Densely Populated Cell and Organelles Segmentation With Multi-Plane Deep Learning Pipeline

Samuel J. Fulton[1], Cameron M. Baenen[1], Brian Storrie[2], Zijie Yang[1], Richard D. Leapman[1], and Maria A. Aronova[1]
[1]*Nation Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Bethesda, Maryland, USA and*
[2]*Department of Physiology and Cell Biology, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA*
(Dated: November 19, 2024)

Here we develop a robust machine vision pipeline for cell and organelle segmentation within vEM datasets. We collect neural network predictions along multiple planes, capturing 3D correlations using only 2D neural networks. We segment and analyze hundreds of platelets and report quantitative morphological measurements. We apply our segmentation pipeline to the CREMI neuron segmentation challenge data and provide state-of-the-art results. Our work paves the way for large-scale platelet studies and lends unique insight into platelet structure. We publish a lightweight Jupyter notebook for initializing and running the neural network. We also provide Amira Avizo protocols for converting neural network outputs to cell and organelle instances and subsequent morphological measurements.

INTRODUCTION

Advances in high-throughput volume electron microscopy (vEM) have enabled the collection of gigavoxel datasets consisting of hundreds of cells and thousands of organelles [1]. An important goal in vEM is to elucidate structure-function relationships, requiring instance-level segmentation [2]. Historically, segmentation was performed by hand. However, for high-throughput vEM, hand segmentation becomes time-intensive, requiring thousands of human hours for tens of cells [3].

We improve upon semantic segmentation work conducted by Gauy et al., to present an instance-level platelet and organelle segmentation pipeline [1]. Our pipeline is based on a 2D U-Net with a ResNet18 backbone, shown in Fig.1. We leverage multi-plane analysis, systematically capturing 3D correlations with only 2D neural networks. Multi-plane analysis improves segmentation results and represents a viable alternative to 3D neural networks. We segment a previously unanalyzed Focused Ion Beam Scanning Electron Microscopy (FIB-SEM) dataset consisting of densely packed platelets within thrombi. This work extends the research of Matharu et al., which analyzed loosely packed platelets [4]. We report precision, recall, F1 score, and accuracy for our platelet dataset as well as CREMI Serial Block Face Scanning Electron Microscopy (SBF-SEM) neuron dataset[5].

PLATELET DATASET AND MULTI-PLANE ANALYSIS

We center the analysis on a FIB-SEM platelet dataset from a mouse thrombus, which is comprised of tightly arranged blood platelets. The FIB-SEM dataset is stored in a 3D tiff stack with dimensions $2796 \times 2388 \times 1012$ and voxel size $5 \times 5 \times 7$ nm. We perform instance-level segmentation of mitochondria, α-granules, and whole platelets. The highly intertwined platelets result in a complex segmentation task, leading to difficulties for traditional 2D neural networks. To improve upon standard 2D segmentation methods, we run a 2D neural network along three orthogonal axes. We start by running the neural network on every 2D slice in the tiff stack. We then rotate the tiff stack and run the network on the new forward face of the data. We repeat the process a second time so that the network has been run on three orthogonal axes. We integrate the predictions into a single 3D probability map, a technique we refer to as multi-plane analysis. This approach is motivated by the geometric observation that when a membrane is parallel to the forward-facing plane, it becomes challenging to distinguish the membrane from the background, often causing the U-Net to produce false negatives. However, because the cell membrane can only be parallel to one of the three orthogonal planes, multi-plane analysis minimizes the likelihood of false negatives to just one plane. A visual depiction of the multi-plane analysis process is provided in Fig.2.

Utilizing 2D neural networks with multi-plane analysis has several advantages over 3D neural networks and captures most of the predictive power of a 3D neural network. The first advantage is the fact that 3D training datasets are far more difficult to generate than 2D datasets. The second advantage is the existence of flexible 2D pre-trained backbones such as ResNet18 and VGG16, which significantly decrease training time and reduce the need for large hand-labeled training data sets [6]. Similar approaches to multiplaner analysis were implemented in [7, 8]. However, here we extend existing work by applying multi-plane analysis to anisotropic datasets. Furthermore, we supply two pipelines tailored for both organelle and whole-cell segmentation.

