Cellulose (2021) 28:2183–2201
https://doi.org/10.1007/s10570-020-03668-8

ORIGINAL RESEARCH

# Semi-automatic image analysis of particle morphology of cellulose nanocrystals

**Sezen Yucel · Robert J. Moon · Linda J. Johnston · Berkay Yucel · Surya R. Kalidindi**

Received: 3 October 2020 / Accepted: 24 December 2020 / Published online: 21 January 2021


**Abstract** Morphology analysis of cellulose nanocrystals (CNCs) using transmission electron microscopy (TEM) and atomic force microscopy (AFM) images is an important step in the design and optimization of the processes employed in the manufacture and utilization of CNCs. Current protocols used in the analyses of CNC particle morphology for such microscopy images are largely manual and time-consuming, and often produce inconsistent results between different researchers. This paper describes a new semi-automated image analysis framework that can reliably and quickly detect and classify CNCs from TEM and AFM images and measure their dimensions. The proposed image analysis framework is named CNC-Standardized Morphology Analysis for Research and Technology (SMART). The viability of this framework is demonstrated in this paper using exemplar images obtained for a National Research Council Canada certified reference material, CNCD-1. The results obtained from the SMART approach presented in this work are compared critically against the results obtained from the conventional manual approaches. These comparisons revealed a good agreement between the manual and SMART approaches. Notably, the SMART approach showed significant potential for consistent CNC identification and dimensional measurements at a much higher throughput (e.g., number of CNCs measured and number of images analyzed) compared to the conventional manual approaches.

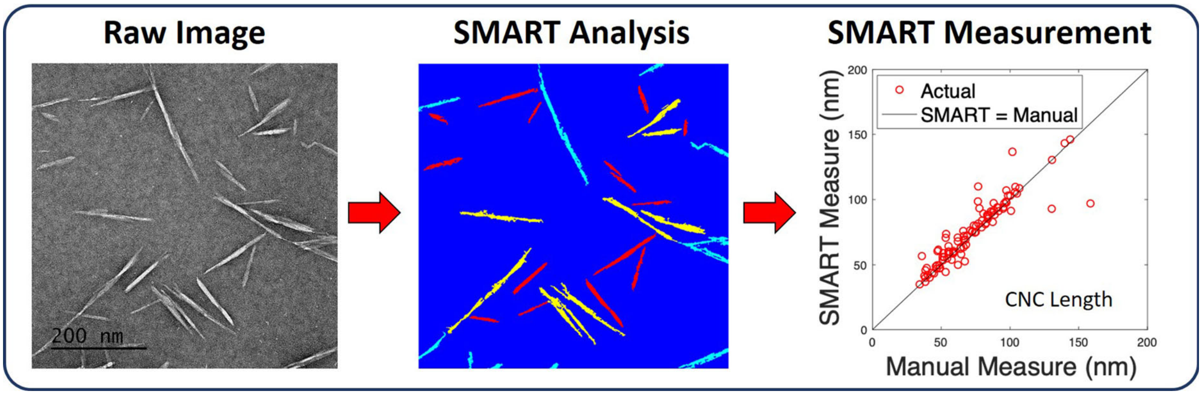
**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10570-020-03668-8.

S. Yucel · B. Yucel · S. R. Kalidindi
George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

R. J. Moon (✉)
The Forest Products Laboratory, U.S. Forest Service, Madison, WI 53726, USA
e-mail: robert.j.moon@usda.gov

L. J. Johnston
Metrology Research Centre, National Research Council Canada, Ottawa, ON K1A 0R6, Canada

**Graphic abstract**




**Keywords** Cellulose nanocrystals · Particle morphology · Atomic force microscopy · Transmission electron microscopy · Chord length analysis

## Introduction

Cellulose Nanocrystals (CNCs) are cellulose based, spindle shaped nanoparticles of typical length 50–350 nm, width 5–20 nm, and aspect ratios of ∼ 3–30, extracted from various cellulose sources such as trees, plants and recycled paper products, by the strong acid hydrolysis method (Moon et al. 2011). Depending on the source and extraction process parameters, CNCs have varying particle morphologies and size distributions. CNCs exhibit various desirable features including high aspect ratio, low density, low thermal expansion, high stiffness and mechanical strength, self-assembly capability, biocompatibility, sustainability, biodegradability, and scalability. These characteristics make CNCs attractive for uses across a wide application space, such as composites, rheology modifiers, flexible electronics, drilling fluids, biomedical, cosmetics, food coatings and packaging (Awan et al. 2018; Ferrer et al. 2017; Goswami et al. 2019; Li et al. 2016; Martínez-Sanz et al. 2013; Valentini et al. 2014; Zhou et al. 2013a).

The addition of CNCs modifies the processing, structure and property/performance of fluids (e.g., as rheology modifiers) and solids (e.g., as polymer reinforcement phase); these effects are strongly related to the CNC particle size and size distribution (Bespalova et al. 2017; Boluk et al. 2011; Sacui et al. 2014). Unfortunately, quantitative measurement of CNC morphology remains challenging due to CNCs' broad size distributions, irregular shapes (i.e., not perfect rectangles or ellipses) and strong propensity for aggregation. The efforts aimed at optimizing CNC production, quality control, and use in various applications have been hampered by the lack of standardization for CNC material characterization (Foster et al. 2018; Jakubek et al. 2018; Kaushik et al. 2015). Several approaches have been used to assess the morphology of CNC nanoparticles, such as small angle X-ray scattering, small angle neutron scattering and dynamic light scattering (DLS) (Mao et al. 2017), field flow fractionation (Mukherjee and Hackley 2018), transmission electron microscopy (TEM) (Brinkmann et al. 2016; Campano et al. 2020; Jakubek et al. 2018), and atomic force microscopy (AFM) (Honorato-Rios and Lagerwall 2020; Reid et al. 2017; Sacui et al. 2014). The most widely used approaches are TEM and AFM image analysis, but both are hindered by manual, time-consuming, and inconsistent procedures that have a high potential for user bias/subjectivity.

Current TEM and AFM image analysis protocols used for CNC particle measurement are subject to variability and error due to the subjectivity in image analysis and data analytics performed by the experimenter. Manual procedures are used to identify individual CNCs, and the recorded length and width are taken as the lengths of the largest line segments through the long and short axes of the CNC (Jakubek et al. 2018). This manual selection is the main reason for the inconsistent measurements from run to run, lab to lab, and analyst to analyst, as the identification of CNC edges can be very challenging, and different analysts may select different target CNCs to measure within a given image. Depending on the image type, TEM or AFM, these procedures are typically performed using different image analysis software such as ImageJ or Gwyddion, respectively (Necas and Klapetek 2012; Schneider et al. 2012). These protocols take a significant amount of time when many CNCs are measured, which may further exacerbate error from analyst bias. Such inconsistencies affect the reliability of the morphology data that is subsequently used in optimization of CNC production, and their utilization for applications. Though there are semi-automated image analysis programs, either commercial or open source (e.g., plugins for ImageJ, such as FibrilJ (Sokolov et al. 2017)), for TEM and AFM image analysis of nanosized particles, application to cellulose nanomaterials has been inconsistent due to difficulties in correctly identifying single particles. To remedy these drawbacks, a semi-automated detection and measurement system specifically designed for CNCs is proposed.

In this work, a semi-automated image analysis framework has been developed (e.g., SMART) to detect and categorize CNC groupings and measure the dimensions of individual CNCs in TEM and AFM images. This framework utilizes basic, widely adopted, digital image processing techniques used in materials science (Campbell et al. 2018; Heilbronner and Barrett 2014; Salem et al. 2014; Sun et al. 2007; Tahoces et al. 2019; Wargo et al. 2013; Yucel et al. 2020) to produce a modular, clear process that can be deployed for CNC image analysis. This system can analyze many images in a short time period via batch processing and can produce CNC morphology data (e.g., particle dimensions and shape) in an accurate, consistent and time-efficient way that is minimally impacted by human bias and variability. Having the ability to measure particle dimensions (i.e., length and width) from a large population of CNC particles in a very short amount of time can allow further investigation into length and width correlations for different CNC processing conditions (Elazzouzi-Hafraoui et al. 2008), different cellulose sources (Elazzouzi-Hafraoui et al. 2008; Sacui et al. 2014), and CNC production from different laboratories (Reid et al. 2017). Lastly, though SMART was designed for CNC materials, it may also be applicable to other similarly shaped nanoparticles (e.g., chitin).

## Methodology

Automation of the detection and quantification of CNCs in TEM and AFM images was completed by utilizing existing image transformation methods such as filtering for image enhancement, segmentation and chord length analysis for measuring particles easily and rapidly. For quantifying CNC particle morphology, it is necessary to: (1) have high quality digital images from which measurements can be made, and (2) have a consistent and reliable method to correctly identify and measure individual CNCs. Obtaining high quality digital images is contingent on several factors, such as homogeneous dispersion of individual CNC particles with minimal CNC–CNC contact or overlapping points, and a uniform contrast across the image, while maintaining high edge contrast between CNC particles and the supporting substrate. Some guidance in this regard can be found in prior work (Foster et al. 2018; Ogawa and Putaux 2019; Stinson-Bagby et al. 2018). The focus of this paper is on the development of a consistent and reliable method to identify and measure individual CNCs while excluding overlapping and clustered particles. In this study, we developed a multi-step process for CNC image analysis that included: (1) pre-processing to remove noise and improve contrast in the image, (2) CNC classification to identify individual CNCs for measurement, and (3) new digital measurements and parameterization tools for CNCs. The proposed image analysis framework is named CNC-Standardized Morphology Analysis for Research and Technology (SMART).

