CNC-SMART TEM image analysis tutorial

# To download code (.m file)

1. Go to the GitHub repository: <https://github.com/seyucel/CNC-SMART> and download ZIP via the ‘Code’ button. (Fig 1)
2. Zip file includes two .m files:

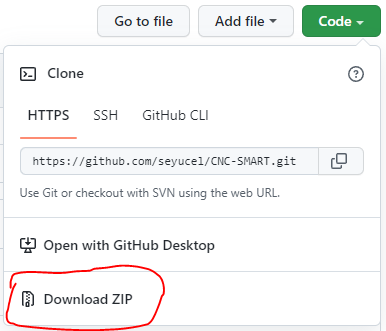


Figure 1 Download ZIP

* 1. ***SMART\_TEM\_measurements.m*** file is used to obtain particle size measurements without any visualization or further CNC identification verification analysis options. (**Caution**: black box!!)
  2. ***SMART\_TEM\_analysis.m*** file is a more interactive code where the user can check (and visualize) each step of the SMART analysis.

1. Unzip the ‘CNC-SMART-TEM.m’ file you wish to use and move it either to desktop or to the folder containing the TEM images to be analyzed.
2. You will need MATLAB software license (2018 or more recent versions) to run the CNC-SMART.

# SMART Analysis: A Seven Step Process

1. The SMART image analysis is a process that includes 7 steps as illustrated in the flow chart in fig 2. Each step is necessary for successfully applying SMART to TEM image analysis of CNCs. The description of how to use SMART follows this flow chart, and provides specific details and examples, and highlights aspects to watch out for (noted with “CAUTION”).

A picture containing graphical user interface

Description automatically generated

Figure 2 SMART seven step process for image analysis

1. Practical advice when using SMART:
   1. When using SMART for the first time or analyzing new TEM image data set, it is strongly recommended to start with a single image that is typical of your image data set. This allows you to do a pre-run with SMART through all 7 steps and you can learn some of the nuances of the parameter settings and how they affect segmentation, grouping, CNC identification and measurement. This also allows one to adjust SMART parameters, and complete visual inspections of segmentation, grouping, CNC identification and measurement, so that one can gain some level of confidence that the program and the parameters used are appropriate for analysis of your images.
   2. This process should then be repeated with a batch of 10 images, and we strongly recommend examining ALL the visulations of segmentation, grouping and CNC identification for all 10 images so that you have a better understanding of whether SMART is correctly identifying individual CNCs and what it is missing (see papers Yucel 2021 and 2022 for more details). Addjust parameters to try to improve the number of correctly identified CNCs.
   3. CAUTION: Even if you do the 7 steps perfectly, and step i & ii above, SMART analys is not perfect as discussed in Yucel 2022. Even with ideal TEM images and parameter settings, SMART will miss some fraction of isolated CNCs, and it may also incorrectly identify fragments as CNCs and cut off the wispy ends of CNCs or “tails”. As long as the ratio of correctly identifed CNCs to incorrect objects remains high, SMART measurements have been shown to be as good as manual measurements. However, it is prudent to have some level of skepticism on the results that SMART reports.
   4. SMART can generage many files: i) Images for visulation of segmentation, object identification, and object grouping, ii) Excel file with CNC measurements (length, width, aspect ratio, area fraction), iii) excel file with cumulative length and width measurement for assessing critical number of measurments for reliable measurement, and iv) for batch image processing plot histograms for length, width, aspect ratio. All files are saved within a subfolder (‘\results’) which is created by SMART automatically before saving the files.

