

Self-Supervised Orientation-Guided Deep Network for Segmentation of Carbon Nanotubes in SEM Imagery

Nguyen P. Nguyen¹, Ramakrishna Surya², Matthew Maschmann², Prasad Calyam¹, Kannappan Palaniappan¹, Filiz Bunyak^{1*}

¹Department of Electrical Engineering and Computer Science

²Department of Mechanical and Aerospace Engineering

University of Missouri-Columbia, MO, USA

Email: {npntz3, rst7b}@mail.missouri.edu

{maschmannm, calyam, pal, bunyak}@missouri.edu

Abstract. Electron microscopy images of carbon nanotube (CNT) forests are difficult to segment due to the long and thin nature of the CNTs; density of the CNT forests resulting in CNTs touching, crossing, and occluding each other; and low signal-to-noise ratio electron microscopy imagery. In addition, due to image complexity, it is not feasible to prepare training segmentation masks. In this paper, we propose *CNTSegNet*, a dual loss, orientation-guided, self-supervised, deep learning network for CNT forest segmentation in scanning electron microscopy (SEM) images. Our training labels consist of weak segmentation labels produced by intensity thresholding of the raw SEM images and self labels produced by estimating orientation distribution of CNTs in these raw images. The proposed network extends a U-net-like encoder-decoder architecture with a novel two-component loss function. The first component is dice loss computed between the predicted segmentation maps and the weak segmentation labels. The second component is mean squared error (MSE) loss measuring the difference between the orientation histogram of the predicted segmentation map and the original raw image. Weighted sum of these two loss functions is used to train the proposed CNTSegNet network. The dice loss forces the network to perform background-foreground segmentation using local intensity features. The MSE loss guides the network with global orientation features and leads to refined segmentation results. The proposed system needs only a few-shot dataset for training. Thanks to its self-supervised nature, it can easily be adapted to new datasets.

Keywords: - *semantic segmentation, self-supervised learning, carbon nanotubes (CNT), electron microscopy*

1 Introduction

Carbon nanotubes (CNTs), discovered in 1991, [20] have an intriguing combination of mechanical, thermal, electrical, and chemical properties [13, 21]. The

* Corresponding author.

unique physical properties of CNTs are a result of the hexagonal sp^2 -bonded graphene sheets that comprise their walls. Single-walled CNTs (SWNTs) may exhibit metallic or semiconducting properties depending on their chirality. SWNT transistor devices fabricated with sub-10 nm channel lengths have exhibited high current density of 2.4 mA/micron and low operating voltage of 0.5 V. Multi-walled CNTs (MWNTs) are metallic in nature, with diameters ranging from approximately 2-40 nm. MWNTs have been spun together to form yarns [43] that are electrically conductive and strong, yet capable of being tied into a knot.

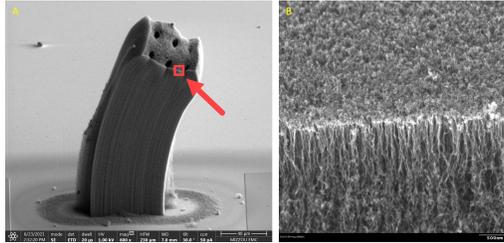


Fig. 1. Carbon nanotube (CNT) pillar imaged using scanning electron microscope (SEM). (A-left) Full pillar view. (B-right) Zoomed side view of the CNT pillar.

Isolated, individual CNTs are difficult to synthesize and are often impractical for device-level integration. CNTs are more frequently grown as CNT forests – high density CNT populations synthesized on a support substrate. Crowding within a growing CNT population forces CNTs to orient vertically within a forest, normal to the growth substrate. Interactions between contacting CNTs generate persistent attractive van der Waals bonds which resist mechanical loads generated within the forest during synthesis. Individual CNTs within a forest in response to mechanical loading, leading to a CNT forest morphology that resembles an open-cell foam. The properties of the CNT forests are vastly diminished when compared to that of an individual CNT. For example the elastic modulus of an individual CNT is in the order of 1TPa, while the compressive elastic modulus of a CNT forest may be as low as 10 MPa [32] – similar to that of natural rubber. The deformation mechanisms of compressed CNT forests are highly variable [7, 8, 18, 31–33, 39, 42] and are thought to result from variations in CNT forest morphology generated during cooperative synthesis [3, 28, 37]. CNT forests are candidates for the dry spinning of conductive, high-strength fibers [23, 43], piezoresistive sensing [29, 30, 36], electrochemical energy storage [9, 12], and thermal interface materials [10, 11]. Testing for physical properties of CNT forests often requires destruction of the forest which prevents further data collection. A method to determine physical properties of CNT forests indirectly using images of the said forest would help overcome the problem (Figure 1). Thus a thorough analysis of CNT images is a critical step in determining the CNT forests’ physical characteristics. The data obtained could then be used to determine how growth parameters can be modified to obtain a CNT forest with favorable

properties. CNT image analytics aims to quantify CNT attributes such as orientation, linearity, density, diameter etc. The first step towards CNT feature characterization is segmentation. Earlier works on CNT image analytics relied on classical image processing approaches. In [14] thresholding was used to produce partial CNT masks to determine CNT diameters. In [41] class-entropy maximization was used to segment CNT images with modest magnification levels (800X-4000X). [40] thresholded image pixels into three classes: background, CNT, and uncertain areas. Feature vectors of class background and CNT extracted from small image patches were used to train a multi layer perceptron neural network. The network then classified pixels of uncertain area as either background or CNTs. However, this strategy was only effective with extremely sparse, non-overlapping CNTs in small patches. In [15, 16] synthetic CNT forest images obtained from physics-based simulation have been analyzed using machine learning approaches to predict mechanical properties. While not developed for CNT image segmentation, recent works on detection and segmentation of other curvilinear structures such as fibers may be of interest for CNT image analytics. In [27] a 3D deep neural network pipeline was proposed for segmentation of short and thick glass fibers with acceptable density levels. The network was trained with a combination of synthetic data and real CT scan data with associated ground truths. In [4] an improved pipeline with deep center regression and geometric clustering was proposed for this type of glass fiber data.

