Supplementary Materials: CollagenFitJ, a FIJI plugin for the quantification of collagen in polarization-resolved second harmonic generation image sets

Radu Hristu[a], Farzad Fereidouni[b], Stefan G. Stanciu[a,c], Lucian G. Eftimie[d,e], Oana C. Voinea[f,g], Danielis Rutkauskas[h], Laura Sironi[i], Giuseppe Chirico[i], Denis E. Tranca[a], Alberto Diaspro[j,k], and George A. Stanciu[a]*

[a]National University of Science and Technology Politehnica Bucharest, Center for Microscopy-Microanalysis and Information Processing, Bucharest, Romania
[b]Emory University School of Medicine, Department of Pathology and Laboratory Medicine, Atlanta, Georgia, United States
[c]National University of Science and Technology Politehnica Bucharest, CAMPUS Research Institute, Bucharest, Romania
[d]Central Military Emergency University Hospital "Dr. Carol Davila", Department of Pathology, Bucharest, Romania
[e]National University of Physical Education and Sports, Faculty of Physical Therapy, Bucharest, Romania
[f]Cantacuzino National Medical-Military Institute for Research and Development, Bucharest, Romania
[g]"Carol Davila" University of Medicine and Pharmacy, Pathology Department, Faculty of Medicine, Bucharest, Romania
[h]Center for Physical Sciences and Technology, Vilnius, Lithuania
[i]Università Degli Studi di Milano-Bicocca, Physics Department, Milan, Italy
[j]Istituto Italiano di Tecnologia, Nanoscopy and NIC@IIT, Genoa, Italy
[k]University of Genoa, Department of Physics, Genoa, Italy

*Correspondence to: stanciu@physics.pub.ro

Summary:

1. **CollagenFitJ installation and operation procedure**
2. **Comparison of the fitting accuracy of CollagenFitJ and a slow iterative algorithm**
3. **Dependence between the number of images in a PSHG stack and the quality of the fit estimated by ERR, SNR, R2**
4. **Collagen segmentation strategies using CollagenFitJ outputs**

1. CollagenFitJ installation and operation procedure

Installation of the CollagenFitJ plugin:

1) Copy the "Collagen_fit.jar" file in the *FIJI/plugins* folder or in a new folder such as: *FIJI/plugins/CollagenFitJ*.
2) Restart FIJI if it was open while performing step 1. The plugin can be accessed by selecting the "Collagen fit" command in the *Plugins* menu bar. This command should appear either directly in the *Plugins* menu, or under *CollagenFitJ* if the plugin was placed in the *CollagenFitJ* folder during step 1.

Installation of the CollagenFitJ macro:

3) Copy the file "CollagenFitJ_macro.ijm" in the same location used in step 1.
4) Restart FIJI if it was open while performing step 3. The macro has now been automatically installed as a "CollagenFitJ macro" command visible in the *Plugins* menu or under *CollagenFitJ* in case the macro was copied in the *CollagenFitJ* folder in step 3.
5) (optional) To re-open the macro, use the *File* > *Plugins* > *Macros* > *Edit* command, or drag and drop it on the FIJI window.

Installation of the CollagenFitJ macros for data export:

6) Copy the files "CollagenFitJ_export.ijm" and "CollagenFitJ_NaN to 0.ijm" to the same location as in step 1.
7) Restart FIJI if it was open while performing step 6. The macros are now installed as "CollagenFitJ export" and "CollagenFitJ NaN to 0" commands, both visible in the *Plugins* menu, or under *CollagenFitJ* if the macros were copied in the *CollagenFitJ* folder in step 6.

First run of the CollagenFitJ plugin:

8) Open a PSHG image stack or a collection of PSHG images. If you have a set of PSHG images, use the *Image* > *Stacks* > *Images to Stack* command to create a new stack from the images opened in separate windows. Ensure that the images are acquired at various laser beam polarizations ranging from 0° to 180°, regardless of the angle step. In the newly created stack, the PSHG images should be arranged in increasing order of polarization angle, from 0° to 180°, including images taken at both 0° and 180° polarization directions. Be aware that manually dragging and dropping individual images in FIJI may cause unexpected image arrangement. Before running the plugin, verify that the images in the PSHG image stack are correctly ordered according to the increasing polarization angle.
9) Run the plugin accessing the *Plugins* > *(CollagenFitJ >) Collagen fit* command.

