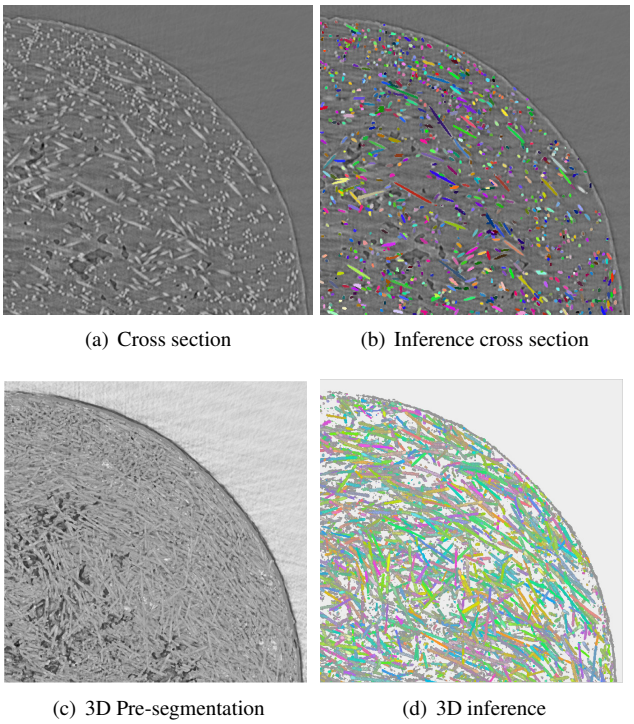
3D Fiber Segmentation with Deep Center Regression and Geometric Clustering

Camilo Aguilar, Mary Comer
School of Electrical and Computer Engineering
Purdue University
{aguilarh,comerm}@purdue.edu

Imad Hanhan, Ronald Agyei, Michael Sangid
School of Aeronautics and Astronautics
Purdue University
{ihanhan,ragyei,msangid}@purdue.edu

Abstract

Material and biological sciences frequently generate large amounts of microscope data that require 3D object-level segmentation. Often, the objects of interest have a common geometry, for example spherical, ellipsoidal, or cylindrical shapes. Neural networks have became a popular approach for object detection but they are often limited by their training dataset and have difficulties adapting to new data. In this paper, we propose a volumetric object detection approach for microscopy volumes comprised of fibrous structures by using deep centroid regression and geometric regularization. To this end, we train encoder-decoder networks for segmentation and centroid regression. We use the regression information combined with prior system knowledge to propose cylindrical objects and enforce geometric regularization in the segmentation. We train our networks on synthetic data and then test the trained networks in several experimental datasets. Our approach shows competitive results against other 3D segmentation methods when tested on the synthetic data and outperforms those other methods across different datasets. The reference codes and pytorch dictionaries are available at https://github.com/camilo-aguilar/3D-Fiber-Segmentation.



(a) Cross section (b) Inference cross section

(c) 3D Pre-segmentation (d) 3D inference

Figure 1. Sample results of our method displayed with Fiji[18] software. Each color represents a different instance of an object. Image courtesy of the ACME Lab at Purdue University.

1. Introduction

Experiments in material and biological sciences often use non-invasive image modalities that generate gray-scale point clouds representing the system's microstructure. One example of these experiments is the development of fiber reinforced composites shown in Fig. 1(a). Performance improvement of these materials relies on an accurate structural characterization in order to predict the material's response to external forces and to prevent material failures. This structural characterization requires of instance-level segmentation in order to extract information such as object location, dimensions, interactions, and orientation[3].

During recent years, the area of instance segmentation experienced a rapid growth manifested in popular object proposal techniques such as Mask-RCNN[7], PANet [14], Box2Pix[19]. While these approaches have proven successful in popular benchmark datasets, the instance segmentation paradigm becomes challenging when transitioning to microscopy volumetric datasets: microscopy images tend to have large numbers of clustered objects, the jump in dimensions represents a significant increase in memory requirements, and thin 3D geometries oriented in arbitrary orientations pose challenges for bounding box characterizations.

Another popular technique for approaching instance segmentation is based on proposal-free methods that rely on instance grouping learning. These methods learn to group

instance pixels into clusters living in an embedded space. Same-instance pixels are grouped near each other and different-instance pixels are discriminated from each other. Subsequently, a clustering algorithm assigns a label to each cluster to finally map each label to an instance in the initial image space. The drawbacks of this method are that it needs prior background/foreground pixel classification, the choice of the clustering algorithm including parameter settings, is of paramount importance in obtaining a correct segmentation, and the clustering algorithm is normally independent of the object's geometry.

Numerous 3D instance segmentation techniques borrowed the embedding learning concept to segment 3D point clouds of natural images and microscopy images. Nevertheless, their clustering algorithm is applied in an abstract embedded space and is generally independent of geometric constraints of the original image space. This issue results in time consuming parameter tuning, faulty segmentations, and difficulties adapting to unseen datasets. Microscopy images tend to have repeated geometries that can provide useful information when clustering pixels. For example, Fig. 1(c) shows a sample microscopy system that is comprised mainly of cylindrical geometries.

In this paper, we propose an instance segmentation approach based on centroid regression and regularized clustering. Our method detects foreground pixels and learns to regress the foreground pixels to their instance centroid. In addition, we propose a clustering algorithm that considers both the network output as well as the system's geometric constraints in order to ensure the recovered instances preserve a consistent shape.

