

Deep learning-based hybrid reconstruction algorithm for fibre instance segmentation from 3D X-ray tomographic images

Mengqi Fang^a, Aurélien Sibellas^a, James Drummond^a, Yankai Cao^a, Andre Phillion^b, Mark Martinez^a, Vijay Kumar Pediredla^a, Bhushan Gopaluni^a

^aDepartment of Chemical and Biological Engineering, University of British Columbia, Vancouver, V6T 1Z3, BC, Canada

^bDepartment of Materials Science and Engineering, McMaster University, Hamilton, L8S 4L8, ON, Canada

Abstract

3D X-ray tomography is a powerful scanning technique used for generating images of complex fibre structures. A novel machine-learning algorithm to identify and separate individual fibres using 3D images is proposed in this article. The developed four-step hybrid 3D fibre segmentation algorithm involves deep-learning aided semantic segmentation that slices 3D images to create 2D images for fibre extraction, elliptical contour estimation combined with the marker-controlled watershed algorithm for separating fibres from the background area, identifying individual fibres through 3D reconstruction, and, lastly, the 3D object refining approach based on outlier object detection and replacement. The proposed methodology is implemented on a real-time sample of nylon fibre bundle under compression and its 3D X-ray image volume to validate the performance. The results show its superior performance compared to off-the-shelf image processing algorithms in terms of precision, that is, with a validation accuracy greater than 90%, and efficiency, that is, preventing the need for a huge data set and reducing the

complexity.

Keywords: 3D reconstruction, deep learning, fibre segmentation, X-ray tomographic images

1. INTRODUCTION

The quality of paper handsheets manufactured in the pulp and paper industry depends on the characteristic properties of the pulp fibre and the operating conditions of the pulping process [1]. Understanding the effect of the pulping process operating conditions on the fibrous structure of the final product would enable the manufacture of paper with the desired properties. The images are used from 3D X-ray tomography in conjunction with advanced deep learning-based image processing techniques to correlate the micro-structure of the fibres in paper handsheets with the process operating conditions. An efficient 3D image segmentation algorithm is proposed and implemented to solve this problem. The study focuses on nylon-fibre bundles to initiate the primitive study. Nylon fibres are synthetic fibres made of polymers that are close contrasts to the wood fibres in the images allowing them to deal with similar levels of noise and artifacts to the paper handsheets but on simpler shapes.

In recent years, there have been significant advances in image processing algorithms and tools, particularly those that use advanced deep learning algorithms. These algorithms often provide highly accurate models for image processing. Our goal is to use these advanced algorithms for image segmentation to identify specific properties of fibres in paper handsheets. This information can eventually be used to determine appropriate process operating conditions. For example, a soft sensor can be developed to infer the correlations between process operating con-

ditions and paper handsheet properties using handsheet images. This way, process operating conditions can be optimized to obtain desired product qualities. Image segmentation comprises semantic and instance segmentation, where the latter extends the former by further separating individual objects [2]. Many existing deep learning segmentation algorithms are developed using basic convolutional neural networks (CNN), such as the widely known AlexNet [3], ResNet [4], GoogLeNet [5], and so on. Moreover, in order to incorporate specific practical considerations, various extensions of the basic CNN structure have been implemented. Examples include the recurrent neural network (RNN) [6] and its improved version, long short term memory network (LSTM) [7], generative adversarial networks [8], fully convolutional networks [9], and so on. Most recently, motivated by the successful applications of transformers in the natural language processing field, vision transformer [10] and its variants [11] have been applied on image recognition tasks and demonstrated better performance than convolutional models on a very large dataset. Many of these algorithms can be adapted to different environments and perform well in describing complicated real-world scenarios. However, these algorithms generally result in extreme model complexity, and require huge datasets and significant maintenance efforts after a model is built.

Some existing literature focuses on fibre segmentation and 3D micro-structure reconstruction from tomographic image volumes. For example, Agyei et al. [12] proposed a four-step sequential 2D segmentation approach, followed by a 3D volume rendering algorithm for object matching. Viguie et al. [13] designed an image analysis method to identify the fibres with irregular cross sections and quantify the fibre contacts. Emerson et al. [14] developed a centre point detection and tracking approach to segment individual fibres from X-ray 3D tomography. To the best

of the authors' knowledge, deep learning-based segmentation algorithms have not yet been used in the X-ray based nylon fibre segmentation problems, motivating us to develop a hybrid 3D reconstruction algorithm.

Using efficiency and simplicity as criteria, the encoder–decoder structured deep neural networks, called U-Nets [15], are chosen to perform semantic segmentation. Other deep learning-based semantic segmentation algorithms, such as Mask R-CNN [16] and Faster R-CNN [17], require heavier annotation, training, and tuning workloads. However, for post-processing the segmentation results, the traditional instance segmentation algorithms, such as connected component analysis [18] and watershed segmentation [19], are considered. Therefore, the proposed work aims to develop an efficient segmentation algorithm that addresses these challenges.

