HHS Public Access

Author manuscript

Methods Mol Biol. Author manuscript; available in PMC 2019 January 23.

Published in final edited form as:

Methods Mol Biol. 2017 ; 1627: 429–451. doi:10.1007/978-1-4939-7113-8_28.

Methods for Quantifying Fibrillar Collagen Alignment

Yuming Liu[#1], **Adib Keikhosravi**[#1,2], **Guneet S. Mehta**[1,2], **Cole R. Drifka**[1,2,3], and **Kevin W. Eliceiri**[4,5,6]

[1]Laboratory for Optical and Computational Instrumentation, University of Wisconsin, Madison, WI, USA.

[2]Department of Biomedical Engineering, University of Wisconsin, Madison, WI, USA.

[3]Morgridge Institute for Research, Madison, WI, USA.

[4]Laboratory for Optical and Computational Instrumentation, University of Wisconsin, Madison, WI, USA. eliceiri@wisc.edu.

[5]Department of Biomedical Engineering, University of Wisconsin, Madison, WI, USA. eliceiri@wisc.edu.

[6]Morgridge Institute for Research, Madison, WI, USA. eliceiri@wisc.edu.

[#] These authors contributed equally to this work.

Abstract

Recent evidence has implicated collagen, particularly fibrillar collagen, in a number of diseases ranging from osteogenesis imperfecta and asthma to breast and ovarian cancer. A key property of collagen that has been correlated with disease has been the alignment of collagen fibers. Collagen can be visualized using a variety of imaging techniques including second-harmonic generation (SHG) microscopy, polarized light microscopy, and staining with dyes or antibodies. However, there exists a great need to easily and robustly quantify images from these modalities for individual fibers in specified regions of interest and with respect to relevant boundaries. Most currently available computational tools rely on calculation of pixel-wise orientation or global window-wise orientation that do not directly calculate or give visible fiber-wise information and do not provide relative orientation against boundaries. We describe and detail how to use a freely available, open-source MATLAB software framework that includes two separate but linked packages “CurveAlign” and “CT-FIRE” that can address this need by either directly extracting individual fibers using an improved fiber tracking algorithm or directly finding optimal representation of fiber edges using the curvelet transform. This curvelet-based framework allows the user to measure fiber alignment on a global, region of interest, and fiber basis. Additionally, users can measure fiber angle relative to manually or automatically segmented boundaries. This tool does not require prior experience of programming or image processing and can handle multiple files, enabling efficient quantification of collagen organization from biological datasets.

Keywords

Fibrillar collagen; Collagen quantification; Collagen alignment; Collagen organization; Curvelet transform; Fiber tracking; Second-harmonic generation microscopy; Tumor-associated collagen signatures; Cancer imaging; Extracellular matrix

1. Introduction

Fibrillar collagens are one of the prominent extracellular matrix components. Studies have shown that specific changes in collagen alignment play an important role in contributing to the mechanical properties of tissues [1–4] and are associated with disease progression as shown in breast [5–7] and pancreatic [8–10] cancers.

Image analysis methods suitable for quantifying fibrillar collagen alignment and orientation are generally based on the intensity derivatives [11–16] or intensity variation [17], Fourier transform [18–22] or Hough transform [23, 24] or directional filters [25–27], or fiber tracking algorithm [28–31]. Derivative-based methods can usually yield pixel-wise orientation and require additional steps to detect fiber edges. The filter or transform-based methods can usually only yield global measurements for a window with a given size and would be time-consuming to yield pixel-wise orientation. Fiber tracking algorithms use fiber propagation criteria based on, for example, local fiber center location and orientation to form individual fibers and are usually computationally demanding and have been challenged in extracting complex fiber networks.

Given the importance of fiber alignment and the need for transparency on how it is being quantified, an emphasis of the collagen research community has been to develop open-source approaches rather than closed-source proprietary tools. Besides being freely available with fewer barriers to use and adaption, open-source tools allow for one to examine the algorithms being used to assure the best results and also allow programmers to adapt or reuse code in their specific workflows. Also, software licenses of most open-source scientific programming efforts are commercially friendly opening up opportunities for adaptation by commercial imaging tools.

Currently freely available open-source software tools for collagen alignment quantification include the ImageJ [32] plug-ins “FibrilTool” [15] and “Orientation J” [13, 33], the MATLAB (MathWorks, Natick, MA, USA) stand-alone applications “CytoSpectre” [34, 35], “CurveAlign” [25, 31, 36], and “CT-FIRE” [37, 38]. FibrilTool is based on the first derivative and borrows the concept of local nematic tensor as applied in the field of liquid crystal physics. It calculates the main orientation and alignment of the fibers within the region of interest (ROI) by using the average local nematic tensors over an ROI. This tool needs manual involvement to evaluate the results and to choose ROIs that are suitable for analysis. Hence, it can be time-consuming for high-throughput processing. Orientation J uses the first derivative and a Gaussian weighting function to construct a structure tensor commonly used in the field of image processing. It calculates the orientation and anisotropy information for each pixel and has additional steps to identify pixels belonging to a fiber and assign a weight to elongated structures for orientation calculation. This tool involves manual operation to annotate the ROI and is convenient for processing a single image or an image stack, but it is not trivial to process a large number of images. CytoSpectre uses Fourier transform-based spectrum analysis to extract the orientation and its anisotropy within a specified window. It does not require additional pre-or post-processing steps, is highly robust

to background noise, and can be easily used for processing multiple images. However, this tool can only provide global information rather than the localized fiber measurements.

Here, we describe the latest version of CurveAlign (CurveAlign 4.0), which provides a powerful and comprehensive platform for quantifying fibrillar collagen on a global, ROI, and fiber basis. The development of this program was driven by our active work in characterizing the role of collagen organization in disease [4–10, 31, 38–45]. This software platform uses two fiber analysis methods, CT-FIRE [38] and curvelets fiber representation (CFR) [25]. While these analysis methods can be downloaded and used as separate tools, we will here refer to the "CurveAlign" platform where CurveAlign can be run in both the CT-FIRE and CFR modes. CT-FIRE first uses the curvelet transform (CT) [46] to de-noise the image and enhance the fiber edges and then uses a fiber tracking algorithm (FIRE) [29] to form individual fiber networks, which includes major steps of fiber boundary detection, fiber center point identification, fiber extension from the center points, short fiber removal, and fiber network creation. While CFR does not extract individual fibers, it uses the curvelets in CT to directly represent the edges of fiber segments since curvelets have high-resolution information about their scale, location, and orientation and can optimally represent the local curve-like shapes. In contrast to other comparable methods, CFR does not need to add additional steps to detect fiber edges as in many pixel-wise approaches or to assume that spatial frequency of pixel intensity change can represent the fiber distribution as in the Fourier transform-based approach. In addition, measuring relative angle of the collagen fibers with respect to a boundary of interest, such as the tumor-stroma interface [5], is an important image analysis capability but is not available in the tools briefly mentioned above except for CurveAlign. Specifically, alignment can be calculated as angle relative to manually drawn boundaries or boundaries segmented through computational means [31]. Moreover, since fiber-wise orientation is calculated, CurveAlign also calculates alignment related to each fiber or fiber segments besides the alignment for a full image and an ROI. That is, each fiber or fiber segment has features related to alignment of specified surrounding rectangular regions (called box alignment) or nearest number of fibers (named nearest alignment) as described in detail in our previous publication [31].

CurveAlign capabilities have been successfully applied to numerous studies of collagen alignment. For example, the CT-FIRE module has been used to quantify fiber alignment in SHG images acquired from human pancreatic histology tissues. This enabled Drifka and colleagues to determine that stromal collagen alignment is significantly higher in pancreatic cancer compared to normal and chronic pancreatitis tissue [9]. Moreover, it was discovered that the degree of collagen alignment in pancreatic tumors correlates with patient survival [10]. A demonstration of CurveAlign in CT-FIRE mode, without a boundary for both a full image and ROI-based analysis, in pancreatic tissue [9] is shown in Fig. 1. CurveAlign has also been used in breast cancer research: for example, in CFR mode to quantify collagen alignment in breast cancer tissue where it indicated significantly less aligned fibers after treatment of a tumor [47]. It was also used in an in vitro study of breast carcinoma cells in 3D culture to show how the angle of fibers with respect to the boundaries formed by groups of cells differs under varying culture conditions and pharmacological perturbations [44]. In CT-FIRE mode, it was used to automatically quantify both the angle with respect to a boundary and localized fiber-based relative alignments, which are important factors in

human breast carcinoma prognosis [31]. A demonstration of using both CT-FIRE mode and CFR mode for quantification of relative angle and ROI-based alignment is shown in Fig. 2.

There are three kinds of alignments calculated in CurveAlign: (1) alignment with respect to the horizontal named "angle" or "orientation," ranging from 0 to 180°; (2) alignment with respect to a boundary named "relative angle" or "relative orientation," ranging from 0 to 90°; and (3) alignment of fibers with respect to each other named "alignment coefficient," ranging from 0 to 1. In CurveAlign, to calculate the alignment coefficient, the angle is first multiplied by 2 to map to the orientation range of [0–2π], and then the alignment coefficient is calculated as the normalized vector sum of orientation vectors or mean resultant vector length in circular statistics. The alignment coefficient ranges from 0 to 1, with 1 indicating perfectly aligned fibers, and smaller values representing more randomly distributed fibers. To be noted, the alignment coefficient indicates the dispersity of the fiber orientations and has been defined in different ways or terms, such as anisotropy index [21] based on eigenvalues of a second rank tensor, coherence [13] based on structure tensor eigenvalues, anisotropy based on circular variance [15] in circular statistics, or mean vector length [12, 31] in circular statistics (as in CurveAlign). Hereafter, the name "alignment" will be used to simplify the term of "alignment coefficient" unless stated otherwise.

The schematic workflow for quantifying collagen alignment in a single image with CurveAlign is shown in Fig. 3. The details of each step will be described in Subheading 3.1. The graphical user interface (GUI) of CurveAlign is modular, and some functions were implemented in different modules, which are separate from but can be called from the main GUI. To process multiple images in a batch mode, if needed, the CT-FIRE (CTF), boundary creation, and ROI manager modules can be used to prepare the required CTF output file, boundary mask file, or ROI annotation file in advance, respectively. Once all the required files associated with each image are provided, one can simply follow the same procedure as that for a single image analysis to run the quantification for all of the images. Output can be in the form of overlaid images of the extracted fibers or curve-lets on the original image, table of summary statistics, heatmap showing the alignment or relative angle distribution, Excel files of the fiber alignment coefficient and angles, as well as others. Postprocessing module can combine output of CurveAlign to show combined summary statistics of the quantified fiber features including collagen fiber alignment and orientation.