### Sample preparation and imaging

The image analysis method was developed using images obtained as a part of the characterization of a CNC certified reference material (CNCD-1) produced by the National Research Council, Canada (http://www.nrc.ca/crm). CNCs were generated by sulfuric acid hydrolysis of softwood pulp, followed by neutralization with NaOH. This material has been fully characterized for several physical and chemical properties, including size measurements by dynamic light scattering, TEM and AFM. Methods for CNC dispersion by ultrasonication and preparation of samples for microscopy measurements are reported in Jakubek et al. (2018). TEM samples were prepared by depositing 0.02 wt% CNC aqueous suspension onto a plasma-exposed carbon-coated copper grid and then stained with uranyl acetate to improve contrast. The use of negatively stained CNCs allows higher magnification with lower defocusing, resulting in a higher resolution of the CNC edges and a higher accuracy in the measurements of the particle dimensions. AFM samples were prepared by spin coating 0.004 wt% CNC aqueous suspension onto a poly-L-lysine coated mica, a method that provides more homogenous samples with a larger fraction of individual CNCs per image. The initial characterization of CNCD-1 and a subsequent interlaboratory comparison (Bushell et al. 2021; Meija et al. 2020) that utilized this material have optimized the sample preparation and image acquisition/analysis methods and provide the baseline for comparison of manual image analysis with the semi-automated approach developed here.

## Pre-processing and segmentation

The TEM and AFM images were analyzed using a sequence of image processing algorithms that included: (1) contrast enhancement, (2) noise removal (smoothing), (3) sharpening, and (4) segmentation. The contrast enhancement algorithm helps to separate CNCs from the background and is based on a histogram stretching method (Jain et al. 1995). A grayscale image generally has pixel values ranging between 0 and 255. However, most raw images do not use the full range of grayscales. The histogram stretching expands the range to utilize the full range of grayscales by mapping the pixel values of background regions to the values closer to 0 and the pixel values of CNC regions to the values closer to 255. The level of contrast enhancement is adjusted automatically by this algorithm for the different raw images. TEM images typically have much lower contrast than AFM images, and thus require a higher level of enhancement. The noise removal step applies a smoothing filter on the images to adjust undesirable pixel values (i.e., lighter pixels in the background region or darker pixels in the lighter CNC region). TEM images typically have noisier backgrounds and edges than AFM images, and thus require a higher level of noise removal. However, since edges must be preserved, an edge-preserving median filter is preferred over a more standard smoothing algorithm such as the Gaussian filter (Huang et al. 1979). The sharpening algorithm makes the edges and interfaces (e.g., CNC–CNC) more distinct (Davies 2004). Lastly, a segmentation method called adaptive thresholding (Bradley and Roth 2007) is used to segment the processed grayscale image into a binary image of only 0's (background) and 1's (CNCs). In this step, each pixel value is compared with a threshold value and assigned either 0 or 1 depending on whether the pixel value is smaller or larger than the threshold, respectively. A detailed explanation of this approach is given in the supporting document, and Fig. S1 in the supporting document illustrates the explained method over a small representative region from an AFM image. An example set of raw images (Fig. 1a, e), gray scale pre-processed images (Fig. 1b, f) and the final segmented images (Fig. 1c, g) are shown for TEM and AFM images, respectively, in Fig. 1. The final segmented images are used in the next step of the analysis, namely grouping.

## CNC groupings

The distribution of CNCs within the image influences the ability to obtain the desired geometrical measurements. The ideal configuration with uniformly distributed, well separated, CNCs is difficult to achieve; most CNCs are often found contacting, overlapping, and/or parallel stacked with each other. In this work, it was decided to categorize CNCs into three groups: isolated CNCs (i.e., no contact with other CNCs in the image), border CNCs that touch an edge of the image, and agglomerated CNCs (this group includes all the CNCs not included in the other two groups, e.g., touching, overlapping, and parallel stacked). The agglomerated CNCs can be further sub-divided based on the different types of spreading/alignment as touching, overlapping, lateral parallel stacking and linear parallel stacking (see Fig. 2). In this work, we will focus only on the higher level of groupings, namely, isolated, border, and agglomerated groups.

The CNC groupings are identified using a multi-step procedure. The first step in this process involves labeling individually all of the unconnected foreground objects in the segmented images (Fig. 1c,g). This step will identify isolated (individual) CNCs and

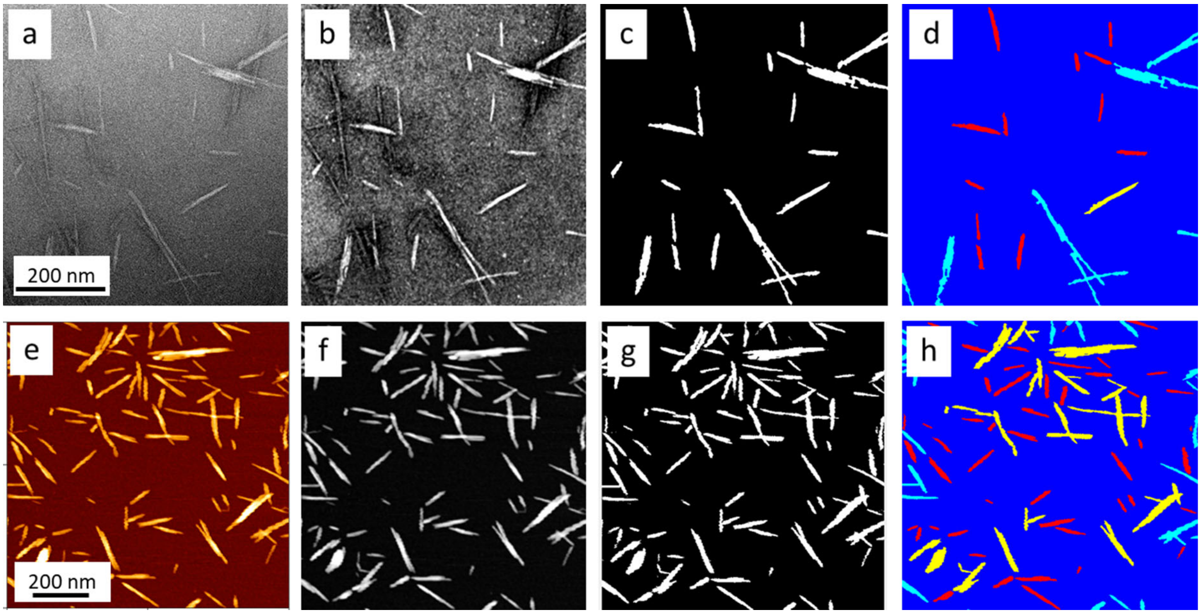



**Fig. 1** Workflow of the SMART image analysis system. For TEM images, **a** raw image, **b** contrast enhanced, denoised and sharpened (pre-processed) image, **c** segmented image, and **d** image with color coded CNC groupings. For AFM images, **e** raw image, **f** pre-processed image, **g** segmented image, and **h** image with color-coded CNC groupings. The color coding: red for isolated CNCs, light blue for border CNCs, and yellow for agglomerated CNCs

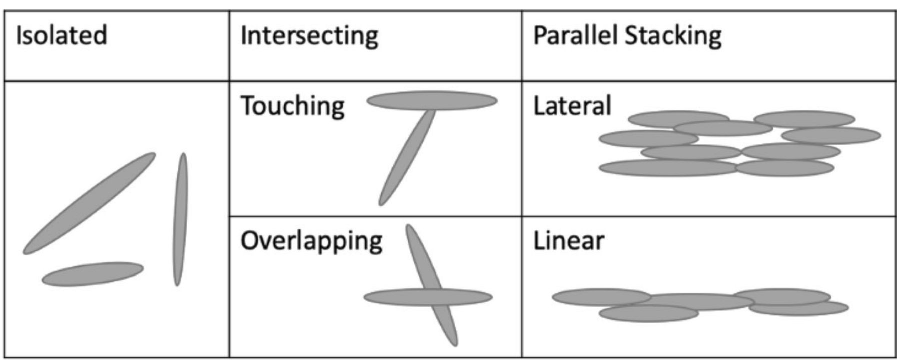



**Fig. 2** Schematic showing the five different CNC groupings typically observed in TEM or AFM images. In the current study, only isolated CNCs were analyzed

agglomerated CNCs as uniquely labelled objects that could be further analyzed individually. The labeling of all unconnected foreground objects is accomplished using a built-in MATLAB function *bwlabel* (MATLAB 2020). Next, the labels of all border CNCs (i.e., unconnected foreground objects with pixels touching any of the edges of the image) are detected and saved under a single group identification using the built-in MATLAB function *imclearborder* (MATLAB 2020). The border CNCs are shown with light blue color in Fig. 1d, e. The remaining unconnected foreground objects will then be analyzed individually and assigned group identification corresponding to either the isolated CNCs or the agglomerated CNCs. This grouping is accomplished using the built-in MATLAB function *regionprops* (MATLAB 2020), which estimates the major and minor axis lengths of each unconnected object by encapsulating it within an ellipse (Fig. 3). To identify/separate isolated CNCs from agglomerates, a unique combination of ranges for aspect ratio (i.e., the ratio of the length of the major axis over the length of the minor axis) and size

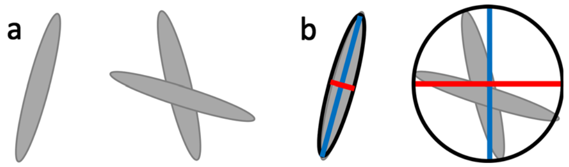



**Fig. 3** Schematic showing **a** an isolated CNC and overlapping CNCs, and **b** isolated and overlapping CNCs with their corresponding elliptical construct overlaid (black line). The aspect ratio (i.e., major axis length (blue line)/minor axis length (red line)) is used as one of the primary parameters for determining the classification of CNC groupings

limitation (e.g., minimum and maximum lengths of major and minor axis) are used. Since the CNC morphology is strongly dependent on cellulose source material and the production process, the parameters that best suit a given batch of CNCs will vary. To accommodate this, the SMART system accepts user input for these parameters. Unless otherwise stated, the parameters used in this study are those typical for isolated CNCs produced from wood pulps and are as follows. The parameters used for TEM images included aspect ratio greater than 2.5, major axis length between 15 and 250 nm, and minor axis length of less than 15 nm. The parameters used for AFM images included aspect ratio greater than 2, major axis length between 15 and 300 nm, and minor axis length less than 30 nm. These parameters were selected based on prior reports in the literature (Brinkmann et al. 2016; Jakubek et al. 2018), and should be adjusted as needed. The current study focused exclusively on isolated CNCs; future analysis will explore approaches to analyze the other CNC groupings. The fit to an ellipse described above is used only for the grouping of CNCs into the different classes. Subsequently, other advanced approaches are used for accurately quantifying the geometry of the individual CNCs.