# Data Preparation: Key points for batch image processing

Before doing anything with the MATLAB code, be sure that all the TEM images meet the following:

1. All TEM images are in the same folder.
2. All TEM images are the same size (e.g. 2048 x 2048 pixels).
3. All TEM images are recorded with the same magnification, which ensures the pixel length is the same for all images.
   1. Magnification defines the pixel length (e.g., nm/pixel)
   2. Pixel length information can be obtained from the imaging software during or the saved images.
   3. The magnification used for TEM image collection should be adjusted to give a pixel size compatible with the size of the CNCs that are measured. For example, with a pixel dimension of 0.30 nm, a 1-pixel error in measuring width will lead to an uncertainty of 1.5% for the width of a 6-nm particle. For wood-based CNCs, a pixel size of ≤ 0.3 nm is recommended. [Meija et al. 2020, <https://doi.org/10.1021/acs.analchem.0c02805> ].
   4. Note that the pixel length value will be entered at the beginning of the SMART code.
4. All TEM images should have the following; (see papers Yucel et al. 2021, 2022):
   1. Well-dispersed CNCs and few aggregates
   2. High contrast between the background regions and CNC (ie., lighter CNCs and darker backgrounds), the key is to have high contrast along the CNC edges.
   3. Low background noise (ie., background pixels should have low scatter of grayscale values)

# Run and Upload images: Single or batch image processing

Click on ‘Run’ button to run the entire code. (Fig. 3)

Graphical user interface, application

Description automatically generated

Figure 3 Run button

1. After clicking on the ‘run’ button, next step is to select images to be analyzed. If the ‘CNC-SMART-TEM.m’ file is on the desktop, an image selection window will automatically open up, then search for the folder where you have your image data set. If the ‘CNC-SMART-TEM.m’ file is located within the folder containing the image data set an image selection window will automatically appear (fig. 4) from which you can select which images are to be analyzed:

Graphical user interface, application, PowerPoint

Description automatically generated

Figure 4 Image selection

1. Select and upload the image(s) to be analyzed.

The user can either analyze a single image or a batch of images. If just starting out with a new set of TEM images, it is suggested to do the steps i and ii, first. Alternatively, if you already have confirmed that the current SMART parameters can analyze your images property, then it is fine to start with batch image analysis (step iii).

1. First run: When using SMART for the first time or analyzing a new TEM image data set, it is strongly recommended to upload and analyze a single image and do a pre-run with SMART. This will save computation time while you are working on the initial adjustments to the SMART parameters so that the program can analyze your images correctly.
2. Second run: analyze more TEM images, a batch of 10 images is recommended. Run SMART and then verify that the parameters identified in the single image analysis work. If they do not, go back and adjust parameters and rerun the analyses. Repeat this step as needed.
3. Batch image analysis: Batch image analysis happens when you select multiple images from the folder and click “open” (see fig 4.). **CAUTION:** The computer RAM and storage will dictate how many images can be analyzed per batch. Thus, it is suggested to be mindful of how many images you select as the system may slow down or crash if too many images are selected. It may be worthwhile to systematically assess the number of images that can be analyzed in a batch, be it 10, 20, 50, 100, etc..
4. Once you click “open” in the upload window... (see fig 4) the SMART program stars pre-processing.

# Pixel Length: Ensures proper calibration

Before running the code the pixel length information is required. The code will not start if pixel length information is not given.

1. **CAUTION**: If the pixel length is incorrect, all aspect of SMART analysis will be erroneous. Everything is based on pixel length (from image pre-processing, CNC-selection criteria, to the final CNC measurements).
2. Enter the pixel length information (input box will pop up) in nanometers. (Fig.5)

Graphical user interface, application

Description automatically generated

Figure 5 Pixel length info input box

# Pre-processing check: Segmentation and image quality

Once pixel length information is entered by the user, the SMART program starts pre-processing automatically, by smoothing, sharpening and binarizing (segmentation) the images. The smoothing and sharpening parameters can be adjusted from 1 to 5. Initially SMART has these preset to 5 and 1, for smoothing and sharpening (to minimize the background noise effect), respectively, which is based on TEM images of wood CNCs as reported in Yucel 2021, 2022. So, your initial run will use these parameters (lines 12-13 in the code).