2 Materials

2.1 Images

Collagen is typically imaged using SHG microscopy, polarized light microscopy, or dyes and stains in conjunction with bright-field or fluorescence microscopy [48–54]. CurveAlign can analyze any image file readable by MATLAB regardless of the imaging modality used to collect it. As SHG imaging is a label-free modality offering optical sectioning and high specificity and resolution, it has become a popular fibrillar collagen imaging method *(see* Chap. 27 for protocols and tips). Our group and others have widely used SHG microscopy to investigate collagen in a variety of tissue types including human breast cancer tissue [38] and human pancreatic cancer tissue [9]. Data from these two studies in.tif file format will be used here for a demonstration.

2.2 Software

Download the MATLAB Runtime Compiler (MCR) and the CurveAlign software from [36] and follow the detailed requirements and installation instructions from the above web page. A user guide is also available for download: it provides detailed information for installation and operation.

Both stand-alone and source-code versions are freely available. The stand-alone version does not require a full MATLAB installation, but the user needs to install the freely available MATLAB MCR package (download link available at [36]). As CurveAlign requires significant memory use especially for big image processing or parallel computing, we only provide stand-alone applications for 64-bit Windows or Mac systems. These programs can be run on Unix as well but require code compilation for this operating system including Linux. The source-code version requires a full MATLAB installation including toolboxes of Signal Processing, Image Processing, Statistics Analysis, and Parallel Computing.

For the stand-alone version, open the software by double clicking (in Windows system) or right-clicking (in Mac) on the CurveAlign icon. This will launch the main GUI shown in Fig. 4a. Figure 4b shows the GUI for CT-FIRE individual analysis that can be launched by clicking on the “CT-FIRE” button (labeled as 6 in Fig. 4a).

3 Methods

3.1 Collagen Fiber Analysis with CurveAlign

1. Download and install MATLAB MCR and CurveAlign *(see* Subheading 2.2).
2. *Fiber analysis mode selection* (step 1 of Fig. 3): Select the fiber analysis method from the drop-down menu (labeled as 1 in Fig. 4a) among the following four options from CFR mode (CT) and CT-FIRE mode (segments, fibers, and endpoints) *(see* **Notes 1-3**):
   - **(a)** *CT:* The curvelet transform is performed on the image(s), and each curvelet corresponds to an observation in the feature list. Each curvelet has a unique angle and position. In this case, no individual fiber is extracted and the curve-lets that indicate the representative locations and directions of fiber segments are calculated as shown in Fig. 5a.
   - **(b)** *CT-FIRE segments:* The output from CT-FIRE as illustrated in Fig. 5 b is used as the input, providing information about the fibers in the image. Each fiber is represented as segments with individual orientations and locations, and each fiber segment in the CT-FIRE output file corresponds to an observation in the CurveAlign feature set. Each segment has equal length (set at 5 pixels by default) and has a unique angle and position.
   - **(c)** *CT-FIRE fibers:* The output from CT-FIRE as illustrated in Fig. 5 b is used as the input, providing information about the fibers in the image. Each fiber center point in the CT-FIRE output file corresponds to an observation in the feature set. Each center point has a unique angle

corresponding to the average angle of the fiber. The position is the midpoint between the fiber endpoints by default.

(d) *CT-FIRE fiber endpoints:* The output from CT-FIRE as illustrated in Fig. 5 b is used as the input, providing information about the fibers in the image. Each fiber endpoint in the CT-FIRE output file corresponds to an observation in the feature set. Each endpoint has a unique angle and position. The angle corresponds to the angle of the entire fiber.

3. *Determine boundary conditions* (step 2 of Fig. 3): Select the boundary processing mode from drop-down menu (labeled as 2 in Fig. 4a) among the following options: No Boundary, CSV Boundary, and TIFF Boundary (see **Notes 4** and **5**). For calculations of fiber angles with respect to a boundary, a corresponding boundary file for each image or each stack has to be provided. The user can follow the instructions on the screen to manually or gautomatically annotate either a "CSV" (comma-separated values format based, x-y coordinates) boundary file or a "TIFF" boundary file. The boundary files created in CurveAlign will be saved according to the file directory and file naming conventions described below.

(a) *CSV boundaries:* Previously stored boundary files can be used that contain the x-y positions of points along the boundary. CSV boundary files can also be manually annotated in later steps. CSV files must adhere to the following naming convention if they are defined outside CurveAlign: If the image is named "TACS-3a.jpg," the CSV file must be named "boundary for TACS-3a.jpg.csv." The boundary file must be in a sub-directory named "CA_Boundary."

(b) *TIFF boundaries* allow for identifying whether a fiber is inside an ROI. Mask files are used to indicate where region boundaries are located. The mask files should be 8-bit image files where the inside of ROIs is given a value of 255 or 1 and outside a value of 0. Mask files must be registered with the collagen image files and therefore have the same dimensions as the image files. TIFF boundary files must adhere to the following naming convention if they are defined outside CurveAlign. If the image is named "1B-a5.tif," the TIFF boundary file must be named "mask for 1B-a5.tif.tif." The boundary file must be presented in a sub-directory named "CA_Boundary". Boundary creation module can be launched by clicking on the "BD Creation" button labeled as 7 in Fig. 4a. An example of automatic TIFF boundary creation is shown in Fig. 6.

4. *Load image(s) or data file(s)* (step 3 of Fig. 3): Click on "Get Image(s)" button (labeled as 3 in Fig. 4a) and browse to an image. Images may be single or stacks. If needed, select multiple images in the dialog using the technique appropriate for your operating system (e.g., in Windows, hold CTRL while selecting multiple files).

5. *Annotate the ROIs* (step 4B of Fig. 3) by clicking on the "ROI Manager" button (labeled as 4 in Fig. 4a). Two new windows ("ROI Manager" and "CurveAlign ROI analysis output table") will open. Options to delineate and annotate different shapes of ROI as shown in Fig. 7b are provided in the dropdown menu of the "ROI Manager" window (Fig. 7a) and include rectangle, polygon, ellipse, freehand drawing arbitrary shapes, or specified rectangle. To create an ROI, select the desired shape in the drop-down menu, draw the shape directly on the image with a drag-and-drop mouse gesture, and click on "Save ROI(s)" (or press the "s" key). The newly created ROI then becomes listed on the left column of the ROI window with the default name ROI1, ROI2, etc. (Fig. 7a). ROIs can be renamed, deleted, combined with other ROIs, and saved into or loaded from a ".mat" file, mask image file, or ".txt" file. After drawing and saving an ROI, press the "d" key to continue to draw ROIs, or drag the previous ROI to a new position to create a new ROI. If original image name is "ImageName.tif," ROI file is named "ImageName_ROIs. mat" and saved in a subfolder of the image's directory named "ROI_management" *(see* **Note 6**).

6. *Set running parameters* in the CurveAlign main window (steps 4A and 4B of Fig. 3):

   (a) If fiber extraction mode "CT" is selected, enter the "fraction of coefs to keep" (labeled as 9 in Fig. 4a) to use the fraction of the largest coefficients of CT for curvelets calculation.

   (b) When calculating the angle of fibers with respect to a boundary (relative angle), the user can spatially restrict the area over which fibers are analyzed. To define this radius from the boundary line, enter the distance in pixels (labeled as 10 in Fig. 4a).

   (c) Select the desired outputs using the check boxes at the bottom of the main window in Fig. 4a (labeled as 11).

7. *Run the program and explore the results* (step 4 of Fig. 3): For full image analysis, click on the "Run" button labeled as 13 at the bottom of Fig. 4a. The progress information is displayed in a window highlighted in green color at the bottom of Fig. 4a, and the output data and images will open and will be automatically saved in a subfolder called "CA_Out" in the directory of the original image(s). For ROI image analysis, click on the "ROI analysis" button (labeled as 5 on Fig. 4a). It will prompt a window to select the ROI analysis mode *(see* **Notes 7** and **8**).

8. Post-processing: This module is used to combine the output of different images in the same folder. In the main window, click on "Post-Processing" (labeled as 8 on Fig. 4) to launch the post-processing module. In the "post-processing CA features" window, click on the "Get CA output folder" button to select the output folder for the processed images. Click on the "OK" button to have the combined output results. Among the Post-Processing Options, the default is to combine

"mean of all features" of each image; another useful option is "Combine feature files" to combine the features of all the fibers in the processed images.

9. The outputs of CurveAlign can be displayed in images, tables, and histograms and are mainly saved in the subfolders of "CA_ Out," "CA_ROI," and "CA_Boundary" in the format of mask image file, overlaid image, heatmap image, ".CSV" file, ".mat" file, and excel file. A demonstration of some typical outputs is shown in Fig. 8.

3.2 Analysis Time Estimates

The estimated time scale for running each step on an image with size of 1024 × 1024 pixels with moderate fiber density is shown below. These times are estimates and depend on the size of the file, the particular CurveAlign features deployed, the central processing unit (CPU) type, and the amount of random-access memory (or RAM) a user has available.

- CurveAlign CRF mode for fiber representation: a few seconds
- CurveAlign CT-FIRE mode for fiber extraction: ~100 s
- CurveAlign boundary analysis: tens of seconds

3.3 Anticipated Results

We conducted tests on computer-generated artificial fibers, different types of SHG images, and polarized light microscope images. By comparing the orientation and alignment calculated by CurveAlign to real values generated by computer program, manually measured values by a panel of experts, or the values computed by Orientation J under certain conditions, we found that, in most cases, CurveAlign yields statistically reliable results that have a good correlation to results yielded from other approaches (comparison not shown here).

3.4 Advantages and Limitations

The main advantages of the CurveAlign software platform include: (1) Its capability to directly track or represent fibers or fiber segments and make the fiber tracking results visible to the user. This trait is particularly useful to provide direct fiber-wise alignment analysis. (2) It allows relative orientation measurements against boundaries that can be manually or automatically created by the tool itself. (3) It is a comprehensive platform that can meet different needs of collagen quantification for images, stacks, and ROIs and is suitable for high-throughput collagen analysis.