As shown in Figure 1B, our dataset contains long, thin, and dense CNT fibers. Clustering or thresholding-based, unsupervised segmentation methods lead to limited success. Because of data complexity and ambiguity, manual labeling is not feasible. Thus, there is a shortage of high-quality datasets with associated labels that can enable use of supervised learning based approaches. Self-supervised learning [22] has emerged as an approach to learn good representations from unlabeled data and to perform fine-tuning with labeled features at the down-stream tasks.

In this paper, we present a self-supervised deep learning network for segmentation of CNT forests in scanning electron microscopy (SEM) images. The imaged CNT forests were grown using an in-situ SEM synthesis process based on chemical vapor deposition (CVD) [26]. The proposed deep segmentation network relies on two complementary training labels. The first label, intensity thresholded raw input image, serves as a weak label that leads the network to perform binary CNT segmentation. The second label, CNT orientation histogram calculated directly from the raw input image, constrains the segmentation process by enforcing the network to preserve orientation characteristics of the original image. Experimental results demonstrate refined segmentation results without supervision and need for manual image annotation.

2 Methods

In order to enable characterization of CNT properties within dense CNT forests, we have developed a self-supervised segmentation method. According to [22],

self-supervised learning trains a model by using pseudo labels that are generated automatically without the requirement for human annotations. The training procedure consists of two steps: a pretext task and a downstream task. Feature representation is learned in the pretext task first, whereas model adaptation is completed in the downstream task. The downstream task also evaluates the quality of features learned by the pretext task. Our proposed system consists of a novel deep neural network with two complementary loss functions which correspond to the pretext task and the downstream task. The following subsections describe the network architecture, loss functions, and generation of training labels.

2.1 Network architecture

We have built a self-supervised deep neural network with two complementary loss functions named CNTSegNet. The architecture of this network is similar to the classical U-Net [38] architecture with the encoder and decoder branches. Input to the CNTSegNet consists of a single channel 2D grayscale image. Output of the CNTSegNet consists of a binary, single channel, 2D image. During training and testing, SEM images of CNT forests are fed to the network to generate binary segmentation masks of CNTs. The network’s encoder uses the ResNet-34 model [17] as the backbone. It’s a fully convolutional deep neural network with shortcut connections to learn residual features. The encoder includes three main layers with 16, 32, and 64 filters respectively. Figure 2 depicts the network architecture and the network training process. The network generates a pixel-wise class likelihood map which is binarized at the level of 0.7 to produce a binary segmentation mask. Orientation histogram for the prediction is computed from the generated class likelihood map. The proposed network is trained with weighted sum of two complementary loss functions described below.

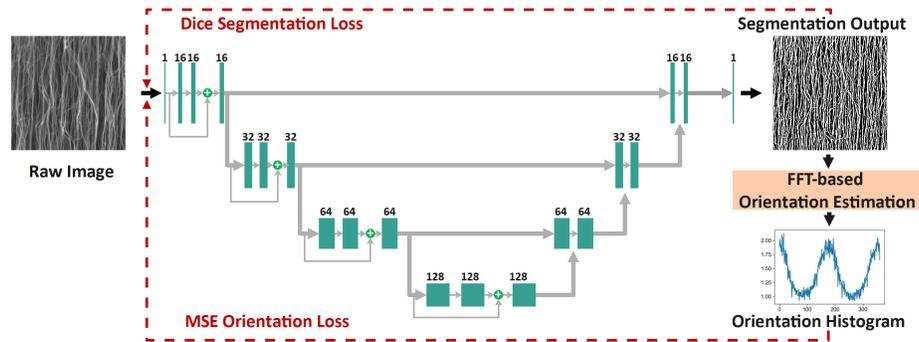


Fig. 2. Network architecture and training pipeline for the proposed dual loss and self-supervised network CNTSegNet. The network is based on an encoder-decoder architecture similar to U-Net segmentation network [17, 19, 38], but includes a second loss function and involves a self-supervised training scheme.

a) CNT segmentation loss: The first loss component aims to drive the network to perform segmentation prediction to match a given binary segmentation mask. Given a prediction mask $\text{Mask}_{\text{pred}}$ and a training mask $\text{Mask}_{\text{train}}$, dice loss is computed using the following equation

$$\text{Loss}_{\text{Dice}}(\text{Mask}_{\text{pred}}, \text{Mask}_{\text{train}}) = 1 - 2 \times \frac{|\text{Mask}_{\text{pred}} \cap \text{Mask}_{\text{train}}|}{|\text{Mask}_{\text{pred}}| + |\text{Mask}_{\text{train}}|} \quad (1)$$

Dice loss is a pixel-wise function that matches local image features in spatial domain. In this case, the dice loss by itself is not sufficient to generate reliable segmentation masks since the network is trained with automatically generated coarse weak labels rather than precise ground truth segmentation masks. This weakly supervised learning step plays the role of a pretext task.

b) CNT orientation loss: To compensate for weak labels, we introduced a second loss component. This second loss aims to drive the network to refine the predicted segmentation masks by enforcing the output to preserve the orientation patterns of the input. CNT forest orientation patterns is encoded using orientation histograms calculated without human annotation using the frequency domain methods proposed in [6, 24] and briefly described in Section 2.2. Given the input and output orientation histograms h_{in} and h_{pred} , orientation loss is computed as the following equation

$$\text{Loss}_{\text{MSE}}(h_{\text{pred}}, h_{\text{in}}) = \sum_{b=1}^n (h_{\text{pred}}(b) - h_{\text{in}}(b))^2 \quad (2)$$

where n and b refer to number of bins in the histogram and index for an histogram bin. Optimizing this orientation loss will push the arrangement of foreground pixels in the segmentation mask towards the orientation of corresponding pixels in the raw images. This operation plays the role of a downstream task in self-supervised learning approach.

c) Total loss: The proposed network is trained with a total loss computed as the weighted sum of the dice segmentation and MSE orientation losses

$$\text{Loss}_{\text{Total}} = k_1 \times \text{Loss}_{\text{Dice}} + k_2 \times \text{Loss}_{\text{MSE}} \quad (3)$$

where k_1 and k_2 refer to scalar weights. This total loss combines local and global image features extracted using spatial and frequency domain operations to preserve CNT morphology, particularly orientation patterns, during segmentation.