A four-step hybrid 3D fibre segmentation algorithm that makes use of both deep learning and conventional image processing algorithms for image segmentation is proposed in this article. Each of these steps has challenges associated with training data set creation, instance segmentation, 3D reconstruction, and refining. To address these challenges, some improvements to the existing algorithms are proposed that include augmentation of the U-Net training data set and development of an improved version based on a conventional marker-controlled watershed algorithm. More importantly, motivated by the centroid object tracking algorithm, a novel multivariate Gaussian modelling and Kullback–Leibler (KL) divergence oriented fibre tracking algorithm is developed, followed by a novel 3D object refining approach based on outlier detection and replacement.

The rest of this paper is organized as follows. Section 2 highlights the 3D image data set and the challenges associated with fibre segmentation. Section 3

illustrates the four steps in the proposed algorithm for segmentation, and Section 4 demonstrates the superior performance of the proposed algorithm on a compressed nylon fibre bundle sample.

2. PROBLEM STATEMENT AND CHALLENGES

2.1. 3D individual nylon fibre segmentation

A typical nylon fibre bundle sample studied in this work is shown in Figure 1. It shows both the integrated 3D volume and the 2D slice representations of the same bundle. These images are generated using a non-destructive X-ray tomography technique to reveal the internal fibre microstructure [12]. A 3D image can be decomposed into a series of 2D slices through any of the x , y , or z axes. For instance, Figure 1B is sliced along z axis, where the ellipse-like structures denote the fibre cross sections.

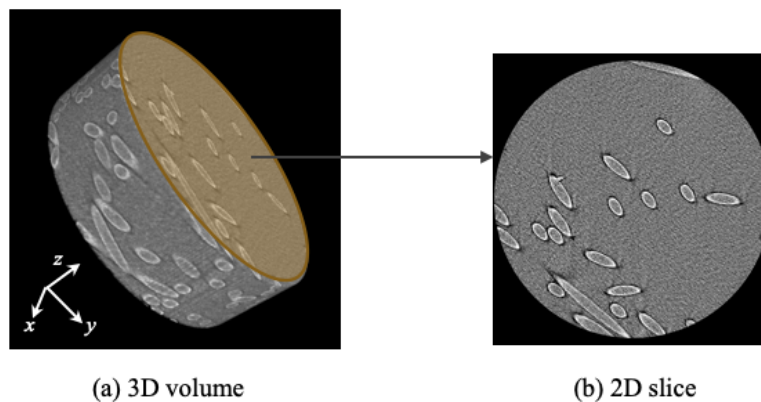


Figure 1: An illustration of 3D volume and a 2D slice of the same nylon fibre bundle sample.

In the X-ray tomography images, the nylon fibres are embedded in a grey background, from which the nylon fibres need to be initially extracted and then

separated from each other. Good quality images with high contrasts (distinct grey values for the different phases, e.g., air and fibres) are challenging to obtain for such light and porous structures. Moreover, the effect of beam hardening creates black rings and shadows on fibres' cross-sections, preventing the utilization of an easy classical intensity-based segmentation. The nylon fibres are uniform in concentration. The variation in grayscale across their diameter results from beam hardening, an artifact of image acquisition. In order to accurately identify and segment the nylon fibres, a 2D fibre segmentation is performed on all single 2D slices, followed by a 3D reconstruction using the 2D segmentation results.

2.2. Challenges

As mentioned in the introduction section, deep learning algorithms can provide highly automatic and accurate instance segmentation on some real-world problems. However, in this fibre segmentation problem, the efficiency of the advanced deep learning algorithms is restricted by several factors listed below, which result in degraded performance. Therefore, instead of solely relying on the deep learning approaches, a hybrid algorithm combining both deep learning and traditional machine learning algorithms is proposed. The developed methodology addresses several challenges unique to the segmentation of fibres in tomographic images:

- *Lack of labelled samples*: Each 3D image is sliced into hundreds of 2D tomograms, and therefore, manual annotations on any 3D sample will require significant effort. This restricts the use of deep learning algorithms that need a large, fully annotated training data set.
- *Limited number of 3D tomographic samples*: The lack of a sufficiently large

data set to train a full-fledged deep learning model. This challenge can be addressed through transfer learning algorithms such as VGGNet [20]. However, these algorithms must be initially trained by big open-source data sets and then fine-tuned with relevant smaller data sets. There are no existing labelled large data-sets of tomographic images that can be utilized to train such a network.

- *Low accuracy of deep learning-based instance segmentation:* The dense fibre occurrence, arbitrary orientations, and locations of fibres make the training of deep neural networks, such as Mask-RCNN [16], extremely difficult. Deep neural networks will take several hours to train and provide lower segmentation accuracy than traditional approaches, such as watershed segmentation [21].
- *3D reconstruction errors:* The errors in 2D tomogram instance segmentation will also impact the 3D reconstruction performance.

We address the above challenges through a hybrid 3D fibre instance segmentation algorithm that integrates deep learning and traditional image segmentation algorithms. The detailed workflow of this algorithm is explained and illustrated in the next section.

3. THE PROPOSED METHODOLOGY

Our ultimate objective is to automatically segment and label the individual nylon fibres in the 3D images. Starting with the 2D tomogram slices as inputs, four sequential image processing blocks are developed as shown in Figure 2. These

include: (1) deep learning aided semantic segmentation for fibre extraction; (2) elliptical contour estimation combined with marker-controlled watershed algorithm for fibre separation; (3) Gaussian modelling and distribution similarity based 3D reconstruction; and (4) outlier detection and replacement to accommodate for the 2D segmentation errors. Among these steps, supervised learning occurs only in the first step, and the remaining steps do not require labelled data sets. The following section explains the four sequential function blocks in detail.