10) Conduct a preliminary review of the output images by examining their histograms (*Analyze > Histogram*). For instance, verify if the histogram of the R2.tif image shows values clustered near 1, which would indicate a good fit. Additionally, assess the pixel values in the CHI3115.tif image; the median value should be close to 1.

11) (Optional for advanced users. For an alternative automated option, refer to step 16A) Create a custom collagen mask using one of the following methods:

A. Select one of the fitting quality assessment images (i.e., ERR, SNR, R2) to generate a collagen mask with the command: *Process > Binary > Make Binary*. Alternatively, for a more tailored approach, use the command *Image > Adjust > Threshold…*. Suggested threshold values are as follows: for R2, consider foreground pixel values between 0.8 and 1; for SNR, values above 6; and for ERR, values less than 0.1. These values are approximate and may not be suitable for all image sets. Inspect the histograms of the fitting assessment images (*Analyze > Histogram*) to determine the optimal threshold values.

B. Apply different thresholding techniques to create a mask from the PSHG image stack. For example, binarize the average intensity image derived from the stack (access *Image > Stack > Z Project…* and select Average Intensity).

Advanced users may refine these masks further using binary operations. Refer to step 16B for instructions on how to utilize these collagen masks.

Setting the CollagenFitJ macro

12) Run the CollagenFitJ macro (**Figure S1a**) by accessing the *Plugins > (CollagenFitJ >) CollagenFitJ macro* command.

13) In the "Input / Output" dialog box (**Figure S1b**), choose one of the following input options:

- select "Input data is a series of images in a folder" and specify the folder containing the PSHG image series.
- select "Input data is an image stack" and provide the location of the PSHG image stack file.

If neither option or both options are selected, the dialog will reappear until one option is chosen.

