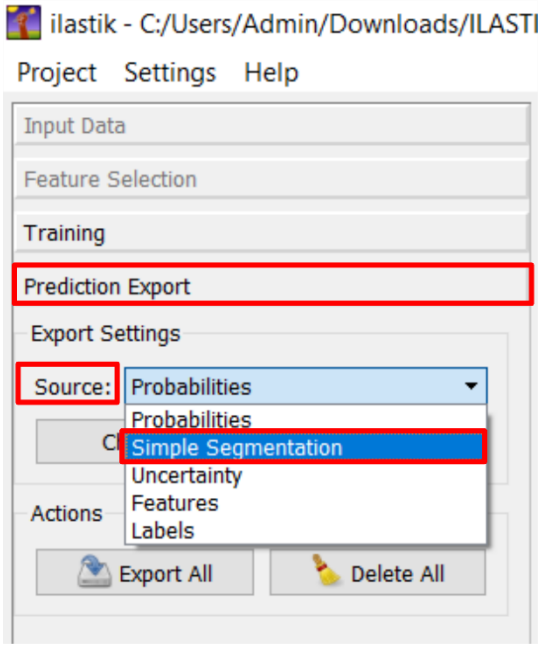
Proceed to click the “Live Update” button to view a segmented wing using the pre-trained classifier.



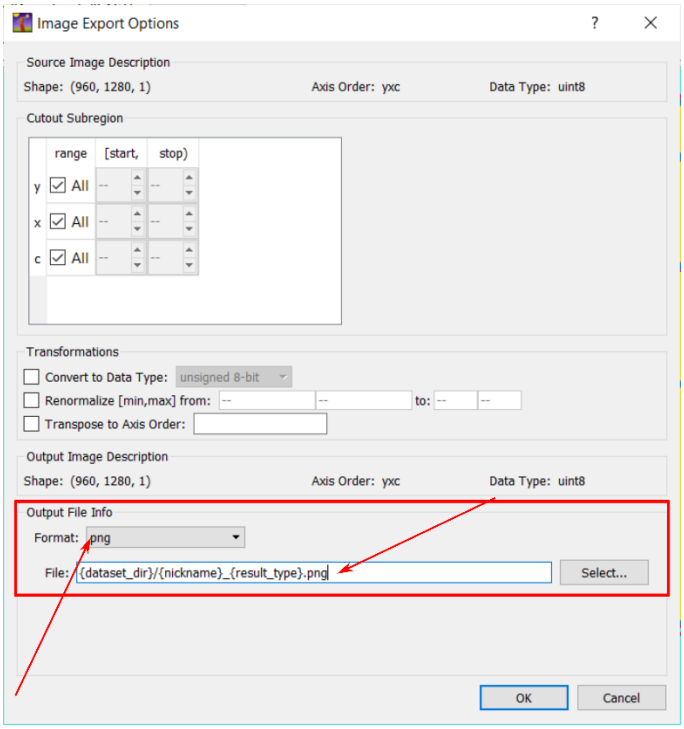
A progression bar can be found in the bottom-left corner of the page. Once the progression bar has reached 100%, ILASTIK will update the view of the image to be a color-coordinated segmentation mask of the pre-trained wing image. The “Batch Processing” and “Prediction Export” toolbars will also be available to click.



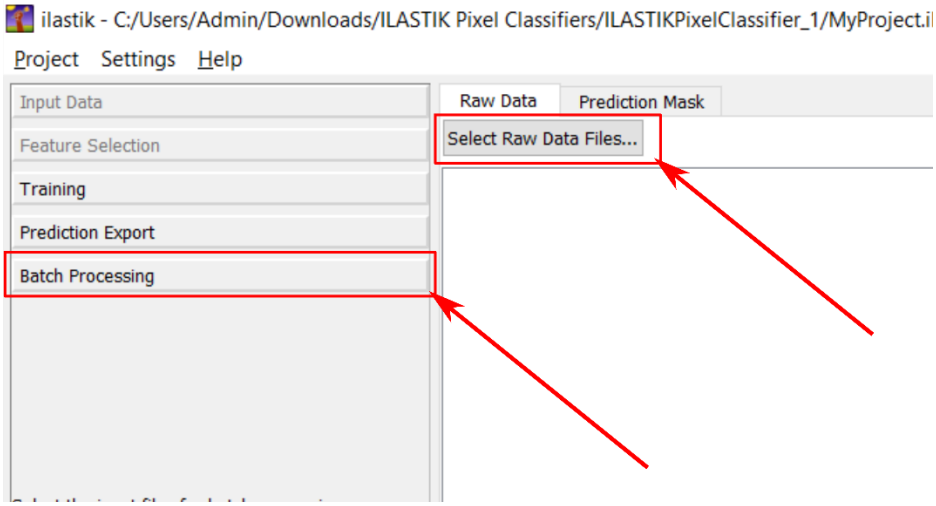
Proceed to click the “Prediction Export” toolbar to bring up the export settings of the output files from ILASTIK. In the “Source” dropdown menu, select “Simple Segmentation.”