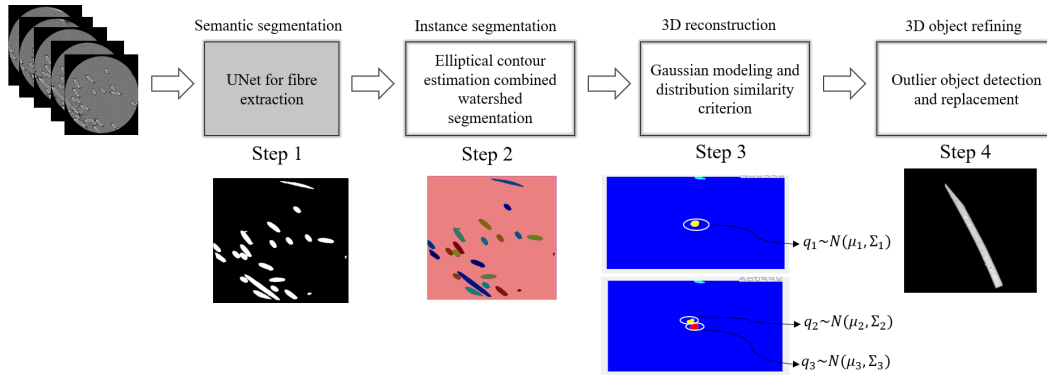


Figure 2: The workflow of the proposed methodology, where the shaded block is based on supervised learning and others are unsupervised learning.

3.1. U-Net aided semantic segmentation

U-Net is a simpler version of convolutional neural networks (CNNs) and has been widely used to segment biomedical images that share similar texture with 2D nylon fibre tomograms. The U-Net model is composed of two major components, namely the encoder and decoder. The encoder half of the model is utilized to identify image features that carry out a down-sampling process, creating a lower dimensional feature embedding space using the input image. The decoder half

performs an up-sampling operation to recover the spatial information of the input image before building the model output using the features as input. Due to its features such as simple structure and ability to train efficiently and provide good accuracy, U-Net is selected as the front-end semantic segmentation algorithm for 2D fibre extraction. In this problem, the compressed nylon fibre bundle is the target sample to be segmented. To segment the compressed sample, a U-Net model needs to be trained on fully annotated compressed sample images. However, in our case, only fully labelled uncompressed nylon fibre samples are available. To address this challenge, in addition to the uncompressed samples, partially annotated images from the compressed Nylon fibre bundle are added to the training dataset, in which labels are added manually. The remaining uncompressed nylon fibre images are sent to the test dataset. Labelled samples in the data are denoted by $\mathcal{D}_s^l = \{X_{1:N_s}, Y_{1:N_s}\}$ and the rest are denoted by $\mathcal{D}_t = \{X_{1:N_t}\}$. Here, $X_n \in \mathbb{R}^{m_1 \times m_2}$ denotes the n^{th} 2D tomogram, which consists of image slices and $Y_n \in \mathbb{R}^{m_1 \times m_2}$ is the corresponding pixel-wise label map, with 1 indicating nylon and 0 indicating the background.

The characteristic features of X-ray tomographic images depend on experimental conditions and the sampling procedure used. As such, 2D tomogram images of only certain nylon fibre specimens can be easily segmented by naive thresholding approaches. The carefully selected data set \mathcal{D}_s belongs to this category, and the labels $Y_{1:N_s}$ are generated using a simple thresholding method. In contrast, the target data set \mathcal{D}_t is difficult to segment by applying any thresholding algorithms. The following two steps are used to create an effective data set for the U-Net model for training:

- Step 1: Scale the images in \mathcal{D}_s based on the mean and standard deviation of

the image grayscale intensity in \mathcal{D}_t as shown in Equation (1), thus creating new scaled data set $\mathcal{D}_S = \{Z_{1:N_s}, Y_{1:N_s}\}$.

$$Z_{n_s} = \frac{X_{n_s} - E(X_{1:N_s})}{std(X_{1:N_s})} \cdot std(X_{1:N_t}) + E(X_{1:N_t}) \quad (1)$$

where $X_{n_s} \in \mathcal{D}_s$ and $Z_{n_s} \in \mathcal{D}_S$, $n_s = 1, \dots, N_s$, represent the original and converted image grayscale intensities, respectively; $X_{n_t} \in \mathcal{D}_t$, $n_t = 1, \dots, N_t$, denotes the image grayscale intensity in the target data set; and $E(\cdot)$ and $std(\cdot)$ denote the mean and standard deviation operations, respectively. The advantage of scaling the \mathcal{D}_s is to guarantee the training and test data sets follow the same distribution.

- Step 2: Manually annotate n_t images, $n_t \ll N_t$, in the target data set, and perform image augmentation on these labelled images, forming a new data set $\mathcal{D}_T = \{X_{1:N_a}, Y_{1:N_a}\}$.

Then, we create an integrated new data set for U-Net training by combining \mathcal{D}_S and \mathcal{D}_T , after which the trained U-Net model is employed to label the entire target data set \mathcal{D}_t .