1. Visualization of the first segmented image will be displayed (Fig. 6). Using this segmented image SMART calculates the contrast and the background noise in terms of grayscale intensity values (i.e., background noise is quantified as the standard deviation of the background pixels’ intensity values where a pixel can have values between 0-255 and contrast is quantified as the difference between the mean of CNC pixel intensity values and the mean of background pixel values).
2. Check the command window for the warning message ‘WARNING: noise or contrast might affect the particle identification and measurements.’ (the message appears if the noise value is above 40 out of 255).
   1. If the message appears, a higher smoothing parameter and a lower sharpening parameter might give better segmentation results, but still check the segmented image for the step c).
   2. If the message doesn’t appear, the segmentation of the first image can be good enough to continue, but still check the segmented image for the step c).
3. Then, the code will stop and ask for your approval to apply the same smoothing and sharpening parameters for the rest of the images (‘Yes’ option) or to modify the parameters (‘No’ option):
   1. If the segmentation results are good enough (ie, most of the CNC pixels are white and most of the background pixels are black), click on ‘Yes’ and the code will proceed to the next step on the flowchart (e.g. Grouping).
   2. If the pixels are not properly segmented into 2 phases, click on ‘No’ and try to adjust the image processing parameters (increase l\_smooth and/or decrease l\_sharp on lines 12-13 in the code).

Graphical user interface, application

Description automatically generated

Figure 6. Visualization of the first segmented image, with the segmentation quality check notification overlayed.

1. Batch image analysis: When selecting “yes” in part c, SMART will automatically segment all the images in the batch and save these segmented images in the ‘results/image\_analysis\_figures’ folder. It is suggested that you should check some of these files and see if the segmentation has occurred properly. If not then go back to part c, ii, which will allow you to try different parameter combinations (increase lsmooth and/or decrease lsharp on lines 12-13 in the code).
2. When the segmentation process is completed for all the images, the noise, contrast, and the CNC area fraction values of each image is stored as the variables called noise\_std, contrast, and CNC\_percentage.
3. Example for user: Fig 7 shown below is an example of improved segmentation results by changing the image processing parameters. It is typical that segmentation will not be perfect, the goal is just to be “better”. The middle image is obtained when smoothing parameter was 1, while the right image is obtained when smoothing parameter was 5. The image on the right is considered “better” as it has less background noise captured, either as groups, or as add objects attached to the CNC making them wider, longer, and more jagged perimeter.