Meanwhile, there have been some limitations or challenges for each of the two fiber analysis modes CT-FIRE and CFR. While the CT-FIRE mode can provide the most complete information for each individual fiber and can be applied to most images with large heterogeneities in terms of contrast and noise level, there are three major challenges for the CT-FIRE mode: (1) The software has difficulty to extract an intact fiber (especially those having curvy shapes and varying intensities along their length) from a dense fiber network. The tracking of such a long curvy fiber may end up with a few fiber segments along the fiber propagation direction. In order to overcome this issue, the local pixel-wise orientation and

anisotropy information (that can be readily calculated as described in other approaches [11–16]) can be combined to form a fiber. Alternatively, other techniques [28] such as those based on energy minimization along the fiber extension can be adapted to the current fiber extracting algorithm. In practice, if the orientation of individual fibers is not of interest, this issue is not actually a problem. In CurveAlign, the CT-FIRE segment mode can be selected as the fiber analysis method, which theoretically would yield comparable results to other applicable methods. (2) If fiber thicknesses vary a lot in an image, either the thinnest or the thickest fibers have to be compromised to some extent since only fibers whose thicknesses are within a fixed searching range can be accessed in the current fiber propagation algorithm. The fiber tracking strategy algorithm can be adapted to consider fiber thickness at different scales by adding fiber edge or ridge detection with automatic scale selection [55]. (3) The fiber tracking algorithm is computationally demanding, and it may take a much longer time than most derivative-or transform-based tools. The speed issue will need to be properly addressed before the program can perform real-time image processing or analysis of very large images (e.g., 10,000 × 10,000 pixels). Options include simplifying the fiber tracking algorithm to make it faster and combining ROI analysis (by automatically dividing image to smaller tilts) and parallel computing. These options will be investigated in future versions of the software. Of note, we have enabled parallel computing in CT-FIRE mode. Although this addition cannot reduce the processing time for each image, it allows the computer to process multiple images simultaneously depending on how many CPU cores in the computer are available for running CurveAlign, allowing for more rapid batch analysis of many images.

The CFR mode can directly find optimal representation of fiber edges using curvelets without any additional edge detection algorithms (that might be needed for derivative-or transform-based methods). The CFR mode operates at a similar time scale to derivative-or transform-based methods, which means it runs much faster than fiber tracking algorithms like CT-FIRE. However, in the current version of CT-FIRE (version 2.0), there are two major limitations. (1) It cannot be applied to images with a large variation in contrast. The program tends to bias toward fibers with highest intensities. A practical solution is using the ROI manager to annotate regions with even contrast for analysis and check the overlay results to make sure the extracted fiber edge matches the visual inspection. (2) It only uses the curvelets in one scale to characterize fiber edges and may not be able to faithfully represent all the fibers at different scales of thickness. In the current tool, we can set the scale relative to a specific thickness and set different group radius to group curvelets belonging to the same fiber, which can partially resolve this issue. A future solution would be to develop a strategy to combine the fiber information at different scales or do automatic scale selection [12] when calculating local curvelets orientation.

3.5 Troubleshooting and Practical Tips

Parallel computing techniques can be used for the CT-FIRE mode, which can save significant fiber extraction time for a batch of images. In some situations (e.g., for images with acceptable contrast), the CFR mode is equivalent to “CT-FIRE segments” option.

3.5.1 Adjust Parameters in CFR Mode—Some advanced settings can be selected by clicking on the "Advanced" button (labeled as 12 in Fig. 4a) and make changes to corresponding values, including:

1. The radius of grouping the curvelets, which is largely determined by the range of fiber thickness. The larger the grouping radius, the fewer curvelets can be displayed.
2. Scale: the default scale to be analyzed in CT-FIRE mode is the second finest scale. However, depending on the primary thickness the fiber and the image size, the scale can be set to a coarser scale, like the third finest scale, the fourth finest scale, etc.
3. Control of the alignment or relative angle heatmap. There are three related parameters, i.e., heatmap standard deviation filter for the no-boundary case, heatmap square max filter size, and heatmap Gaussian disc filter sigma. To do an accurate quantitative comparison of the heatmap, e.g., counting the number of pixels of different colors, those three parameters have to be set constant between analyses.

3.5.2 Adjust Parameters in CT-FIRE Mode—The primary parameters to be adjusted are the background intensity (named "thresh_im2" in the parameter setting window) and the searching radius (named "s_xlinkbox" in the parameter setting window) to find the center line of a fiber. The background threshold is associated with the noise level: the larger the threshold, the less collagen information remains. The searching radius is associated with the range of fiber thicknesses: the larger the value, the fewer of the smaller fibers information can be extracted. There are other parameters that can be used to find the optimal fiber extraction or representation. However, their effects on the final assessment of the alignment are minor and in most cases are negligible.

3.5.3 Using Boundaries—The automatic boundary creation is based on corresponding hematoxylin and eosin (H&E)-stained slide data (and not the primary SHG data) to best visualize the tumor boundary (Fig. 6). This gives the best result to calculate the angle of collagen fibers with respect to a tumor boundary (relative angle). The current version is optimized for H&E-stained human breast samples (with the corresponding SHG) but can be used for other epithelial cancers. The automatic boundary creation is needed to have a map of tumor locations, which includes image color enhancement, color segmentation, finding nuclei and grouping them, and removing unwanted regions. Due to having two steps of color enhancement in this algorithm, it should be robust for variations in stain hue or imaging systems. To improve the accuracy of the segmentation, a white balancing as a pre-processing step sometimes can help.

The segmentation process requires the bright-field H&E image be registered to the SHG image in the beginning. The registration algorithm includes several steps: first, enhancing the colors using decorrelation stretch [49]; second, extracting the collagen-like structure from the bright-field image by color separation; and third, registering the image from step two to the SHG image using an iterative algorithm at different scales. As a rule of thumb, for

the best results from the registration algorithm, the fields of view of the two images shouldn't differ by more than 10%, and the SHG signal intensity shouldn't be too weak.

When boundaries are used as part of the analysis, CurveAlign computes up to two relative angles per observation. The first is the nearest point angle and the second is the intersection point angle. The nearest point angle is the relative angle between the observation point and the tangent angle of the region boundary at the point nearest to the observation point. The intersection point angle is the relative angle between the observation point and the intersection point of the fiber, interpolated out a user-defined distance, and a boundary. For most applications in breast cancer, for example, nearest point angle is the preferred one.

3.5.4 Addttional Tips.—For the alignment quantification, in particular for fiber-based or ROI-based ones, the number of fiber or fiber segments should be larger than ten to yield a statistically robust measurement. Although the alignment coefficient and mean of orientations or angles are the main metrics used for quantifying fibrillar collagen alignment, CurveAlign also outputs other orientation-related metrics such as the variance, standard deviation, median, skewness, Kurtosis, etc. to characterize the distribution or dispersion of collagen fiber orientations. These provide additional useful information for some potential applications.

Visual inspection of the output images, which includes extracted fibers or curvelets overlaid on the original image, enables the user to assess the results for any image without the need for comparing against a ground truth. In addition, we can compare the statistics change by adjusting the primary parameters mentioned above. In many cases we tested, the alignment or orientation did not have significant change if the visual inspection is properly conducted.

3.6 Future Directions

For the current version of the CurveAlign platform (which includes software versions CurveAlign 4.0 and CT-FIRE 2.0), there are several improvements that we would like to make, many of them already mentioned above. Much of future development will be focused on usability, particularly in making it easier to run each mode and to be more modular. This will entail combining the functionality of CurveAlign and CT-FIRE more than it currently is, ideally into one common program. We also need to continue to improve the speed and performance of analysis. This will include exploring the better use of parallelization and possible deployment of graphics processing units. We also have plans to improve the image processing capabilities of CurveAlign by linking it to our Bio-Formats [56] and ImageJ efforts [32, 57, 58]. This need for image processing will also address a biological need to better track cells and tumor changes with the collagen alignment. This type of cell and fiber analysis can be used with machine learning approaches such as the WEKA toolkit in ImageJ [59]. We also are actively using and developing the CurveAlign platform to quantitate other properties of collagen organization such as length, width, straightness, and density. Another future development based on biological imaging need is adding support for three-dimensional quantitation of collagen. To date, the majority of the reported collagen analysis studies are of 2D sections out of 3D image stacks. As imaging work continues to characterize the role of collagen, there are likely to be studies that characterize collagen in

3D, including changes in collagen fiber between focal planes. We plan on adding true 3D support to our platform to accommodate such analysis. Additional feature requests can be made by any user at the CurveAlign website.

4. Notes

1. CFR mode is usually applicable to images without a large variation in contrast. Options in CT-FIRE mode are applicable to situations where individual fibers are most concerned and can be used to analyze images of any contrast. Importantly, the accuracy of the CT-FIRE mode is influenced by the complexity of fiber networks and not as efficient as the CFR mode (also *see* Subheading 3.5). CFR mode can be used to screen images with any type of collagen network.
2. There are default parameters for both CFR mode and CT-FIRE mode. Although most of those parameters are suitable for many images, optimal parameters can be set according to the image contrast, image size, fiber thickness distribution, and collagen density distribution as discussed in Subheading 3.5.
3. For CFR mode, the fraction of coefficient to keep (labeled as 9 in Fig. 4a) is usually the only parameter to be adjusted to represent the most prominent fibers. In general, the smaller the image size is, the larger this parameter is. For CT-FIRE mode, the .mat file generated by CT-FIRE must be in the subfolder "ctFIREout" of the original image. To launch the CT-FIRE module as shown in Fig. 4b, click the CT-FIRE button (labeled as 6 in Fig. 4a). After clicking on "Open File(s)" button (labeled as 3 in Fig. 4b) to load image, click the "Update" button (labeled as 4 in Fig. 4b) to confirm or update default parameters. Some practical tips to adjust the parameters can be seen in Subheading 3.5. Click on "RUN" button (labeled as 5 in Fig. 4b) to run CT-FIRE fiber extraction, and the progress information will be displayed in both the information window highlighted in green color at the bottom of Fig. 4b and in the command window. To run CT-FIRE on multiple images, first check on the "Batch" box (labeled as 1 in Fig. 4b), and optionally, check on the "Parallel" box (labeled as 2 in Fig. 4b) to enable parallel computing.
4. No boundaries: Fibers will be compared to each other, but not to a boundary.
5. CSV boundary is an open boundary and is only used for CT mode, and only one boundary can be defined in this boundary condition. Alternatively, the TIFF boundary condition can contain one or more boundaries and can replace the CSV boundary if it is carefully annotated and thus is a preferred option for most applications.
6. If ROI analysis is desired, after images are loaded, click on the "ROI manager" button (4 in Fig. 4a) to launch ROI manager window (Fig. 7a) to define an ROI for each image, update a previously defined ROI, or load ROI files defined in other tools (such as Fiji or ImageJ) in the format of mask image file or ".txt" file.
7. If the full image has already been analyzed by either CT-FIRE mode or CFR mode, choosing "ROI post-processing" will extract the features from the

processed output. (CT-FIRE mode only supports ROI post-processing in the current version).