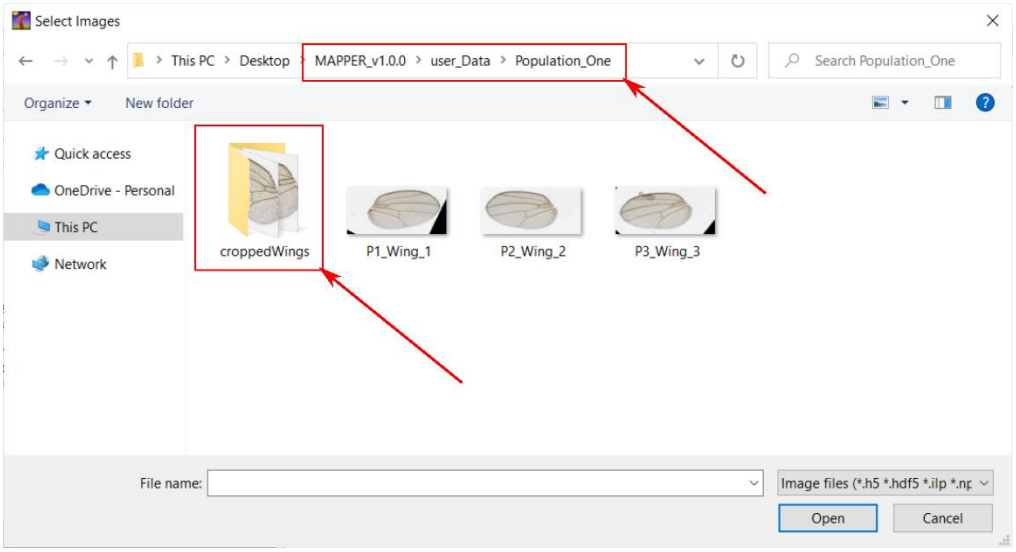
Below the dropdown box, click the “Choose Export Image Settings…” button. In the “Image Export Options” window, ensure that the “Output File Info” is formatted to be a “.png” file, with the file name as “{dataset_dir}/{nickname}_{result_type}.png”. Once this is done, close the export options by clicking “OK”. This will ensure that your output files are named and formatted correctly to be able to be processed by MAPPER.



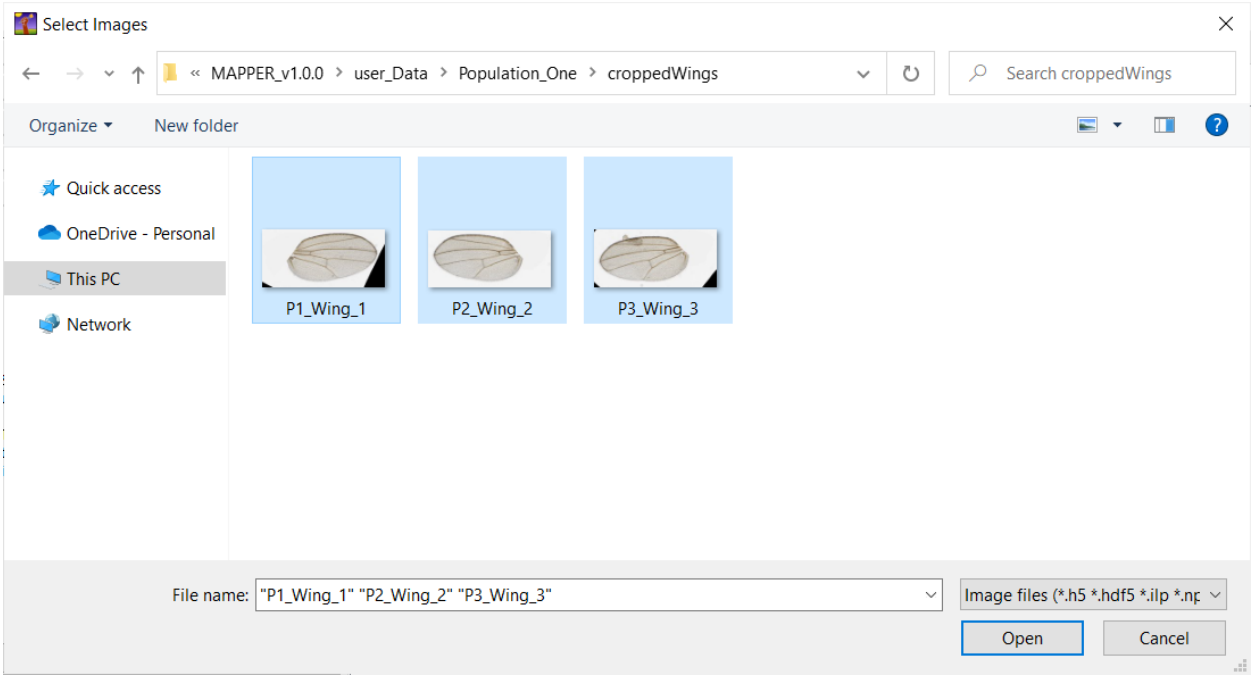
Next, you will need to click the “Batch Processing” toolbar on the left side of the available toolbars. Upon clicking the toolbar, you will need to click the “Select Raw Data Files…” button to select the images you would like to process with ILASTIK.