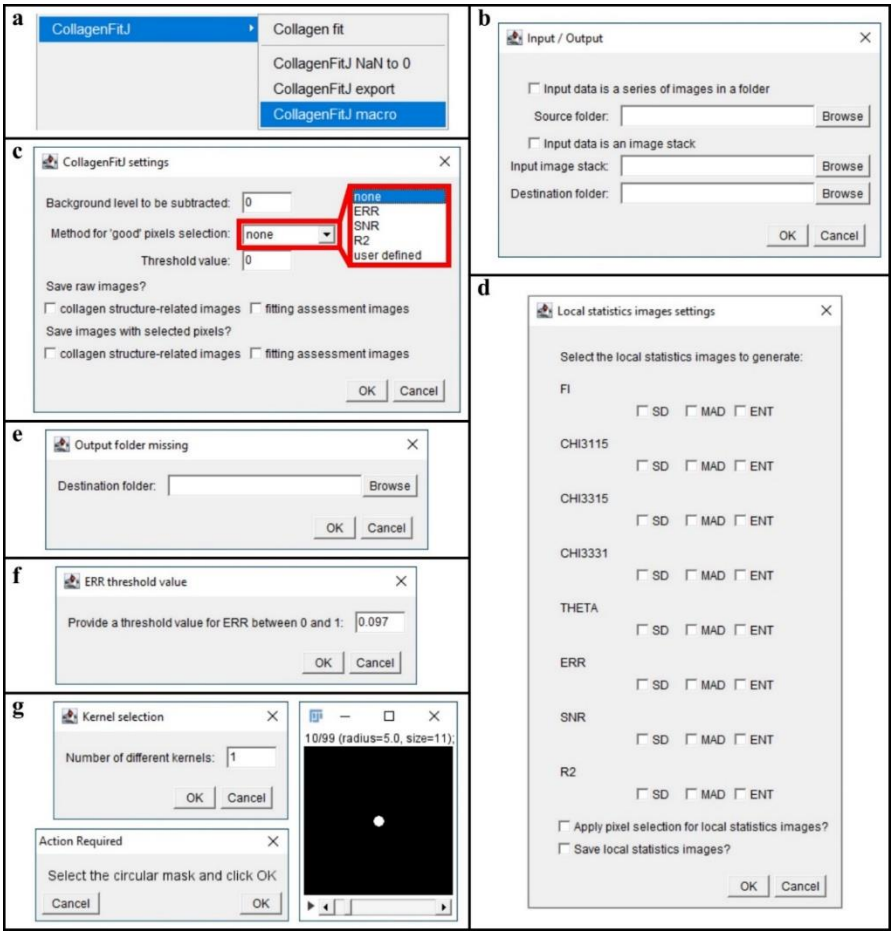

Figure S1. An overview of the CollagenFitJ plugin and macro for FIJI, designed as a step-by-step wizard: a) Launching CollagenFitJ: After installing the plugin and macros in FIJI, the user selects the "CollagenFitJ macro" command to start. b) Input Dataset Selection: Choose either a stack or a series of polarization-dependent images from a folder. c) Plugin Settings: Configure options for background subtraction, pixel selection, and specify which images to save. d) Local Statistics Settings: Select which images should have local statistics (SD, MAD, ENT) computed. e) File Saving Reminder: If any "Save" checkboxes are selected but no output folder is specified, this dialog box will prompt the user to provide a destination for saving the images. f) Threshold Value Check: If the threshold values provided for ERR, SNR, or R2 are outside the recommended range from step 16, the dialog box will prompt the user to enter a new value, showing the mean value as a reference. g) Kernel Selection: Choose the number of kernels and select specific kernels from ImageJ's "Circular Masks" for generating local statistics images. Each kernel is defined by two parameters: radius and size.

14) In the "Input / Output" dialog box, specify a destination folder where you would like to save the results, if desired.

15) In the “CollagenFitJ settings” dialog box (**Figure S1c**), enter a background value. This value will be subtracted from each pixel in the original image stack before running the plugin. This step is particularly recommended for widefield PSHG image sets to eliminate the average dark current, which can contribute to a non-zero image background.
16) In the “CollagenFitJ settings” dialog box, choose one of the strategies for selecting “good” pixels, which are those considered to contain relevant collagen information. Pixels not meeting the criteria will be assigned a NaN value and excluded from further processing. The available options are (red box in **Figure S1c**):

    A. Select one of the fitting assessment images (e.g., ERR, SNR, R2) from the dropdown menu to create a binary mask, and specify a threshold value. Thresholds should fall within the following ranges: for ERR, between 0 and 1; for SNR, between 0 and 30; and for R2, between 0 and 1. The generated mask will be applied to each of the eight CollagenFitJ output images.

    B. Choose “user defined” from the dropdown menu to use a custom-created binary mask. Refer to step 11B for suggestions on how to generate such a mask.

    C. Select “none” from the dropdown menu if you do not want to exclude any pixels.
17) In the “CollagenFitJ settings” dialog box, choose which output images to save. The images will be saved in the destination folder specified in step 14. You can select to save the collagen structure-related images (i.e., FI, CHI3115, CHI3315, CHI3331, THETA), the fitting quality assessment images (i.e., ERR, SNR, R2), all images, or none. Regardless of your selection, all images will be provided as an image stack when the macro finishes.
18) In the "CollagenFitJ Settings" dialog box, select the output images that will include the chosen "good" pixels to be saved. These images will be saved in the output folder specified in step 14, but only if one of the pixel selection methods was chosen from the dropdown menu in steps 16A or B. Options include saving versions with selected pixels for the collagen structure-related images, the fitting quality assessment images, all images, or none. Regardless of your selection, all images will be provided as an image stack for visual inspection and analysis when the macro is completed.
19) (Optional: if “user defined” is selected in the “Method for ‘good’ pixels selection” dropdown menu) Specify the file location for the binary collagen mask in the “User defined collagen mask” dialog box. Ensure that background pixels are set to 0 and foreground pixels to 255.
20) In the “Local statistics images settings” dialog box (**Figure S1d**), select the randomness and dispersion maps (i.e., SD, MAD, and ENT) to be computed for each of the eight images

generated by the CollagenFitJ plugin (i.e., FI, CHI3115, CHI3315, CHI3331, THETA, ERR, SNR, R2). You can choose the statistics to compute for each CollagenFitJ image individually.

