User guide for:

MASQH

Multi-scale Algorithm for Segmentation-free Quantification of Homogeneity

Developed by Fiona Milano

Members of the Biomaterial and Cartilage Laboratory, a research group of Polytechnique Montréal

Version 1 – Août 2020

Table des matières

General information

MASQH is a MATLAB standalone application autonomously quantifying homogeneity in images. This algorithm is based on a straightforward statistical, multi-scale approach. The key idea behind the proposed method is that it does not require a prior segmentation of the image. The numeric results obtained with the application can be save either in excel or text format. The MASQH algorithm is further explain in the following article:

- F. Milano, A. Chevrier, G. D. Crescenzo, and M. Lavertu, "Robust Segmentation-Free Algorithm for Homogeneity Quantification in Images," *IEEE Transactions on Image Processing*, vol. 30, pp. 5533-5544, 2021, doi: 10.1109/TIP.2021.3086053.

This software is freely available for both MAC and WINDOWS at https://github.com/Biomaterials-and-Cartilage-Laboratory/SAM-Scratch. The application requires the MATLAB component runtime (MCR) version 2019b to run. This version of the runtime is freely available at https://www.mathworks.com/products/compiler/matlab-runtime.html. To install the software, please refer to the "installation" section. The MATLAB code of the algorithm behind SAMScartch is also given in our Github.

The present document describes the algorithm behind MASQH, details the installation procedures for MAC and WINDOWS and explains how to use the interface. If you need any help to use the software or if you have ideas to improve it, feel free to contact Fiona.milano@polymtl.ca.

Legal information

By downloading and/or installing the MASQH software, referred to bellows as the SOFTWARE, you agree to the following license agreement.

- The BCL group of Polytechnique Montréal, referred to bellow as BCL, agrees that the LICENSEE uses the SOFTWARE for *non-profit* and *non-commercial* purposes only. SOFTWARE shall mean all SOFTWARE provided by BCL, including any derivatives thereof. The software is intellectual property of BCL and all rights, title and interest, including copyright, remain with BCL.
- Nothing in this agreement shall be construed as granting any license under any intellectual property right vested into BCL or any right to use the SOFTWARE or any BCL information other than herein expressly specified.
- In any publications resulting from research done under the use of the SOFTWARE, the LICENSEE will properly cite either party's contribution. Any report or publication of results obtained with the SOFTWARE will acknowledge its use by stating that SOFTWARE « was developed by the Biomaterial and Cartilage Laboratory (BCL), at Polytechnique Montréal ». Any published work that utilizes the SOFTWARE shall include the following reference:
    - Ahunon, L., Milano, F., Chevrier, A., & Lavertu, M. (2020). A novel image analysis algorithm reveals that media conditioned with chitosan and platelet-rich plasma biomaterial dose dependently increases fibroblast migration in a scratch assay. *Biomedical Physics & Engineering Express*.
- Proper citation is essential to continued BCL funding for research, as it is the primary way in which we demonstrate the value of our research to the scientific community. Thank you for your attention to these details.
- The BCL does not assume any responsibility for the SOFTWARE and the use of it. BCL will not assume any liability for damages or loss of data occurring through the use of the SOFTWARE. BCL does not guarantee the suitability of the SOFTWARE for any application. The LICENSEE will hold BCL harmless for any claims on damages or loss of data which occur during the LICENSEE's use of the SOFTWARE.
- BCL does not guarantee that the SOFTWARE does not infringe any third party's intellectual property rights (including copyrights, patents, patent applications, publications or any other type of intellectual property right). The LICENSEE has to acquire on his/her own all necessary licenses and permissions for the use of the SOFTWARE if not otherwise agreed in writing.
- In case the SOFTWARE is or will be under the physical control of the LICENSEE before this agreement is signed, BCL gives consent to the use of the SOFTWARE under the condition of the LICENSEE's prior consent to this agreement.
- The use and distribution of the MATLAB Compiler Runtime (MCR) libraries is subject to the MATLAB license agreement (see MCR_license.txt). In particular, the MCR may not be redistributed in any way separate from the SOFTWARE.