2.2 Generation of training labels

The two sets of training labels, the weak CNT segmentation masks and the CNT orientation histograms are generated as follows to enable self-supervised network training.

a) Weak CNT segmentation masks: To produce weak segmentation labels we explored two thresholding-based unsupervised methods. The first method, Otsu [35], is a global thresholding strategy that calculates an "ideal" threshold by maximizing the inter-class variance between background and foreground classes. The results of the Otsu approach are shown in the second columns of Figure 3. Because of the increasing illumination, CNTs in the lower image regions are prominent while in the upper regions fade into the background. The second method, adaptive thresholding, generates improved segmentation results if the background intensity varies widely. Adaptive thresholding determines threshold values in local regions by computing the weighted mean of the local neighborhood minus an offset value [2]. The third column of Figure 3 illustrates the outcomes of adaptive thresholding. These segmentation masks provide a greater level of detail. For training of the proposed deep neural network, we used weak segmentation masks generated using adaptive thresholding.

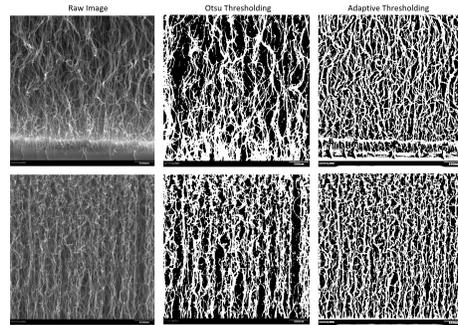


Fig. 3. Thresholding of CNT forest SEM images. Raw SEM image (first column), Otsu thresholding [35] (second column), adaptive thresholding [2] (last column).

b) CNT forest orientation histograms: CNT forests' physical properties are strongly affected by orientation and alignment of the CNTs forming them. Alignment of the CNTs can enhance the mechanical, thermal, and electrical characteristics. Birefringence and linear dichroism, fluorescence polarization, and polarized Raman spectroscopy are some of the most commonly used methods to evaluate CNT alignment. [6, 24].

In this study, we employed an image-based CNT forest orientation distribution estimation approach using radial sum method described in [6, 24]. As the orientation feature is extracted by summing operation, it is possible to incorporate this feature to the computational pipeline and perform back-propagation to train our deep neural network. This approach can estimate the orientation of CNTs in either raw images or their binary masks. Figure 4 illustrates the processes required to estimate the orientation distribution of CNTs. Initially, the input image (raw image or binary mask) (Figure 4A-B) is transformed into

Fourier space (Figure 4D). The output is masked by a circle (Figure 4E) divided into a number of bins (e.g. 360 bins, corresponding to 360 degrees). The total intensity in each bin indicates count of image pixels of that associated angle (Figure 4C). This histogram represents the practical orientation distribution of CNTs in the input image. It’s possible to fit this practical distribution as a mixture of several theoretical distributions as shown in Figure 4F.

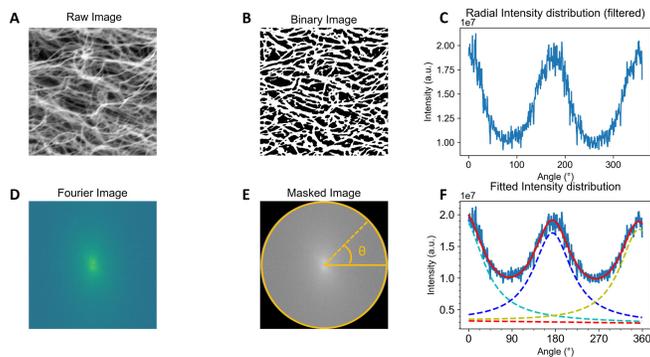


Fig. 4. CNT forest orientation distribution estimation using robust frequency-domain orientation estimation method described in [6, 24].

3 Experimental results

3.1 Datasets

The CNT forests used in this study were grown using chemical vapor deposition (CVD) [26] and were imaged using a FEI Quanta environmental SEM. The collected images had a resolution of 1536×1094 pixels and were acquired with a pixel dwell time of $10 \mu s$. A CNT forest can grow up to a height of several millimeters and often has different morphology at different locations of the forest pillar [5], thus it is important to collect images from different locations to obtain data encapsulating all the morphologies in the CNT forest. SEM images in this study were collected at a magnification of 50,000X. Care was taken to prevent a large overlap between images. 110 image patches of size of 768×768 were used, with 34 of them going into the training set and the rest 76 patches going into the test set.

3.2 Training process

We trained the proposed deep neural network with the dice and MSE loss functions separately and with the combined (dice + MSE) loss function to explore the effects of each loss component. During training, we maintained the same