NEURAL NETWORK ARCHITECTURE

We design our network based on the U-Net architecture originally proposed by Olaf Ronneberger et al. in 2015 [9, 10]. The encoder of our U-Net consists of layers from

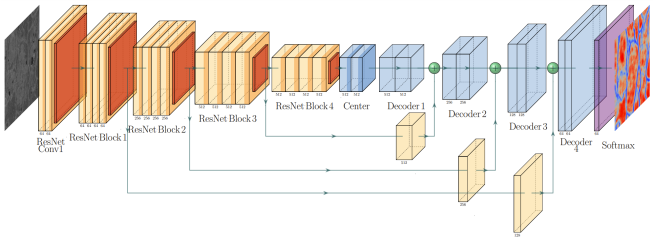

FIG. 1. U-Net architecture with ResNet18 pre-trained backbone and decoder layers, utilizing three residual connections. Network input is a greyscale image and output is a probability map.

ResNet18, which are split into a convolution layer and four subsequent ResNet18 blocks, as shown in 1. The bottleneck consists of two convolution layers and ReLU activations. The decoder consists of four convolution and upsampling blocks. We utilize three residual connections, which help stabilize the training process. Lastly, a 1×1 convolution layer with sigmoid activation produces the segmentation probability maps.

U-NET TRAINING

The training dataset was labeled by hand using Amira Avizo commercial software and consists of fourteen randomly selected images each 2796×2388 pixels. We split our labeled data into 90 percent training and 10 percent validation data. We augmented the training data by applying elastic deformations, rotations, brightness shifts, and contrast shifts.

We randomly cropped each image into 400 patches of 512×512 pixels with each patch receiving a random combination of augmentations. We train for twenty epochs with a batch size of 32 and use early dropout with a patience of five epochs. We train using a single A100 GPU on NIH's computing cluster (Biowulf), where training required thirty minutes [11]. We also train using a local consumer-grade GeForce GTX 1080 GPU which required a longer training time of two hours.

SEGMENTATION PIPELINES

The output of the U-Net is a probability map representing the likelihood of each pixel belonging to a specific material. We use Amira Avizo's watershed segmentation tools to convert the probability maps to cell and organelle instances [12]. The watershed segmentation technique interprets pixel intensities as a topographic surface, "flooding" valleys from predefined markers to segment regions based on their topological boundaries [13]. Numerous software packages and products offer implementations of the watershed algorithm; however, despite sharing a common conceptual basis, their performance can differ significantly [14]. Amira Avizo, for instance, features three primary watershed implementations: H-Extrema, Hierarchical, and Marker-Based, each suited for unique topographical landscapes. Due to differences in the landscape of the organelle and cell membrane probability maps, we propose two segmentation pipelines based on different watershed techniques. Both segmentation pipelines are outlined in Fig.2

Whole-Cell Segmentation Pipeline

The whole-cell segmentation pipeline is a two-step process. First, we train a U-Net to identify cell membranes. We perform multi-plane analysis, following the outline in Fig.2. We use H-Extrema watershed to convert the probability maps into cell instances. H-Extrema watershed uses areas of low-cell-membrane-likelihood as starting seeds for cells and expands outward until the cells bump into each other. This form of watershed is quite robust in the case of uniformly dense cells and produces quality results for the segmentation of platelets and neurons. However, mitochondria and alpha granules, which greatly vary in density throughout space, H-extrema watershed performs poorly. For H-extrema watershed we tune a single parameter that sets the distance between starting seeds. This parameter allows for the merging of starting seeds that are within a defined distance from each other, which is easily tuned for uniform platelet density.

Organelle Segmentation Pipeline

To segment organelles, we employ a dual network pipeline. We train one neural network to identify the organelle membrane and one neural network to identify the entire organelle. We run both neural networks on the data, creating two sets of probability maps. We subtract the membrane map from the whole organelle map to create a probability map of the organelle contents. We use hierarchical watershed to segment only organelle contents (rather than full organelle) ensuring that neighboring organelles are not merged together. We then use marker-based watershed expanding the organelles to include their membranes.

BENCHMARKING ON CREMI DATASET

We benchmark the whole-cell segmentation pipeline on CREMI neuron segmentation challenge Sample_A. The CREMI dataset comprised 125 hand-labeled slices each with a dimension of 1250×1250 pixels. Unlike our FIB-SEM dataset with a near-isotropic voxel shape, the CREMI datasets have a voxel shape of $4 \times 4 \times 40$ nm. To utilize multi-plane analysis, we train two neural networks. The first neural network is trained using 12 slices