We recommend performing an initial run with optional pixel selection before generating local statistics images, as this process can be time-consuming depending on the kernel size. Use the R2.tif and CHI3115.tif images (see step 10) for an initial evaluation of the fitting.

21) (Optional: if a pixel selection method was chosen in steps 16A or B) In the "Local statistics images settings" dialog box, check the "Apply pixel selection for local statistics images" box to generate the SD, MAD, and ENT images only for the "good" pixels. This checkbox will be unavailable if "none" was selected in the dropdown menu in step 16.
22) In the "Local statistics images settings" dialog box, decide whether to save the randomness and dispersion maps. Regardless of your choice, all generated maps will be included in the CollagenFitJ output image collection displayed at the end of the macro for visual inspection and analysis.
23) In the "Kernel selection" dialog box (**Figure S1g**), specify the number of kernels to use for generating the randomness and dispersion maps. This dialog box will only be accessible if at least one local statistics image is selected in step 20. The available kernel options include ImageJ's predefined circular masks, which consist of 99 different kernels (accessed via *Process > Filters > Show Circular Masks…*).
24) Choose a kernel from the circular masks and confirm your selection by clicking OK. This step must be repeated for each kernel number selected in the previous step.
25) The CollagenFitJ macro will begin execution. Once it is completed, an image stack containing all the CollagenFitJ images and local statistics maps will be displayed.

Option A: the "CollagenFitJ NaN to 0" macro

26) Select the "CollagenFitJ NaN to 0" command installed in step 7 to assign a value of 0 to pixels with NaN values.
27) In the "CollagenFitJ NaN to 0" dialog box (**Figure S2a**), choose one of the following input options: single file or all images in the folder. A single file can be either an individual image or an image stack. If neither or both options are selected, the dialog will reappear until one option is chosen.
28) In the "CollagenFitJ NaN to 0" dialog box, specify a destination folder where the results will be saved.
29) Click OK, and the macro will process all the provided images, convert NaN pixels to 0, and save the modified images with their original filenames in the output folder.