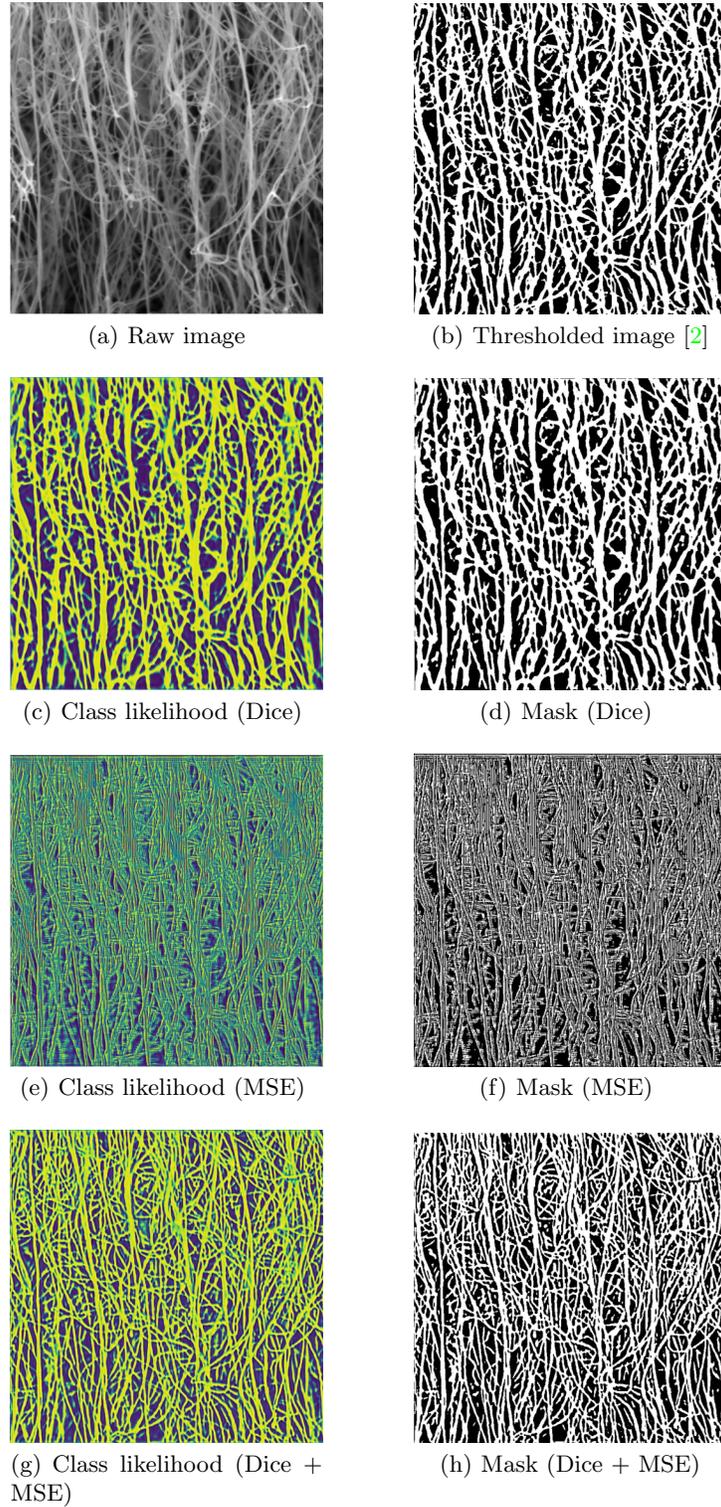


Fig. 5. Segmentation results for CNTSegNet with different loss functions. Raw SEM image of a sample CNT forest (a). Thresholded image [2] used as a weak segmentation label (b). Class likelihood maps and binary masks predicted using only the dice loss (c-d), only the MSE loss applied to orientation histogram (e-f), the combined dice and MSE loss functions (g-h).

learning rate and the same type of optimizer but varied the number of iterations for each loss function due to their distinct convergence processes.

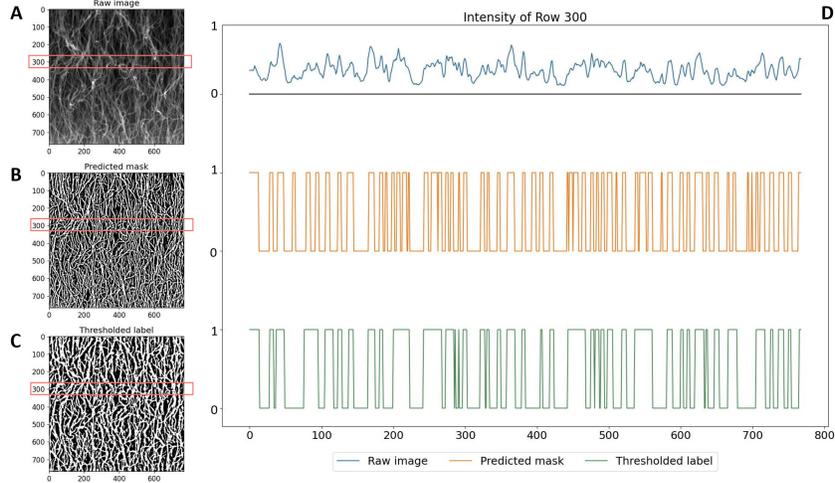


Fig. 6. Inspection of intensity profiles. (a) Raw SEM image; (b) mask predicted using CNTSegNet; (c) weak-label used to train the network; (d) associated intensity profiles at a sample image row.

a) Training with dice segmentation loss. The network was first trained with just the dice segmentation loss component for 10 epochs. We used Adam optimizer with a learning rate of $5e-4$. The final average dice score compared to the weak labels was 0.86.

Figure 5C & D show this network’s prediction as class likelihood map and as binary mask. The binary mask produced by the network is smoother and has fewer tiny debris compared to the thresholded image due to the effects of convolution filters and upsampling layers of the network. However, since the network was trained with thresholded images rather than ground truth masks, prediction was coarse resulting in merging of many neighboring CNTs.

b) Training with MSE orientation loss. In an other task, the network was trained using just the MSE loss between the orientation vectors of the raw input image and the network prediction. For 8 epochs, we ran the Adam optimizer with a learning rate of $5e-4$. In the absence of segmentation loss, the final average dice score for the network predictions was only 0.53 compared to the weak segmentation labels. This is due to the fact that the orientation vector includes just the summed global features. The loss function lacked the necessary local (pixel-wise) information needed to drive the segmentation mask.

Figure 5E & F show this network’s prediction as class likelihood map and as binary mask. As expected, this approach dropped the dice scores significantly from 0.86 to 0.53.

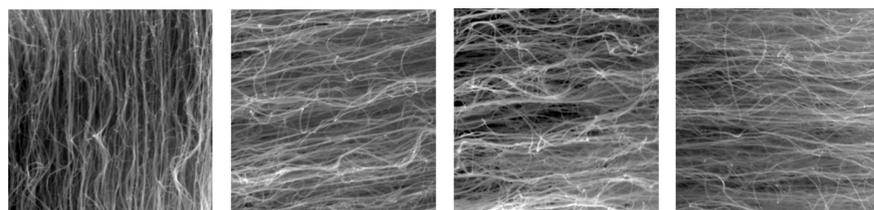
c) Training with two-component loss function. In the final task, we trained the proposed network with a combined loss function consisting of dice segmentation and MSE orientation losses. We utilized an Adam optimizer with a learning rate of $5e-4$. In the first 6 epochs, we initialized the segmentation output by training only with the dice loss. To regularize this output mask, for the following 3 epochs, we continued to train the network with a weighted sum of the dice segmentation and MSE orientation losses. As the maximum value of the dice loss is 1 but the magnitude of the orientation loss reaches much larger values at the levels of $1e+7$ (as seen in Figure 4C & F), we set the weight of the dice loss to 0.6, and the weight of the MSE loss to $1e-7$ for a more balanced influence. This scheme resulted in an average dice score of 0.84 compared to the weak labels.