8. If the full image has not been analyzed, CurveAlign can be applied to the specified ROI for analysis in CFR mode. If the ROI is of rectangular shape, which is the most commonly used shape, we recommend choosing the option "CA on cropped rectangular ROI." For all other shapes, choose the option "CA on mask with ROI of any shape." As described in Subheading 3.5, full image analysis in CFR mode tends to be biased toward fibers with higher intensity. ROI analysis in CFR mode can focus fiber alignment analysis on a smaller region with improved fiber representation.

Acknowledgments

We thank many users and contributors to CT-FIRE and CurveAlign over the years including Rob Nowak, Carolyn Pehlke, Jeremy Bredfeldt, Matt Conklin, Paolo Provenzano, Patricia Keely, and Jayne Squirrell. This work was supported by NIH grants R01CA179556 and U54DK104310 to K.W.E.

References

1. Lanir Y (1983) Constitutive equations for fibrous connective tissues. J Biomech 16(1):1–12. doi: 10.1016/0021-9290(83)90041-6 [PubMed: 6833305]
2. Martin RB, Lau ST, Mathews PV et al. (1996) Collagen fiber organization is related to mechanical properties and remodeling in equine bone. A comparison of two methods. J Biomech 29(12):1515–1521. doi:10.1016/S0021-9290(96)80002-9 [PubMed: 8945649]
3. Chen C-H, Cheng C-H.(1996) Effective elastic moduli of misoriented short-fiber composites. Int J Solids Struct 33(17):2519–2539. doi:10.1016/0020-7683(95)00160-3
4. Tian L, Wang Z, Liu Y et al. (2016) Validation of an arterial constitutive model accounting for collagen content and crosslinking. Acta Biomater 31:276–287. doi:10.1016/j.actbio.2015.11.058 [PubMed: 26654765]
5. Provenzano PP, Eliceiri KW, Campbell JM et al. (2006) Collagen reorganization at the tumorstromal interface facilitates local invasion. BMC Med 4(1):38. doi:10.1186/1741-7015-4-38 [PubMed: 17190588]
6. Conklin MW, Eickhoff JC, Riching KM et al. (2011) Aligned collagen is a prognostic signature for survival in human breast carcinoma. Am J Pathol 178(3):1221–1232. doi:10.1016/.ajpath.2010.11.076 [PubMed: 21356373]
7. Keikhosravi A, Bredfeldt JS, Sagar AK et al. (2014) Second-harmonic generation imaging of cancer. Methods Cell Biol 123:531–546. doi:10.1016/B978-0-12-420138-5.00028-8 [PubMed: 24974046]
8. Drifka CR, Eliceiri KW, Weber SM et al. (2013) A bioengineered heterotypic stroma-cancer microenvironment model to study pancreatic ductal adenocarcinoma. Lab Chip 13(19):3965–3975. doi:10.1039/c3lc50487e [PubMed: 23959166]
9. Drifka CR, Tod J, Loeffler AG et al. (2015) Periductal stromal collagen topology of pancreatic ductal adenocarcinoma differs from that of normal and chronic pancreatitis. Mod Pathol 28(11):1470–1480. doi:10.1038/modpathol.2015.97 [PubMed: 26336888]
10. Drifka CR, Loeffler AG, Mathewson K et al. (2016) Highly aligned stromal collagen is a negative prognostic factor following pancreatic ductal adenocarcinoma resection. Oncotarget 7(46):76197–76213. doi:10.18632/oncotarget.12772 [PubMed: 27776346]
11. Meijering E, Jacob M, Sarria JC et al. (2004) Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. Cytometry A 58(2):167–176. doi:10.1002/cyto.a.20022 [PubMed: 15057970]
12. Rubbens MP, Driessen-Mol A, Boerboom RA et al. (2009) Quantification of the temporal evolution of collagen orientation in mechanically conditioned engineered cardiovascular tissues. Ann Biomed Eng 37(7):1263–1272. doi:10.1007/s10439-009-9698-x [PubMed: 19415496]

13. Rezakhaniha R, Agianniotis A, Schrauwen JT et al. (2012) Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy. Biomech Model Mechanobiol 11(3– 4):461–473. doi:10.1007/s10237-011-0325-z [PubMed: 21744269]
14. Kabir MM, Inavalli VV, Lau TY et al. (2013) Application of quantitative second-harmonic generation microscopy to dynamic conditions. Biomed Opt Express 4(11):2546–2554. doi: 10.1364/B0E.4.002546 [PubMed: 24298415]
15. Boudaoud A, Burian A, Borowska-Wykret D et al. (2014) FibrilTool, an ImageJ plug-in to quantify fibrillar structures in raw microscopy images. Nat Protoc 9(2):457–463. doi:10.1038/nprot. 2014.024 [PubMed: 24481272]
16. Sun M, Bloom AB, Zaman MH (2015) Rapid quantification of 3D collagen fiber alignment and fiber intersection correlations with high sensitivity. PLoS One 10(7):e0131814. doi:10.1371/ journal.pone.0131814 [PubMed: 26158674]
17. Quinn KP, Georgakoudi I (2013) Rapid quantification of pixel-wise fiber orientation data in micrographs. J Biomed Opt 18(4):046003. doi:10.1117/1.JBO.18.4.046003 [PubMed: 23552635]
18. Chaudhuri S, Nguyen H, Rangayyan RM et al. (1987) A Fourier domain directional filtering method for analysis of collagen alignment in ligaments. IEEE Trans Biomed Eng 34(7):509–518. doi:10.1109/TBME.1987.325980 [PubMed: 3610201]
19. Pourdeyhimi B, Dent R, Davis H (1997) Measuring fiber orientation in nonwovens part III: fourier transform. Text Res J 67(2):143–151. doi:10.1177/004051759706700211
20. Marquez JP (2006) Fourier analysis and auto-mated measurement of cell and fiber angular orientation distributions. Int J Solids Struct 43(21):6413–6423. doi:10.1016/j.ijsolstr.2005.11.003
21. Sander EA, Barocas VH (2009) Comparison of 2D fiber network orientation measurement methods. J Biomed Mater Res A 88(2):322–331. doi:10.1002/jbm.a.31847 [PubMed: 18286605]
22. Sivaguru M, Durgam S, Ambekar R et al. (2010) Quantitative analysis of collagen fiber organization in injured tendons using Fourier transform-second harmonic generation imaging. Opt Express 18(24):24983–24993. doi:10.1364/OE.18.024983 [PubMed: 21164843]
23. Pourdeyhimi B, Kim HS (2002) Measuring fiber orientation in nonwovens: the hough transform. Text Res J 72(9):803–809. doi:10.1177/004051750207200909
24. Bayan C, Levitt JM, Miller E et al. (2009) Fully automated, quantitative, noninvasive assessment of collagen fiber content and organization in thick collagen gels. J Appl Phys 105(10):102042. doi: 10.1063/1.3116626 [PubMed: 24803683]
25. Schneider CA, Pehlke CA, Tilbury K et al. (2013) Quantitative approaches for studying the role of collagen in breast cancer invasion and progression In: Pavone FS, Campagnola PJ (eds) Second harmonic generation imaging. Taylor & Francis, Boca Raton, FL, pp 373–390
26. Wen BL, Brewer MA, Nadiarnykh O et al. (2014) Texture analysis applied to second harmonic generation image data for ovarian cancer classification. J Biomed Opt 19(9):096007. doi: 10.1117/1.JBO.19.9.096007 [PubMed: 26296156]
27. Puspoki Z, Storath M, Sage D et al. (2016) Transforms and operators for directional bioimage analysis: a survey. Adv Anat Embryol Cell Biol 219:69–93. doi:10.1007/978-3-319-28549-8_3 [PubMed: 27207363]
28. Mori S, van Zijl PC (2002) Fiber tracking: principles and strategies - a technical review. NMR Biomed 15(7–8):468–480. doi:10.1002/nbm.781 [PubMed: 12489096]
29. Stein AM, Vader DA, Jawerth LM et al. (2008) An algorithm for extracting the network geometry of three-dimensional collagen gels. J Microsc 232(3):463–475. doi:10.1111/j. 1365-2818.2008.02141.x [PubMed: 19094023]
30. Jacques E, Buytaert J, Wells DM et al. (2013) MicroFilament Analyzer, an image analysis tool for quantifying fibrillar orientation, reveals changes in microtubule organization during gravitropism. Plant J 74(6):1045–1058. doi:10.1111/tpj.12174 [PubMed: 23489480]
31. Bredfeldt JS, Liu Y, Conklin MW et al. (2014) Automated quantification of aligned collagen for human breast carcinoma prognosis. J Pathol Inform 5(1):28. doi:10.4103/2153-3539.139707 [PubMed: 25250186]
32. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9(7):671–675. doi:10.1038/nmeth.2089 [PubMed: 22930834]