The area fraction of the image occupied by each of the three groups can be easily calculated. In Fig. 1d and h, the three CNC groupings are represented with different colors (e.g., red—isolated CNCs, blue—border CNCs, yellow—agglomerated CNCs). Area fractions are calculated by dividing the number of colored pixels for each group (e.g., number of red pixels for isolated CNCs) by the total number of all CNC pixels (i.e., sum of red, yellow and blue pixels) from the color-coded image. This metric can be used to assess the ratio of isolated CNCs to agglomerated CNCs.

## CNC size analysis

The focus of this paper is the size analysis of isolated CNCs; the fundamental measurement approach used in SMART follows the established TEM and AFM image analysis protocols, and thus allows a direct comparison with prior reports in literature. For each isolated CNC identified within the segmented image (Fig. 4a), SMART analysis rotates the object such that the major axis lies parallel to the horizontal x-axis (Fig. 4b). Line segments (chords) in x and y directions are then measured for both major and minor axis directions, respectively (Fig. 4b). These axes are perpendicular to each other, and the chords along these x and y axes represent length and width profiles of the CNC, respectively. All of the x and y chord lengths for each isolated CNC are collected to construct chord length profiles that are used for analysis (Fig. 4c, d). For the example shown in Fig. 4, the length profile is based on 45 separate horizontal chord lengths, while the width profile is based on 272 separate vertical chord lengths. The number of chord lengths used to build the profile is based on the particle aspect ratio, and the pixel density of the image. The protocols to efficiently obtain chord length information have been established in previous studies in material science (Turner et al. 2016; Yucel et al. 2020).

For TEM images, the resulting CNC length and width are obtained by averaging the chord values (black data points in Fig. 4c and d) above a certain threshold (i.e., maximum chord minus 10% and 50% of the range of the x and y direction chords, respectively). Subtracting the range-based values from the maximum chords makes the length and width calculations specific to the morphology of each CNC since this range is the difference between maximum and minimum chords. Taking the average of the selected longest chords helps to remove noise associated with noisy background pixels, or pixel clusters, becoming "attached" to the periphery of the CNC and artificially increasing the corresponding chord length by several nanometers. Selection of the averaging regions (e.g., the chord values above longest chord minus 10% and 50% of range of all chords for length and width, respectively) was empirically based on

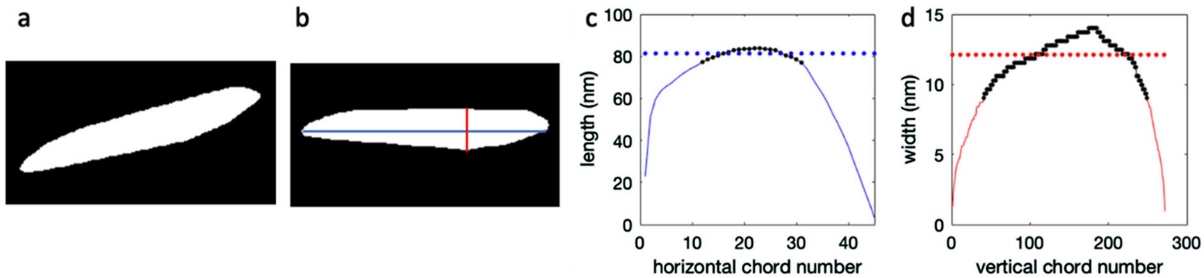



**Fig. 4** **a** A typical segmented isolated CNC. **b** The CNC is rotated so that the long axis is parallel to the horizontal axis; two example chords are shown, one horizontal (blue) and one vertical (red). **c** Profile of horizontal chord lengths. Black points represent the chords with a length larger than the maximum chord value minus the 10% of range of all chords. The blue dashed line is the average value of the black points. **d** Profile of the vertical chord lengths. Black points represent the chords with a length larger than the maximum chord value minus 50% of range of all chords. The red dashed line is the average value of the black points

comparisons between SMART and manual measurements of many CNCs. The noise effect is more dominant for width chords, as the size of the noise can be a significant fraction of the width of a CNC. As a result, a higher percentage value for the range summation, 50%, for averaging the width was used. This averaging approach, though empirical, helps the SMART analysis to minimize the effect of noise on the measurement of CNC length and width.

For AFM image analysis, an adjustment to the approach was needed due to AFM tip convolution (e.g., dilation) effects that artificially increased dimensions of the object being imaged (Canet-Ferrer et al. 2014). The tip convolution effect is significant for CNC width as it can be a high fraction of the total width, but is small for CNC length, as it is a small fraction of the total length. Due to the difficulty of applying deconvolution approaches to obtain an accurate measurement of CNC width, most AFM studies do not measure width, but rather measure the height of the cross-sectional dimension. Height measurements are not affected by tip convolution effects (Brinkmann et al. 2016). The SMART analysis utilizes the height information of every pixel in the CNC image to estimate the measured height for each CNC. For a given CNC, the average of top 20% region of its height values (shown in green in Fig. 5b) is used to estimate the height of the CNC. The parameter for averaging the top region is empirically selected through comparisons between SMART – 3D height profiles and manual – 2D height profiles of several CNCs to provide higher consistency between these sets of height measurements. It should be noted that the CNC height, though similar, is not necessarily

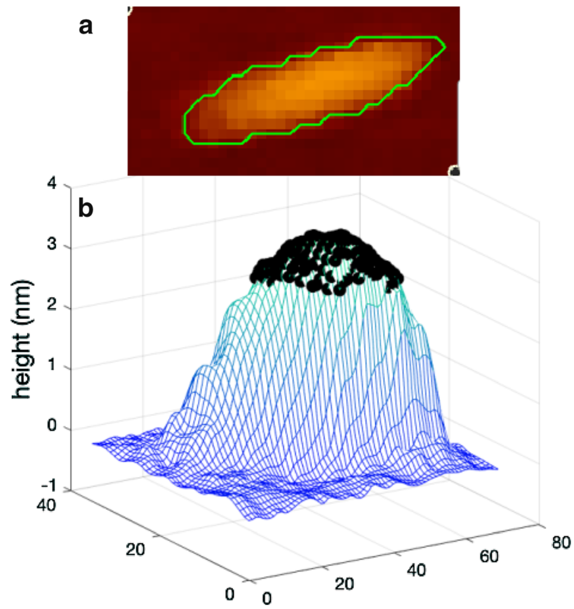



**Fig. 5** AFM height profile analysis for **a** single CNC, and **b** its 3D height profile. The height values with the maximum height value minus 20 percent of the range of all points (black data points on top of the profile) are used to assess an average height for this particle

equivalent to the CNC width (Brinkmann et al. 2016; Jakubek et al. 2018; Postek et al. 2010).

Chord length plots are also used in the identification and selection of isolated CNCs. Using the chord length plot for width measurements, a specific width ratio, the ratio maximum width over the average width value (e.g., $w_{ratio} = w_{max}/w_{average}$), is defined to remove parallel stacked CNCs that were initially selected as single CNCs (Fig. S2). The width-based criterion used for TEM and AFM image analysis removed CNC with $w_{ratio} > 1.5$. Introducing this selection step in addition to the criteria of aspect ratio and the lengths of major and minor axis differences improves the accuracy of isolated CNC selection process in a systematic way and is more representative of current manual CNC selection approaches. This chord length analysis step was used on both TEM and AFM images.

Similar to the width-based method explained above, a height-based selection is also defined for the removal of linearly parallel stacked CNCs that are initially selected as single CNCs. To detect and remove these CNCs, the maximum horizontal chord is found (Fig. S3a and S3b), and height values of the pixels on this longest chord are extracted to plot a height profile along the long axis of the CNC (Fig. S3c). Using this height profile, the presence of local minima (red star in Fig. S3c) is checked first. Following the detection of one or more local minima, a height-based ratio is calculated (e.g., $h_{ratio} = h_{localmin}/h_{average}$). The height-based criterion used for AFM image analysis removed CNCs with $h_{ratio} > 0.75$. Introducing this selection step also improves the accuracy of the isolated CNC selection process. This removal method is applied only on AFM images, since TEM images do not provide height information.

### Reliability testing

The reliability of the SMART approach was investigated by conducting a series of image analysis validation runs to assess how well the system identified isolated CNCs and measured their dimensions. Two different validation approaches were used: parameter testing and comparison testing. Parameter testing was used to gain insight on how changes in the inputs used in the program (e.g., smoothing, sharpening, etc.) translate to changes in CNC identification and dimension measurements. The comparison testing was undertaken to assess how well the automated system can identify and measure CNCs as compared to manual measurements. For these tests, data sets of 30 TEM and 15 AFM images were used (see "Sample preparation and imaging" section for details); the details of the testing are provided in the corresponding sub-section of the "Results" section.

## Results

The SMART image analysis requires high quality images (e.g., good contrast between CNCs and the background region, distinct CNC edges, and a low concentration of CNCs homogeneously distributed on the imaging substrate), requiring special care in TEM and AFM sample preparation and imaging as explained in references (Brinkmann et al. 2016; Jakubek et al. 2018; Stinson-Bagby et al. 2018). Once this has been achieved, the utility of the SMART image analysis framework can be demonstrated. First, the idealized parameter setting is described, which is used for pre-processing of TEM and AFM images (i.e., segmentation and binary image processing) to prepare them for subsequent image analysis. Next, the identification of individual CNCs is assessed by comparing the SMART analysis to manual identification, which is considered the current state of the art approach. Following the isolated CNC selection, the SMART results are validated by comparisons with manual measurements. Finally, a small study is presented demonstrating the effect of different image pre-processing parameters on the isolated CNC identification and dimension measurements.