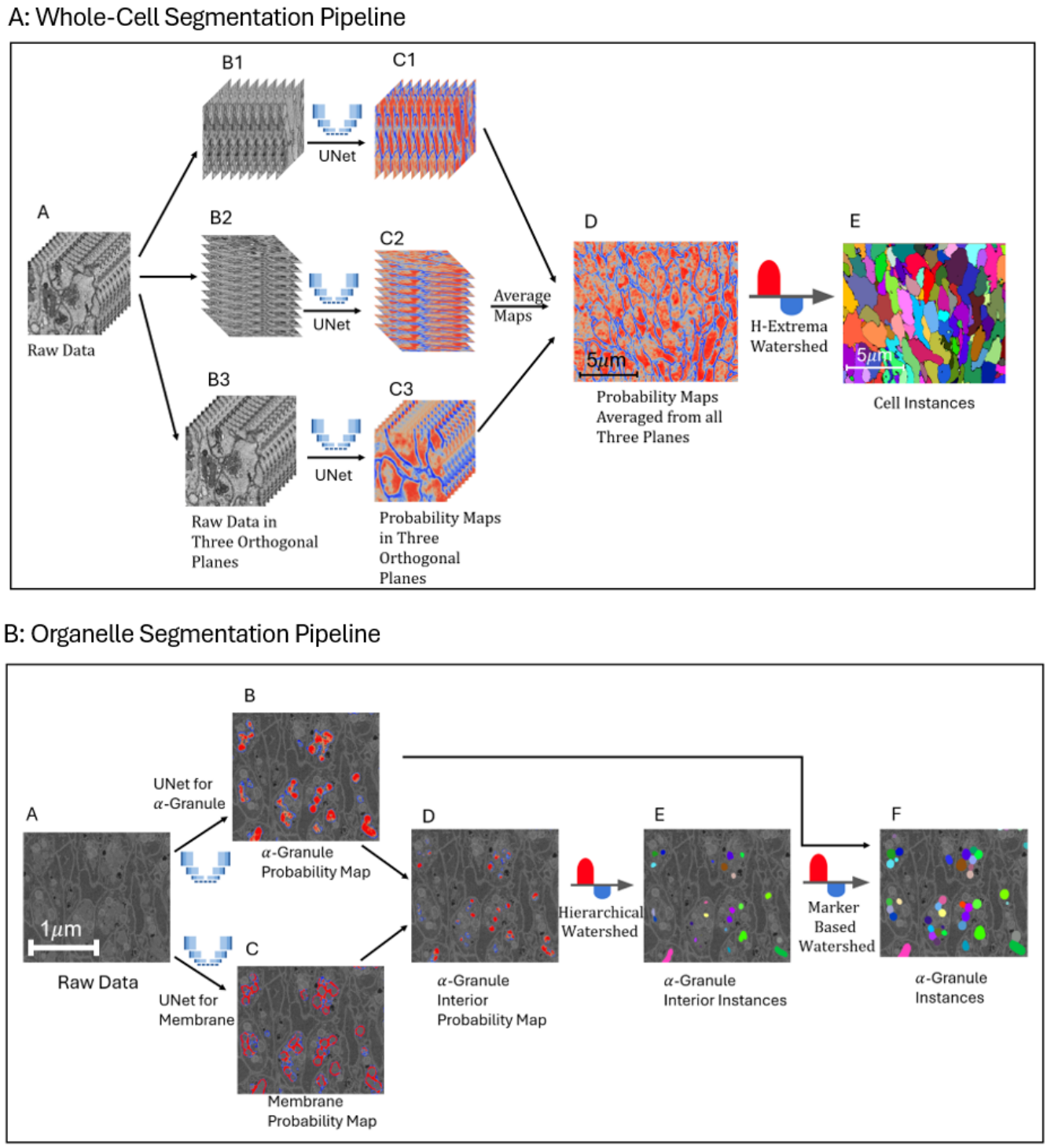

FIG. 2. We propose two complementary pipelines for cell and organelle segmentation. The **Whole-Cell Segmentation Pipeline** performs much better for cases in instances where density is high. Part **(A)** through **(C3)** shows multi-planar analysis. In parts **(D)** and **(E)**, we apply H-Extrema watershed to create final cell instances. The **Organelle Segmentation Pipeline** performs much better for cases of varying instance density. We employ two neural networks: one to segment the entire organelle **(B)** and one to segment organelle membrane **(C)**. We subtract the membrane from the full organelle results, **(D)**, and use Hierarchical watershed to create instance segmentation of the interior of the organelles **(E)**. We then used the interiors and the full organelles probability mask to perform marker-based watersheds, creating instances of full organelles **(F)**.