33. BIG > OrientationJ. http://bigwww.epfl.ch/demo/orientation/#soft Accessed 23 Nov 2016
34. CytoSpectre: Spectral analysis of microscopy images. http://www.tut.fi/cytospectre/ Accessed 23 Nov 2016
35. Kartasalo K, Polonen RP, Ojala M et al. (2015) CytoSpectre: a tool for spectral analysis of oriented structures on cellular and subcellular levels. BMC Bioinformatics 16:344. doi: 10.1186/s12859-015-0782-y [PubMed: 26503371]
36. CurveAlign | LOCI. http://loci.wisc.edu/software/curvealign Accessed 26 Nov 2016
37. CT-FIRE | LOCI. http://loci.wisc.edu/software/ctfire Accessed 26 Nov 2016
38. Bredfeldt JS, Liu Y, Pehlke CA et al. (2014) Computational segmentation of collagen fibers from second-harmonic generation images of breast cancer. J Biomed Opt 19(1):16007. doi: 10.1117/1.JBO.19.1.016007 [PubMed: 24407500]
39. Provenzano PP, Eliceiri KW, Yan L et al. (2008) Nonlinear optical imaging of cellular processes in breast cancer. Microsc Microanal 14(6):532–548. doi:10.1017/S1431927608080884 [PubMed: 18986607]
40. Ajeti V, Nadiarnykh O, Ponik SM et al. (2011) Structural changes in mixed Col I/Col V collagen gels probed by SHG microscopy: implications for probing stromal alterations in human breast cancer. Biomed Opt Express 2(8):2307–2316. doi:10.1364/BOE.2.002307 [PubMed: 21833367]
41. Barcus CE, Keely PJ, Eliceiri KW et al. (2013) Stiff collagen matrices increase tumorigenic prolactin signaling in breast cancer cells. J Biol Chem 288(18): 12722–12732. doi:10.1074/jbc.M112.447631 [PubMed: 23530035]
42. Riching KM, Cox BL, Salick MR et al. (2014) 3D collagen alignment limits protrusions to enhance breast cancer cell persistence. Biophys J 107(11):2546–2558. doi:10.1016/j.bpj.2014.10.035 [PubMed: 25468334]
43. Bauman TM, Nicholson TM, Abler LL et al. (2014) Characterization of fibrillar collagens and extracellular matrix of glandular benign prostatic hyperplasia nodules. PLoS One 9(10):e109102. doi:10.1371/journal.pone.0109102 [PubMed: 25275645]
44. Barcus CE, Holt EC, Keely PJ et al. (2015) Dense collagen-I matrices enhance pro-tumorigenic estrogen-prolactin crosstalk in MCF-7 and T47D breast cancer cells. PLoS One 10(1):e0116891. doi:10.1371/journal.pone.0116891 [PubMed: 25607819]
45. Corsa CA, Brenot A, Grither WR et al. (2016) The action of discoidin domain receptor 2 in basal tumor cells and stromal cancer-associated fibroblasts is critical for breast cancer metastasis. Cell Rep 15(11):2510–2523. doi:10.1016/j.celrep.2016.05.033 [PubMed: 27264173]
46. Candes E, Demanet L, Donoho D et al. (2006) Fast discrete curvelet transforms. Multiscale Model Simul 5(3):861–899. doi:10.1137/05064182X
47. Walsh AJ, Cook RS, Lee JH et al. (2015) Collagen density and alignment in responsive and resistant trastuzumab-treated breast cancer xenografts. J Biomed Opt 20(2):26004. doi: 10.1117/1.JBO.20.2.026004 [PubMed: 25700233]
48. Junqueira LC, Bignolas G, Brentani RR (1979) Picrosirius staining plus polarization microscopy, a specific method for collagen detection in tissue sections. Histochem J 11(4):447–455. doi:10.1007/BF01002772 [PubMed: 91593]
49. Gillespie AR, Kahle AB, Walker RE (1986) Color enhancement of highly correlated images. I. Decorrelation and HSI contrast stretches. Remote Sens Environ 20(3):209–235. doi: 10.1016/0034-4257(86)90044-1
50. Campagnola PJ, Wei MD, Lewis A et al. (1999) High-resolution nonlinear optical imaging of live cells by second harmonic generation. Biophys J 77(6):3341–3349. doi:10.1016/S0006-3495(99)77165-1 [PubMed: 10585956]
51. Campagnola PJ, Millard AC, Terasaki M et al. (2002) Three-dimensional high-resolution second-harmonic generation imaging of endogenous structural proteins in biological tissues. Biophys J 82(1 Pt 1):493–508. doi:10.1016/S0006-3495(02)75414-3 [PubMed: 11751336]
52. Ayala G, Tuxhorn JA, Wheeler TM et al. (2003) Reactive stroma as a predictor of biochemical-free recurrence in prostate cancer. Clin Cancer Res 9(13):4792–4801 [PubMed: 14581350]
53. Allon I, Vered M, Buchner A et al. (2006) Stromal differences in salivary gland tumors of a common histopathogenesis but with different biological behavior: a study with picrosirius red and

polarizing microscopy. Acta Histochem 108(4):259–264. doi:10.1016/j.acthis.2006.05.007 [PubMed: 16899283]

54. Arun Gopinathan P, Kokila G, Jyothi M et al. (2015) Study of collagen birefringence in different grades of oral squamous cell carcinoma using picrosirius red and polarized light microscopy. Scientifica (Cairo) 2015:802980. doi:10.1155/2015/802980 [PubMed: 26587310]
55. Lindeberg T (1998) Edge detection and ridge detection with automatic scale selection. Int J Comput Vision 30(2):117–156. doi:10.1023/A:1008097225773
56. Linkert M, Rueden CT, Allan C et al. (2010) Metadata matters: access to image data in the real world. J Cell Biol 189(5):777–782. doi:10.1083/jcb.201004104 [PubMed: 20513764]
57. Schindelin J, Arganda-Carreras I, Frise E et al. (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9(7):676–682. doi:10.1038/nmeth.2019 [PubMed: 22743772]
58. Schindelin J, Rueden CT, Hiner MC et al. (2015) The ImageJ ecosystem: an open platform for biomedical image analysis. Mol Reprod Dev 82(7–8):518–529. doi:10.1002/mrd.22489 [PubMed: 26153368]
59. Arganda-Carreras I, Kaynig V, Eliceiri KW et al. (2016) Trainable weka segmentation: a machine learning tool for microscopy pixel classification. J Bioinformatics. Submitted

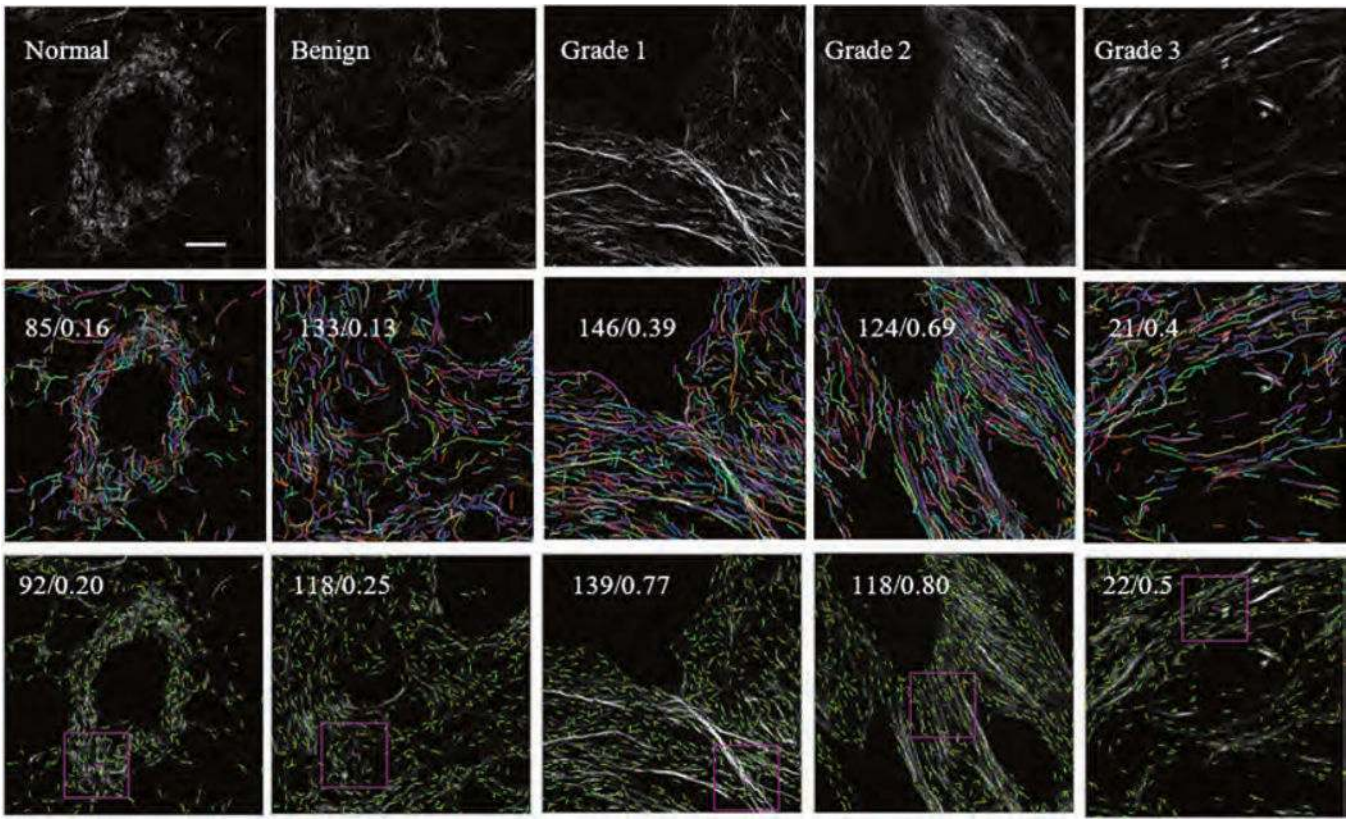

Fig. 1.