Algorithm description

The MASQH algorithm is based on the following mathematical definition for homogeneity for grayscale images: « *a grayscale image is homogeneous when it is made of sub images having identical gray level cumulative distribution function*. [1]»

Two principles arise from this definition: 1) our perception of homogeneity is dependent on scale. Images can seem homogeneous at one scale but not at another. 2) Two sub-images taken from a homogeneous image should "look" the same. In a more rigorous way, this means that the histograms of the sub-image's gray-level should come from the same distribution.

From 1) and 2), we understand that two large sub-images taken from a chosen image could be derived from the same definition, while two small sub-images taken from the same image could be derived from different distributions.

The MASQH algorithm takes advantage of this principle. It works on a single-channel image (chose the one with the highest contrast if your initial image is RGB, the same channel should be used for every image to be compared).

The image is analyzed at *k* different scales. At each scale, *n* squares are randomly placed in the image, where *n* is an even number. At each of the *k* scales, the size of these *n* squares is decreaseed. *x* pixels are then randomly chosen in each square and used to form a gray-levels histogram for each square. The squares are then randomly paired, and their histogram compared through a Kolmogorov-Smirnov (KS) test. The null hypothesis (H_0) is the population from each couple comes from the same statistical distribution (*p*-value threshold of 0.05).

At each scale, MASQH computes the proportion of p-value > 0.05, which corresponds to the proportions of squares couples for which H_0 was not rejected (i.e. the KS test does not reject that those couples originate from the same statistical distribution). Referring to the homogeneity's definition, the more the image is homogeneous at the analyzed scale, the bigger this proportion will be. This proportion will vary at each analyzed scale. The mean of the *k* proportions obtained in the end is taken as the index quantifying the overall image's homogeneity (homogeneity index). When close to 0, this index indicates that the image is not uniform at most scales and is therefore considered not homogeneous. When close to 1, it indicates that the image is uniform at most scales and is therefore considered homogeneous.

MASQH relies on 5 parameters described in Table I (from [1]). One of these parameters, l_{min}, depends on the field from which the image originates and must be set by the user. The 4 other parameters were empirically fixed, and the algorithm was shown to be robust to a change in their values.

For more detailed information on the homogeneity definition, the MASQH algorithm and the robustness of its parameter, please refer to the following article:

- F. Milano, A. Chevrier, G. D. Crescenzo, and M. Lavertu, "Robust Segmentation-Free Algorithm for Homogeneity Quantification in Images," *IEEE Transactions on Image Processing*, vol. 30, pp. 5533-5544, 2021, doi: 10.1109/TIP.2021.3086053.

TABLE I
DESCRIPTION AND VALUE OF THE PARAMETERS. FROM [1]

Parameter	Definition	Value	Description
l_{min}	Minimal analysis length	100 pixels	Linked to the characteristics of the objects of interest. Field dependent. Once chosen, should be constant for every image being compared.
l_{max}	Maximal analysis length	Half the smallest image dimension	Defined as half the smallest image dimension to avoid the prioritization of one image's dimension. Could be smaller. Once chosen, should be constant for every image compared.
k	Number of square sizes used in the analysis (i.e. number of iterations)	100	For a *kn* product sufficiently high, *k* does not influence the index's value. Moreover, it is inversely proportional to the index variance (at the expense of a longer execution time). A *kn* product of 40000 for example leads to a variance in the order of 10^{-5}.
x	Number of points taken per square to form a histogram	5000	Linked to the sensitivity of the KS test. Recommended range of values: between 5000 and 10000. Once chosen, should be constant for every image being compared.
n	Number of squares used at each iteration of the analysis	400	For a *kn* product sufficiently high, *n* does not influence the index's value. Moreover, it is inversely proportional to the index variance and directly proportional to the execution time. A *kn* product of 40000 for example leads to a variance in the order of 10^{-5}.