This paper is organized as follows: in Section 2 we present popular neural network approaches used for 3D object detection. We discuss the details of our proposed method in Section 3. We show experiments on 3 different datasets and compare our approach to other methods in Section 4 and we provide insights and a conclusion in Section 5.

2. Related Work

Our literature review focuses on approaches that have been successful or that can be adapted without significant changes to perform instance segmentation on 3D systems comprised of nearby thin structures such as 3D fibers.

2.1. Instance Embedded Learning

These methods consist of segmenting and grouping foreground pixels into clusters in abstract high dimensional embedded spaces. These networks learn to group same-instance pixels into clusters, while separating the cluster centers from each other. Sequentially, an unsupervised clustering algorithm such as DBSCAN[5] assigns labels to pixels in each cluster. This work was proposed by [4] but has been borrowed to extract fibers in 3D microscopy volumes[11].

The embedded learning concept has been shown to be memory efficient and promising; however, this method relies on finding an arbitrary embedded space that does not have a direct interpretation in the image space and the choice of both the clustering algorithm and its parameters influence greatly the segmentation results. For example, a large eps parameter in the DBSCAN clustering algorithm merges nearby clusters and a small eps parameter splits clusters or does not detect them.

2.2. Joint-Task Learning

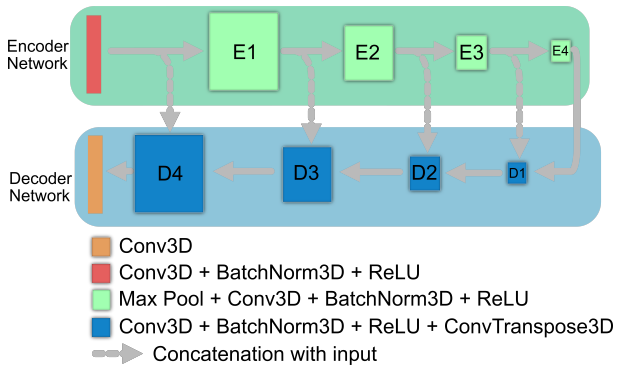
Several papers (for both 2D and 3D) have proposed joint-task approaches. These methods rely on single encoder-multiple decoder networks. For example, Neven et al.[15] trained an encoder-double decoder network to jointly learn instance centers and the cluster bandwidth. However, this method uses a loss function that relies on a weighted sum of the center regression and bandwidth loss. This issue could result in time-consuming parameter tuning when training the networks for several days. Kendall et al. proposed an efficient ensemble training for both tasks[9]. This approach estimates the weight uncertainties for each task to find the optimal weight parameter. These approaches are promising; however, the combination of two decoders and one encoder can worsen the results of each decoder. In fact, in our experiments, the implementation of two separate networks obtained better results across all the tests.

3. Method

We propose a two-network approach, one for segmentation, one for centroid regression. In addition, we propose to use the regression vectors to obtain information about the original properties of each instance, such as the instance's orientation and length. We use a modification of the architecture presented in [17] due to its effectiveness in microscopy volumes and robustness to scarce training data. The two networks are the semantic segmentation network, which has two outputs denoting the probability of each class, and the regression network, which has three outputs denoting a 3D vector pointing to the instance center. Fig. 2 denotes the detailed architecture of our networks. The regression network uses the output of the semantic segmentation to regress to regress only the foreground pixels.

3.0.1 Pixel Classification

For semantic segmentation, we use the dice loss with two classes, foreground to represent fibers and background to represent the non-fiber pixels. The dice loss function is ex-



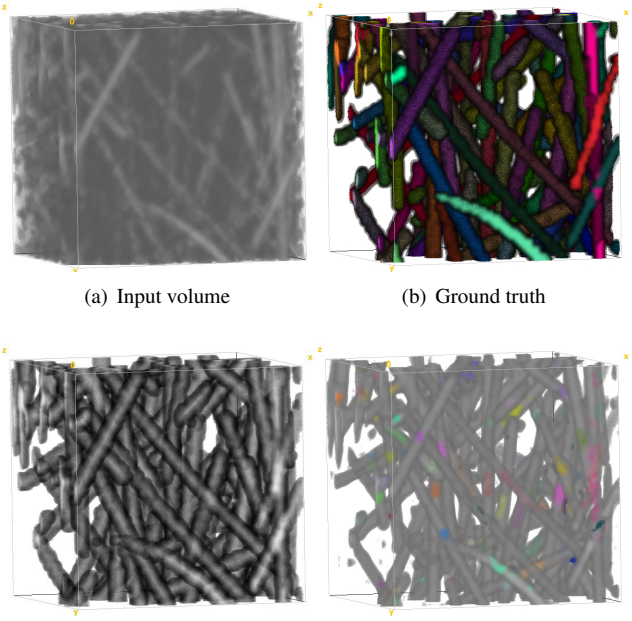
(a)

Figure 2. Architecture: we used a variation of the architecture proposed in [17] where each encoder and each decoder consist of 4 components.

pressed as:

$$\mathcal{L}_{Seg}(X,Y) = 1 - \frac{2|X \cap Y|}{|X| + |Y|}, \tag{1}$$

where X and Y are the input and training vectors respectively. The output of this network is depicted in Fig. 3(c).