3.2. Elliptical contour estimation with watershed algorithm for instance segmentation

The U-Net model uses the grayscale 2D image slices and generates a categorical output that separates fibres from their background areas. Once the binarized images are created, the next step is to separate the fibres from each other through a process called instance segmentation. The elliptical fused objects are rather difficult to separate compared to circular fused objects using conventional marker-controlled watershed algorithms [22]. This observation is illustrated in the first two sub-figures of Figure 3.

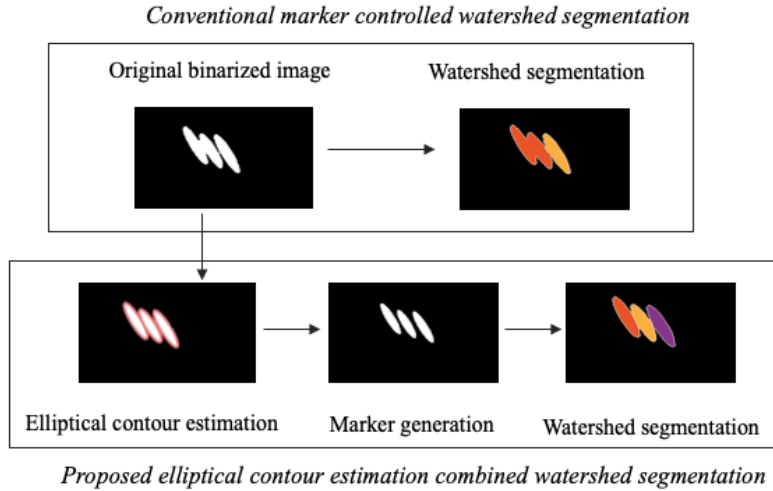


Figure 3: An illustrative comparison of the conventional and proposed marker-controlled watershed instance segmentation.

As an alternative, the concave points supported elliptical contour estimation algorithm [23] that provides better separation between two closely connected elliptical fibres is used in this methodology. However, the estimated elliptical contours sometimes share small overlapped areas between each other. In order to obtain smoother instance segmentation, a simple but effective combination of the following three steps is proposed: (1) apply the elliptical contour estimation algorithm on the original figure to generate multiple elliptical contours; (2) reduce the size of the elliptical contours until they appear as separate objects; and (3) use the obtained markers in step 2 and apply the marker-controlled watershed segmentation. As demonstrated in the lower three sub-figures of Figure 3, in most cases, the estimated elliptical contours can provide reliable marker locations, resulting in higher accuracy of the watershed segmentation algorithm compared to the existing non-elliptical contour-based segmentation algorithms [14]. Although the

elliptical contour estimation recognizes shapes such as circles and ovals, it cannot differentiate when the cross-sections are close to each other. Therefore, steps 3 and 4 are proposed to address these concerns.

3.3. Gaussian modelling based 3D reconstruction

In the 3D reconstruction functional block of Figure 2, the 2D slices with separated nylon fibres are integrated into 3D space by matching and linking the labelled cross sections belonging to the same nylon fibre. The segmented 2D image slices are stitched together to generate a 3D image while accounting for the 3D spatial and geometric layout of the nylon fibre bundle. The 3D reconstruction involves tracking the location of every fibre that appears in two adjacent 2D slices and identifying new fibres in those slices. There are several existing object-tracking algorithms that can be used to track the location of the same fibres between adjacent 2D slices. Among them, the centre point matching algorithm has been successfully applied to glass and carbon fibre tracking due to its simplicity, high processing speed, and satisfactory accuracy [14]. However, without considering the shape and size of individual nylon fibres and focusing only on the centre, it tends to lose its accuracy when the nylon fibres on each slice are densely distributed or when there exist instance segmentation biases from the previous watershed segmentation block.

In order to compensate for the weakness of the centroid matching algorithm, a novel fibre tracking approach that uses a Gaussian probability distribution model is proposed. Samples from a two-dimensional multivariate Gaussian distribution automatically fall in an elliptical shape. The centre of the ellipse is the mean value of the Gaussian distribution, and the corresponding covariance matrix defines the shape and orientation of the ellipse. Therefore, each separated fibre object is

modelled with a Gaussian distribution, as follows:

$$\{x_i^{(l)}, y_i^{(l)}\}_{i=1}^m \sim N(\boldsymbol{\mu}^{(l)}, \boldsymbol{\Sigma}^{(l)}) \quad (2)$$

where l represents the l^{th} detected nylon fibre object on the current slice, and $\{x_i^{(l)}, y_i^{(l)}\}_{i=1}^m$ indicates the total m pixel coordinates of this fibre object. $\boldsymbol{\mu}^{(l)}$ and $\boldsymbol{\Sigma}^{(l)}$ are the Gaussian parameters that are estimated by maximizing the following log-likelihood function.

$$\begin{aligned} \boldsymbol{\mu}^{(l)}, \boldsymbol{\Sigma}^{(l)} &= \arg \max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} \log P(x_{1:m}^{(l)}, y_{1:m}^{(l)}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) \\ &= \arg \max_{\boldsymbol{\mu}, \boldsymbol{\Sigma}} \sum_{i=1}^m \left\{ -\log 2\pi \sqrt{|\boldsymbol{\Sigma}|} - \frac{1}{2} \cdot \right. \\ &\quad \left. ([x_i^{(l)}, y_i^{(l)}]^T - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} ([x_i^{(l)}, y_i^{(l)}]^T - \boldsymbol{\mu}) \right\} \end{aligned} \quad (3)$$