Figure 5G & H show the prediction class likelihood map and associated binary mask obtained using this dual loss network. In this scheme, dice loss guides the network to match the segmentation mask of the weak labels, while the MSE loss regularizes the segmentation process and guides the network to preserve the orientation properties of its input resulting in finer segmentation details.

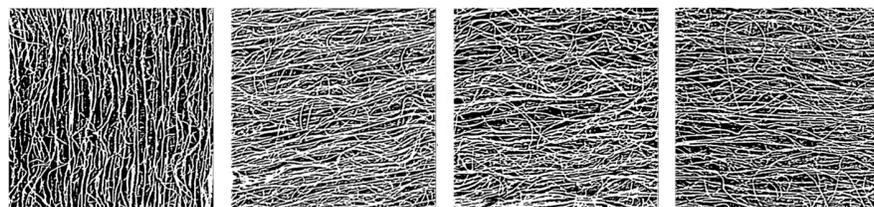
Figure 6 aims to provide further visual insight. We plotted single-row intensity profiles (row=300) for a sample raw SEM image, the corresponding weak segmentation label used for network training, and the binary mask predicted using the proposed CNTSegNet network. The intensity peaks in the original signal correspond to individual CNTs. These plots show that the weak label tends to merge the neighboring peaks in the original signal resulting in wider/thicker foreground blocks. We can observe that the CNTSegNet prediction can better detect these intensity peaks resulting in narrower/thinner foreground blocks and more refined segmentation masks.

3.3 Segmentation evaluation

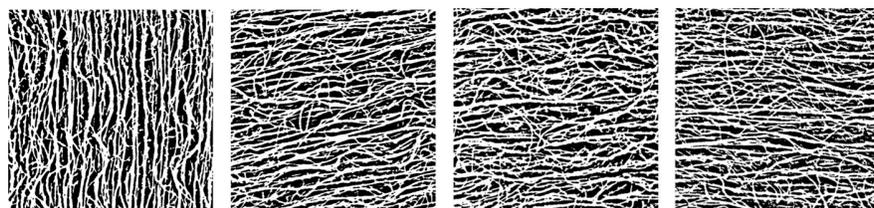
We conducted inference on the test set after training the proposed CNTSegNet network with the combined loss function. We compared the performance of the CNTSegNet to three other segmentation methods, adaptive intensity thresholding [2], k-means clustering [1], and a recent unsupervised deep learning-based segmentation method [25]. K-means [1] is an unsupervised clustering method that partitions a dataset into K clusters where each data item belongs to the cluster with the closest centroid. We used k-means clustering to cluster the intensity levels into two clusters corresponding to CNTs and background. [25] is a novel unsupervised convolutional neural network using differentiable feature clustering to enable unsupervised image segmentation. The network includes a normalizing function to generate a response map from the network output, and an argument max function to select a cluster for each pixel. For a single test image, this method first trains the network to minimize the difference of network



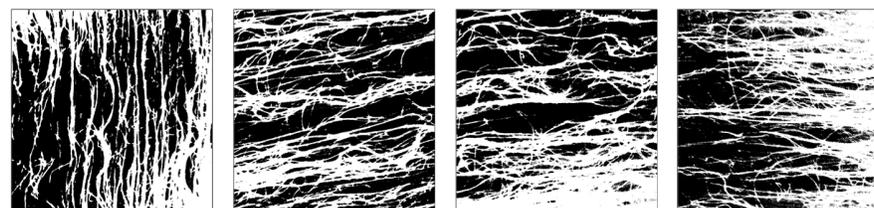
(a) Raw SEM images of CNT forest



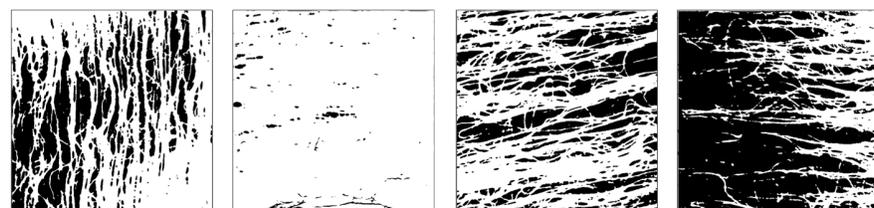
(b) Proposed CNTSegNet segmentation results



(c) Adaptive thresholding results [2]



(d) K-means clustering results [1]



(e) Unsupervised deep segmentation results [25]

Fig. 7. Segmentation results for four sample images (columns 1-4). (a) Raw SEM images. (b) Segmentation masks obtained using the proposed CNTSegNet network. (c) Adaptive thresholding results [2]. (d) K-means clustering results [1]. (e) Unsupervised deep segmentation results using [25].