### Optimization of image processing parameter settings

The image analysis can be influenced by several variables: image enhancement processes (e.g., smoothing, sharpening), the type of image (e.g., pixel resolution, contrast level, TEM, AFM), and the CNC material itself (e.g., length and level of agglomeration). Within the image analysis program, there are two parameters that can be adjusted: smoothing and sharpening. In general, smoothing is used to reduce noise, while sharpening is used to increase edge contrast, and adjustments can be made within the range of 1 to 5, which represents the size of the applied filter (i.e., 1 represents a 1 pixel neighborhood, giving a 3 by 3 pixel filter size, where the new value of center pixel is calculated using the values of 8 neighboring pixels and the center pixel; 5 is for a 5 pixel neighborhood, giving a 11 by 11 pixel filter size, where the new value of center pixel is calculated using the values of 120 neighboring pixels and the center pixel; see supporting documentation, Fig. S4 & S5, for additional description). The determination of appropriate parameter settings is based on an iterative assessment of how well individual CNCs are identified and the level of overlap of the SMART predicted outline of the CNC to that in the raw image. The ideal parameter settings are significantly influenced by the pixel density and contrast of the images, and as a result, these settings are considerably different for TEM and AFM images.

Typically, TEM images of CNCs have poor contrast, poor edge definition, and high levels of noise, despite best attempts in sample preparation and imaging protocols (Foster et al. 2018). The TEM images used in this study have a pixel size of $\sim$ 0.3 nm/pixel and a high pixel density (2048 by 2048 pixels), thus allowing considerable capability for smoothing and sharpening. The ideal parameter settings were determined to be at the maximum levels for smoothing and sharpening, of 5 and 5, respectively. In contrast, the AFM images have high contrast, high edge definition, and low levels of noise. The AFM images used in this study have a pixel size of $\sim$ 2 nm/pixel and a low pixel density (512 by 512 pixels), limiting the filter size that can be used since larger filter sizes approach the size of the CNC particle, resulting in artifacts. For the AFM images used in this study, ideal parameter settings were determined to be at minimum levels for smoothing and sharpening, of 1 and 1, respectively. The effects of changing parameter levels on the isolated CNC identification and measurements are discussed in the section "Parameter Testing".

## Identification of individual CNCs

The identification of individual CNCs is a critical component for size measurements, however, it is not always straightforward, and many times can be exceedingly challenging. The SMART image analysis uses the protocols developed previously for manual measurements of TEM and AFM images as a guide (Jakubek et al. 2018) and (Brinkmann et al. 2016). Subsequent analysis and description are based on the use of TEM and AFM images that were obtained in a previous study (Jakubek et al. 2018). The SMART image analysis approach using the pre-set parameters and CNC size criteria defined in "CNC groupings" section, had a high level of success in the identification of individual CNCs and CNC edge boundaries. When comparing the SMART system to manually measured datasets, the majority of the same CNCs were identified. Examples of CNC identification by SMART and manual approaches for a TEM image and an AFM image are shown in Fig. 6. The TEM image in Fig. 6a shows several CNCs with those that were manually identified as single CNCs highlighted with blue ellipses. Figure 6b shows the outcome of the SMART identification of individual CNCs as well as the identification of the CNC edges with green outlines. For the AFM image, the white lines and individual number labels in Fig. 6c show the manually identified isolated CNCs, while the green outlines in Fig. 6d show SMART identification of individual CNCs and their corresponding edges. When the contrast between CNCs and background is high enough and the CNCs are well dispersed, SMART identification is consistent with manual selection. There is a high overlap in the identification of the same isolated CNCs between SMART and manual approaches (e.g., greater than 50% for TEM and 70% for AFM). This is a consequence of the high image contrast in these images, which allows the CNCs to be easily distinguished from the dark background. Overall, the typical examples in Fig. 6 indicate that the SMART and manual approaches can identify similar isolated CNCs for both TEM and AFM images. Having complete agreement with the manual approach is not to be expected, nor is it the goal; there is ambiguity with manual identification of CNCs since the selection of single particles is subjective and different analysts often select different sets of particles for the same image (Brinkmann et al. 2016; Jakubek et al. 2018).

As shown in Fig. 6 the SMART system has some false CNC identifications and some missed identifications; although these issues are not ideal, they occur at a low level. Such errors are primarily related to image quality and skewed contrast when CNCs are in very close proximity to each other, and more sophisticated approaches for improved selection protocols are being developed to address such issues. The CNC identification via the SMART approach can be adjusted by changing the CNC selection criterion (e.g., aspect ratios, major and minor axis lengths of the fitted ellipses and additional width-based and height-based ratios that are also used for the detection of parallel stacked CNCs). For example, lowering the minimum value for the major or minor axis length identifies smaller objects as CNCs, but non-CNC objects such as

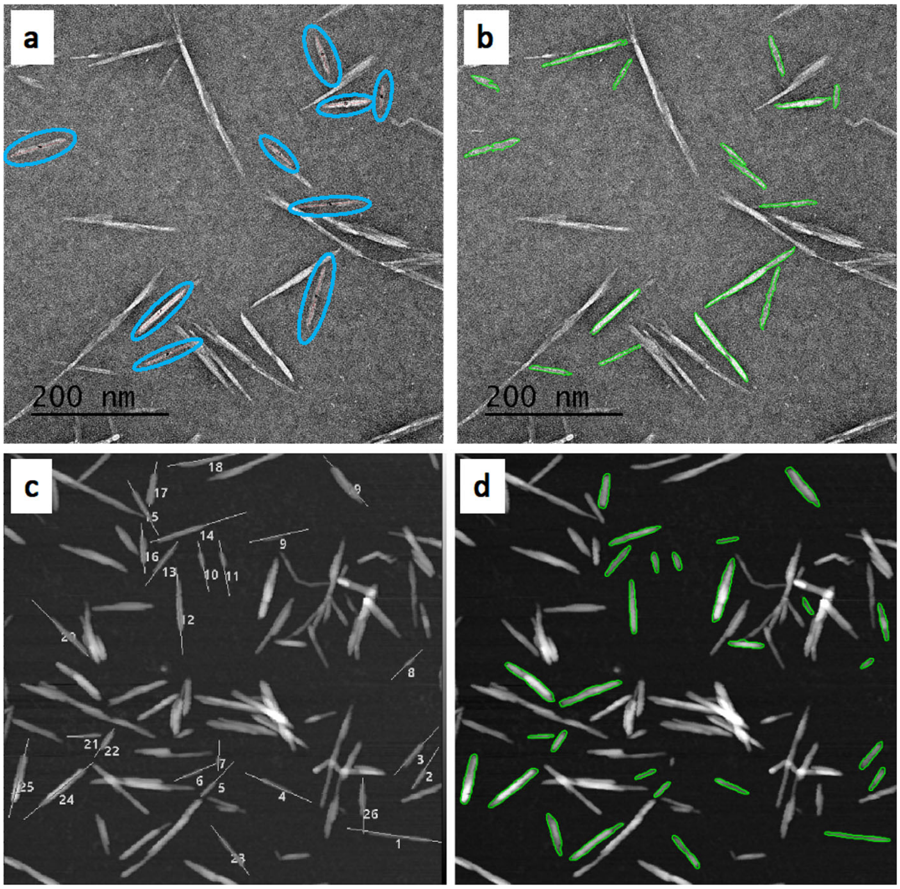



**Fig. 6** Example of CNC identification by manual and SMART approaches for TEM and AFM images **a** TEM image, manual approach (9 CNCs identified) **b** TEM image, SMART approach (14 CNCs identified) **c** AFM image, manual approach (26 CNCs identified) **d** AFM image, SMART approach (29 CNCs identified). Blue ellipses are superimposed on the micrograph in part **a** to highlight CNCs that were manually selected; white lines in part **c** are used to manually identify and measure CNCs. Green outlines in parts **b** and **d** are the CNC edges identified by the SMART approach. Images are from (Jakubek et al. 2018)

accumulated background noise pixels are more likely to be incorrectly identified as CNCs. Increasing these values can allow the identification of larger objects; however, CNC agglomerates (e.g., either end-to-end, or side-by-side stacking of multiple CNCs) might be incorrectly identified as larger individual CNCs. Overall, isolated CNCs could be readily identified with the parameters specified earlier (Fig. 7a, b), where SMART was able to track the CNC edges correctly in TEM (Fig. 7a) and AFM (Fig. 7b) images. However, there were problems for some TEM images due to high levels of noise in the image, which affected the CNC edge resolution, resulting in either wider, longer or shorter CNC objects depending on the situation. The broadening of CNCs was due to two factors; the capture of noisy pixels on the CNC edges, or reduction of edge contrast (due to noise pixels) at the interface between two or more isolated CNCs that were parallel stacked and were subsequently identified as a single CNC with larger dimensions (Fig. 7c). The shortening of CNCs occurred at the CNC ends when the noise pixel background was similar to the CNC contrast (Fig. 7d). The lengthening of CNCs occurred by two methods; the capturing of noisy pixels at the CNC tails or ends, or when noisy pixels lowered edge contrast at the interface between two or more isolated CNCs that are stacked end-to-end, resulting in their identification as a single CNC with longer dimensions (Fig. 7e).