As a result, every individual fibre object is compressed into a Gaussian distribution on each 2D slice. When performing object tracking in a subsequent slice, instead of using Euclidean distance to trace the centroid movement of a fibre, we use KL divergence as a matching criterion by measuring the similarity between two Gaussian distributions. Given two fibre objects with distributions $q_1(\mathbf{x}^{(1)}, \mathbf{y}^{(1)}) \sim N(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}^{(1)})$ and $q_2(\mathbf{x}^{(2)}, \mathbf{y}^{(2)}) \sim N(\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}^{(2)})$, the KL divergence is computed as follows:

$$\begin{aligned} D_{KL}(q_2||q_1) &= \frac{1}{2} \left[\log \frac{|\boldsymbol{\Sigma}^{(2)}|}{|\boldsymbol{\Sigma}^{(1)}|} - 2 + \text{tr}(\{\boldsymbol{\Sigma}^{(2)}\}^{-1} \boldsymbol{\Sigma}^{(1)}) + \right. \\ &\quad \left. (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)})^T \cdot \{\boldsymbol{\Sigma}^{(2)}\}^{-1} (\boldsymbol{\mu}^{(2)} - \boldsymbol{\mu}^{(1)}) \right] \end{aligned} \quad (4)$$

While traversing the fibre objects in the current 2D slice, the KL divergence criterion between the current object distribution and the distributions of suspected objects in the previous slices are computed. When the target object with minimal

KL divergence is identified, and the calculated D_{KL} is less than a predefined small threshold τ , this identified fibre object is connected to the one on the current slice. Therefore, q_1 and q_2 fibre objects are coming from the same fibre, as shown in step 3 of Figure 2.

3.4. 3D object refining based on probability density function

The Gaussian distribution and KL divergence contain a lot of information of the segmented nylon fibres. As explained in Section 3.3, once the 3D reconstruction is completed, the individual fibres in the 3D images can be identified. However, because of the inaccuracies in the instance segmentation, some extracted 3D fibres may include attached parts of adjacent fibres. To fix this problem, an outlier detection and replacement approach is developed in the final step of the workflow in Figure 2.

Once the fibres are extracted from a 3D image, an indicator of the fibre cross-section area in every slice is computed and stored as $\mathbf{A}^{(l)} = [A_{k_1}^{(l)}, A_{k_2}^{(l)}, \dots, A_{k_V}^{(l)}] \in \mathbb{R}^{k_V}$, where l denotes the l^{th} nylon fibre and $\{k_1, \dots, k_V\}$ are the 2D slice identities containing this object. The individual element $A_{k_v}^{(l)}$ in array $\mathbf{A}^{(l)}$ is computed as follows:

$$A_{k_v}^{(l)} = \text{tr}(\Sigma_{k_v}^{(l)}), \quad v = 1, 2, \dots, V \quad (5)$$

where $\Sigma_{k_v}^{(l)}$ is the estimated covariance matrix of the fibre cross section on the k_v^{th} 2D slice.

Normally, considering the continuous extension of nylon fibre, the variation of fibre cross-section area is expected to be smooth. Any abrupt increment in the area implies a potential occurrence of additional fibre pieces introduced by erroneous instance segmentation. Compared with the actual area of the fibre cross-section,

the above indicator can better highlight the abnormal increase in the fibre cross-section areas. Therefore, it is selected for outlier detection. Based on the indicator $\mathbf{A}^{(l)}$, the upper quartile, denoted by τ_u , is used to formulate the detection threshold. Because $\mathbf{A}^{(l)}$ sometimes exhibits a non-stationary characteristic, the outliers in $\mathbf{A}^{(l)}$ are finally detected by satisfying the following two conditions: We can detect the fibre cross-section changing rapidly, which contains redundant components of adjacent fibres. The detection threshold is chosen to be 1.5 times the upper quartile based on the empirical analysis.

$$A_{k_v}^{(l)} > 1.5\tau_u, \quad \Delta A_{k_v}^{(l)} > std(A_q), \quad v = 1, 2, \dots, V \quad (6)$$

where $\Delta A_{k_v}^{(l)}$ denotes the increment of the detected outlier compared with the nearest normal point, and A_q includes all the $\mathbf{A}^{(l)}$ elements within the upper and lower quartile range.

When outliers are detected, the relevant slices are labelled as erroneous instance segmentation and need further corrections. In order to trim the redundant fibre pieces, linear interpolation is performed on the mean values of normal fibre objects. According to the correlation between k_v and $\mu_{k_v}^{(l)} = [\mu_{k_v,x}^{(l)}, \mu_{k_v,y}^{(l)}]^T$, for $v = 1, \dots, V$, a linear model can be formulated as in Equation (7).

$$\begin{bmatrix} \mu_{k_v,x}^{(l)} \\ \mu_{k_v,y}^{(l)} \end{bmatrix} = \begin{bmatrix} a_x & b_x \\ a_y & b_y \end{bmatrix} \cdot \begin{bmatrix} k_v \\ 1 \end{bmatrix}, \quad for \ v = 1, \dots, V \quad (7)$$

where the model parameters a_x , b_x , a_y , and b_y are estimated from the normal data set $\{k_{v_n}, \mu_{k_{v_n}}^{(l)}\}_{v_n \in \mathbf{V}_{normal}}$.