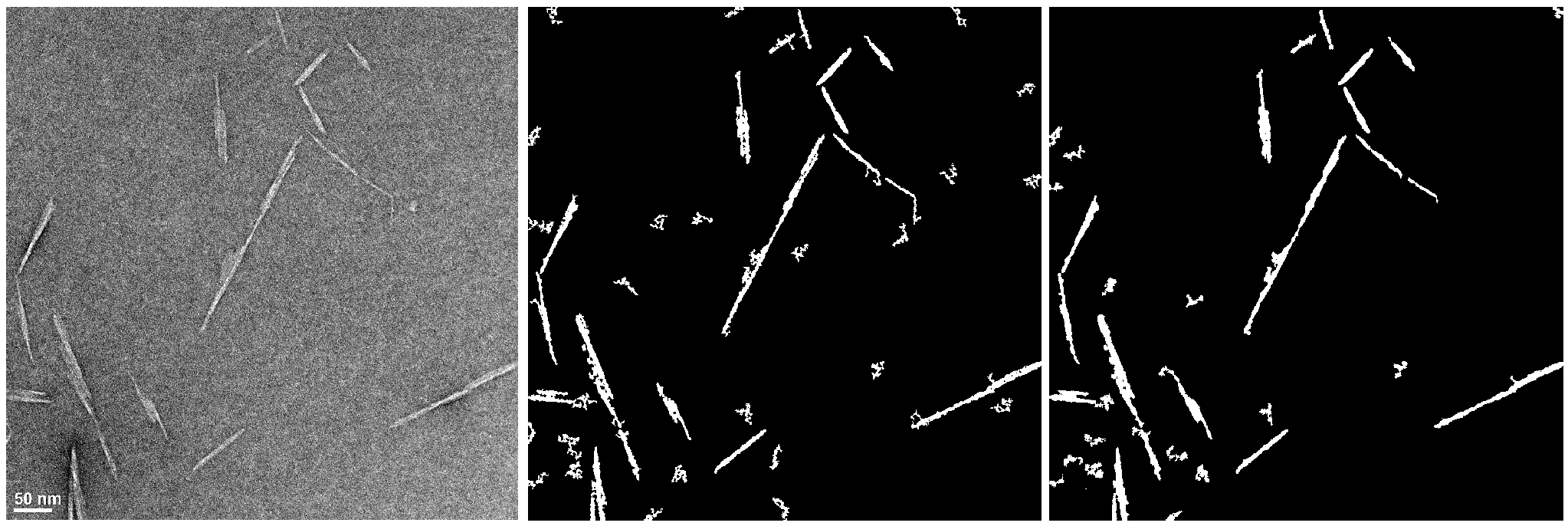


Figure 7. Example of the influence of changes in smoothing parameter on the resulting segmentation. Left is raw image, middle is with Smoothing =1, and right is with Smoothing =5.

# Grouping

Once pre-processing segmentation of all images is completed, CNC grouping and identification function is initiated. SMART then automatically groups every object on the segmented images either as “individual”, “border” or “aggregates”. Border identification is based on Matlab command, while individual and aggregates objects are based on a size-based criteria.

Individual CNC objects are identified via a 2 level selection process. In the first level SMART fits ellipses around each object and calculates the major and minor axis lengths for each fitted ellipse. These limits also can be modified by the user if needed (edit lines 130-132 in the code). Your initial run will use this size-based criteria listed below, which is based on TEM images of wood CNCs as reported in Yucel 2021, 2022:

* + max\_minor parameter sets the limit for the thickest particle selection. (<15 nm)
  + max\_major and min\_major parameters control the particle length. (<300 nm and >15 nm, respectively)
  + ar\_min parameter sets the limit for the smallest aspect ratio. (>2.5)

SMART further analyzes the first level selected individual particles by completing a second level refinement to detect stepped-like features, which would be indicative of parallel stacked CNCs. These “**stepped**” particles are considered as aggregates. The selection criteria, which is based on TEM images of wood CNCs as reported in Yucel 2021, 2022.:

* A specific width ratio, the ratio of the maximum width over the average width (e.g., wratio = wmax/wave), is defined (wratio > 1.5) and used to remove parallel stacked (275th line in the code).

1. Visualization of the grouped image will be displayed only for the first image (fig 8).

A picture containing text

Description automatically generatedGraphical user interface, application

Description automatically generated

Figure 8. Visualization of the first grouped image, with the grouping quality check notification. The red objects are isolated CNCs, the yellow objects are agglomerated CNCs, while the blue objects are boarder CNCs.

1. If the user approves, the same selection criteria will be applied to the remaining images.
2. Batch image analysis: When selecting “yes” in part a, SMART will automatically group all the images in the batch and save these grouped images in the ‘results/image\_analysis\_figures’ folder. It is suggested that you should check some of these files and see if the group has occurred properly. If not, then go back to lines 130-132 in the code and adjust selection criteria ranges.
3. To perform intentional identification of CNC aggregates it is possible to activate relevant sections in the code. If the size information of parallel stacked CNCs (which were identified using the width-based selection explained above) is needed, the section in the code between lines 450 and 460 should be commented off by removing the ‘%’ sign in front of the lines. If the size information of thicker aggregates (which were removed in the level 1 selection step) is needed, the user needs to specify the new selection criteria first (minimum aspect ratio – ar\_min, maximum and minimum thickness – max\_minor and min\_minor) in the lines 465-467. Once the criteria is defined, the section of code between 462 to 480 should be commented off by removing the ‘%’ sing in front of the lines. Yucel 2022 article presents a relevant case study about the size information of additional CNC aggregates.

# CNC identification and measurements

Once grouping of all images is completed, single CNC identification and measurement function is initiated. SMART then automatically encircles the perimeter of every SMART identified CNC in the images and obtains the morphology information of each CNC.