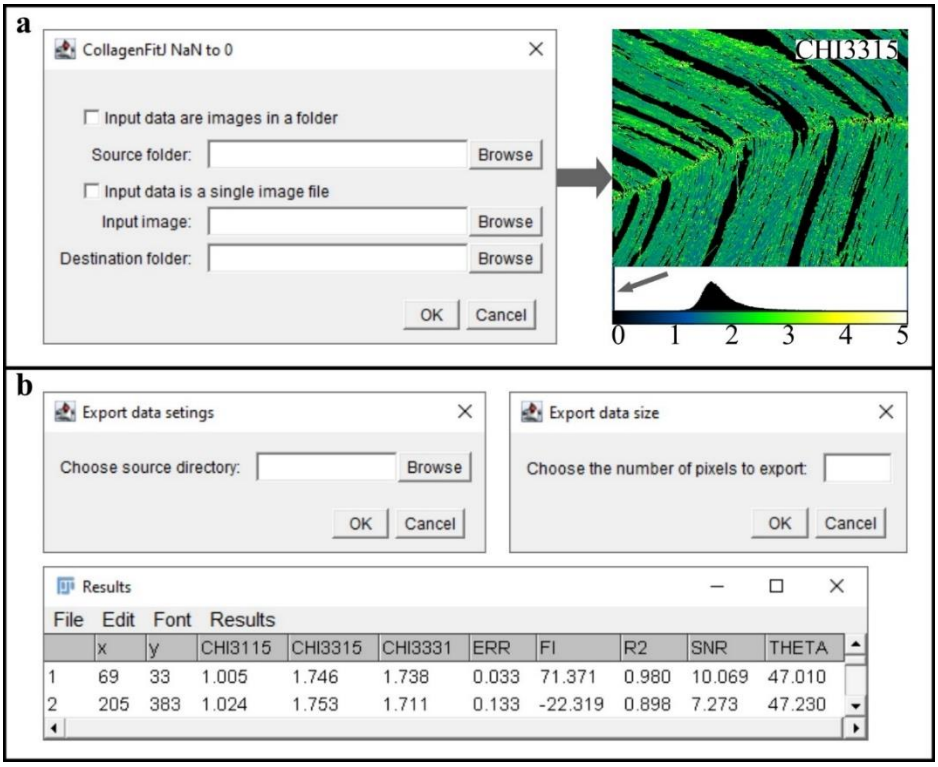

Figure S2. Optional features of CollagenFitJ: a) A macro for converting NaN pixels to 0 (workflow B). b) A macro for exporting randomly selected pixels for statistical analysis (workflow C).

Option B: the "CollagenFitJ export" macro

30) Click the "CollagenFitJ export" command installed in step 7 to start the macro, which will extract a random sample of pixel values from the entire set of generated images.

31) In the "Export data settings" dialog box (**Figure S2b**), specify the path to the folder containing the images from which the pixel sample will be extracted.

32) In the "Export data size" dialog box (**Figure S2b**), enter the desired sample size. Ensure the sample size does not exceed the maximum number of non-NaN pixels across all images in the selected folder. This maximum value will be displayed in the dialog box.

33) Click OK to display the macro output in the Results window (**Figure S2b**). The first column shows the number of entries corresponding to the selected pixel sample size. The next two columns present the pixel coordinates (x, y), and subsequent columns list the pixel values for each image, with the image titles provided in the column headings. To analyze or plot the data, copy and paste the contents of the Results window into a spreadsheet or directly into statistical software (e.g., GraphPad Prism).

2. Comparison of the fitting accuracy of CollagenFitJ and a slow iterative algorithm

We tested both CollagenFitJ and the slow iterative algorithm on two PSHG image sets acquired from malignant (**Figure S3a**) and benign (**Figure S3b**) thyroid tissue sections. Comparing computational speed, the same image sets are processed in approximately six hours using the iterative algorithm, whereas the typical runtimes for CollagenFitJ are presented in Section 3.3 of the manuscript and are ranged betweeen several seconds and sereval minutes depending on the processing parameters provided by the user.