(a) Input volume (b) Ground truth

(c) Semantic segmentation (d) Centroid regression

Figure 3. Network Outputs: The networks detect the fiber pixels and the fiber centers. Each color represents an instance, and pixels are clustered around their centers. Gray fibers are for display purposes. Figures generated with Fiji Software.

3.0.2 Centroid Regression

The second network receives the semantic segmentation results and learns to cluster foreground pixels around their instance centroid μ_c. The network outputs a vector $v_i \in \mathbb{R}^3$ for each foreground pixel's coordinate $s_i \in S_f$ such that $v_i = s_i - \mu_c$. The loss function for centroid regression is defined as:

$$\mathcal{L}_{Inst}(S_f) = \sum_{c=1}^{C} \sum_{s_i \in S_c} (||s_i - \mu_c|| - \delta_v)_+^2, \tag{2}$$

where S_f is the set of foreground pixels coordinates, $S_c \subset S_f$ is the subset of pixel coordinates belonging to fiber c and C is the total number of fibers (obtained from the ground truth). The term $(a)_+ = \max(a, 0)$ is inspired by the Hinge Loss function, with δ_v as a hyper-parameter represents the maximum distance between neighboring points. In all our experiments, we set this hyperparameter $\delta_v = 1$. We use the result from this clustering network to guide our geometric clustering. The results of applying the offset output by this network are depicted in Fig. 3(d). Note that both networks are trained separately using the ground truth labels.

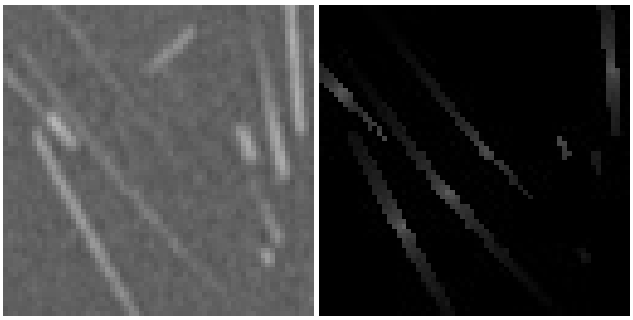
3.1. Geometric Constrained Clustering

Common instance segmentation approaches perform clustering on the output of the regression network[1, 11, 15]. However, these approaches do not impose shape constrains on the resulting segmentation and can merge nearby fibers. Therefore, we impose a geometric regularization on the clustering approach by finding a set of clusters $\mathbf{w} = \{\omega_1, \omega_2, \ldots, \omega_{\hat{C}}\}$ that represent a set of cylinders in the volumetric space, where $\hat{C}$ is the final estimate of the number of cylinders. We follow the convention used in the connected tube marked point process[13] to describe objects with marks describing their properties. We use cylindrical clusters with the marks $m_c = (\mu_c, r_c, l_c, \theta_c, \phi_c)$ where μ_c denotes the center coordinate of the cylinder, $r_c \in [r_{min}, r_{max}]$ denotes the cylinder radius and r_{min}, r_{max} denote the minimum and maximum possible radii respectively. The parameter $l_c \in [l_{min}, l_{max}]$ denotes the fiber length, and l_{min}, l_{max} denote the minimum and maximum possible fiber lengths. The parameters $\theta_c \in [\theta_{min}, \theta_{max}]$ and $\phi_c \in [\phi_{min}, \phi_{max}]$ denote the cylinder orientation with respect to the positive xy axis and with respect to the positive z axis, with their respective possible minimum and maximum values.

3.1.1 Birthmap Computation

We use the centroid regression vectors v_i to shift all the foreground pixels $s_i \in S_f$ to their estimated instance center o_i i.e., $o_i = s_i - v_i$, hence generating a set of offset pixels O_f. The offset pixels tend to be concentrated around

a fiber's estimated center, $\hat{\mu_c}$, thus, generating a heat-map near the true instance center μ_c. This map is shown in Fig 4(b), where brighter pixels represent a higher count of regressed pixels to that location. This map is used as a "birthmap" to propose clusters in section 3.1.2.

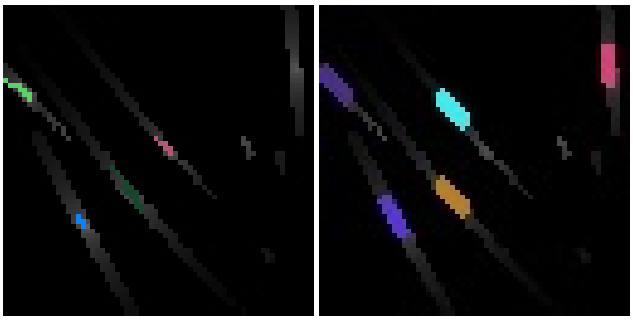


(a) Original volume (b) Birthmap

Figure 4. Cropped original image and its birthmap.

3.1.2 Cluster Proposal

We propose an estimated cluster center $\hat{\mu}_c$ at the location with the highest count of unlabeled offset pixels o_i in the birthmap (brightest spots in Fig. 4(b)). We then gather those offset pixels that are within a distance of r_{min} from the estimated center, $\hat{w}_c = \{o_i \in O_f : ||o_i - \hat{\mu_c}|| < r_{min}\}$. These pixels contain their original coordinate information and hence, generate a cloud of points that can be fitted with a cylinder. We use a GPU adaptation of [16] to estimate the marks $\hat{m_c}$: radius $\hat{r}_c$, length $\hat{l}_c$, and orientation $\hat{\theta}_c, \hat{\phi}_c$ of the cloud of points to propose a cylindrical cluster. Finally we add to the proposed cluster set $\hat{\omega}_c$ all the offset pixels that lie inside the proposed cylinder. Fig. 5(b) shows a sample cluster proposal result.