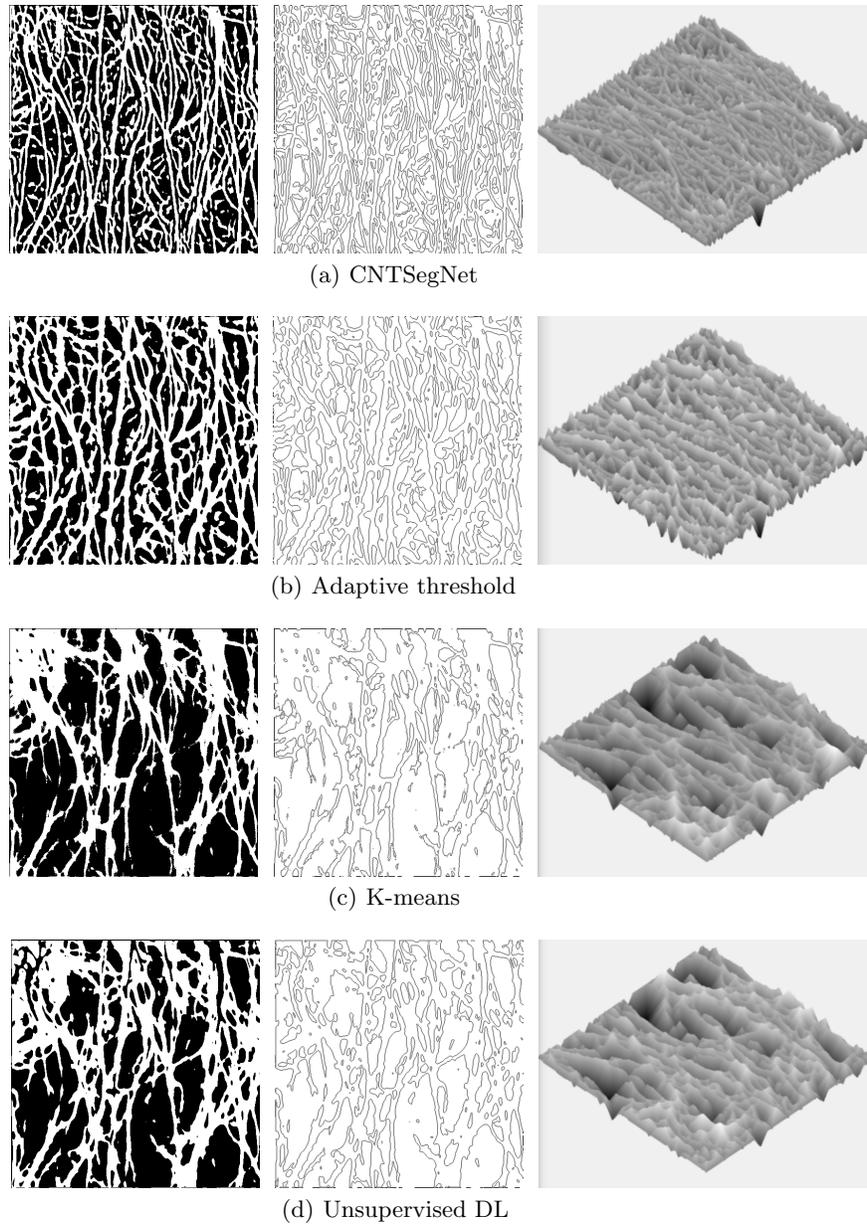


Fig. 8. Segmentation masks (left) and associated edge maps (middle), associated signed distance transform maps (right) for the proposed CNTSegNet network, adaptive thresholding [2], k-means clustering [1], and unsupervised deep segmentation [25] results.

output and the argument-max output, then it uses the argument-max output as the segmentation mask.

Figure 7 depicts the segmentation results of four aforementioned methods for four sample images. It can be observed that CNTSegNet results in more refined segmentation masks with better recall of the individual CNTs compared to all three methods. Compared to k-means clustering [1] and unsupervised deep segmentation [25], CNTSegNet is also more robust to illumination variations.

To measure and compare segmentation quality we computed five unsupervised measures: orientation loss, edge coverage, average thickness, average separation, and distance entropy. In order to compute these measures two intermediate representations, edge maps $E(x, y)$ and signed distance transforms $D(x, y)$ were generated as shown in Figure 8. An edge map is a binary image identifying foreground-background transitions. Signed distance transform [34] assigns to each pixel of the foreground its distance to the closest background point, and to each pixel of the background the opposite of its distance to the closest foreground point. Signed distance transform can be used to measure thickness and spatial separations of foreground structures in an image.

a) *Orientation loss* is computed between the original raw image and the corresponding segmentation mask as in Eq 2. Lower values indicate more similar orientation patterns.

b) *Edge coverage* is measured as the ratio of edge pixels to total image area. n_{FG} , n_{BG} indicate number of foreground and background pixels respectively. Higher edge coverage is an indication of more details in the segmentation mask.

$$Edge\ coverage = \frac{100}{n_{FG} + n_{BG}} \sum_{x,y} E(x, y) \quad (4)$$

c) *Average thickness* is measured as average distance on foreground pixels where $D(x, y)$ takes positive values.

$$Average\ thickness = \frac{1}{n_{FG}} \sum_{D(x,y)>0} D(x, y) \quad (5)$$

d) *Average separation* is measured as average distance on background pixels. For CNT forest images containing dense clusters of thin CNTs, lower average thickness and lower average separation indicate finer segmentation and higher recall of CNTs in a segmentation mask.

$$Average\ separation = -\frac{1}{n_{BG}} \sum_{D(x,y)<0} D(x, y) \quad (6)$$

e) *Distance entropy* is the last statistic we utilized to assess our segmentation outcomes. Entropy, which derives from thermal dynamics, is a measure of the disorder and uncertainty of a piece of information in information theory. It is computed by multiplying an event likelihood by its log probability

$$Distance\ entropy = -\sum_x p_D(x) \log p_D(x) \quad (7)$$

for our case p_D refers to probability distribution of signed distance transform. Lower distance entropy indicates lower variations in CNT thickness and CNT separation which is an indication of good segmentation for CNT forest images consisting of dense layouts of CNTs with similar diameters. Table 1 presents results of these unsupervised segmentation quality measures for the proposed CNTSegNet method and compared adaptive intensity thresholding [2], k-means clustering [1], and unsupervised deep learning-based segmentation [25] methods. The table indicates that the proposed CNTSegNet outperforms the compared methods in terms of all measures.

Measurement	CNTSegNet	Adaptive Threshold [2]	K-Means [1]	Unsup. DL [25]
Orientation Loss ↓	<u>9922</u>	9977	10009	10038
Edge Coverage ↑	<u>15.64%</u>	10.85%	9.37%	8.46%
Average Thickness ↓	<u>1.9965</u>	2.8043	3.8478	7.9640
Average Separation ↓	<u>2.4233</u>	3.2337	5.4858	11.2432
Distance Entropy ↓	<u>4.1677</u>	4.9932	5.6466	5.7731

Table 1. Unsupervised evaluation of segmentation quality for the proposed CNTSegNet method and compared adaptive intensity thresholding [2], k-means clustering [1], and unsupervised deep learning-based segmentation [25] methods. The values indicate average for 76 test images. Underlined number in each row indicates the best result.