Installation

Installation on WINDOWS is straightforward but installation on MAC is more complicated. If you have access to a WINDOWS computer, we therefore recommend you use it rather than a MAC.

Mac

Windows

To install the MASQH app on WINDOWS, please follow the instructions below:

1. Unzip the file `MASQH.zip` and extract `MASQH.exe`.
2. Double-click on the file `MASQH.exe` to unpack the program files.
3. Upon unpacking, the file `MRCInstaller.exe` should be automatically executed, starting the installation of Matlab Component Runtime (MCR) 2019b. You can cancel this installation of this runtime is already on your device.
4. Once the runtime is installed, you can start the app by clicking the `MASQH.exe` file. You can manually add shortcut for the app on your desktop and use the file `MASQH_logo.ico` file as an icon for the app if you like.

Interface use

When launching the MASQH application, a window similar to Figure 1 appears. SAMScratch supports all the graphic files formats supported by MATLAB: JPG, PNG, TIFF, TIF, BMP, GIF, HGF, PCX, XWD and others. Once the application is launched, the image(s) can be analyzed through the following steps:

1. First, in the "images directory" panel, select the type of analysis you wish to perform (number 1 in Figure 1). It can either be a single image, a folder containing a batch of images to be analyzed separately or a folder containing several images

originating from a single object that you want to analyze as if they were a single image.

2. Click the “Browse” button to select a directory (an image or a folder depending on the type of analysis).
3. Before launching the analysis, verify (for RGB) images that the selected channel is the desired one (number 2 in Figure 1) and that you entered the smallest scale of analysis desired (number 3 in Figure 1). If you want to exclude some zone of an image from the analysis (blurred or damaged zones for example), just thick the “remove zones from the analysis” for the corresponding images (number 4 in Figure 1). If you did, after launching the analysis, a new window will open with the corresponding image (see Figure 5), and you will be able to circle the zones to exclude.
4. Launch the analysis by clicking on the “Launch analysis” button (number 5 in Figure 1). The program will then analyze the selected image, or all the images contained on the selected folder. During this time, the “Launch analysis” button will change color and indicate the progression of the analysis.
5. After the analyses is completed, the evolution of the proportions of p-values > 0.05 will be shown, as in Figure 6, with the computed index displayed in the “computed index” field (number 6 in Figure 1). You will then be able to save these numeric results in a text file or in an excel file for further use. The graph can be saved as an image (number 7 in Figure 1).
6. When analyzing several images, Click the “<” and “>” buttons to browse through the analyzed images. You can also click “>>” and “<<” to move to the end or the beginning of the dataset (example in Figure 6).
7. If the algorithm fails to give a satisfying result for your application with the pre-set parameter’s value, the simplest way is first to try to modify the chosen channel (if your image is RGB) and to be sure you chose a minimal size making sense for your application (number 2 and 3 in Figure 1).If you are not sure how to set this parameter, please refer to the core article [1]. After checking those two parameters, if no satisfying results are obtained, you can modify each of the 5 parameters through the “More parameters” button (number 8 in Figure 1). This will open a new window that will be discussed in the “parameters” section.
8. When your analysis is completed and your results are saved, you can perform a new analysis by clicking on the “new analysis” (number 9 in Figure 1).