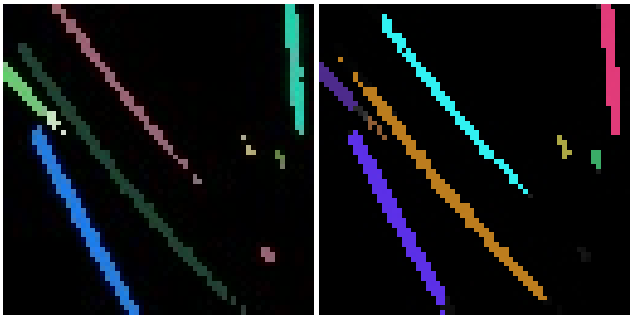


(a) Regressed ground truth pixels (b) Cluster proposals

Figure 5. Sample cluster proposal and ground truth labels.

3.1.3 Cluster Evaluation

Finally, we evaluate the proposed cylinder in the original volume space by calculating the volume percent of segmented pixels covered by the cylinder V_d and the volume percent of overlap with other cylinders V_p. We set two threshold parameters for volume percent of segmented pixels T_d and for volume percent of overlap, T_p. If $V_d > T_d$ and $V_p < T_p$, we accept the cylinder, otherwise we leave the set of pixels as unlabeled. We choose $T_d = 0.5$ to keep consistent with the IoU metrics explained in Section 4 and T_p = 0.2 to allow the proposed cluster 20% of overlap with other clusters. The volume percent V_p ensures the proposed object fits the semantic segmentation and the overlap percent V_p prevents from proposing multiple clusters for the same fiber. These evaluations were inspired by the success of the overlap prior from the marked point process model [13]. The detailed procedure of the clustering algorithm is listed in Algorithm 1.



(a) Ground truth (b) Sample inference

Figure 6. Sample ground truth labels and sample inference.

3.2. Volume Tiling and Merging

We tile the initial volumetric data into overlapping subvolumes of size 64×64 in order to circumvent large memory demands from some datasets. We detect fibers locally and we implement a merging procedure to detect fibers in the full volume. We extend the approach presented by Konopczynski[11], which consists of merging nearby fibers in overlapping tiles. However, this procedure does not consider the fiber properties and can merge nearby local fibers even if they are perpendicular to each other. We implement an additional constraint based on the orientation information provided by the cluster properties. If two fibers in overlapping tiles are nearby, we merge them if the angle between between them is less than a threshold $T_{\Delta\theta,\Delta\phi}$. The value of this threshold is set based on prior knowledge of the system, such as the fiber expected curvature. The detailed merging procedure is included in Algorithm 2.

Algorithm 1: Geometric Constrained Clustering

Result: Set of clusters $\mathbf{w} = \{\omega_1, \omega_2, \ldots, \omega_{\hat{C}}\}$
initialization;
$\mathbf{w} = \{\}$;
$S_f \leftarrow$ original foreground pixels coordinates;
$O_f \leftarrow$ center regressed foreground pixels coordinates;
$k \leftarrow 1$;
Birthmap Computation;
counts $\leftarrow$ histogram of O_f;
birthmap $\leftarrow$ sorted elements of O_f by decreasing order of counts;
for *each pixel s in birthmap* **do**
  if $s \in w_i : w_i \in \boldsymbol{w}$ **then**
    continue (pixel has a label);
  end
  Cluster Proposal;
    $\hat{\mu}_k \leftarrow s$;
    $\hat{w}_k \leftarrow \{s_k \subset O_f : ||s_f - \hat{\mu}_k|| < r_{min}\}$;
    $\hat{m}_k \leftarrow \text{marks}(\hat{w}_k, S_f, O_f)$;
  Evaluate Cluster;
    $V_d \leftarrow$ percent of segmented pixels inside cylinder with marks $\hat{m}_k$;
    $V_p \leftarrow$ percent of overlapping pixels between the cylinder with marks $\hat{m}_k$ and other cylinders ;
  if $V_d > T_d$ *and* $V_p < T_p$ **then**
    $\mathbf{w} \leftarrow \mathbf{w} \cup w_k$;
    $k \leftarrow k + 1$;
  end
end

Algorithm 2: Merging Algorithm with orientation constrains

Result: $\mathbf{w}_V$ clusters for volume V
Initialization;
  $S_{or} \leftarrow$ set of partitions for volume V with overlapping ratio or;
  $\mathbf{w}_V = \{\}$;
Volume Merging;
for *partition s in* S_{or} **do**
  $\mathbf{w}_s \leftarrow$ clusters at partition s (using Algorithm 1);
  for *cluster* ω_i *in* $\boldsymbol{w}_s$ **do**
    if *Overlap*$(\omega_i, \omega_j) > 0$, $\omega_j \in \boldsymbol{w}_V$ *and* *AngleDiff*$(\omega_i, \omega_j) < T_{\phi,\theta}$ **then**
      $\omega_j \leftarrow$ Merge Clusters(ω_j, ω_i);
    else
      $\mathbf{w}_V \leftarrow \mathbf{w}_V \cup \omega_i$;
    end
  end
end