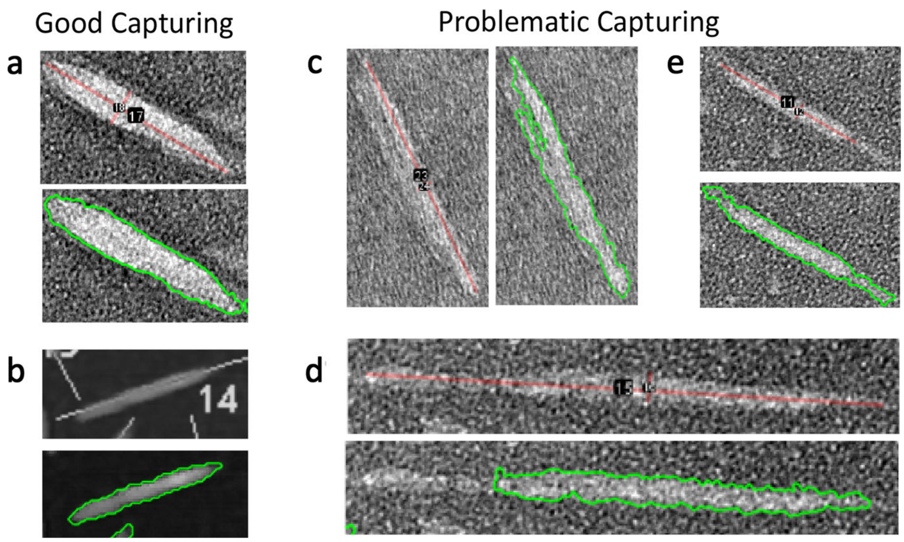



**Fig. 7** Comparisons of the same CNC identified in both manual and SMART approaches, showing examples where the SMART approach had good results (parts **a** and **b**) and substantial errors (parts **c**, **d**, and **e**) in the capture of the CNC shape. Green outlines are the CNC edges identified by the SMART approach. The larger errors are caused by the effects of noise within the image, and can result in CNCs that are wider (part **c**), shorter (part **d**) or longer (part **e**) than those measured by the manual approach. TEM images are shown in parts **a**, **c**, **d**, and **e**, while an AFM image is shown in part **b**

## CNC dimension measurements: SMART versus manual measurements

To assess how well the SMART system can identify viable CNCs to measure, as well as how well it can measure the length, width, and height of CNCs, its measurements were compared to manual measurements from both TEM and AFM images from a previous study (Jakubek et al. 2018). To facilitate a side-by-side comparison of CNC dimensional measurements (e.g., length, width, height), the CNCs commonly identified and measured by both the manual and SMART approaches were selected; thus any differences in measurements and distributions can be attributed directly to the accuracy of the analysis/measurement approaches. The results demonstrate that the SMART approach can identify CNCs and measure their morphologies at least “as good as a human”.

For a comparison of the TEM measurements, 14 TEM images were selected and 156 individual CNCs were identified and analyzed manually, while the SMART system re-analyzed the same images and identified 174 individual CNCs, of which 99 were identical to those measured manually. The comparison of the dimensional measurements of these 99 CNCs from SMART versus manual approaches is summarized with histograms (Fig. 8a, c) and point-wise comparisons (Fig. 8b, d). While the length distributions from both approaches are quite similar to each other, the width distribution obtained by the SMART approach is somewhat narrower than that obtained by the manual approach. Additionally, the mean widths measured by the manual approach appear to be slightly smaller than the mean widths measured by the SMART approach. One reason the SMART system may be measuring a smaller mean width could be attributed to the averaging protocols used (see “CNC size analysis” section). Note that the manual method measured the width from the widest part of the CNC. For length comparison, there are only a few outliers. This was caused by errors in the SMART approach in capturing the CNCs as discussed in the previous section (see Fig. 7d, e). The mean and standard deviation values for both length $\bar{L}_{SMART} = 74.9\,\text{nm}$, $\sigma_{L\_SMART} = 24.0\,\text{nm}$) and width ($\bar{W}_{SMART} = 8.1\,\text{nm}$, $\sigma_{W\_SMART} = 1.5\,\text{nm}$) are very close to those obtained from the manual measurements ($\bar{L}_{manual} = 72.5\,\text{nm}$, $\sigma_{L\_manual} = 25.1\,\text{nm}$ and $\bar{W}_{manual} = 8.5\,\text{nm}$, $\sigma_{W\_manual} = 2.1\,\text{nm}$).

For the comparison of the AFM measurements, 5 AFM images were analyzed by both approaches. The manual approach identified 128 individual CNCs

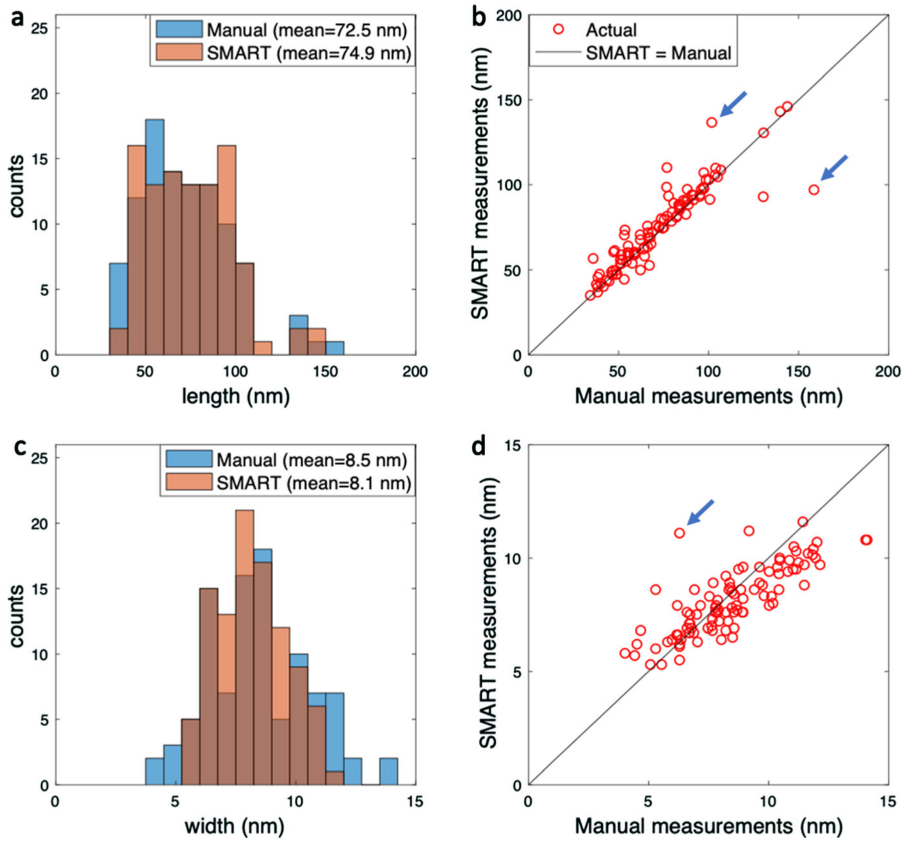



**Fig. 8** TEM image analysis comparison, SMART versus manual approaches for the 99 commonly identified CNCs: **a** histogram of length measurements, **b** corresponding one-to-one comparison of identical CNCs (the two blue arrows identify measurements with the highest differences and are discussed in more detail in Fig. 7e and d), **c** histogram of width measurements, **d** corresponding one-to-one comparison of identical CNCs (the blue arrow identifies a measurement with the highest discrepancy and is discussed in more detail in Fig. 7c). Overlapping regions of histograms are shown in brown

(Jakubek et al. 2018), while the SMART system identified 148 individual CNCs, of which 101 were commonly identified by both approaches. The dimensional measurements of these 101 CNCs from SMART and manual approaches are presented as histograms (Fig. 9a, c) and point-wise comparisons (Fig. 9b, d). The SMART measurements were very consistent with the manual measurements in terms of histograms, point-wise comparisons, mean and standard deviation values of length ($\bar{L}_{SMART} = 75.2$ nm, $\sigma_{L\_SMART} = 31.9$ nm, and $\bar{L}_{manual} = 77.0$ nm, $\sigma_{L\_manual} = 31.9$ nm) and height measurements ($\bar{H}_{SMART} = 3.5$ nm, $\sigma_{H\_SMART} = 1.1$ nm and $\bar{H}_{manual} = 3.5$ nm, $\sigma_{H\_manual} = 1.1$ nm).

## Parameter testing

Parameter testing studies provided insights on the sensitivity of the results (e.g., length, width measurements) obtained by the SMART approach to various parameters used in the individual image analyses functions of the SMART system. The length and width measurements were affected by the parameters embedded in the image processing functions (e.g., filter sizes in sharpening and smoothing). The optimum pre-set image processing parameters were determined for high resolution TEM images to be 5-sharpening and 5-smoothing (i.e., level 5 represents a 5-pixel neighborhood having a 11 by 11 filter size), while for low resolution AFM images they were 1 and 1 (i.e., level 1 represents a 1-pixel neighborhood having a 3 by 3 filter size), respectively. The 'ideal'

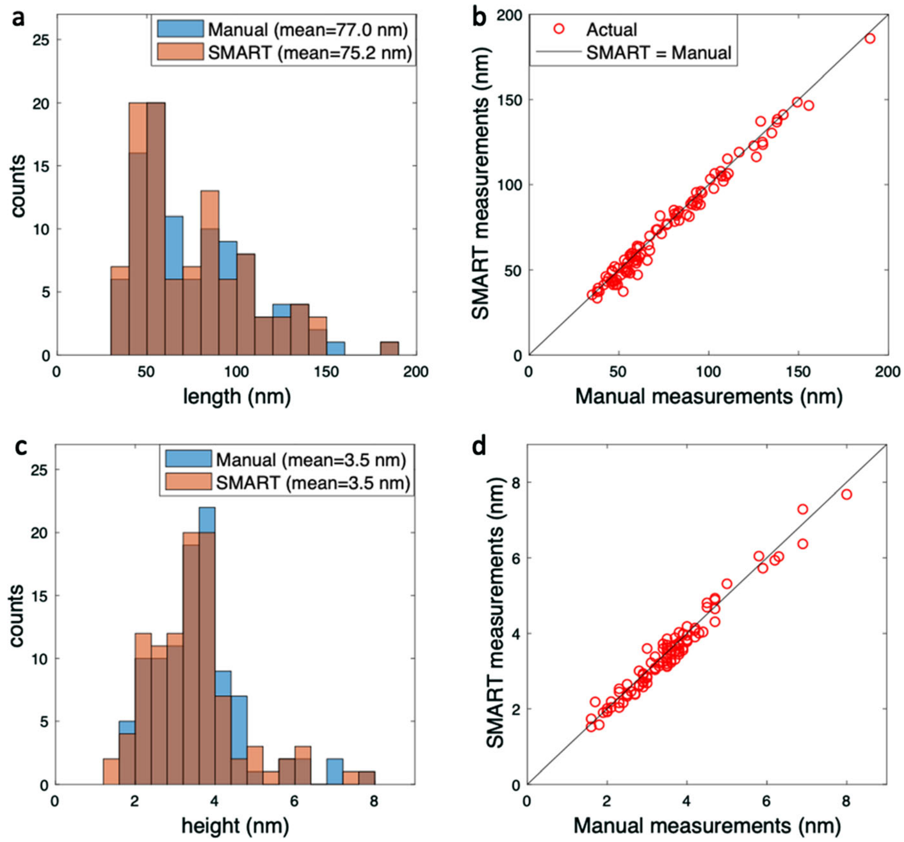



**Fig. 9** AFM image analysis comparison, SMART verse manual approaches for the 101 commonly identified CNCs: **a** histogram of length measurements, **b** corresponding one-to-one comparisons, **c** histogram of width measurements, **d** corresponding one-to-one comparisons. Overlapping regions of histograms are shown in brown