4 Conclusions

In this paper, we proposed a self-supervised deep neural network, CNTSegNet, with two complementary loss functions for segmentation of CNT forests in SEM imagery. Despite lack of supervision, the proposed network was able to generate more detailed segmentation masks indicated by various unsupervised segmentation quality measures. The network was also able to better preserve orientation characteristics as indicated by lower orientation losses. This was an important feature since CNT forest physical properties are strongly affected by orientation and alignment of CNTs forming them. Thanks to its self-supervised nature, the proposed network is highly suitable for complex segmentation tasks where manual annotation is not practical or even feasible. The proposed network can easily be retrained using new datasets to improve performance or to adapt to new image characteristics. This study is our first step towards effective quantification of CNT forest characteristics from SEM imagery. Imaging and automated image analysis will be critical steps towards our ultimate goal of human out of the loop material discovery.

Acknowledgement. This work was partially supported by the National Science Foundation under award number CMMI-2026847. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. OpenCV: K-Means Clustering in OpenCV (Aug 2022), https://docs.opencv.org/3.4/d1/d5c/tutorial_py_kmeans_opencv.html, [Online; accessed 12. Aug. 2022]
2. Thresholding — skimage v0.13.1 docs (Jun 2022), http://devdoc.net/python/scikit-image-doc-0.13.1/auto_examples/xx_applications/plot_thresholding.html#id4, [Online; accessed 28. Jun. 2022]
3. Abadi, P.P.S.S., Maschmann, M.R., Mortuza, S., Banerjee, S., Baur, J.W., Graham, S., Cola, B.A.: Reversible tailoring of mechanical properties of carbon nanotube forests by immersing in solvents. *Carbon* **69**, 178–187 (2014)
4. Aguilar, C., Comer, M., Hanhan, I., Agyei, R., Sangid, M.: 3D Fiber Segmentation with Deep Center Regression and Geometric Clustering. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 3741–3749. IEEE (Jun 2021). <https://doi.org/10.1109/CVPRW53098.2021.00415>
5. Bedewy, M., Meshot, E.R., Guo, H., Verploegen, E.A., Lu, W., Hart, A.J.: Collective mechanism for the evolution and self-termination of vertically aligned carbon nanotube growth. *The Journal of Physical Chemistry C* **113**(48), 20576–20582 (2009)
6. Brandley, E., Greenhalgh, E.S., Shaffer, M.S.P., Li, Q.: Mapping carbon nanotube orientation by fast fourier transform of scanning electron micrographs. *Carbon* **137**, 78–87 (Oct 2018). <https://doi.org/10.1016/j.carbon.2018.04.063>
7. Brieland-Shoultz, A., Tawfick, S., Park, S.J., Bedewy, M., Maschmann, M.R., Baur, J.W., Hart, A.J.: Scaling the stiffness, strength, and toughness of ceramic-coated nanotube foams into the structural regime. *Advanced Functional Materials* **24**(36), 5728–5735 (2014)
8. Cao, A., Dickrell, P.L., Sawyer, W.G., Ghasemi-Nejhad, M.N., Ajayan, P.M.: Super-compressible foamlike carbon nanotube films. *Science* **310**(5752), 1307–1310 (2005)
9. Carter, R., Davis, B., Oakes, L., Maschmann, M.R., Pint, C.L.: A high areal capacity lithium–sulfur battery cathode prepared by site-selective vapor infiltration of hierarchical carbon nanotube arrays. *Nanoscale* **9**(39), 15018–15026 (2017)
10. Cola, B.A., Xu, J., Cheng, C., Xu, X., Fisher, T.S., Hu, H.: Photoacoustic characterization of carbon nanotube array thermal interfaces. *Journal of applied physics* **101**(5), 054313 (2007)
11. Cola, B.A., Xu, X., Fisher, T.S.: Increased real contact in thermal interfaces: A carbon nanotube/foil material. *Applied physics letters* **90**(9), 093513 (2007)
12. Davis, B.F., Yan, X., Muralidharan, N., Oakes, L., Pint, C.L., Maschmann, M.R.: Electrically conductive hierarchical carbon nanotube networks with tunable mechanical response. *ACS Applied Materials & Interfaces* **8**(41), 28004–28011 (2016)
13. De Volder, M.F.L., Tawfick, S.H., Baughman, R.H., Hart, A.J.: Carbon Nanotubes: Present and Future Commercial Applications. *Science* **339**(6119), 535–539 (Feb 2013). <https://doi.org/10.1126/science.1222453>
14. Gommès, C., Blacher, S., Masenelli-Varlot, K., Bossuot, Ch., McRae, E., Fonseca, A., Nagy, J.B., Pirard, J.P.: Image analysis characterization of multi-walled carbon nanotubes. *Carbon* **41**(13), 2561–2572 (Jan 2003). [https://doi.org/10.1016/S0008-6223\(03\)00375-0](https://doi.org/10.1016/S0008-6223(03)00375-0)
15. Hajilounezhad, T., Bao, R., Palaniappan, K., Bunyak, F., Calyam, P., Maschmann, M.R.: Predicting carbon nanotube forest attributes and mechanical properties using simulated images and deep learning. *npj Computational Materials* **7**(1), 1–11 (2021)