4. Experiments

We use two metrics to evaluate the instance segmentation results. First, we follow the guideline provided in [12] and use the Adjusted Rand Index[8] (R_a) metric for evaluation. This metric evaluates the agreement between two sets of clusters $C = \{c_1, c_2, \ldots, c_k\}$, and $C' = \{c'_1, c'_2, \ldots c'_l\}$ and ranges from 0 to 1 depending on the degree of similarity between C and C'. The R_a criteria is given by:

$$R_a(C, C') = \frac{\sum_i^k \sum_j^l \binom{m_{ij}}{2} - t_3}{\frac{1}{2}(t_1 + t_2) - t_3} \tag{3}$$

Where $m_{ij} = |c_i \cap c_j|$, $t_1 = \sum_{i=1}^{i=k} \binom{|c_i|}{2}$, $t_2 = \sum_{j=1}^{j=l} \binom{|c_j|}{2}$, and $t_3 = \frac{2t_1 t_2}{n(n-1)}$.

In addition, we use the intersection over union evaluation (IoU) where we say a fiber is detected correctly if its IoU with a ground truth fiber is greater than 0.50. We label fibers whose IoU is less than 0.5 as false positives (broken fiber segments) and we label as false negatives all the ground truth fibers that were not captured by fibers that had an IoU more than 0.5.

We average the results obtained from multiple tiles of size $64 \times 64 \times 64$ voxels and compare our method with our implementation of the method proposed by Kendall et. al[9] , with the method proposed by Aguilar et al [1], and with using center regression and DBSCAN for clustering. We do not use the angle criteria for fiber merging when we use the DBSCAN algorithm since we do not have fiber orientation information.

We implemented all the networks and trained with the same training parameters as described in Section 4. We followed the parameters denoted in each work and chose $\delta_v = 0.2$ and $eps = 0.4$ for the embedding learning parameters shown in [1].

4.1. Training Information: Synthetic Fibers

This dataset was generated from a computational model by Konopczynski et al.[12] and it is a simulation of a short glass fibers embedded in a reinforced polymer. The fibers have a radius of 6.5 μm and a mean length of 500 μm with a deviation of 100 μm and are oriented in arbitrary directions. We used the dataset named "2016-S-HR-5.35p$''$ depicted in Fig. 7 for training the network and the dataset named "2016-S-HR-5.38p$''$ for testing.

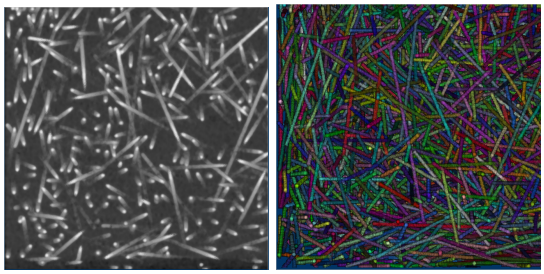
We trained all the networks with the Adam [10] optimizer with a learning rate of lr=0.001 over 2000 epochs. We normalized the data to unit variance and zero mean and we trained cropping subvolumes of sizes $64 \times 64 \times 64$ cropped at uniformly random locations of the full volume, and we performed volume rotations for data augmentation. All the models were trained on an NVIDIA-Titan RTX GPU with

Method	R_a	f1
Embedding Learning[1]	0.756	0.983
Multitask Learning[9]	0.622	0.977
Centroid Regression	**0.767**	**0.993**
Proposed	0.719	0.973

Table 1. f1 and R_a results for synthetic dataset

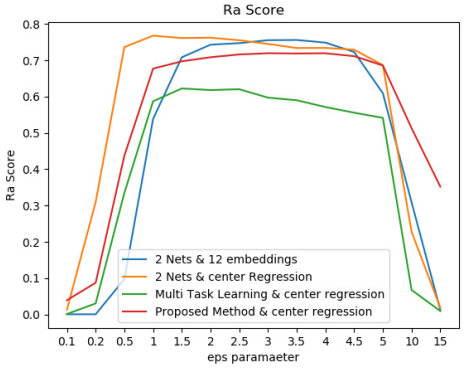
25GB of memory with a training time of approximately two days.

Fig. 8 shows the average results over all tiles when varying the eps parameters for different methods and Table 1 shows the best scores obtained for each method over the tested values of eps. The implementation that performs only centroid regression with DBSCAN achieved the best scores with 0.993 f1 score and 0.767 R_a score. Our approach obtained 0.973 for the object-wise f1 score and 0.719 for the R_a score.