image processing parameters were determined based on three criteria: (1) which one resulted in most consistent and accurate SMART measurements compared to manual data, (2) which one had the least amount of false CNC identification, and (3) which one minimized noise in the system. The sensitivity of the selected parameter values on CNC grouping categorization, isolated CNC detection, and the variability of CNC dimension measurements was evaluated to assess the robustness of the SMART image analysis framework. Note that with these comparisons we are not necessarily claiming we have the "correct" capture of all CNCs and their dimensional measurements, however, we demonstrate that certain parameters result in less noise, and thus improve CNC identification and measurement.

The sensitivity of the parameter values on the SMART measurements of length (TEM, AFM), width (TEM), and height (AFM) was evaluated by repeating the measurements on a 5 by 5 matrix of possible values for the sharpening filter (1–5) and the smoothing filter (1–5), resulting in a total of 25 different parameter combinations on a randomly selected TEM image and a randomly selected AFM image. Two types of investigations were conducted: (1) specific tracking of one individual CNC through the entire parameter matrix, and (2) general tracking of all CNCs identified on the image. For the tracking of the same CNC it was important to choose an isolated CNC from a region of the image with well-dispersed CNCs, so that the CNC would not attach to another particle as a result of dilation effects from poor parameter settings.

For the TEM image, a summary of the percent changes to length and width of a single CNC tracked through all 25-parameter settings is shown in Fig. 10. The ideal image processing parameters for TEM images are level 5 for both smoothing and sharpening filters and were used as the reference point in

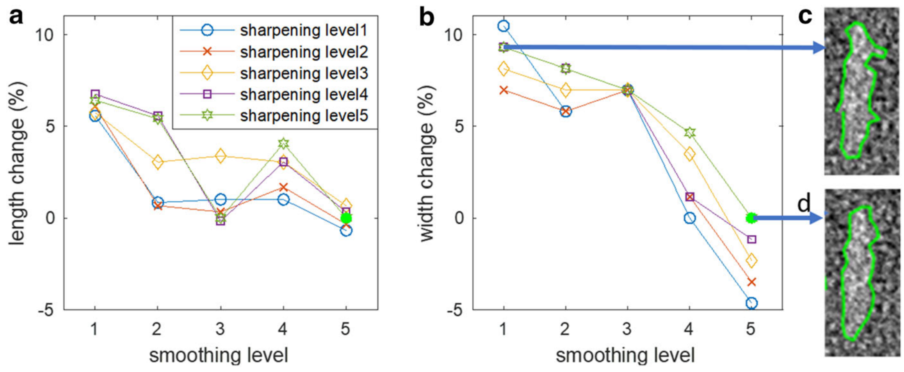



**Fig. 10** Effect of the different parameters in the SMART approach on the CNC dimension measurements for a TEM image, where the ideal parameter setting (smoothing level 5 and sharpening level 5) is highlighted with the green dot. **a** Percentage change of length for different parameter values. **b** Percentage change of width for different parameter values. **c** SMART identification of CNC edges for a non-ideal parameter combination (smoothing level 1 and sharpening level 5), in which less smoothing results in more background noise pixels attached to the CNC. **d** SMART identification of CNC edges for the ideal parameter combination, in which higher smoothing results in less noise

calculating the percent change (green data point in Fig. 10a). Overall, the deviation from the ideal settings resulted in changes of less than 10% in the length and the width measurements. The effect of the smoothing parameter was relatively more dominant. When decreasing the smoothing level from 5 to 1 the length and width measurements of a single CNC increased. When decreasing the smoothing level, less noise was removed and noisy pixels attached to the edge of the CNC, which consequently increased the dimension measurements. Figure 10c shows an example where the CNC outline includes attached noisy pixels, effectively increasing the width of the CNC as compared to the outline at the idealized settings shown in Fig. 10d. In general, the width measurements were affected more than length in terms of percentage change, which is due to the fact the lengths of the CNCs are significantly larger than their widths. Length and width measurements of the CNCs were less affected by the changes in the sharpening filter parameter. Lowering the sharpening level from 5 to 1 increases the degree of smoothing which generally resulted in less defined edges. While length measurements were somewhat less affected by the sharpening level changes, width measurements usually decreased with lower sharpening levels. The apparent length or width change discontinuity at smoothing level 3 (Fig. 10a, b) is a random artifact of the noisy pixel attachment described above.

For the general tracking of all CNCs within the TEM image, the smoothing and sharpening parameter values can modify the shapes of the objects, which then subsequently alter their groupings (e.g., isolated CNCs). When the idealized parameter levels are used (5 for both smoothing and sharpening), the SMART system can consistently identify various CNC groupings and individual CNCs (Fig. S6a, b). For poor selections of parameter values, parallel stacked CNCs or fragments of overlapping CNCs or noisy pixels can blend together during image processing and are subsequently falsely identified as individual CNCs. In general, changes in sharpening had a small effect on CNC shape, CNC grouping classification and CNC identification, while changes in the smoothing parameter to level 2 or 1 had a pronounced effect. Fig. S6c shows an example of CNC grouping classification when the smoothing parameter is decreased from the ideal level of 5 down to 1. By comparing Fig. S6b to S6c, it is noticed that a much lower number of isolated CNCs are identified (i.e., from 14 down to 2, “red” objects), which is a result of objects blending together and no longer meeting the selection criteria used in the identification of isolated CNCs. The comparison of the resulting dimensional measurements of isolated CNCs across the various parameter settings demonstrated that for moderate changes from the ideal parameter setting, the change in dimensions was less than 10%. Larger differences in measurements were only observed when the smoothing parameter was at level 2 or 1 and resulted from a much lower number of individual CNCs that were identified.

For the AFM image, a summary of the percent changes to length of a single CNC tracked through all 25 parameter combinations is shown in Fig. 11. The ideal image processing parameter for AFM images is level 1 for both smoothing and sharpening, and this was used as the reference point in calculating the percent change (green data point in Fig. 11a). For any deviation from the ideal settings (e.g., any combination of smoothing or sharpening) the change in length measurements was small (less than 11%), while the height measurements were unaffected (i.e., less than 0.1 nm change). The effect of the smoothing parameter was more dominant than the sharpening parameter. With increasing smoothing parameter from 1 to 5 the length measurements of a single CNC decreased up to 11%. When the smoothing parameter was increased, the zone over which smoothing is calculated increased and with the low pixel density of the AFM images this zone increasingly included more of the CNC; as a result: (1) the tails of CNCs can be cut, decreasing the length of the CNCs (Fig. 11c), (2) some short and thin CNCs are removed, and (3) regions along CNC length or CNC agglomerates can be cut, making it possible to produce false isolated CNCs. The effect of sharpening level was minimal on CNC length measurements, for which the length changed by less

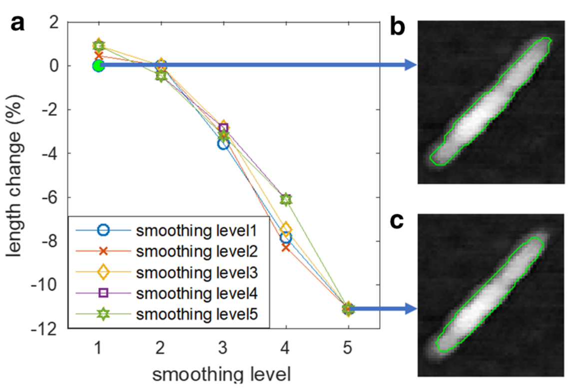



**Fig. 11** Effect of different SMART parameter values on the CNC dimension measurements for an AFM image, where the ideal parameter setting (smoothing level 1 and sharpening level 1) is highlighted with the green dot. **a** Percentage change of length for different parameter levels. **b** SMART identification of CNC edges for the ideal parameter combination. **c** SMART identification of CNC edges for a non-ideal parameter combination (smoothing level 5 and sharpening level 1). Excessive smoothing results in the ends (or tails) of the CNC blending into background noise pixels and being trimmed off in SMART analysis

than 3% even when sharpening was increased to the maximum of 5. This small effect of changes in parameters on the final CNC identification and measurement is likely due to the sufficiently high contrast between the CNC edge and the substrate for the AFM images.

When all objects within the AFM image were tracked across all parameter settings, the amount of interaction between neighboring CNCs and noisy pixels as the parameters shifted further from ideal settings was much lower than what was observed in the TEM images. This resulted in more consistent identification and dimensional measurements of isolated CNCs across all parameter settings. The SMART image analysis was able to consistently identify various CNC groupings and individual CNCs (Fig. S7). In general, changes in sharpening and smoothing had a small effect on object shape, classification and CNC identification. By comparing the various CNC groupings and individual CNCs at ideal parameter settings (Fig. S7b), to that of the worst-case parameter settings (e.g., 5 for smoothing) it is noticed that most groupings are similar, and many of the same isolated CNCs are identified (Fig. S7c). There are small differences in the numbers of the isolated CNCs that are identified, from 20 down to 16, with 12 being the same CNCs. The changes were associated with the disappearance of small CNCs either when they were "smoothed" into the background noise or when the object aspect ratio was lowered to the point where it no longer satisfied the selection criteria used in the identification of isolated CNCs. The comparison of the resulting dimensional measurements of isolated CNCs across the various parameter settings demonstrated that for most changes from the ideal parameter setting, the change in length was less than 6%. Note that the height measurements were unaffected. Larger differences in length of 8–11% were only observed when the smoothing parameter was at levels 4 or 5.