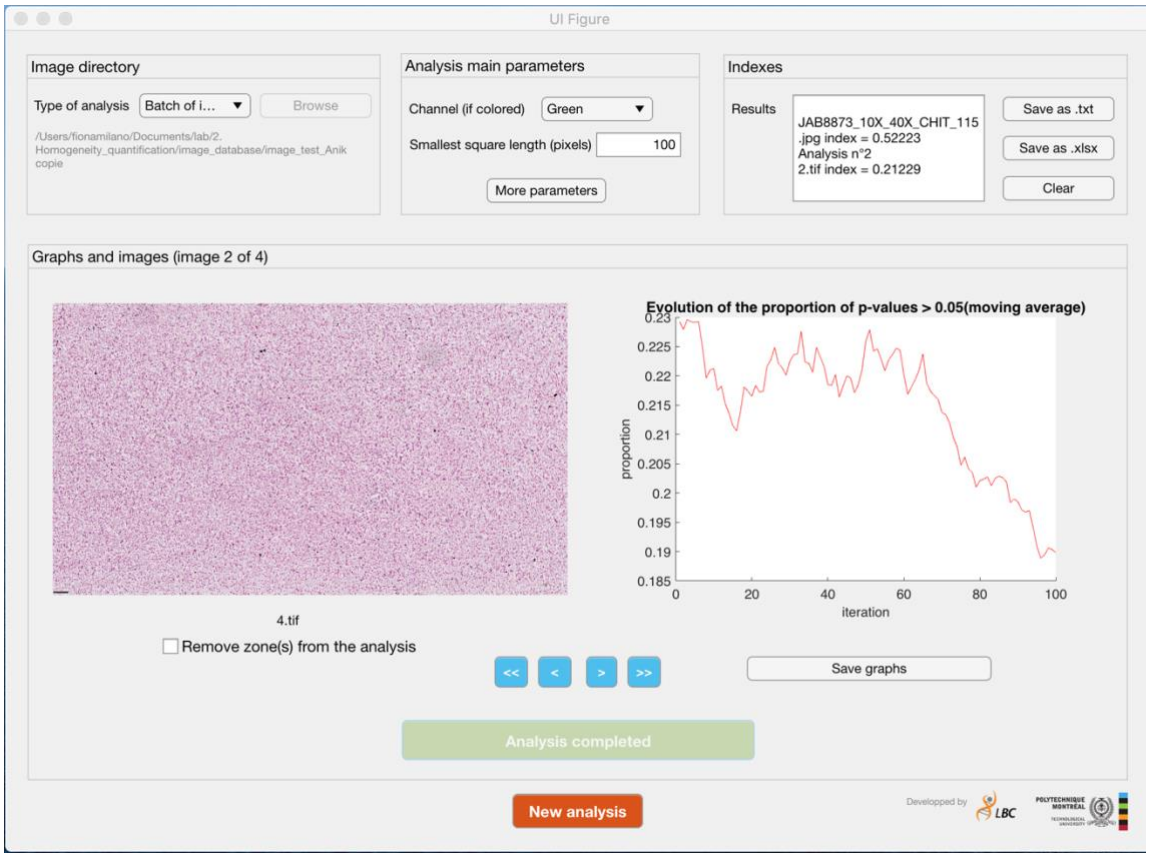

Figure 1. MASQH window after launching the application

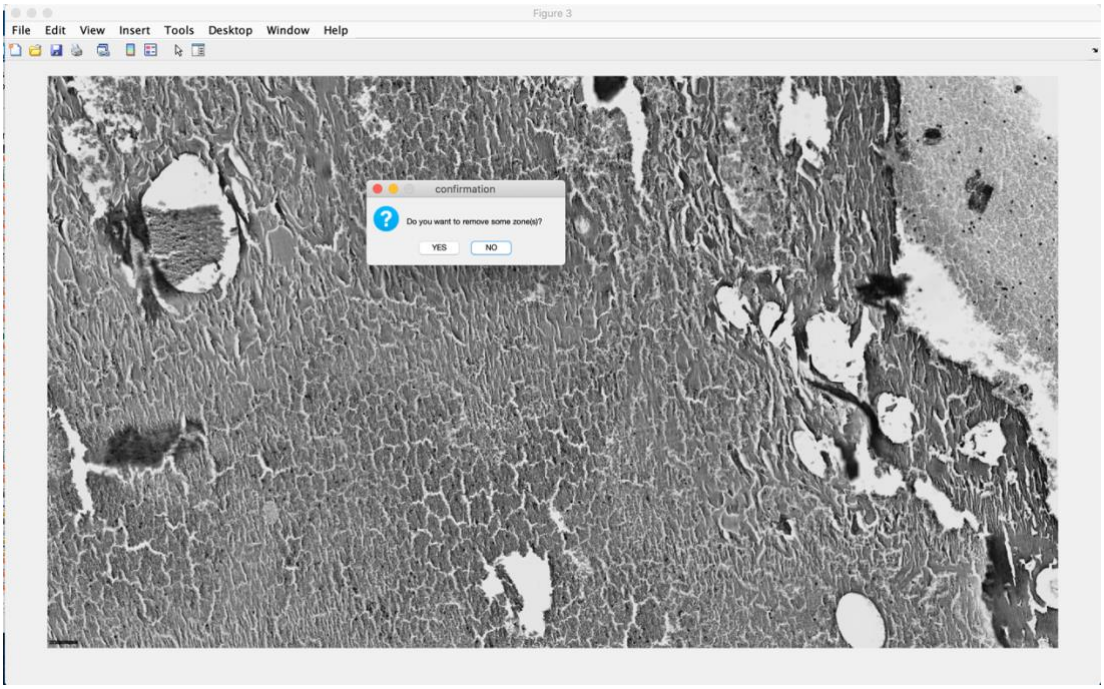

Figure 5. MASQH allows