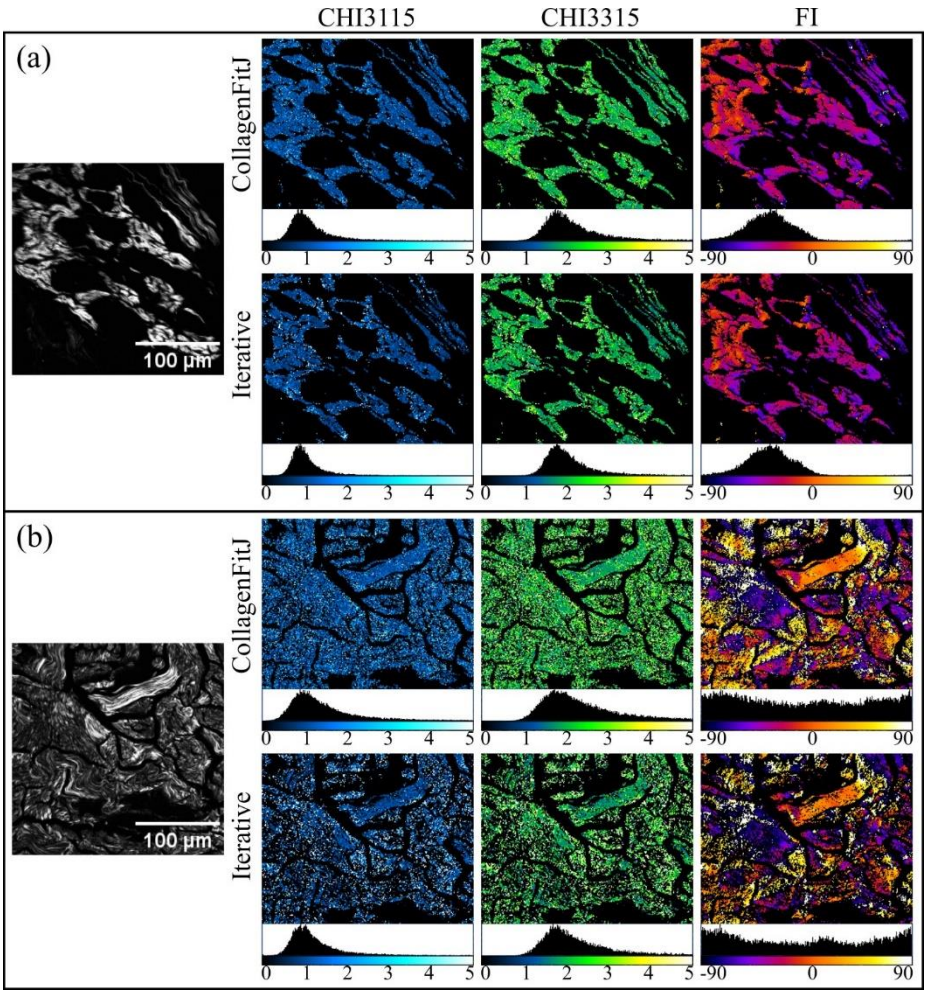

Figure S3. Comparison of fitting outputs (CHI3115, CHI3315, and FI images) obtained for malignant and benign samples using CollagenFitJ and the slow iterative algorithm. Histograms illustrate the distribution of pixel values in the images, and the corresponding summary statistics are provided in Table S1 for clearer comparison.

To compare the quality of fitting between the two methods, we use the fitting efficiency, previously defined as the ratio between the number of pixels with $R^2 > 0.8$ and the total number of pixels with non-NaN values after fitting. For both image sets, the fitting efficiency is slightly higher with the slow iterative algorithm. However, the iterative algorithm produces a

considerably lower number of meaningful fitted pixels compared to the CollagenFitJ plugin. This results in darker regions visible in the CHI3115 and CHI3315 images in **Figure S3**.

Beyond fitting accuracy evaluated by fitting efficiency, it is also important to compare the fitting results for the χ_{31}/χ_{15} and χ_{33}/χ_{15} ratios, represented by the CHI3115 and CHI3315 images, respectively. We compare the distribution of these values considering the mode, mean, and SD (Table S1). While all these values are slightly higher for the iterative algorithm compared to the CollagenFitJ outputs, the higher SD indicates that additional pixels with meaningful fitted results include higher values. We propose correcting these differences by filtering the output images of CollagenFitJ and including only pixels with corresponding $R^2 > 0.8$ in the estimation. For the filtered output, SD values are lower than those in the raw CollagenFitJ output, with both mode and mean values for the parameters approaching those obtained from the iterative algorithm.

Table S1. Comparison of fitting results obtained for malignant and benign samples using CollagenFitJ, a filtered output of CollagenFitJ ($R^2 > 0.8$), and the slow iterative algorithm.

		CollagenFitJ	$R^2 > 0.8$	Iterative
malignant	No. of fitted pixels	16464	11151	13516
	Fitting efficiency	0.68	-	0.73
	χ_{31}/χ_{15} mode/mean/SD	0.85/1.2/0.64	0.86/1.1/0.54	0.87/1.1/0.5
	χ_{33}/χ_{15} mode/mean/SD	1.75/2.19/0.73	1.75/2.15/0.66	1.79/2.1/0.62
benign	No. of fitted pixels	30372	12365	20198
	Fitting efficiency	0.41	-	0.45
	χ_{31}/χ_{15} mode/mean/SD	0.89/1.46/0.83	0.92/1.3/0.73	0.91/1.27/0.69
	χ_{33}/χ_{15} mode/mean/SD	1.73/2.32/0.86	1.73/2.3/0.8	1.75/2.24/0.79