An example of using CT-FIRE mode without the boundary mode of CurveAlign for both full image and ROI analyses in pancreatic tissue. The *top row* shows the SHG image from different types of pancreatic tissue. The *middle row* shows the CT-FIRE extracted fibers highlighted by different colored lines for the purpose of contrast, which are overlaid upon the SHG image. The *bottom row* is the "Overlay" image generated by CurveAlign showing the location *(red dots)* and orientation *(green lines)* of the extracted fibers as well as the annotated ROI overlaid upon the SHG image. The quantification of the full image and the representative ROI are displayed in the format of "mean orientation/alignment" at the top left corner of the *middle row* and *bottom row,* respectively. Both full image-based and ROI-based quantifications show that collagen fibers are more aligned in malignant ducts (Grade 1–3) than in normal and benign ducts. Default parameters were used here. Image size is 1024 × 1024 pixels, and ROI size is 256 × 256 pixels. The scale bar at the top left column indicates 50 μm, and all images are shown at the same scale

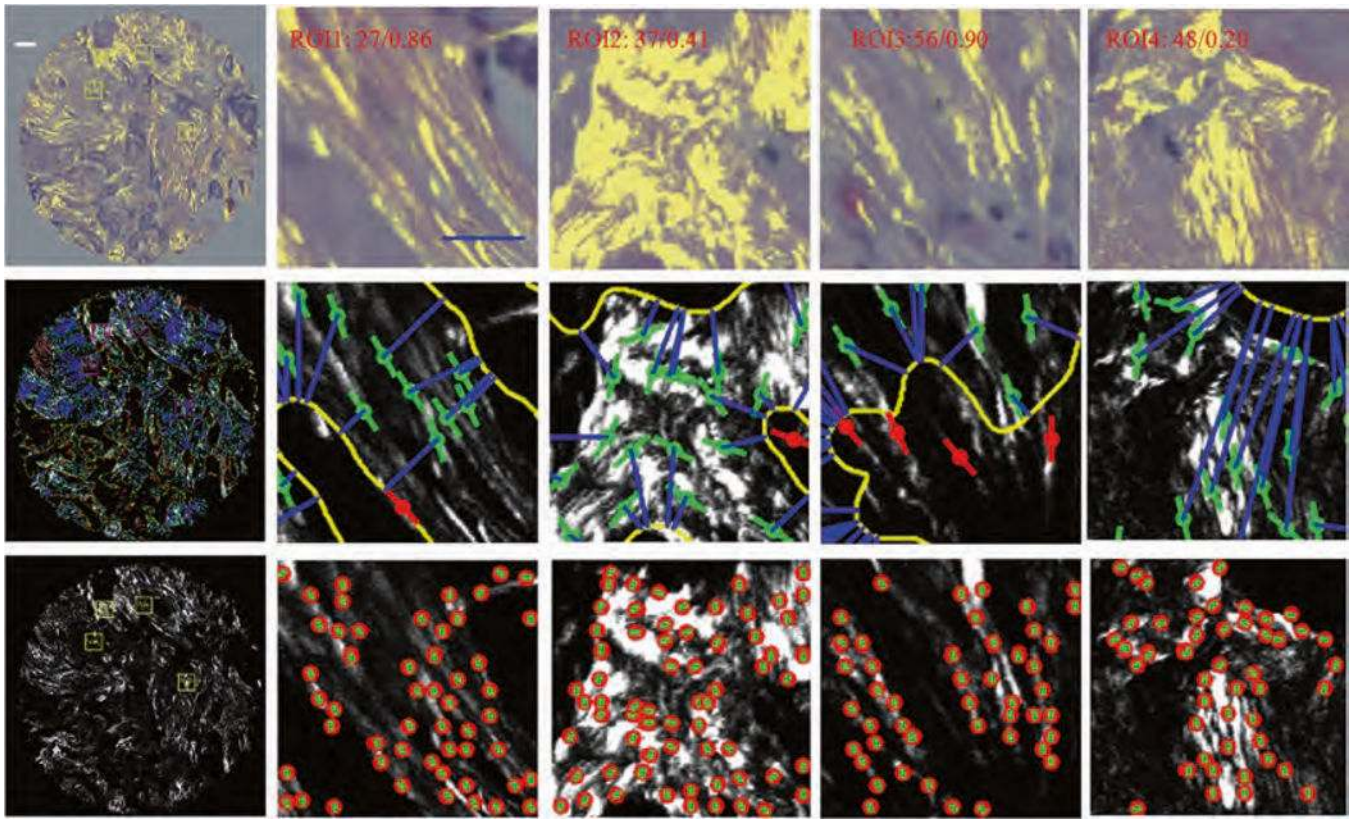

Fig. 2.

An example of using CurveAlign to quantify relative angle and ROI-based alignment on breast cancer tissue microarray. The first column shows the SHG image (*in yellow*) overlaid on the bright-field H&E image with four rectangular ROI annotations *(in yellow) (top),* the orientations and locations of extracted fibers using CT-FIRE mode *(in green color)* with four ROI annotations *(in magenta)* overlaid on the SHG image (in *gray) (middle),* and four rectangular ROI annotations *(in yellow)* overlaid on the SHG image *(in gray) (bottom),* respectively. The remaining four columns from *left* to *right* show the cropped images of the four ROIs, where for the ones in the top row, ROI name as well as its mean relative angle and ROI-based alignment in the format of "ROI name: mean relative angle/ROI alignment." In the middle row, the *blue line* is used to associate the center of fiber extracted by CT-FIRE with the corresponding boundary locations, and the *red line* shows the fibers located either beyond the distance range or within the boundary. In the bottom row, the locations and orientations of fiber segments analyzed by CFR mode were indicated by *red dots* and *green lines.* From the figure, we can see that fibers in ROI 1 and ROI 3 are both well aligned, but there are more fibers in ROI 3 that are more perpendicularly aligned with respect to the tumor boundary. The boundary creation parameter is the same as in Fig. 6, i.e., pixel per micron ratio was set to 2; fiber extraction in CT-FIRE used default settings; the distance parameter was set to 250 pixels; the fraction of kept curvelets is set to 0.06. The image size is 2048 × 2048 pixels, and the sizes of ROI are identically set to 128 × 128 pixels. The scale bar for the full size image at the first column of the first row indicates 100 μm, scale bar for

the ROI image at the second column of the first row indicates 25 μm, and all images either full size images or ROIs are shown at the same scale

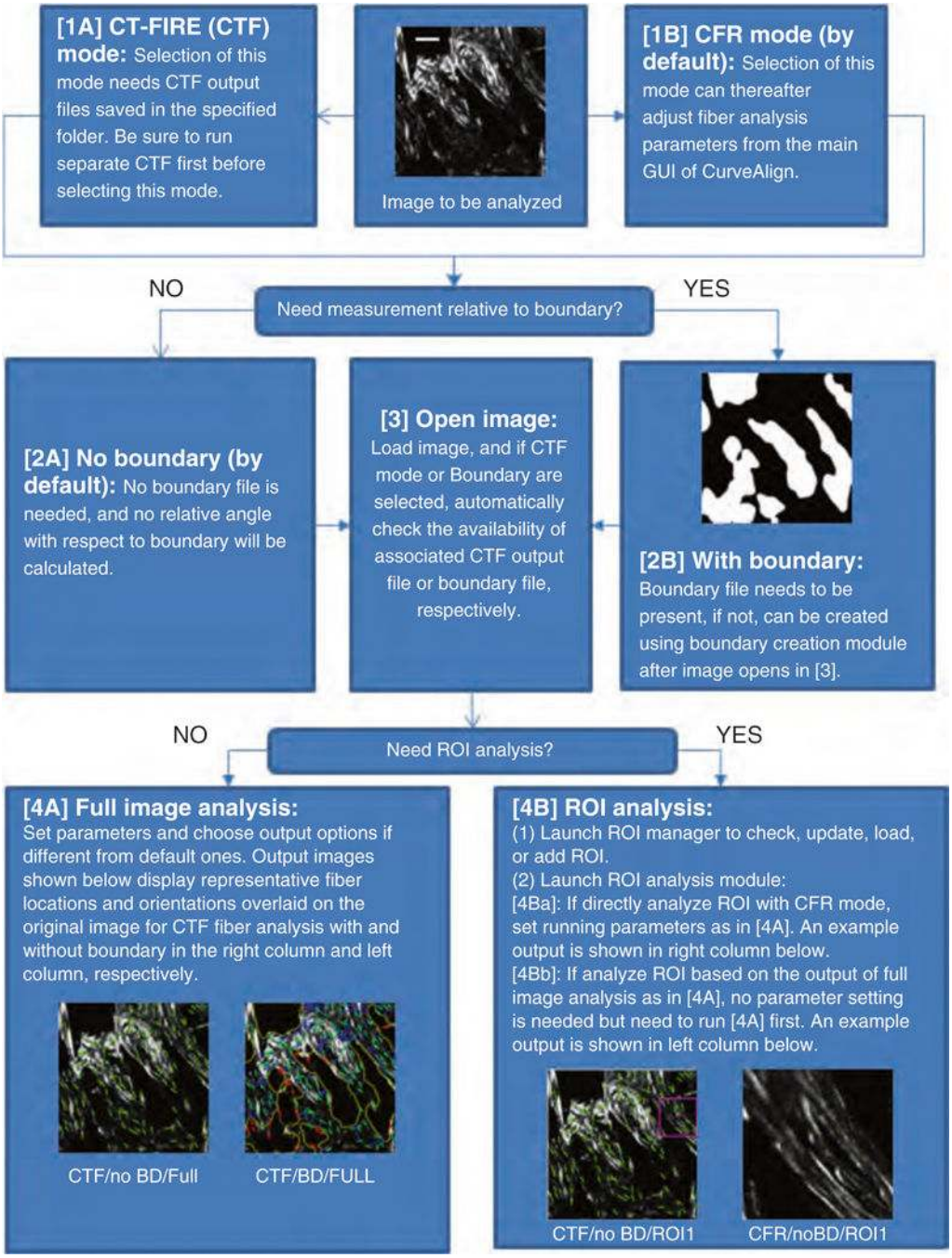

Fig. 3.

Schematic workflow of quantifying collagen alignment using CurveAlign for a single image. For the images in step 4, the fiber location and orientation were highlighted by *red dots* and *green lines,* respectively; the *red lines* indicate the fibers are out of the interested region. For [4B], there are two types of ROI analysis, type 1 ([4Ba]) can directly run CFR mode on the annotated ROI, but type 2 [4Bb] needs to run a full image analysis first either in CFR mode or in CT-FIRE mode. CTF, CT-FIRE; CFR, curvelets fiber representation; BD, with boundary; no BD, without boundary; FULL, full image analysis; and ROI, ROI analysis.