1. Visualization of the image with the identified and numbered CNCs will be displayed only for the first image (fig 9). The remaining identification figures won’t be displayed but they all will be saved in the ‘results/image\_analysis\_figures’ folder. To turn off the numbering, the 310th line in the code should be commented by putting a “%” sign in the beginning of the line.

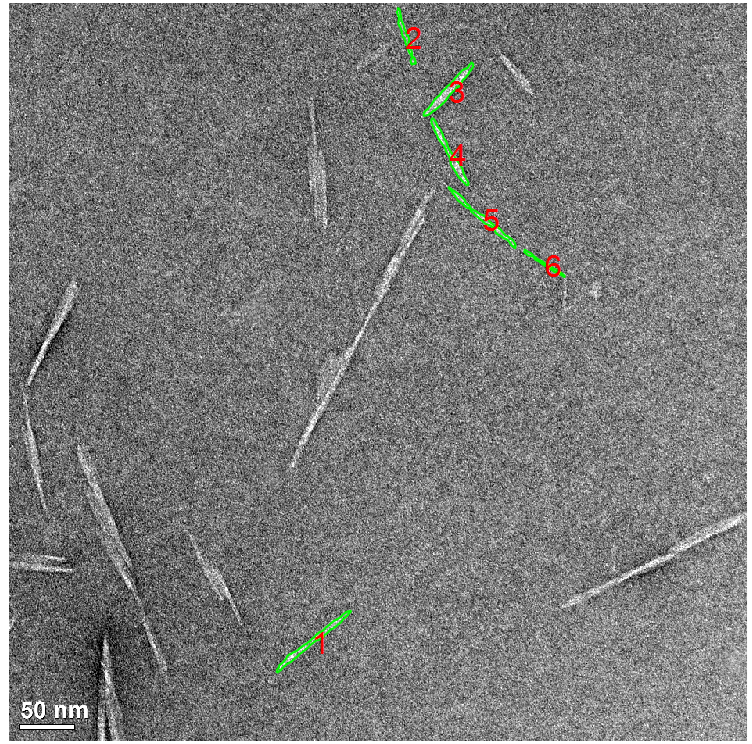


Figure 9. SMART encircled perimeters of every SMART identified CNC, with CNC identification label on. SMART identified 6 CNCs in this image.

1. SMART automatically measures the length and width of each individual CNC identified and creates an excel file to store the length, width and aspect ratio value of each CNC. The excel file has the name of each image (column A) and each identified CNC’s measurements ordered by their numbers (column B). The length and the width measurements of each CNC will be in the column C and column D, respectively. In sheet2, the area fraction data for each grouping (e.g., individual CNCs, board CNCs, and grouped CNCs) is summarized.

# Visualizing and saving results

Once single CNC identification and measurement process is completed SMART saves 2 excel files in the results folder (fig 10). One excel file contains the length, width, aspect ratio measurements in sheet1 and area fraction data in sheet2. The second excel file contains the cumulative mean length and width data for representative measurement analysis).

Graphical user interface, application, Word

Description automatically generated

Figure 10 results folder organization

1. SMART plots 1D probability histograms for length, width, aspect ratio measurements and 2D probability histograms for length and width measurements (figure 11) if the total number of measured CNCs for the batch is more than 40 CNCs. These plots will be saved in “results/measurement\_figures” folder. Graphical user interface

   Description automatically generated

Figure 11 1d and 2d histogram plots

1. SMART performs an additional analysis by calculating cumulative mean length and width measurements with each new CNC measured if the total number of measured CNCs is higher than 40. Then, the percentage change in overall mean (i.e., mean of the total number of measured CNCs) for each cumulative measurement is calculated. Once, the analysis is completed SMART plot the figures for the representative measurement analysis (Fig. 12) and saves it in “results/measurement\_figures” folder.

A picture containing graphical user interface

Description automatically generated

Figure 12 Cumulative mean length and width measurements