3. Dependence between the number of images in a PSHG stack and the quality of the fit estimated by ERR, SNR, R2

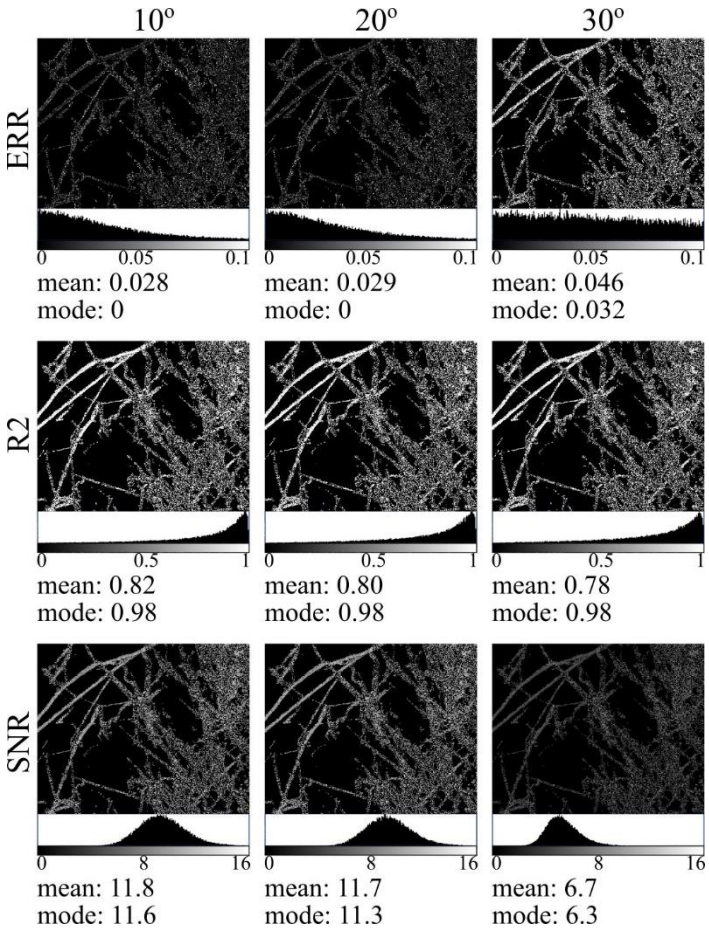

Figure S4. Fitting quality estimation maps and their dependence on the number of images in the PSHG stack. All three stacks include images acquired with laser beam polarization varying from 0° to 180° in steps of 10° (19 images in the stack), 20° (10 images in the stack), and 30° (7 images in the stack), respectively. While R^2 values slightly decrease as the number of images in the stack decreases, ERR and SNR remain nearly constant for stacks acquired with 10° and 20° steps. However, for the stack acquired with 30° steps, ERR increases significantly, while SNR decreases. This outcome is generally expected, but the goodness-of-fit can also be interpreted in relation to the suitability of the collagen model for the sample. While the single-axis molecule model, which assumes cylindrical symmetry, was used here, more general models (e.g., the trigonal model) have been shown to provide better results in certain cases, thus expecting an increase in fitting quality estimators.

4. Collagen segmentation strategies using CollagenFitJ outputs

Using CollagenFitJ fitting assessment images, we evaluate the generation of binary collagen masks using various thresholds: 0.2 and 0.4 for ERR, 2 and 6 for SNR, and 0.8 and 0.9 for R2 images (**Figure S5**). We also test the enhancement of masks generated from CollagenFitJ fitting assessment images using binary operations to reduce noise, refine object boundaries, and improve segmentation results.

In **Figure S5a,b**, because only a few pixels in the R2 images had values above 0.8 and 0.9, respectively, the improved version of these masks was not attempted. For visual comparison, we also include intensity-based segmentation masks generated using FIJI's automated thresholding procedure with the MinError algorithm. These masks were further refined using binary operations.

While intensity-based segmentation is widely used in image analysis, alternative approaches incorporating fitting quality estimation images from CollagenFitJ show promise. Their advantage lies in the fact that SNR, ERR, and R2 values have biophysical significance, as they quantify the goodness-of-fit between experimental data and collagen models.