Based on the above linear model, the centroids of the wrongly segmented fibre pieces can be estimated using their slice identities $\{k_{v_m}\}_{v_m \in \mathbf{V}_{abnormal}}$, and the corresponding covariance matrix $\Sigma_{k_{v_m}}^{(l)}$ value is computed by averaging the Σ values

of the previous several normal slices. After obtaining the pseudo Gaussian model $N(\mu_{k_{vm}}^{(l)}, \Sigma_{k_{vm}}^{(l)})$, the probability density function values of the object coordinates are computed and the ones with extremely low probability density are removed and labelled as background.

The proposed methodology for segmenting and labelling the individual nylon fibres from the 3D X-ray tomographic images is as follows. First, U-Net aided semantic segmentation is used for 2D fibre extraction that labels the target data set. Second, instance segmentation is performed through elliptical contour-based estimation with a watershed algorithm for accurately separating the fibres from each other. Third, the same nylon fibres are integrated into 3D space through a novel 3D reconstruction by Gaussian modelling and distribution. Last, the 3D object refining method is implemented on fibres to eliminate the attached parts of adjacent fibres. In summary, each step in the developed hybrid workflow combining deep-learning and image-processing algorithms enhances the precision of fibre segmentation from 3D X-ray tomographic images.

4. PERFORMANCE VALIDATION ON A CASE STUDY

In this section, a nylon fibre bundle under compression is selected as our sample, and its corresponding 3D X-ray image volume is processed according to the proposed methodology for performance validation. The generated 3D image volume has the dimension of $1013 \times 964 \times 500$ voxels. The resolution is 1 micron, so its physical volume is about $1 \times 1 \times 0.5 \text{ mm}^3$. After slicing along the z axis, 500 2D tomogram slices with dimension 1013×964 pixels are produced. Initially, a randomly selected slice in Figure 4A is analyzed, and the grayscale intensity histograms of the nylon fibres and background, excluding the black area, are depicted

in Figure 4B. It can be observed from this comparison that the grayscale intensity distribution of the nylon fibres significantly overlaps with the background intensity distribution and thus makes it difficult to find an appropriate threshold for segmentation.

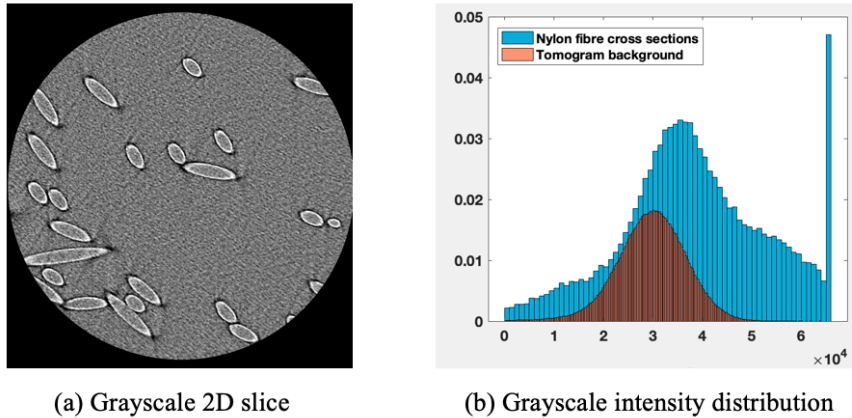


Figure 4: A comparison of grayscale intensity histograms between nylon fibres and the background.

As a result, a U-Net model is initialized with input size 512×512 , four downblocks, and four upblocks, where the kernel size is 3×3 , and activation functions are chosen as ReLU. According to the complexity of the tomography image in this work, the number of convolutional filters of four downblocks, bottleneck block, and four upblocks were chosen to be [16, 32, 64, 128], 256, and [128, 64, 32, 16], respectively. The total trainable number of parameters is 1,962,625. Following the training procedures explained in Section 3.1, the selected U-Net model is well-trained after 20 epoch iterations. The training and validation split rate is 85% and 15% respectively [24; 25]. Finally, the training and validation performances are summarized in Table 1.

Table 1: U-Net training and validation performance.

Training		Validation	
Accuracy	Loss	Accuracy	Loss
97.22%	0.0262	97.50%	0.0313

The U-Net provides semantic segmentation results on all the 500 2D tomograms, with processing times as low as less than 1 s per slice. While completing the fibre extraction, the elliptical contour estimation algorithm is applied on all the binarized 2D images to generate markers for watershed segmentation. In order to demonstrate the performance, the number of fibres on all 500 slices are counted manually and compared with the instance number given by watershed segmentation. The segmentation accuracy of the proposed elliptical contour estimation based watershed is compared with the conventional watershed algorithm. As shown in Figure 5, the proposed watershed algorithm is able to provide higher and more stable segmentation accuracy than the conventional one [26]. Moreover, after investigating the segmentation accuracy of all the 500 slices, the proposed watershed algorithm offers more than 90% accuracy for 77.6% of the 2D slices, and only 6.2% of the slices have segmentation accuracy less than 85%. Given the inevitable manual annotation error, the proposed watershed approach could give satisfactory accuracy at this step. The error of instance segmentation can also be compensated in the 3D reconstruction and refining procedures.