The scale bar shown in the image at the *top middle column* indicates 50 μm, and all images are shown at the same scale

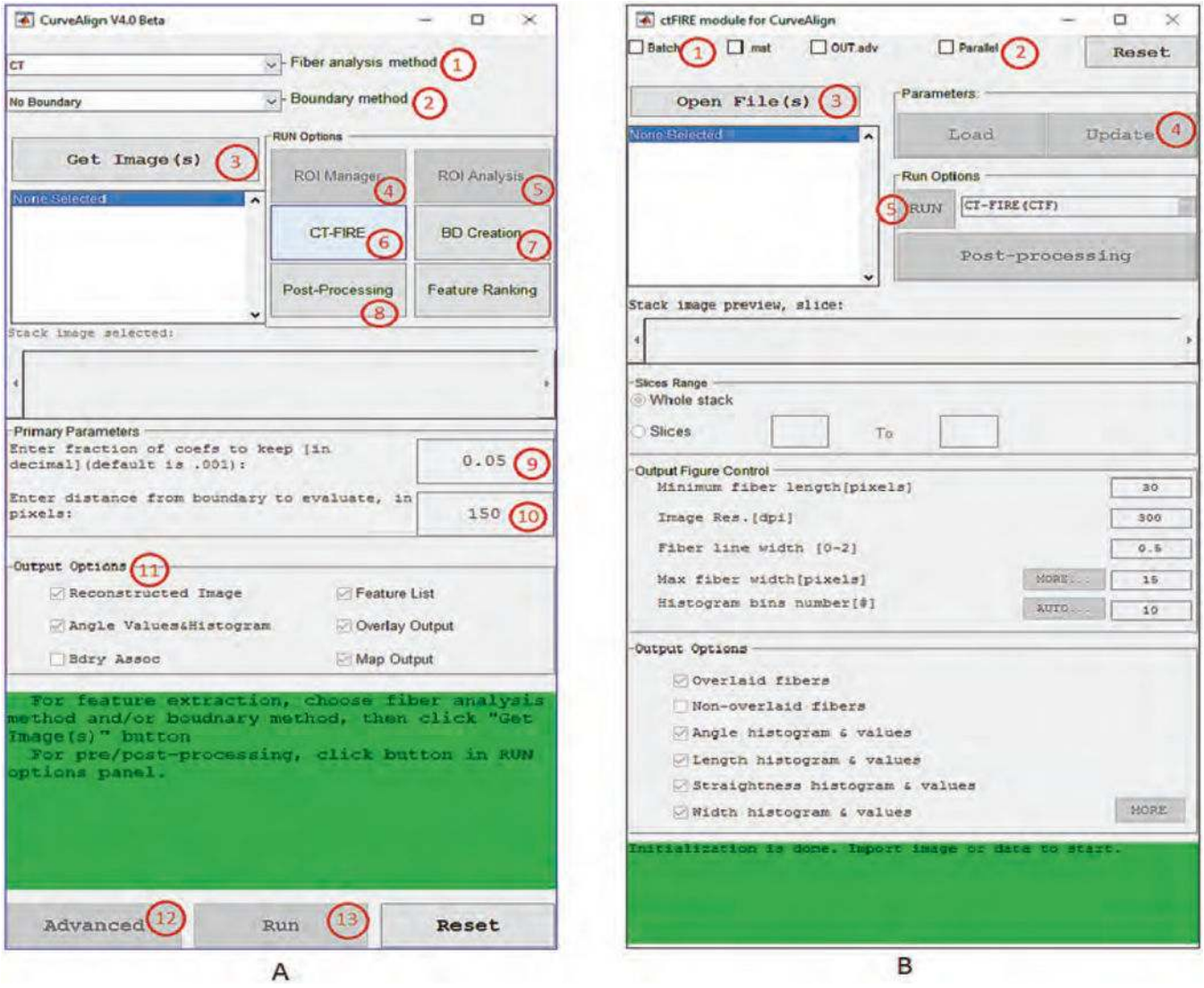

Fig. 4.

Graphical user interface of CurveAlign (**a**) and CT-FIRE (**b**)

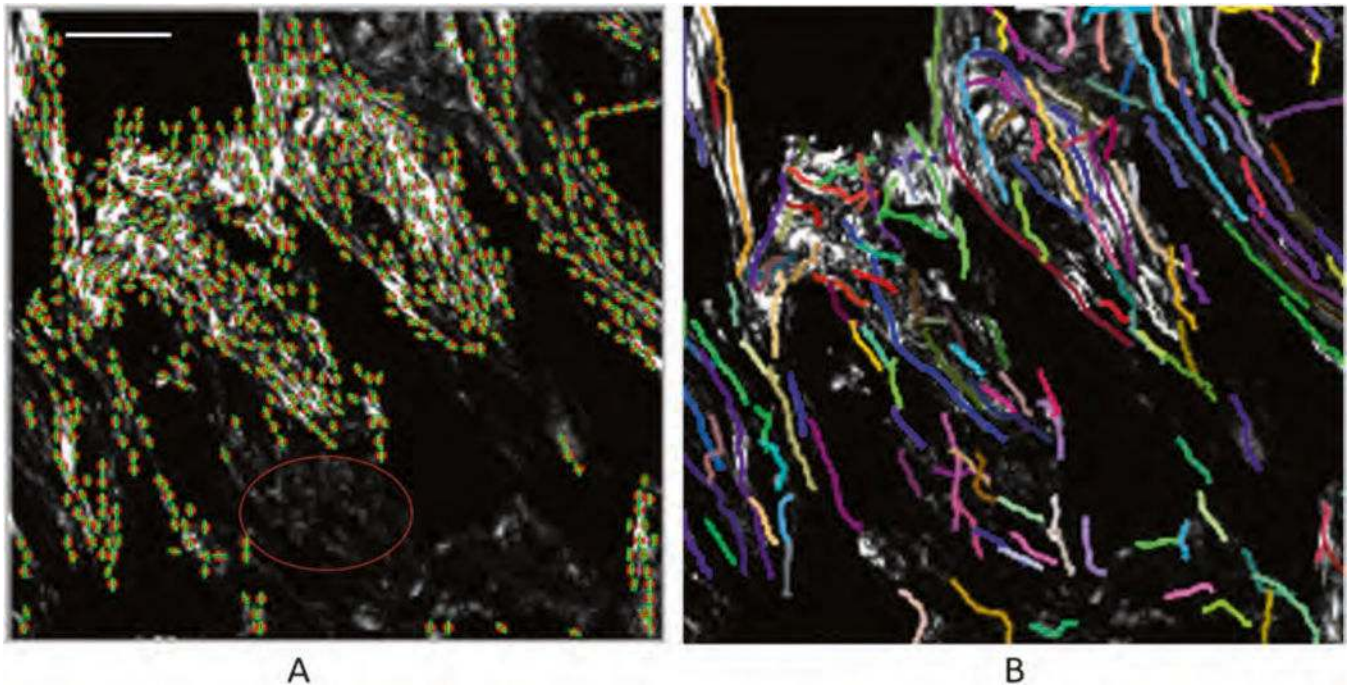

Fig. 5.

Two fiber analysis modes in CurveAlign: (**a**) representative curvelets used in CFR mode; (**b**) extracted individual fibers used in CT-FIRE. This image of a relatively low-resolution portion of a breast tissue microarray illustrates the ability of the CurveAlign platform to extract fibers even when the fibers are relatively low signal to noise ratio but also demonstrates the difference between the two modes in fiber extraction. In (**a**), the *red dots* show the fiber location and green lines show the orientation. In (**b**), extracted fibers are displayed in different colored lines. CFR mode (**a**) indicates the most prominent fibers but does not provide individual fiber extraction as in CT-FIRE mode (**b**), which instead provides the information for entire fibers. CFR mode missed some fibers in the region annotated in *red ellipse* in (**a**) where fibers have relative lower intensity values. For the fiber analysis here, all parameters were kept as default ones except that, in CFR mode, the fraction of kept curvelet was set to 0.06. The image size is 512 pixels x 512 pixels. The scale bar at the *upper left* corner of (**a**) indicates 50 μm, and all images are shown at the same scale

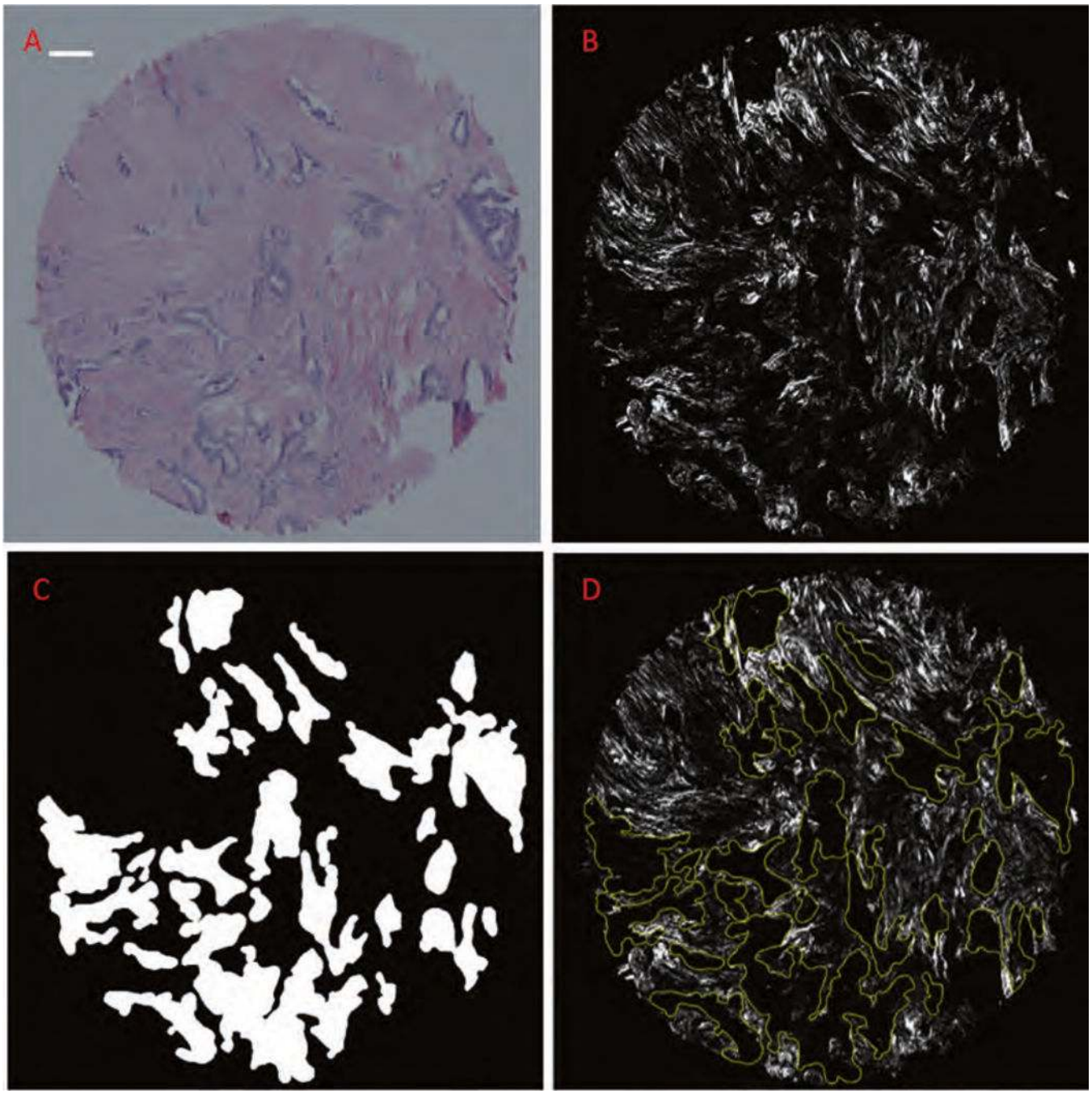

Fig. 6.