The parameter study demonstrates that for small to moderate deviations from the "ideal" parameter settings, the effect on CNC identification is small, and the change in the individual CNC dimensional measurements is less than 10% for both TEM and AFM images. Also, independent of the "true" values of CNC dimensions, the parameter testing study demonstrated that the SMART approach can in principle remove some of the subjectivity associated with deciding whether a feature is one or two particles. AFM height measurements exhibited negligible change since the detection of center region of CNCs was not affected.

## Case study

The SMART system was compared to the previous manual measurement of TEM and AFM images from a large image data set for the CNC certified reference material, CNCD-1 (Jakubek et al. 2018). In the reference paper, dimensional measurements from manual analysis of isolated CNCs from 434 TEM images and 66 AFM images were reported. While the TEM images were analyzed by one analyst, the AFM images were independently analyzed by 2 analysts. The SMART system performed image analysis on the same TEM and AFM images, and the results are reported here. The reported results included the total time to complete the analysis, the CNC dimension measurements, the area fractions of total CNC area over the image area, and the ratio of isolated CNCs over the total CNC area.

SMART analysis on the 434 TEM images (analysis completed in groups of 64–117 images at a time) identified and measured a total of 2177 CNCs using the selection criteria described earlier (aspect ratio greater than 2.5, major axis length larger than 15 nm and lesser than 250 nm, minor axis length lesser than 15 nm, and width-based ratio lower than 1.5). The analysis of 2177 CNCs took 43 min on a desktop computer with i7 processor (16 GB RAM, 3.7 GHz), which is approximately 1.2 s/CNC. An estimate for manual measurements completed by a trained analyst using ImageJ software is 30 s/CNC. The comparison of CNC dimensions from SMART and manual approaches is presented as histograms in Fig. 12. These histograms are suitably normalized to account for the differences in the numbers of measured CNCs (N = 1909 for manual analysis, N = 2177 for SMART analysis). Length histograms of both analyses show similar trends, whereas the width distribution obtained by the SMART approach is narrower than from the manual approach. The difference in width profile may be caused by: (1) the SMART approach using an averaging method to assess width while the manual approach measures a single width line, (2) some of the CNCs identified are different from those identified in the manual approach. The mean values for both length ($\bar{L}_{\mathrm{SMART}}$ = 80.1 nm) and width ($\bar{W}_{\mathrm{SMART}}$ = 7.7 nm) are very similar to the manual measurements ($\bar{L}_{\mathrm{manual}}$ = 82.5 nm and $\bar{W}_{\mathrm{manual}}$ = 7.7 nm). For the 434 TEM images analyzed, the ratio of total CNC area over the image area was 5.6%, and of this CNC area, 13.4% belonged to isolated CNCs (i.e., particle that were measured), indicating that the countable isolated CNCs are only a small portion of all available CNCs on the sample surface.

For SMART analysis on the 66 AFM images (analysis completed in groups of 12–14 images at one time) a total of 1403 CNCs were identified and measured, using the selection criteria (aspect ratio greater than 2, major axis length larger than 15 nm and smaller than 300 nm, minor axis length less than 30 nm, width-based ratio lower than 1.5 and height-based ratio higher than 0.7) described earlier. The analysis of 1403 CNCs took 7.5 min on a desktop computer with i7 processor, which is approximately 0.35 s/CNC. An estimate for manual measurements completed by a trained analyst using ImageJ or Gwyddion software is 30 s/CNC. The comparison of the measured CNC dimensions from SMART and manual approaches is presented as histograms in Fig. 13. SMART results were also compared with the second manual analysis and this comparison is provided in the supplementary material, Fig. S8. SMART analysis shows similar size distribution trends with both analysts for both length and height measurements. The mean values for both length ($\bar{L}_{\mathrm{SMART}}$ = 79.9 nm) and height ($\bar{H}_{\mathrm{SMART}}$ = 3.8 nm) are very similar to the manual measurements ($\bar{L}_{\mathrm{manual-1}}$ = 76.3 nm, $\bar{L}_{\mathrm{manual-2}}$ = 78.0 nm, and $\bar{H}_{\mathrm{manual-1}}$ = 3.4 nm, $\bar{H}_{\mathrm{manual-2}}$ = 3.7 nm). For the 66 AFM images analyzed the ratio of total CNC area over the image area was 14.1%, and of this CNC area, 14.1% belonged to isolated CNCs, indicating that the countable isolated CNCs are only a small portion of all available CNCs on the sample surface. Compared to the TEM images, there are more isolated CNCs in the AFM images although they are still a small portion of all available CNCs on the sample surface.

The CNC image analysis performed by the SMART approach for both TEM and AFM images was in good agreement with the manual approaches reported by Jakubek et al. (2018), attesting that the SMART approach is "as good" as the manual approach. It is

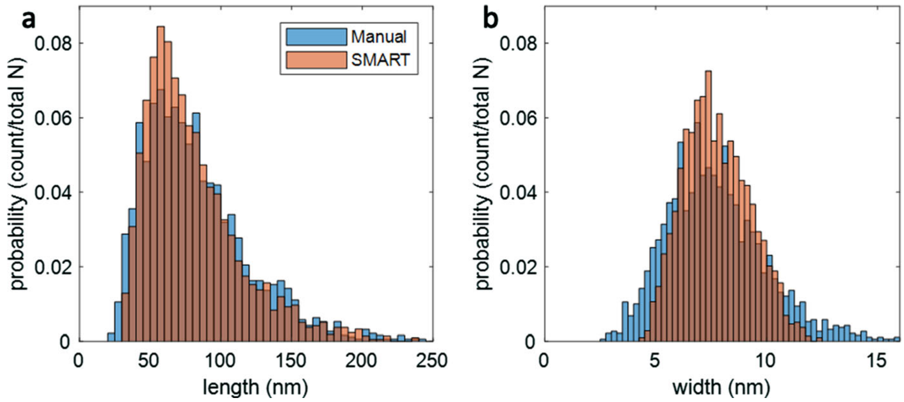



**Fig. 12** TEM image analysis comparison, SMART (N = 2177) versus manual analysis (N = 1909) for the same 434 images. The histograms of SMART and manual analyses are overlaid. **a** Histogram of length measurements, **b** histogram of width measurements. Overlapping regions of histograms are shown in brown

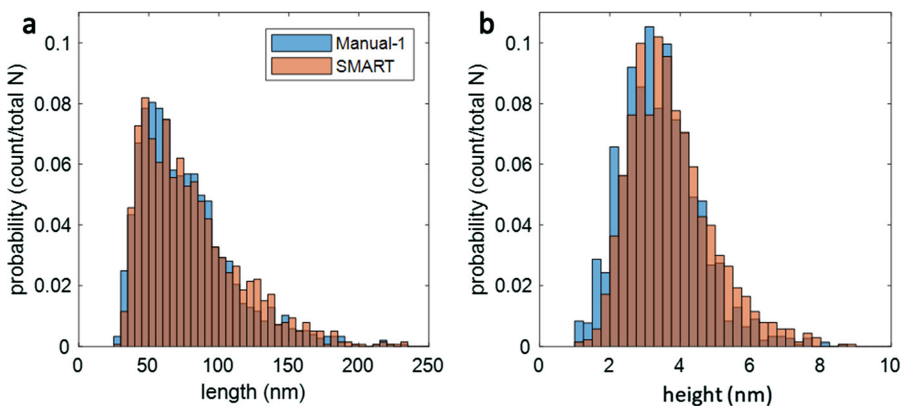



**Fig. 13** AFM image analysis comparison, SMART (N = 1403) versus manual analysis (N1 = 1567) for the same 66 images. **a** Histogram comparison of length measurements, **b** Histogram comparison of height measurements. Overlapping regions of histograms are shown in brown

interesting to note that despite the large differences in image contrast and resolution between TEM and AFM, the small adjustments to the SMART approach were able to accommodate these differences and still identify and measure CNCs as well as the manual approaches. One notable difference between the TEM and AFM image analysis by the SMART approach is the shorter time per CNC required for the AFM image analysis. Two possible reasons for this are, (1) the higher contrast and lower image resolution of the AFM images required less computational time for the SMART approach, and (2) fewer AFM images had to be analyzed due to the higher overall CNC density per image.

The SMART approach indicates that the TEM and AFM measured CNC mean lengths are in excellent agreement, although the AFM measured mean length was lower than the TEM measured mean length in the initial characterization study (Jakubek et al. 2018). The same mean CNC length was also measured in a recent interlaboratory comparison (Bushell et al. 2021). The observation of a larger TEM width than AFM height has been attributed to the presence of laterally aggregated CNCs that are not detected in the TEM images. A detailed discussion of the differences between AFM and TEM measurements is beyond the scope of the current study; the reader is referred to previously published papers using manual analysis (Bushell et al. 2021; Jakubek et al. 2018; Meija et al. 2020).