you to remove zones you don't want to include in an image's analysis.

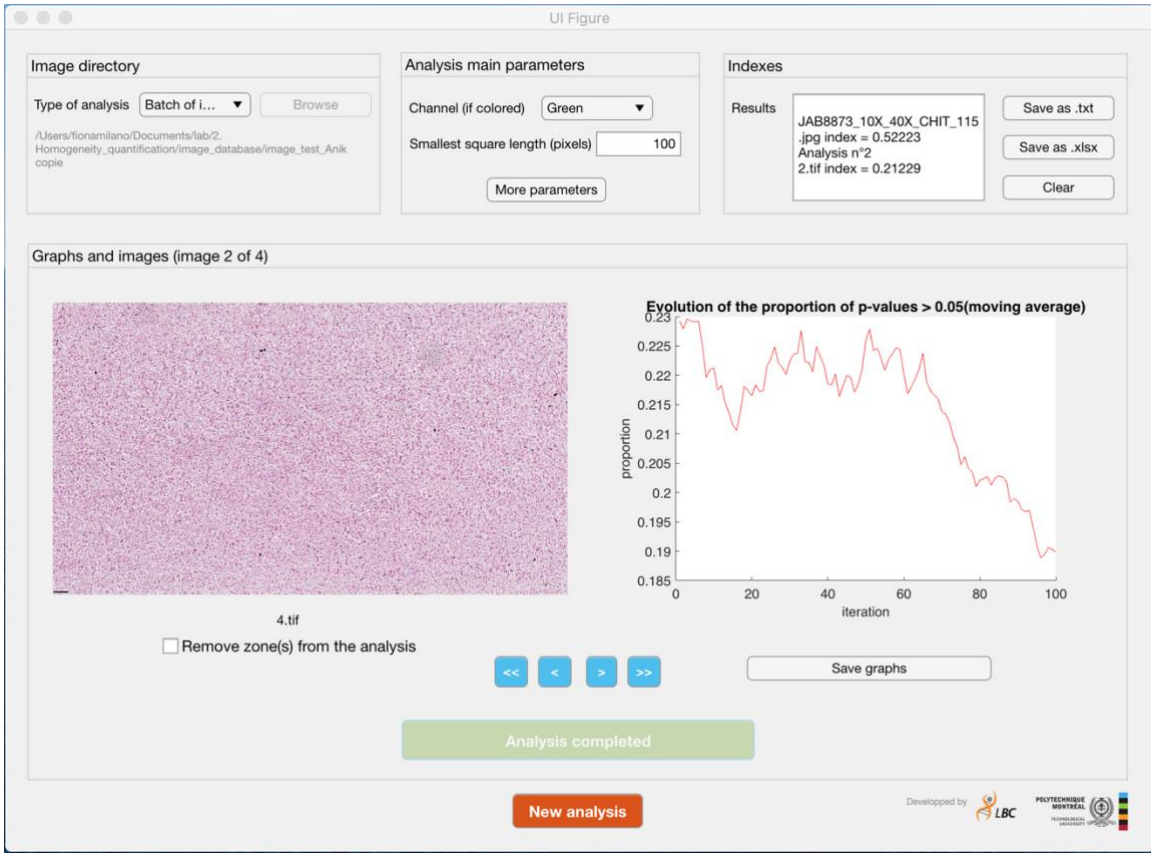

Figure 6. MASQH window after launching the analysis.