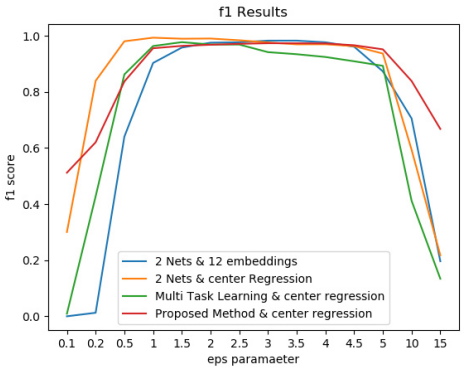


(a) Synthetic data (b) Ground truth

Figure 7. Training dataset: synthetic fibers generated by Konopczynski[12]



(a) Ra score



(b) f1 score

Figure 8. Evaluation of average tile scores vs eps parameter in synthetic data. For the proposed method, we let eps = r_{min}

4.2. Low Resolution SFRP Dataset

This dataset was provided by Konopczynski et al.[12] and it represents two samples of a commercial polybutylene terephthalate PBT reinforced with short glass fibers. The fibers diameters between 10-14μm and 1.1mm of length. The samples were imaged with X-ray CT with an isotropic resolution of 3.9μm and the volumetric dimensions of 200 $\times$ 260 $\times$ 260 voxels. We used the trained dictionaries from section 4.1 and tested all the methods on the low resolution SFRP dataset named "Real MR2". It is worth noting that the ground truth fibers were pre-segmented, hand labeled, and refined with the watershed algorithm. Therefore, the ground truth labels can be biased to the pre-segmentation algorithm followed by the watershed algorithm. The first column of Fig. 10 shows a tiled cross section of sizes 64 $\times$ 64 pixels and each row shows results for different approaches. The third and fourth row show that the DBSCAN approaches encounter difficulties clustering fibers. For example, Fig. 10(c) shows numerous groups of white pixels (unlabeled pixels) that are near two different fibers. Similarly, Fig. 10(d) shows merging of fibers that are close to each other. These issues are translated in the volumetric results as shown in Figs. 10(h) and Figs. 10(i) where fibers are artificially broken or nearby fibers are merged. Our method, shown in Fig. 10(e) can discriminate between nearby fibers and has also the ability to segment curved fibers when using volume merging as shown in Fig. 10(j).

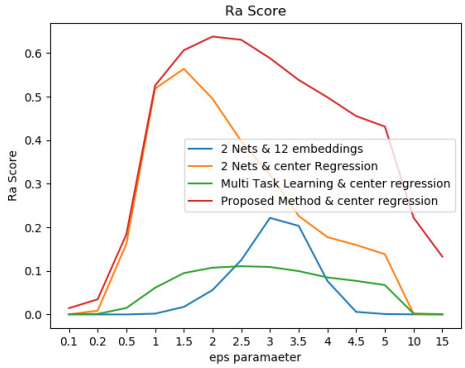
The numerical evaluation in Table 2 shows that our method obtains significantly better scores for both metrics with an R_a value of 0.638 and a $f1$ score of 0.917. Fig. 10 shows that we obtain the highest score when the eps parameter is equivalent to the true fiber radious. However, Fig, 9 shows that our method is more robust to parameter variations than the rest of proposed approaches.

4.3. High resolution SFRP: Polypropylene Matrix

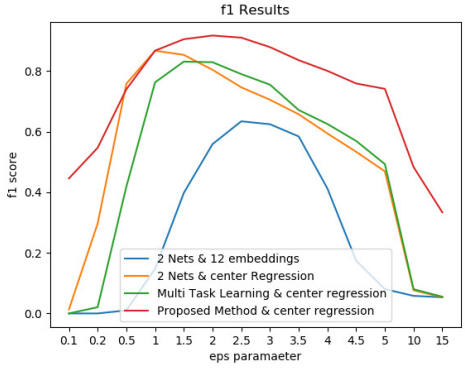
This dataset was provided by the ACME Laboratory at Purdue University[6]. The sample consists of a polypropylene material reinforced with glass fibers, imaged at 1.3μm resolution. The reconstructed volume has dimensions

Method	R_a	f1
Embedding Learning[1]	0.222	0.634
Multitask Learning[9]	0.111	0.831
Centroid Regression	0.563	0.831
Proposed	**0.638**	**0.917**

Table 2. f1 and R_a results for low resolution SFRP dataset.



(a) Ra scores



(b) f1 scores

Figure 9. Evaluation of mean tile scores vs eps parameter in Low Resolution SFRP Dataset. For the proposed method, we let eps = r_{min}.

$2300 \times 2300 \times 1300$ voxels and we used a sub-volume comprised of $950 \times 950 \times 150$ voxels (shown in Fig. 11(a)). The ground truth is from the results of Agyei et al.[2] and is shown in Fig. 11(b).

The first row of Fig. 11 shows the XY-axis cross section of the testing sample, the second row shows a cropped volume of size $256 \times 256 \times 150$, and the third row shows a YZ-axis cross section of the cropped volume. All the networks were trained using the synthetic dataset from section 4.1. This data is significantly different from the training dataset, and hence both the embedding learning and multitask method merge multiple nearby fibers. These results can be noticed in Figs. 11(h), 11(m). and Figs. , 11(i), 11(n). Our results, shown in Figs. 11(j) and11(o) show that our method does not merge nearby fibers thanks to the regularization imposed by cylindrical geometry regularization. We should also point out that our method detects fibers that were not detected in the ground truth dataset. Our method also shows promising results for the merged volume in Fig. 11(e) compared to the merged results shown in Figs.11(c) and 11(d). Our method shows improvement over the provided ground truth in detecting long fibers shown in Fig. 11(b).

Method	R_a	f1
Embedding Learning[1]	0.365	0.604
Multitask Learning[9]	0.051	0.733
Centroid Regression	0.134	0.767
Proposed	**0.422**	**0.855**

Table 3. f1 and R_a results for high resolution SFRP: Polypropylene Matrix.