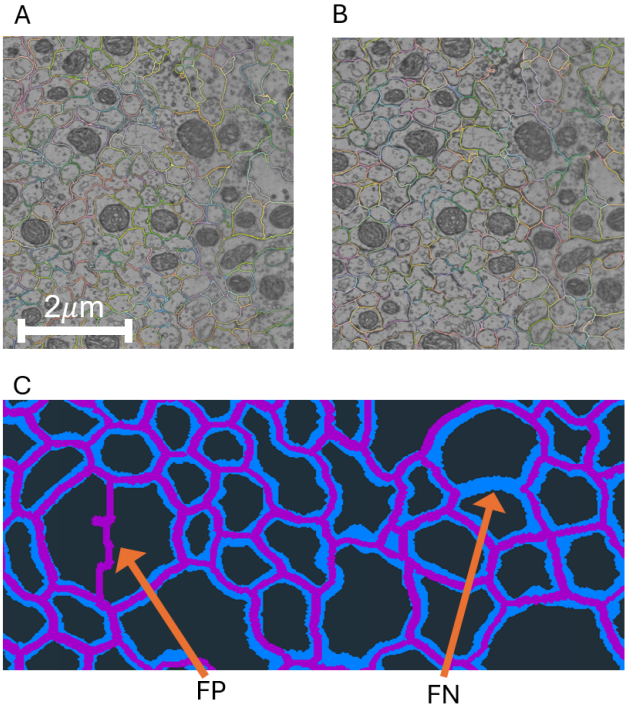

FIG. 3. (A) Ground truth segmentation for CREMI dataset. (B) Segmentation pipeline results for CREMI dataset. (C) Ground truth membranes are shown in blue and whole cell instance membranes are shown in purple. Cases of false positives and falsnegativesve are shown.

from the dataset and can be run in the xy-plane. We create the training dataset for the second neural network by resampling the 12 slices into 120 slices each with pixel shape 4×40 such that the pixel shape matches that of the xz and yz-planes. A slice of our segmentation results is shown next to the ground truth segmentation in Fig.3.

To benchmark the final segmentation output, we assign pixels near the cell instance boundary to the membrane and compare with ground truth membrane. Fig.3 shows ground truth overlaid with membranes from the whole-cell segmentation pipeline. We calculate precision, recall, accuracy, and F1 scores and show results in Table.I. Our pipeline shows quality results, classifying pixels correctly with a precision of 0.923. Some difficulties arise in running watershed on anisotropic datasets. While the U-Net utilizes the high xy-resolution, the watershed algorithm suffers from low z-resolution. Andong Wang et al. implemented an algorithm known as Touching Area-Based Clustering Algorithm (TASCAN) to improve watershed results on anisotropic data [2]. We do not implement TASCAN here. Even with the difficulties with running watershed on anisotropic data, we still see good membrane classification results with accuracy and precision over 0.85.

Metric	**Formula**	**Neuron Membrane**
Precision	$\frac{TP}{TP+FP}$	0.923
Recall	$\frac{TP}{TP+FN}$	0.756
Accuracy	$2\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}$	0.875
F1 Score	$\frac{TP+TN}{TP+TN+FP+FN}$	0.831

TABLE I. The metric formulas are expressed in terms of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

BENCHMARKING PLATELET SEGMENTATION RESULTS

We report the segmentation results of the FIB-SEM platelet dataset for α-granule, mitochondria, and cell membrane detection. Although we produce instance-level segmentation for α-granule, mitochondria, and platelets, we convert all classes to semantic labels for benchmarking. All results show state-of-the-art accuracy and precision in excess of 0.85. Results for analysis can be further improved by tuning parameters to over-divide cells and organelles, effectively creating supervoxels that can be rapidly inspected and re-segmented by hand.

Metric	**Mitochondria**	**Alpha Granules**	**Cell Membrane**
Precision	0.96	0.91	0.87
Recall	0.60	0.79	0.95
F1 Score	0.74	0.85	0.89
Accuracy	0.99	0.97	0.91

TABLE II. The metric formulas are defined in Table.I. All results show state-of-the-art accuracy greater than 0.90.