Parameters

When clicking on the "More parameters" button (number 8 in Figure 1), the "detailed changes" window appears (Figure 2). This window allows you to change all the algorithm's parameters, and so to adapt the MASQH algorithm to you situation. Initially, the value of each parameter was empirically determined and should fulfil most applications. Before changing a parameter, please refer to Table I or Milano *et al.* [1] to make sure you understand the role of this parameter and the implication if its change (e.g. drastically increase the computation time). In the case of a folder of images, after modifying the parameters, clicking on "apply" it will change them for every image analyzed, as the homogeneity index should be computed with the same parameters for all images compared.

When changing the color channel through the detailed change window, you will also be able to see the resulting black and white image, and so to visually determined which channel gives the best contrast.

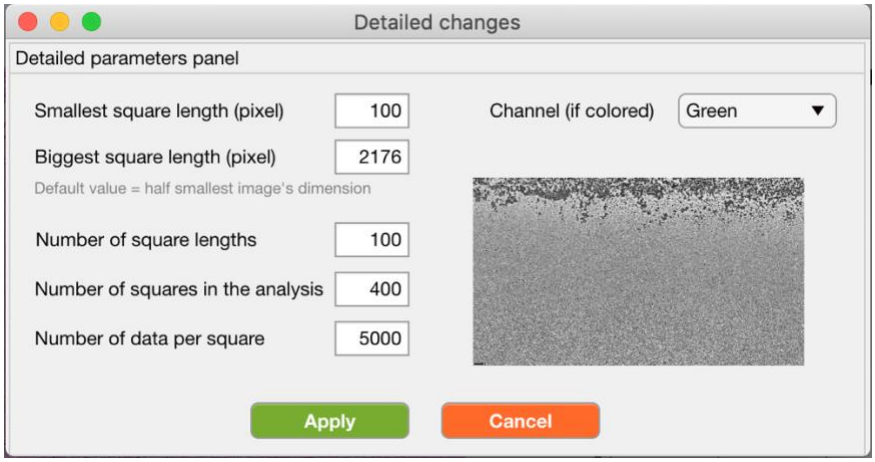

Figure 2. The "detailed changes" window allows to fine tune the scratch segmentation.

Output files

The "save as .txt" button allows you to save all numerical results output by the algorithm in a .txt file, to allow an easy further processing of your data. Similarly, the "save as .xlsx" button allows you to save all numerical results in a .xlsx file.

The "save graph" button allows you to save all outputted graphs.

Feel free to use any numerical result or images given by our application in your article by using the citation information given bellow.

Citation information

When conducting analysis by using MASQH, please include the following citation in all related publications:

- F. Milano, A. Chevrier, G. D. Crescenzo, and M. Lavertu, "Robust Segmentation-Free Algorithm for Homogeneity Quantification in Images," *IEEE Transactions on Image Processing*, vol. 30, pp. 5533-5544, 2021, doi: 10.1109/TIP.2021.3086053.

Contact

If you have any question or suggestion regarding the MASQH software (either source code or application), do not hesitate to contact us.

- Fiona Milano: Fiona.milano@polymtl.ca

We will be pleased to help you use the software, or to discuss with you about any improvement you can think about when using the software.

[1] F. Milano, A. Chevrier, G. D. Crescenzo, and M. Lavertu, "Robust Segmentation-Free Algorithm for Homogeneity Quantification in Images," *IEEE Transactions on Image Processing*, vol. 30, pp. 5533-5544, 2021, doi: 10.1109/TIP.2021.3086053.