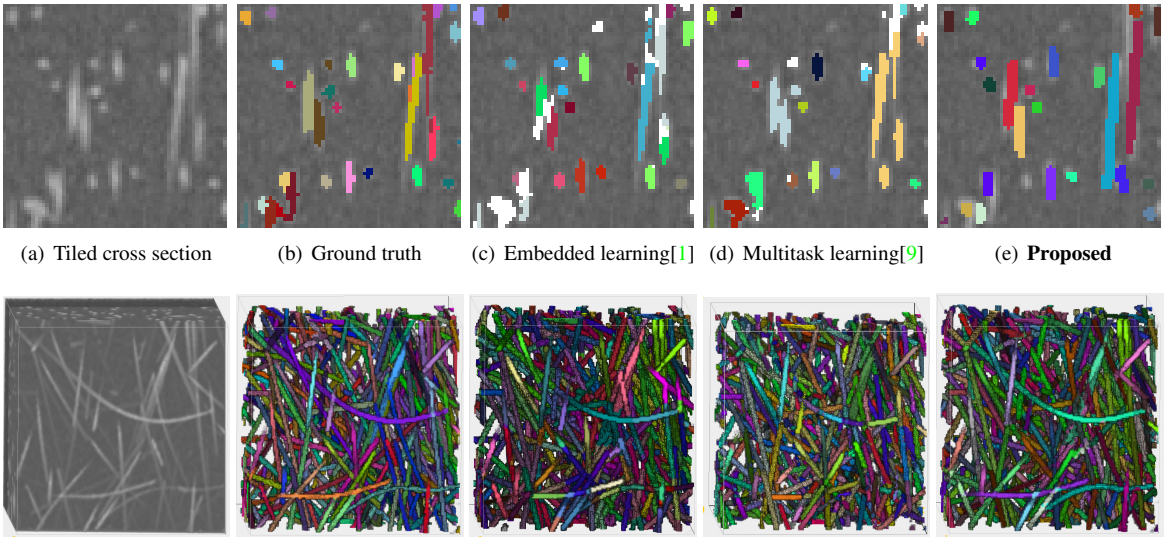
Table 3 shows the R_a and f1 results for each approach. The R_a score is relatively low because it also depends on the segmentation and our method has an over segmentation compared to the provided ground truth, and also because the provided ground truth is not perfect. Fig. 11 shows several visual examples where we believe our approach captured fibers that were not captured in the provided ground truth.

5. Conclusion

We presented a neural network approach to detect fibers in large volumetric datasets by first segmenting and then regressing a vector pointing from each foreground pixel to its instance centroid. Our approach depends on the parameter r_{min} that is related to the fiber minimum radius and shows robustness across several datasets thanks to the geometric constrained clustering and also allows constraining the detected objects with prior image knowledge. Unlike common instance segmentation techniques, we propose a clustering technique that relies on finding objects of specific shapes(cylinders). We showed that our approach outperforms 3D fiber detection in several datasets and we believe that it can contribute to improve fiber-reinforced materials characterization. Similarly, we believe our approach can be extended to other geometries such as using spheres to segment cell nuclei.

Acknowledgments

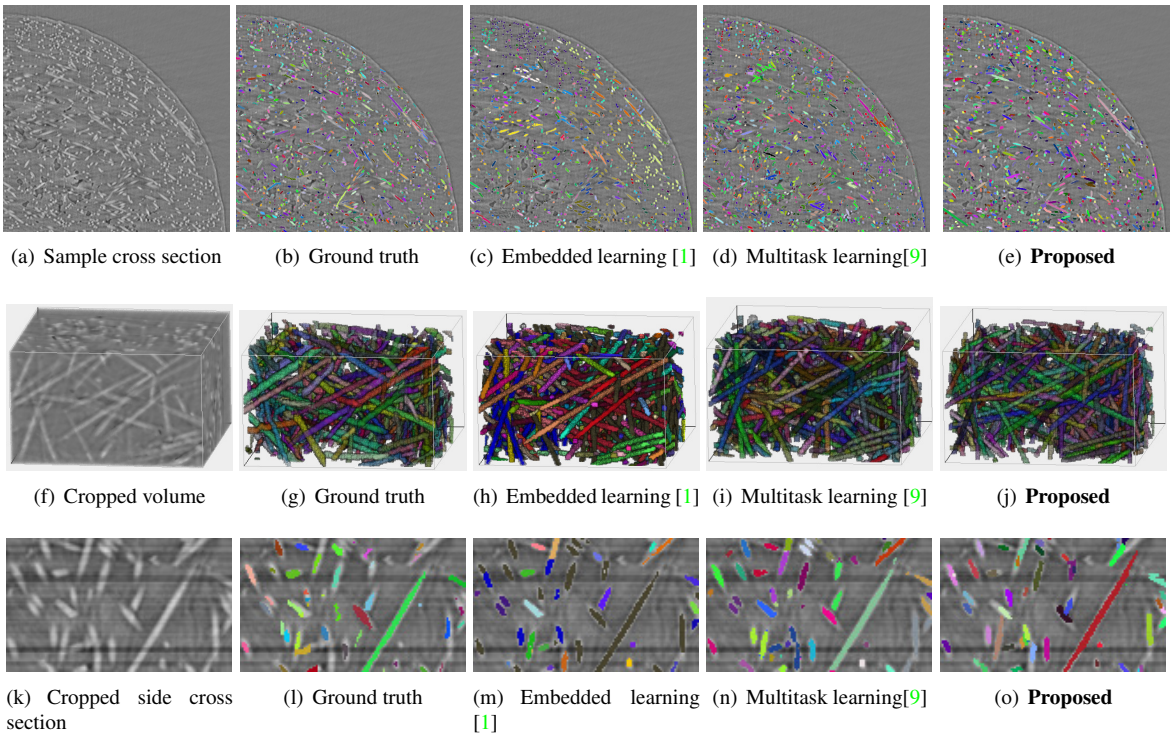
The authors would like to acknowledge the support of the National Science Foundation through grant NSF CMMI MoM 16-62554. Also, the authors would like to thank Dr. Xianghui Xiao for assistance with the tomography characterization at the Advanced Photon Source. The use of the Advance Photon Source is granted by the US Department