Automatic tumor boundary creation for a breast cancer TMA core. The bright-field H&E image (**a**) is coregistered with the SHG image (**b**), and the boundary mask image (**c**) is created from segmentation of the H&E image. The boundaries were detected and overlaid on the SHG image (**d**). The only necessary parameter to be input for the segmentation is the pixel per micron ratio of the image (set to 2 in this image). The scale bar at the *upper left* corner of (**a**) indicates 100 μm, and all images are shown at the same scale

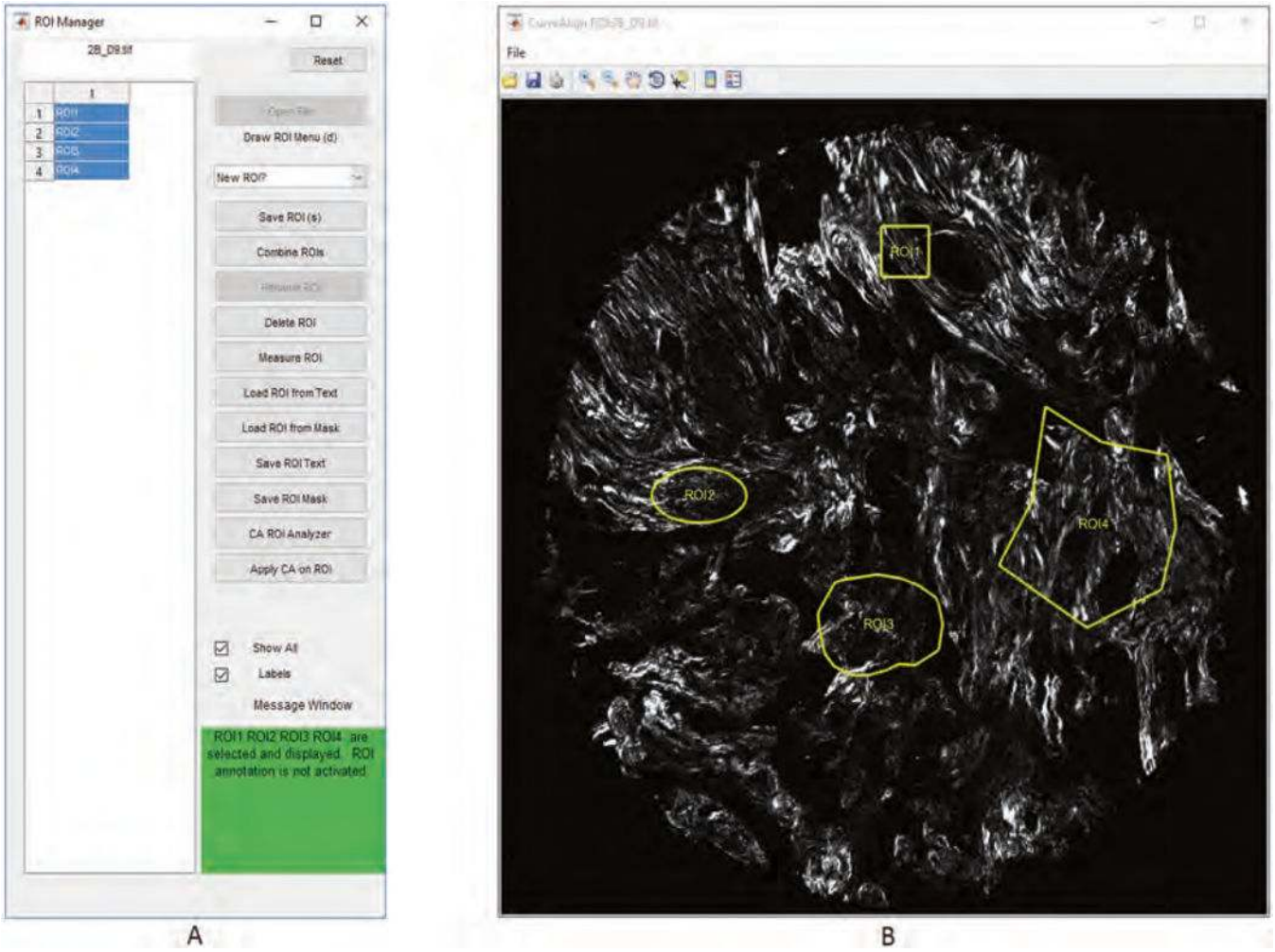

Fig. 7.

Screenshot of the ROI manager being used to annotate different shapes of ROIs. (**a**) ROI manager user interface; and (**b**) ROI output figure

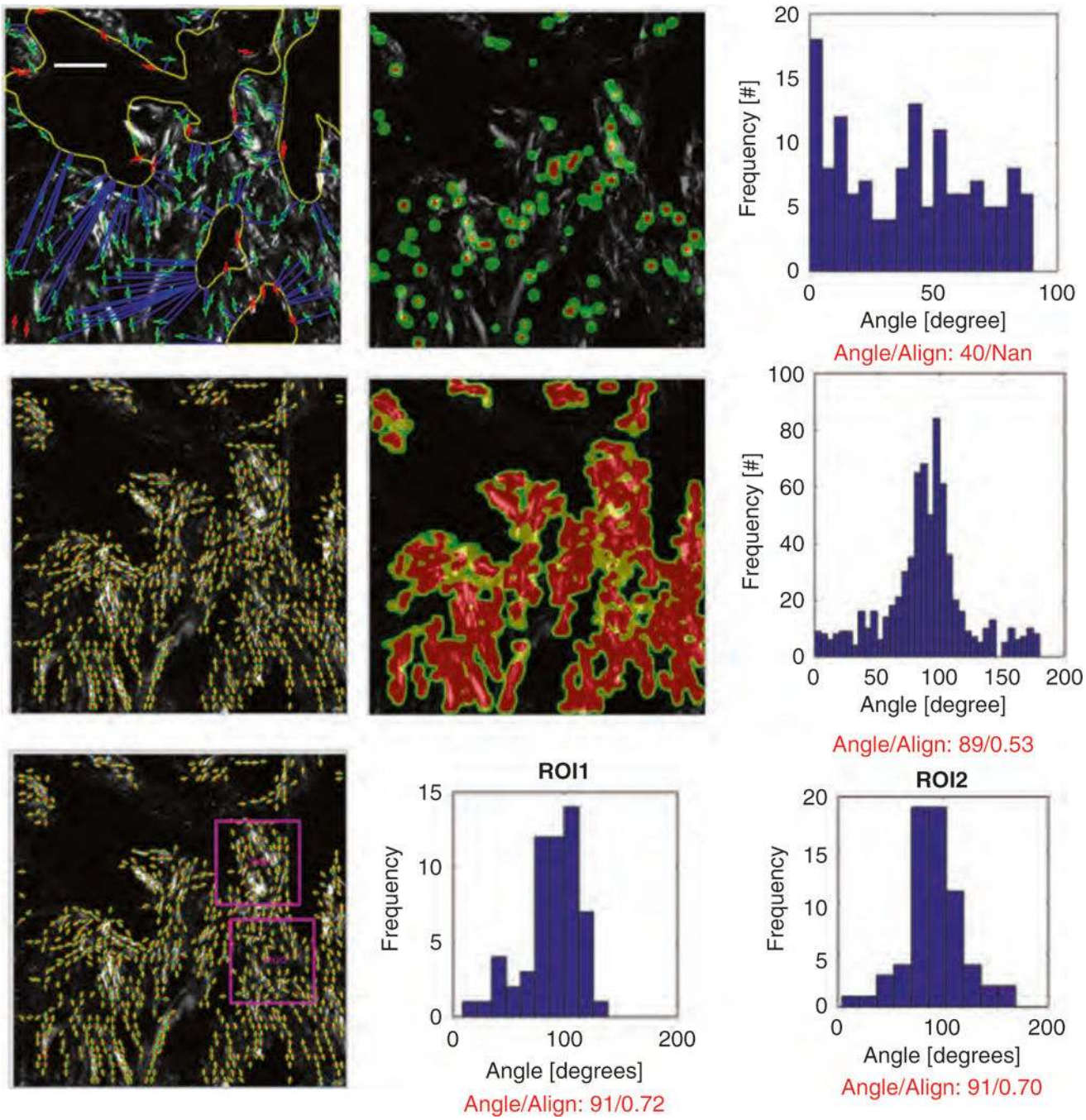

Fig. 8.

Demonstration of typical outputs of CurveAlign: (1) The *top row* shows the relative angle (with respect to a boundary) using fibers extracted in CT-FIRE mode, the *left column* of this row shows that the *blue line* is used to associate the center of the fiber extracted by CT-FIRE with the corresponding boundary location, and the *red line* is used to show the fibers located either beyond the distance range or within the boundary; the heatmap at the *middle column* uses *red color* to indicate larger relative angles, and the histogram at the *right column* shows the distribution of the relative angle of fibers within the range of 0 to 90°; (2) the *middle row*

shows the angle (with respect to the horizontal) measurement with CFR mode, and the *left column* of this row uses a *red dot* to indicate the center of fiber segments and a *green line* to indicate the fiber orientation at that point; the heatmap at the *middle column* uses *red color* to indicate well aligned fiber regions, and the histogram at the *right column* shows the distribution of the angles within the range of 0 to 180°; and (3) the *bottom row* shows ROI analysis based on the result from the CFR results shown in the *middle row.* Two magenta rectangular ROIs are annotated on the SHG image at the first column of this row, the histograms of angles within these two ROIs are displayed at the *middle* and *right columns.* Both the mean angle and alignment coefficient are listed below each histogram except that alignment is not calculated based on the relative angle data. The boundaries were cropped from the mask image shown in Fig. 6; fiber extraction in CT-FIRE used default settings; the distance parameter was set to 250 pixels; the fraction of kept curvelets is set to 0.06, the image size is 512 × 512 pixels, and the sizes of ROI are identically set to 128 × 128 pixels. The scale bar at the *upper left* corner of the *top row* indicates 50 μm, and all images are shown at the same scale