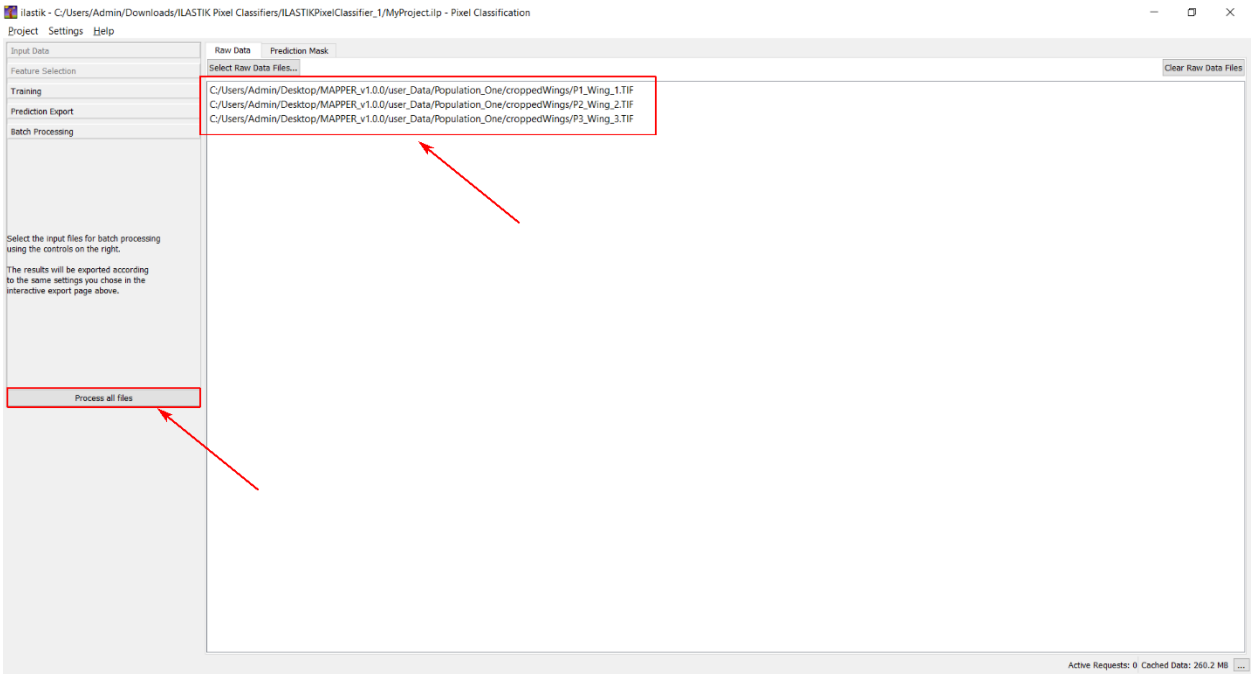
Upon clicking “Select Raw Data Files…”, navigate to the location of your “user_Data” folder containing the folders with your raw data (*e.g.,* “user_Data/Population_One/”) in the pop-up explorer box. Then, select the “croppedWings” folder that contains the images you processed with wing hinge removal step.



Within that folder, highlight and select all of the cropped wings and click “Open.”