## Conclusions

A comprehensive image processing system, SMART, was introduced for the semi-automatic analysis of the CNC morphology from TEM and AFM images. SMART system utilizes a multistep approach that includes: pre-processing to enhance the image, groupings to identify specific classes of CNCs, and computationally efficient algorithms for digital measurements of the CNC dimensions. The proposed protocols provide consistent and reliable results and can in principle remove some of the subjectivity associated with deciding whether a feature is one or two particles. The SMART system's performance and accuracy was validated in multiple comparisons using previously reported data sets of TEM and AFM images of a CNC reference material. When comparing measurements on the same set of CNCs, the SMART system measurements were very similar to that of the trained human analyst. The sensitivity of parameters in various image analyses functions was systematically evaluated. It was demonstrated that for small to moderate deviations from the "ideal" parameter settings, the effect on the CNC groupings is small, and the effect on the individual CNC dimensions was less than 10%. The case study comparing the SMART versus manual approach of a large TEM and AFM image data sets, 434 and 66 images, respectively, demonstrated that the SMART approach could detect a similar number of individual CNCs, and produce length, width, and height distributions of CNCs that were comparable to those obtained from the manual approach. For TEM images, the SMART measured mean length, length distribution and mean width were the same as the manual measurements, but the width distribution was narrower than that of the manual measurements. For AFM images, the SMART measured mean length, length distribution, mean width and width distribution were the same as the manual measurements. The SMART system completed the analysis in 43 minutes for TEM images, and 7.5 minutes for AFM images, which is estimated to be 25 and 85 times faster than the manual approach using ImageJ or Gwyddion software, respectively. Additionally, the SMART system demonstrated that CNCs accounted for 6% and 14% of the image area for TEM and AFM, respectively and that only approximately 15% of the CNCs were isolated particles.

**Acknowledgments** The authors would like to thank the USFS-Forest Products for funding this research (Grant Number: 18-JV-11111129-040) and the Renewable Biomaterials Institute at Georgia Institute of Technology.

## References

Awan F, Islam MS, Ma Y, Yang C, Shi Z, Berry RM, Tam KC (2018) Cellulose nanocrystal–ZnO nanohybrids for controlling photocatalytic activity and UV protection in cosmetic formulation. ACS Omega 3:12403–12411. https://doi.org/10.1021/acsomega.8b01881

Bespalova Y, Kwon D, Vasanthan N (2017) Surface modification and antimicrobial properties of cellulose nanocrystals. J Appl Polym Sci. https://doi.org/10.1002/app.44789

Boluk Y, Lahiji R, Zhao L, McDermott MT (2011) Suspension viscosities and shape parameter of cellulose nanocrystals (CNC). Colloids Surf A Physicochem Eng Asp 377:297–303

Bradley D, Roth G (2007) Adaptive thresholding using the integral image. J Graph Tools 12:13–21

Brinkmann A, Chen M, Couillard M, Jakubek ZJ, Leng T, Johnston LJ (2016) Correlating cellulose nanocrystal particle size and surface area. Langmuir 32:6105–6114

Bushell M et al (2021) Particle size distributions for cellulose nanocrystals measured by atomic force microscopy: an interlaboratory comparison. Cellulose https://doi.org/10.1007/s10570-020-03618-4

Campano C, Balea A, Blanco Á, Negro C (2020) A reproducible method to characterize the bulk morphology of cellulose nanocrystals and nanofibers by transmission electron microscopy. Cellulose 27:4871–4887

Campbell A, Murray P, Yakushina E, Marshall S, Ion W (2018) New methods for automatic quantification of microstructural features using digital image processing. Mater Des 141:395–406. https://doi.org/10.1016/j.matdes.2017.12.049

Canet-Ferrer J, Coronado E, Forment-Aliaga A, Pinilla-Cienfuegos E (2014) Correction of the tip convolution effects in the imaging of nanostructures studied through scanning force microscopy. Nanotechnology 25:395703

Davies ER (2004) Machine vision: theory, algorithms, practicalities. Elsevier, Amsterdam

Elazzouzi-Hafraoui S, Nishiyama Y, Putaux J-L, Heux L, Dubreuil F, Rochas C (2008) The shape and size distribution of crystalline nanoparticles prepared by acid hydrolysis of native cellulose. Biomacromol 9:57–65

Ferrer A, Pal L, Hubbe M (2017) Nanocellulose in packaging: Advances in barrier layer technologies. Ind Crops Prod 95:574–582

Foster EJ et al (2018) Current characterization methods for cellulose nanomaterials. Chem Soc Rev 47:2609–2679

Goswami J, Haque E, Fox DM, Gilman JW, Holmes GA, Moon RJ, Kalaitzidou K (2019) The effect of cellulose nanocrystal coatings on the glass fiber–epoxy interphase. Materials 12:1951

Heilbronner R, Barrett S (2014) Pre-processing. In: Image Analysis in Earth Sciences. Springer, Berlin. https://doi.org/10.1007/978-3-642-10343-8_4

Honorato-Rios C, Lagerwall JP (2020) Interrogating helical nanorod self-assembly with fractionated cellulose nanocrystal suspensions. Commun Mater 1:1–11

Huang T, Yang G, Tang G (1979) A fast two-dimensional median filtering algorithm. IEEE Trans Acoust 27:13–18

Jain R, Kasturi R, Schunck BG (1995) Machine vision. McGraw-hill, New York

Jakubek ZJ et al (2018) Characterization challenges for a cellulose nanocrystal reference material: dispersion and particle size distributions. J Nanopart Res 20:98. https://doi.org/10.1007/s11051-018-4194-6

Kaushik M, Fraschini C, Chauve G, Putaux J-L, Moores A (2015) Transmission electron microscopy for the characterization of cellulose nanocrystals. In: Maaz K (ed) The transmission electron microscope-theory and applications, INTECH. https://doi.org/10.5772/60985

Li M-C, Wu Q, Song K, De Hoop CF, Lee S, Qing Y, Wu Y (2016) Cellulose nanocrystals and polyanionic cellulose as additives in bentonite water-based drilling fluids: rheological modeling and filtration mechanisms. Ind Eng Chem Res 55:133–143

Mao Y, Liu K, Zhan C, Geng L, Chu B, Hsiao BS (2017) Characterization of nanocellulose using small-angle neutron, X-ray, and dynamic light scattering techniques. J Phys Chem B 121:1340–1351

Martínez-Sanz M, Lopez-Rubio A, Lagaron JM (2013) High-barrier coated bacterial cellulose nanowhiskers films with reduced moisture sensitivity. Carbohydr Polym 98:1072–1082

MATLAB (2020) version R2020a. The MathWorks Inc. https://www.mathworks.com/products/matlab.html

Meija J et al (2020) Particle size distributions for cellulose nanocrystals measured by transmission electron microscopy: an interlaboratory comparison. Anal Chem 92:13434–13442. https://doi.org/10.1021/acs.analchem.0c02805

Moon RJ, Martini A, Nairn J, Simonsen J, Youngblood J (2011) Cellulose nanomaterials review: structure, properties and nanocomposites. Chem Soc Rev 40:3941–3994

Mukherjee A, Hackley VA (2018) Separation and characterization of cellulose nanocrystals by multi-detector asymmetrical-flow field-flow fractionation. Analyst 143:731–740

Necas D, Klapetek P (2012) Gwyddion: an open-source software for SPM data analysis. Cent Eur J Phys 10:181–188. https://doi.org/10.2478/s11534-011-0096-2

Ogawa Y, Putaux J-L (2019) Transmission electron microscopy of cellulose. Part 2: technical and practical aspects. Cellulose 26:17–34

Postek MT et al (2010) Development of the metrology and imaging of cellulose nanocrystals. Meas Sci Technol 22:024005

Reid MS, Villalobos M, Cranston ED (2017) Benchmarking cellulose nanocrystals: from the laboratory to industrial production. Langmuir 33:1583–1598

Sacui IA et al (2014) Comparison of the properties of cellulose nanocrystals and cellulose nanofibrils isolated from bacteria, tunicate, and wood processed using acid, enzymatic, mechanical, and oxidative methods. ACS Appl Mater Interfaces 6:6127–6138

Salem AA, Shaffer JB, Satko DP, Semiatin SL, Kalidindi SR (2014) Workflow for integrating mesoscale heterogeneities in materials structure with process simulation of titanium alloys. Integr Mater Manuf Innov 3:24. https://doi.org/10.1186/s40192-014-0024-6

Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9:671. https://doi.org/10.1038/nmeth.2089

Sokolov P, Belousov M, Bondarev SA, Zhouravleva GA, Kasyanenko N (2017) FibrilJ: ImageJ plugin for fibrils' diameter and persistence length determination. Comput Phys Commun 214:199–206

Stinson-Bagby KL, Roberts R, Foster EJ (2018) Effective cellulose nanocrystal imaging using transmission electron microscopy. Carbohydr Polym 186:429–438

Sun W, Chen T, Chen C, Li J (2007) A study on membrane morphology by digital image processing. J Membr Sci 305:93–102. https://doi.org/10.1016/j.memsci.2007.07.040

Tahoces PG, Messina PV, Ruso JM (2019) Quantitative analysis of complex nanocomposites based on straight skeletonization. Colloids Surf A Physicochem Eng Asp 562:71–78. https://doi.org/10.1016/j.colsurfa.2018.11.027

Turner DM, Niezgoda SR, Kalidindi SR (2016) Efficient computation of the angularly resolved chord length distributions and lineal path functions in large microstructure datasets. Model Simul Mater Sci Eng 24:075002. https://doi.org/10.1088/0965-0393/24/7/075002

Valentini L, Bon SB, Cardinali M, Fortunati E, Kenny JM (2014) Cellulose nanocrystals thin films as gate dielectric for flexible organic field-effect transistors. Mater Lett 126:55–58

Wargo EA, Schulz VP, Çeçen A, Kalidindi SR, Kumbur EC (2013) Resolving macro- and micro-porous layer interaction in polymer electrolyte fuel cells using focused ion beam and X-ray computed tomography. Electrochim Acta 87:201–212. https://doi.org/10.1016/j.electacta.2012.09.008

Yucel B, Yucel S, Ray A, Duprez L, Kalidindi SR (2020) Mining the correlations between optical micrographs and mechanical properties of cold-rolled HSLA steels using machine learning approaches. Integr Mater Manuf Innov 9:240–256. https://doi.org/10.1007/s40192-020-00183-3

Zhou C, Shi Q, Guo W, Terrell L, Qureshi AT, Hayes DJ, Wu Q (2013a) Electrospun bio-nanocomposite scaffolds for bone tissue engineering by cellulose nanocrystals reinforcing maleic anhydride grafted PLA. ACS Appl Mater Interfaces 5:3847–3854

Zhou Y et al (2013b) Recyclable organic solar cells on cellulose nanocrystal substrates. Sci Rep 3:1–5