16. Hajilounezhad, T., Oraibi, Z.A., Surya, R., Bunyak, F., Maschmann, M.R., Calyam, P., Palaniappan, K.: Exploration of carbon nanotube forest synthesis-structure relationships using physics-based simulation and machine learning. In: IEEE Applied Imagery Pattern Recognition Workshop (AIPR). pp. 1–8 (2019)
17. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016). <https://doi.org/10.1109/CVPR.2016.90>
18. Hines, R., Hajilounezhad, T., Love-Baker, C., Koerner, G., Maschmann, M.R.: Growth and mechanics of heterogeneous, 3d carbon nanotube forest microstructures formed by sequential selective-area synthesis. *ACS applied materials & interfaces* **12**(15), 17893–17900 (2020)
19. Iakubovskii, P.: Segmentation models pytorch. https://github.com/qubvel/segmentation_models.pytorch (2019)
20. Iijima, S.: Helical microtubules of graphitic carbon. *Nature* **354**(6348), 56–58 (Nov 1991). <https://doi.org/10.1038/354056a0>
21. Iijima, S.: Carbon nanotubes: past, present, and future. *Physica B* **323**(1), 1–5 (Oct 2002). [https://doi.org/10.1016/S0921-4526\(02\)00869-4](https://doi.org/10.1016/S0921-4526(02)00869-4)
22. Jing, L., Tian, Y.: Self-Supervised Visual Feature Learning With Deep Neural Networks: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **43**(11), 4037–4058 (May 2020). <https://doi.org/10.1109/TPAMI.2020.2992393>
23. Jung, Y., Cho, Y.S., Lee, J.W., Oh, J.Y., Park, C.R.: How can we make carbon nanotube yarn stronger? *Composites Science and Technology* **166**, 95–108 (2018)
24. Kaniyoor, A., Gspann, T.S., Mizen, J.E., Elliott, J.A.: Quantifying alignment in carbon nanotube yarns and similar two-dimensional anisotropic systems. *J. Appl. Polym. Sci.* **138**(37), 50939 (Aug 2021). <https://doi.org/10.1002/app.50939>
25. Kim, W., Kanezaki, A., Tanaka, M.: Unsupervised Learning of Image Segmentation Based on Differentiable Feature Clustering. *IEEE Trans. Image Process.* **29**, 8055–8068 (Jul 2020). <https://doi.org/10.1109/TIP.2020.3011269>
26. Koerner, G., Surya, R., Palaniappan, K., Calyam, P., Bunyak, F., Maschmann, M.R.: In-situ scanning electron microscope chemical vapor deposition as a platform for nanomanufacturing insights. In: ASME International Mechanical Engineering Congress and Exposition. vol. 85567, p. V02BT02A052 (2021)
27. Konopczyński, T., Kröger, T., Zheng, L., Hesser, J.: Instance Segmentation of Fibers from Low Resolution CT Scans via 3D Deep Embedding Learning. *arXiv* (Jan 2019). <https://doi.org/10.48550/arXiv.1901.01034>
28. Maschmann, M.R.: Integrated simulation of active carbon nanotube forest growth and mechanical compression. *Carbon* **86**, 26–37 (2015)
29. Maschmann, M.R., Dickinson, B., Ehlert, G.J., Baur, J.W.: Force sensitive carbon nanotube arrays for biologically inspired airflow sensing. *Smart Materials and Structures* **21**(9), 094024 (2012)
30. Maschmann, M.R., Ehlert, G.J., Dickinson, B.T., Phillips, D.M., Ray, C.W., Reich, G.W., Baur, J.W.: Bioinspired carbon nanotube fuzzy fiber hair sensor for air-flow detection. *Advanced Materials* **26**(20), 3230–3234 (2014)
31. Maschmann, M.R., Ehlert, G.J., Park, S.J., Mollenhauer, D., Maruyama, B., Hart, A.J., Baur, J.W.: Visualizing strain evolution and coordinated buckling within cnt arrays by in situ digital image correlation. *Advanced Functional Materials* **22**(22), 4686–4695 (2012)
32. Maschmann, M.R., Zhang, Q., Du, F., Dai, L., Baur, J.: Length dependent foam-like mechanical response of axially indented vertically oriented carbon nanotube arrays. *Carbon* **49**(2), 386–397 (2011)

33. Maschmann, M.R., Zhang, Q., Wheeler, R., Du, F., Dai, L., Baur, J.: In situ sem observation of column-like and foam-like cnt array nanoindentation. *ACS applied materials & interfaces* **3**(3), 648–653 (2011)
34. Maurer, C.R., Qi, R., Raghavan, V.: A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**(2), 265–270 (2003)
35. Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst., Man, Cybern.* **9**(1), 62–66 (Jan 1979). <https://doi.org/10.1109/TSMC.1979.4310076>
36. Park, M., Cola, B.A., Siegmund, T., Xu, J., Maschmann, M.R., Fisher, T.S., Kim, H.: Effects of a carbon nanotube layer on electrical contact resistance between copper substrates. *Nanotechnology* **17**(9), 2294 (2006)
37. Pathak, S., Mohan, N., Decolvenaere, E., Needleman, A., Bedewy, M., Hart, A.J., Greer, J.R.: Local relative density modulates failure and strength in vertically aligned carbon nanotubes. *ACS nano* **7**(10), 8593–8604 (2013)
38. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241. Springer, Cham, Switzerland (Nov 2015). https://doi.org/10.1007/978-3-319-24574-4_28
39. Tawfick, S., Zhao, Z., Maschmann, M., Brieland-Shoultz, A., De Volder, M., Baur, J.W., Lu, W., Hart, A.J.: Mechanics of capillary forming of aligned carbon nanotube assemblies. *Langmuir* **29**(17), 5190–5198 (2013)
40. Trujillo, M.C.R., Alarcón, T.E., Dalmau, O.S., Zamudio Ojeda, A.: Segmentation of carbon nanotube images through an artificial neural network. *Soft Comput.* **21**(3), 611–625 (Feb 2017). <https://doi.org/10.1007/s00500-016-2426-1>
41. Wortmann, T., Fatikow, S.: Carbon nanotube detection by scanning electron microscopy. In: *In Proc. of the Eleventh IAPR Conference on Machine Vision Applications (MVA'09)* (2009)
42. Zbib, A.A., Mesarovic, S.D., Lilleodden, E.T., McClain, D., Jiao, J., Bahr, D.F.: The coordinated buckling of carbon nanotube turfs under uniform compression. *Nanotechnology* **19**(17), 175704 (mar 2008). <https://doi.org/10.1088/0957-4484/19/17/175704>, <https://doi.org/10.1088/0957-4484/19/17/175704>
43. Zhang, M., Atkinson, K.R., Baughman, R.H.: Multifunctional carbon nanotube yarns by downsizing an ancient technology. *Science* **306**(5700), 1358–1361 (2004)