3D RENDERINGS AND MORPHOLOGICAL MEASUREMENTS

We use Amira Avizo to display segmentation results of mitochondria, α-granules, and cell instances in Fig.4. We also show 3D renderings of cells and organelles in Fig.4. All renderings are displayed in Amira Avizo using the "Generate Surface" and "Surface View" modules. We apply Amira Avizo's "Label-Analysis" tool to calculate volumes and axis lengths for all segmented objects and export the results as CSV files to be analyzed with Excel and Python. We report several parameters of interest and compare them with hand-labeled results published in Table.III. Our measured cell volumes, 2.69 ± 0.12 (μm^3) and 2.90 ± 0.34 (μm^3), are within a single error bar of each other, consistent with the understanding that activation will affect platelet shape but not mass. The results differ because of differences in environment and platelet activation levels. The 3D rendering and mor-

phological measurements provide insight into organelle clustering, cellular interactions, and cell polarity. We hope that large-scale segmentation such as the one performed here will help provide insight into the structural underpinnings of blood clot formation and heart disease.

Parameter	**Our Results** ($\chi \pm \sigma/\sqrt{n}$)	**Published** ($\chi \pm \sigma/\sqrt{n}$)
Platelet Volume (μm^3)	2.69 ± 0.12	2.90 ± 0.34
Platelet Minor and Major Axis (µm)	1.74 ± 0.62 3.74 ± 1.67	0.87 ± 0.07 2.70 ± 0.15
Alpha-Granule Number/Platelet	11 ± 1	19 ± 4.3
Alpha Granule Short and Long Axis (µm)	0.364 ± 0.022 0.464 ± 0.005	0.288 ± 0.006 0.537 ± 0.018
Volume/Alpha Granule (μm^3)	0.025 ± 0.00061	0.013 ± 0.00076
Alpha Granule Volume Fraction/Platelet	0.102 ± 0.005	0.088 ± 0.007
Mitochondria Number/Platelet	5 ± 0.29	12 ± 1.8
Volume/Mitochondria (μm^3)	0.012 ± 0.00027	0.0072 ± 0.00015
Mitochondria Volume Fraction/Platelet	0.023 ± 0.003	0.029 ± 0.002

TABLE III. Measured ultrastructure parameters of interest compared with previously published data. Our calculated results are within reasonable uncertainty given the slight differences in sample environment. Published data retrieved from [15].

OUTLOOK

Here we present the lightweight pipeline for cell and organelle segmentation and morphological analysis. We readily apply the pipeline to the segmentation of cells within a thrombi. We lay a foundation for future efforts to improve automated segmentation in biomedical contexts. We are currently analyzing numerous platelet datasets to compare ultrastructure in differing regions of thrombi. We will use this pipeline to provide unique insight into thrombi formation and heart disease. Furthermore, by enabling extensive quantification of large datasets, our framework provides a base for robust research into cellular dynamics, activation, and environments across fields. We design our pipeline for accessibility and hope to bridge the gap between advanced machine-learning methods and practical applications in experimental biology.

ACKNOWLEDGMENTS

We thank Joshua Kim for collecting the FIB-SEM data, Elizabeth R. Driehaus and Sidney W. Whiteheart for performing the surgeries, Irina D. Pokrovskaya for sample prep, Guofeng Zhang for sample embedding, and the Storrie lab for insightful biological discussions motivating this project.

This work is supported by the intramural research program of the National Institute of Biomedical Imaging and Bioengineering. Research in the Storrie laboratory is supported in part by NIH grant R01 HL155519. This work utilized the computational resources of the NIH HPC Biowulf cluster (`http://hpc.nih.gov`).

CODE AVAILABILITY

All code can be found at our public GitHub repository: `https://github.com/fultonsaj/Deep-Learning-Cell-Seg/tree/main`