In the 3D reconstruction step, because of the fibre orientation, fibre tracking from the X direction can be more efficient than from the Z direction. Therefore, in this special case, the binarized 3D image volume is re-sliced from the X direction, generating 1013 slices with dimension 964×500 pixels. The proposed watershed

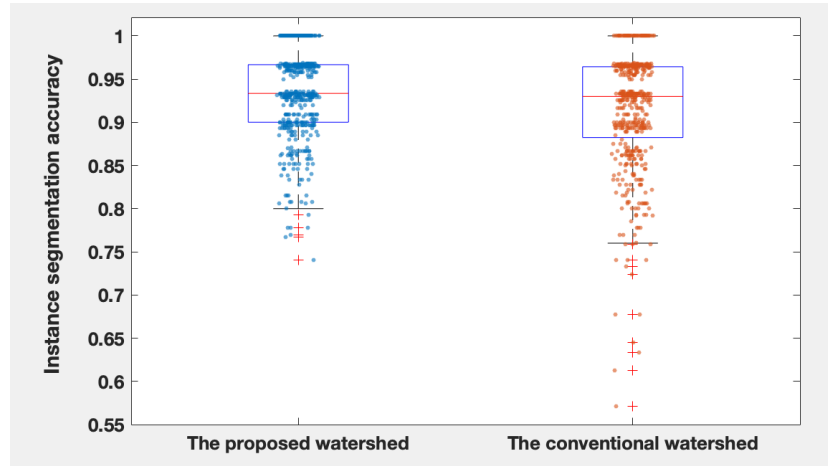


Figure 5: The segmentation accuracy comparison between the proposed and conventional watershed approach.

algorithm is re-applied on each slice to perform instance segmentation. On each slice, every individual object is first represented by a multivariate Gaussian distribution, and then the KL divergence between the current object is compared with all the adjacent objects in the previous six slices. The object in the previous slices with the minimal KL divergence is selected as a potential candidate to connect if the minimal KL divergence is less than 1. Otherwise, the current object will be assigned a new label and treated as a recently emerged one that never appeared before.

After conducting the 3D reconstruction, several extracted fibres can have extra attached pieces due to erroneous watershed instance segmentation. Figure 6A is a typical example. After detecting the abrupt area change and performing trimming based on the probability density function value, the redundant piece has been removed, as shown in Figure 6B.

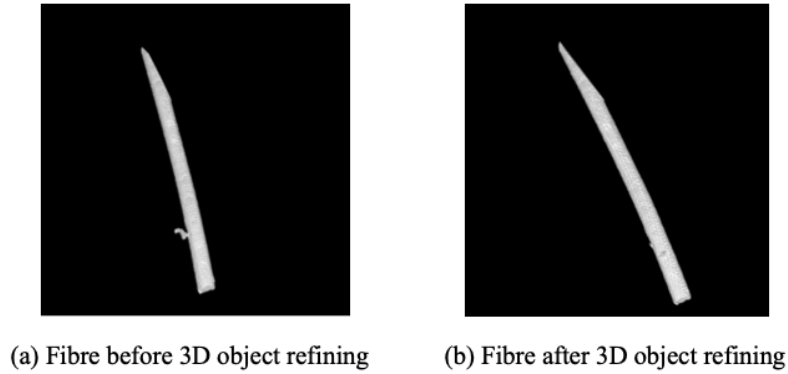


Figure 6: The same fibre before and after 3D object refining.

Finally, after running through the four sequential functional blocks in Figure 2, we have the 3D nylon fibre separation result shown in Figure 7, where one can clearly observe the separated nylon fibres without redundant pieces attached. Integrating deep-learning methods with image processing techniques provided an accurate fibre segmentation even with a lesser training data set.

In summary, since only a limited number of samples exist, none of the 3D instance segmentation approaches can be employed directly. Therefore, the proposed model has its own novelty and efficacy. For the model improvement, the conventional marker-based watershed segmentation is modified by utilizing an elliptical contour estimation. The performance has been compared with the conventional algorithms, and Figure 5 illustrates the efficacy of the proposed approach. Even though there are some similar works on glass fibre segmentation, many of those models started from a thresholding approach to obtain the region of interest, which is not achievable in our case because the nylon fibre and background cannot be separated using this approach. Moreover, the glass fibre orientation

is much more organized than nylon fibres, making the model unusable in this scenario. Therefore, the multivariate Gaussian model is employed for 3D reconstruction followed by novel 3D object refining, which is another novelty resulting in good reconstruction performance, as shown in Figures 6 and 7.

In the future, we plan to investigate two potential avenues for further improvement: (1) a mixture of Gaussian distributions can be used to identify fibres with irregular shapes instead of modelling each object using a single Gaussian distribution; and (2) for better fibre tracking and object refining, fused information can be used from multiple views instead of conducting the reconstruction from a single view of the image.