(a) Tiled cross section (b) Ground truth (c) Embedded learning[1] (d) Multitask learning[9] (e) **Proposed**

(f) Raw volume (g) Ground Truth (h) Embedded learning[1] (i) Multitask learning[9] (j) **Proposed**

Figure 10. Low Resolution SFRP Dataset. Each color represents a different fiber instance and white pixels represent unlabeled pixels.



(a) Sample cross section (b) Ground truth (c) Embedded learning [1] (d) Multitask learning[9] (e) **Proposed**

(f) Cropped volume (g) Ground truth (h) Embedded learning [1] (i) Multitask learning [9] (j) **Proposed**

(k) Cropped side cross section (l) Ground truth (m) Embedded learning [1] (n) Multitask learning[9] (o) **Proposed**

Figure 11. SFRP: Polypropylene Matrix. Each color represents a different instance of a fiber

of Energy, Office of Science, and Office of Basic Energy Sciences under Contract No. DE-AC02 06CH11357.

References

[1] Camilo Aguilar, Imad A. Hanhan, Ronald F. Agyei, Michael D. Sangid, and Mary Comer. Void detection and fiber extraction for statistical characterization of fiber-reinforced polymers. *"Electronic Imaging"*, "2020"("14"):"250–1–250–7", "2020".

[2] Ronald F. Agyei and Michael D. Sangid. A supervised iterative approach to 3D microstructure reconstruction from acquired tomographic data of heterogeneous fibrous systems. *Composite Structures*, 206(January):234–246, 2018.

[3] Ronald F. Agyei, Bhisham Sharma, and Michael Sangid. Investigating Sub-surface Microstructure in Fiber Reinforced Polymer Composites via X-Ray Tomography Characterization. *57th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, (January):1–6, 2016.

[4] Bert De Brabandere, Davy Neven, and Luc Van Gool. Semantic instance segmentation with a discriminative loss function. *ArXiv*, abs/1708.02551, 2017.

[5] Michael Hahsler, Matthew Piekenbrock, and Derek Doran. dbscan: Fast density-based clustering with R. *Journal of Statistical Software*, 91(1):1–30, 2019.

[6] Imad Hanhan and Michael D Sangid. ModLayer: A MATLAB GUI Drawing Segmentation Tool for Visualizing and Classifying 3D Data. *Integrating Materials and Manufacturing Innovation*, 8(4):468–475, 2019.

[7] K. He, G. Gkioxari, P. Dollár, and R. Girshick. Mask r-cnn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017.

[8] Lawrence Hubert and Phipps Arabie. Comparing partitions. *Journal of Classification*, 2(1):193–218, 1985.

[9] Alex Kendall, Yarin Gal, and Roberto Cipolla. Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics arXiv : 1705 . 07115v1 [cs . CV] 19 May 2017. *Cvpr*, pages 7482–7491, 2018.

[10] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[11] Tomasz K. Konopczynski, Thorben Kröger, Lei Zheng, and Jürgen Hesser. Instance segmentation of fibers from low resolution ct scans via 3d deep embedding learning. In *BMVC*, 2018.

[12] Tomasz K. Konopczynski, Jitendra Rathore, Thorben Kröger, Lei Zheng, Christoph S. Garbe, Simone Carmignato, and Jürgen Hesser. Reference setup for quantitative comparison of segmentation techniques for short glass fiber ct data. *ArXiv*, abs/1901.01210, 2019.

[13] Tianyu Li, Mary Comer, and Josiane Zerubia. A Connected-Tube MPP Model for Object Detection with Application to Materials and Remotely-Sensed Images. *Proceedings - International Conference on Image Processing, ICIP*, pages 1323–1327, 2018.

[14] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia. Path aggregation network for instance segmentation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8759–8768, 2018.

[15] Davy Neven, Bert De Brabandere, Marc Proesmans, and Luc Van Gool. Instance segmentation by jointly optimizing spatial embeddings and clustering bandwidth. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[16] Xingjie Pan. Fit a set of 3D points to a cylinder surface, Mar. 2017.

[17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351:234–241, 2015.

[18] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, Jean-Yves Tinevez, Daniel James White, Volker Hartenstein, Kevin Eliceiri, Pavel Tomancak, and Albert Cardona. *Fiji: an open-source platform for biological-image analysis*, 06 2012.

[19] J. Uhrig, E. Rehder, B. Fröhlich, U. Franke, and T. Brox. Box2pix: Single-shot instance segmentation by assigning pixels to object boxes. In *IEEE Intelligent Vehicles Symposium (IV)*, 2018.