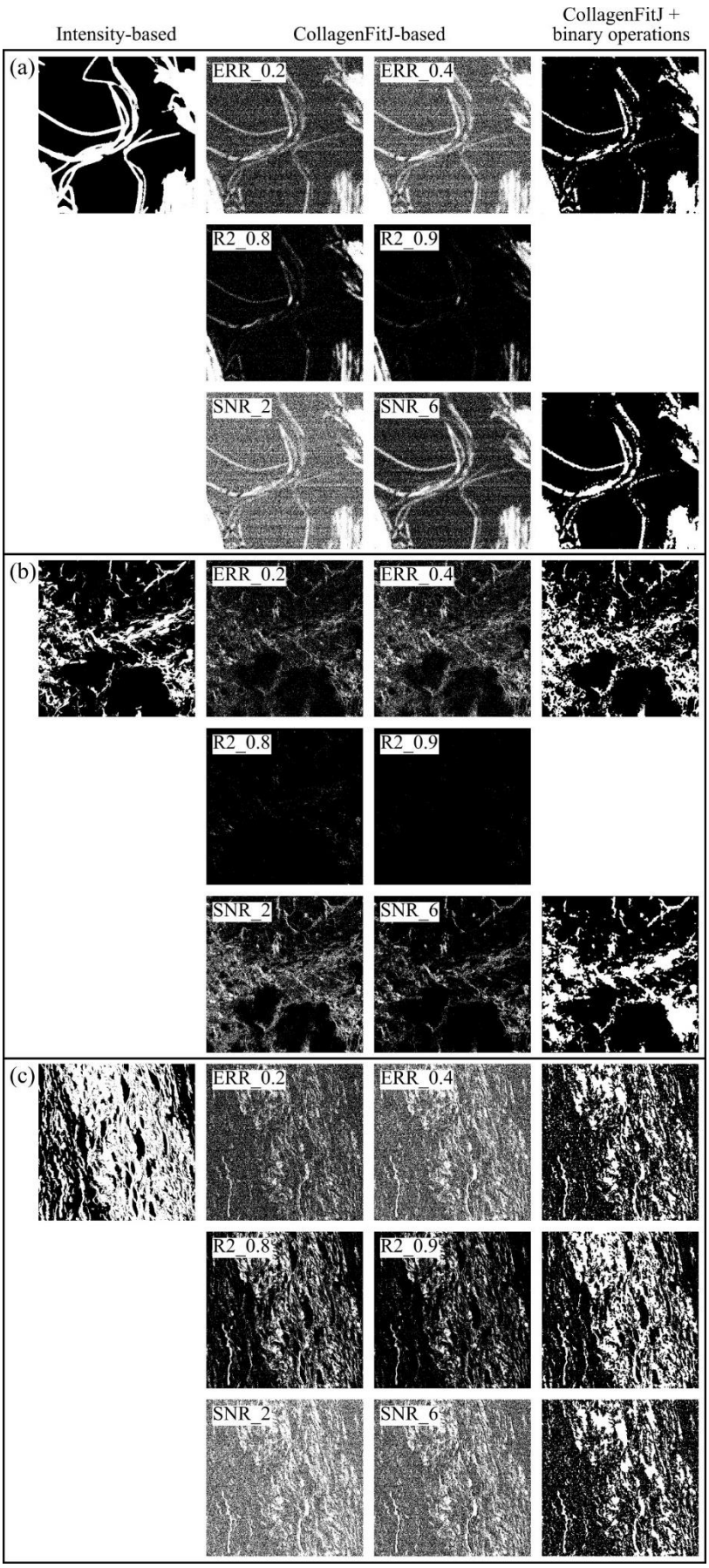

Figure S5. Thresholding strategies based on pixel intensities, CollagenFitJ fitting assessment images with varying threshold values, and the enhancement of these images using binary operations in FIJI on various samples: a) Loose collagen fibrils in rat tail tendon; b) Mouse colon tumor; c) Follicular thyroid carcinoma nodule capsule.