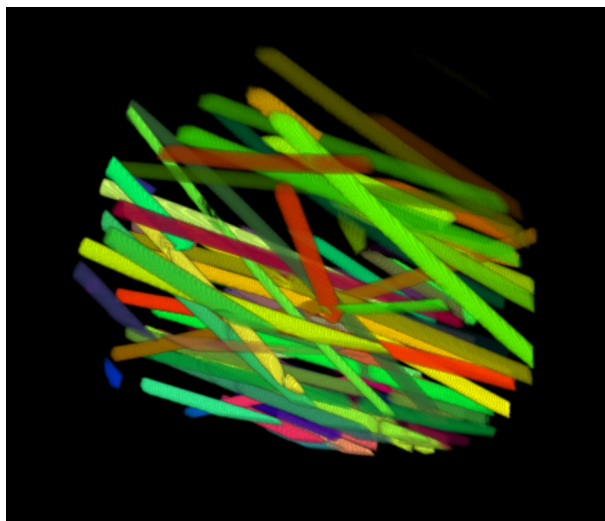


Figure 7: The final 3D nylon fibre segmentation performance.

5. CONCLUSIONS

This paper proposes a 3D hybrid instance segmentation approach to process the 3D tomographic images of nylon fibres. Without loss of generality, we have created a reliable and efficient workflow by considering and improving both deep learning and conventional machine learning algorithms. With the newly proposed Gaussian modelling and KL divergence similarity measurement, a four-step workflow is created consisting of deep-learning aided semantic segmentation, elliptical contour estimation combined with the marker-controlled watershed algorithm, 3D reconstruction, and object refining. The proposed methodology is validated on real-time nylon fibre samples, and the results show satisfactory performance.

References

- [1] R. Wimmer, G. M. Downes, R. Evans, G. Rasmussen, J. French, Direct effects of wood characteristics on pulp and handsheet properties of eucalyptus globulus, *Holzforschung* vol.56, no.3 (2002) 244–252.
- [2] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, D. Terzopoulos, Image segmentation using deep learning: A survey, *IEEE transactions on pattern analysis and machine intelligence* vol.44 (7) (2021) 3523–3542.
- [3] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems* vol.25 (2012) 1097–1105.
- [4] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proc. of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

- [5] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proc. of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.
- [6] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors, *Nature* vol.323 (6088) (1986) 533–536.
- [7] S. Hochreiter, J. Schmidhuber, Long short-term memory, *Neural computation* vol.9 (8) (1997) 1735–1780.
- [8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, *Advances in neural information processing systems* vol.27 (2014).
- [9] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proc. of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.
- [10] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, *arXiv preprint arXiv:2010.11929* (2020), (accessed: January 20, 2023).
- [11] A. Hassani, S. Walton, N. Shah, A. Abuduweili, J. Li, H. Shi, Escaping the big data paradigm with compact transformers, *arXiv preprint arXiv:2104.05704* (2021), (accessed: January 20, 2023).

- [12] R. F. Agyei, M. D. Sangid, A supervised iterative approach to 3d microstructure reconstruction from acquired tomographic data of heterogeneous fibrous systems, *Composite Structures* 206 (2018) 234–246.
- [13] J. Viguié, P. Latil, L. Orgéas, P. Dumont, S. R. du Roscoat, J.-F. Bloch, C. Marulier, O. Guiraud, Finding fibres and their contacts within 3d images of disordered fibrous media, *Composites Science and Technology* vol.89 (2013) 202–210.
- [14] M. J. Emerson, K. M. Jespersen, A. B. Dahl, K. Conradsen, L. P. Mikkelsen, Individual fibre segmentation from 3d x-ray computed tomography for characterising the fibre orientation in unidirectional composite materials, *Composites Part A: Applied Science and Manufacturing* 97 (2017) 83–92.
- [15] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer, 2015, pp. 234–241.
- [16] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask r-cnn, in: *Proc. of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [17] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, *Advances in neural information processing systems* vol.28 (2015) 91–99.
- [18] B.-B. Chai, J. Vass, X. Zhuang, Significance-linked connected component analysis for wavelet image coding, *IEEE Trans. Image Processing* vol.8 (6) (1999) 774–784.

- [19] L. Shafarenko, M. Petrou, J. Kittler, Automatic watershed segmentation of randomly textured color images, *IEEE Trans. Image Processings* vol.6 (11) (1997) 1530–1544.
- [20] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556* (2014), (accessed: January 20, 2023).
- [21] K. Parvati, P. Rao, M. Mariya Das, Image segmentation using gray-scale morphology and marker-controlled watershed transformation, *Discrete Dynamics in Nature and Society* vol.2008 (2008) 1–8.
- [22] M. Faessel, F. Courtois, Touching grain kernels separation by gap-filling, *Image Analysis & Stereology* 28 (3) (2009) 195–203.
- [23] S. Zafari, T. Eerola, J. Sampo, H. Kälviäinen, H. Haario, Segmentation of partially overlapping nanoparticles using concave points, in: *Advances in Visual Computing: 11th International Symposium, ISVC 2015, December 14-16, 2015, Proceedings, Part I 11*, Springer, 2015, pp. 187–197.
- [24] F. Chollet, *Deep learning with Python*, Simon and Schuster, 2021.
- [25] V. R. Joseph, Optimal ratio for data splitting, *Statistical Analysis and Data Mining: The ASA Data Science Journal* vol.15 (4) (2022) 531–538.
- [26] A. Bieniek, A. Moga, An efficient watershed algorithm based on connected components, *Pattern recognition* vol.33 (6) (2000) 907–916.