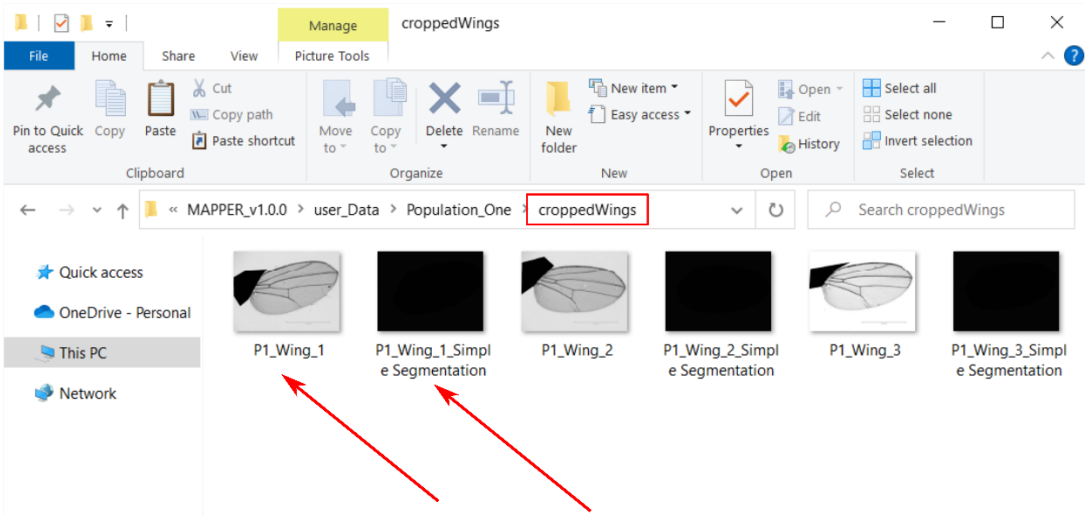


The file location and name of the images will then appear in the ILASTIK raw data files section. Once you confirm you have selected the images you would like to process, click the “Process all files” button.



After ILASTIK has processed the images, you should be able to observe new “.png” files in the “croppedWings” folder you selected your images from. These files will have the naming convention “YOUR_IMAGE_**Simple Segmentation.png**” where “YOUR_IMAGE” indicates the

original name, and "_Simple Segmentation" indicates the type of output from ILASTIK. If the naming scheme of the output files does not contain "_Simple Segmentation" after your original image name, MAPPER **will not** be able to process your images. If this is the case, please review the "Prediction Export" settings detailed earlier.



If you are processing wings from multiple populations, repeat the process of clicking "Select Raw Data Files…", navigating to the "croppedWings" folder of the next population, and selecting the images with the wing hinge removed for each of the populations. The end result will be such that each of the "croppedWings" folder in your population folders will have 1) the cropped wing images and 2) the ILASTIK simple segmentation outputs.

Step Six: Running MAPPER

Understanding the input parameters

In section **II. Input parameters** of the GUI in the upper-right corner, you can modify parameters of MAPPER's image processing pipeline to be custom-fit to your images and output preferences (*i.e.,* images from different magnifications, length-scales, and thresholds). Each of the parameters has the following functionality:

Size of wing filter

CONFIRM WITH NILAY

Size of intervein filter

CONFIRM WITH NILAY

Threshold distance trichome

CONFIRM WITH NILAY

Number of harmonics

CONFIRM WITH NILAY

Conversion factor

CONFIRM WITH NILAY

Labeled intervein Image output name

Changing the text input for this parameter allows you to name the color-coded labeled output image if you are processing a single image. The default is “imageOutput.png” and the labeled output image is saved to the “individual_output_files” subfolder within the main MAPPER folder. Please ensure that the input text has “.png” at the end of your desired name for the file.

Output spreadsheet save name

Changing the text input for this parameter allows you to name the batch processing output “.csv” and “.mat” files. The output files are located in the “batch_output_files” subfolder of the main MAPPER folder. The naming convention of the files will be “TEXT_EFD_Output” for the Elliptic Fourier Descriptor files and “TEXT_Wing_Measurements_Output” for the automated wing measurement data where “TEXT” is the input parameter you specify.

The default input parameters are for *Drosophila* wing images taken at 4X magnification with a length-scale resolution of 0.002 microns per pixel.

Individual wing processing

In section **III. Individual wing analysis** of the GUI in the left-hand side, you will see output images and buttons associated with processing a single image in MAPPER. This is to ensure quality control of your input images if batch processing results in errors. Each of the sections detail the following:

Loaded raw image

Displayed in this figure is the image of the *Drosophila* wing selected that has had the wing hinge removed.

Labeled output image

Displayed in this figure is the color-coded labeled intervein segmentation of your selected wing that has been process by ILASTIK and MAPPER.

Biological axes display

Displayed in this figure is the biological axes measurement of your choice overlayed in a red-dashed line on the cropped wing image.

User operation

These buttons enable the user to use MAPPER to process a single image. This can be done step-wise by clicking “Load Raw Image” > selecting the cropped wing image

Step Seven: Accessing and understanding your output data

(Optional) Train your own ILASTIK module