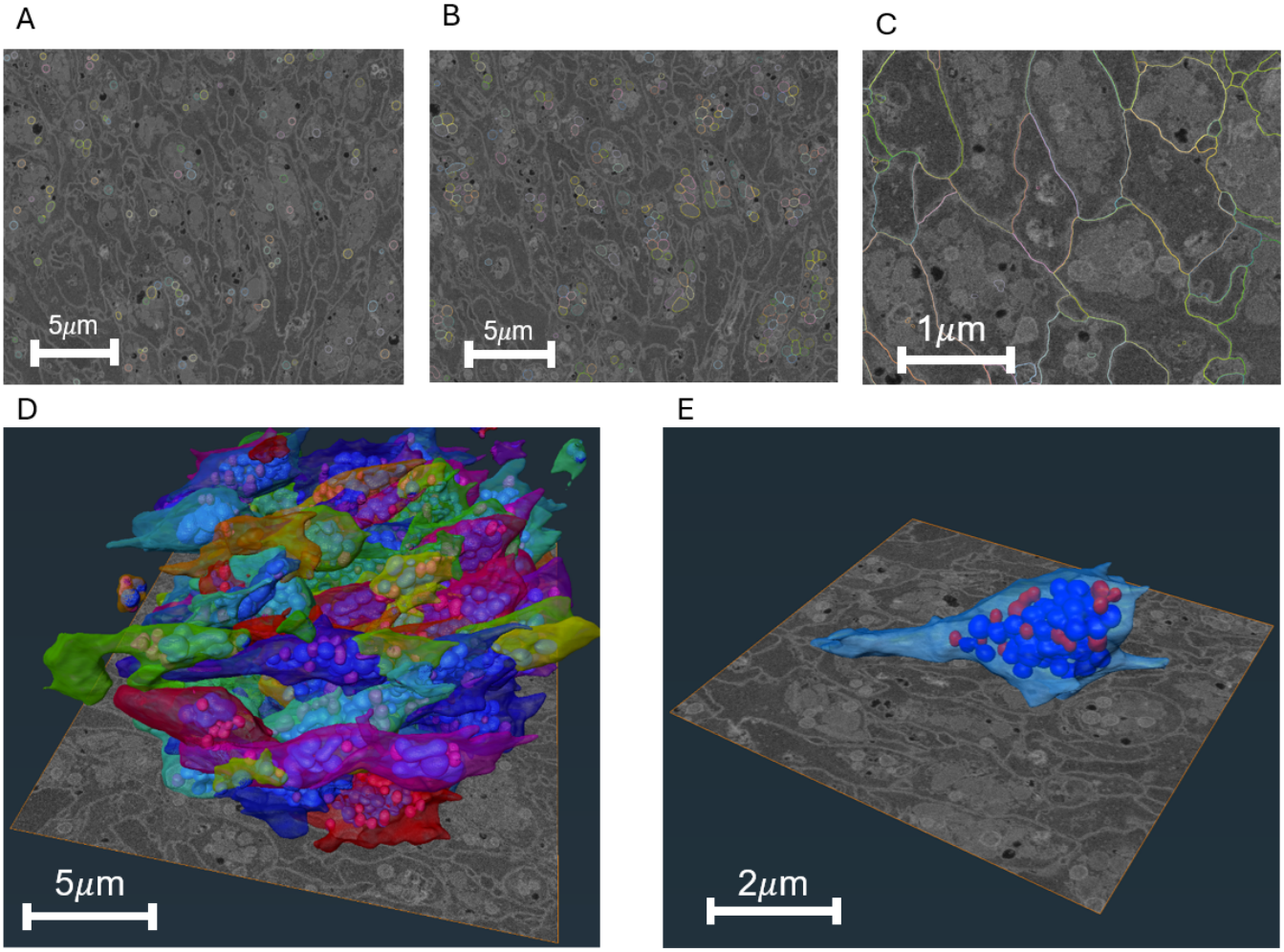

FIG. 4. (A) FIB-SEM data overlaid with corresponding mitochondria instance segmentation results. (B) FIB-SEM data overlaid with α-granule instance segmentation results. (C) FIB-SEM data overlaid with whole-cell segmentation results. (D) 3D rendering showing 127 cells from the interior of a thrombi. (E) 3D rendering showing a single cell from the center of a thrombi.

[1] M. D. Guay, Z. A. S. Emam, A. B. Anderson, M. A. Aronova, I. D. Pokrovskaya, B. Storrie, and R. D. Leapman, Dense cellular segmentation for em using 2d–3d neural network ensembles, Scientific Reports https://doi.org/10.1038/s41598-021-81590-0 (2021).

[2] N. Nešić, X. Heiligenstein, L. Zopf, V. Blüml, K. S. Keuenhof, M. Wagner, J. L. Höög, H. Qi, Z. Li, G. Tsaramirsis, C. J. Peddie, M. Stojmenović, and A. Walter, Automated segmentation of cell organelles in volume electron microscopy using deep learning, Microscopy Research and Technique https://doi.org/10.1002/jemt.24548 (2024).

[3] A. Aswath, A. Alsahaf, B. N. Giepmans, and G. Azzopardi, Segmentation in large-scale cellular electron microscopy with deep learning: A literature survey, Medical Image Analysis (2023).

[4] S. S. Matharu, C. S. Nordmann, K. R. Ottman, R. Akkem, D. Palumbo, D. R. D. Cruz, R. D. Leapman, and B. Storrie, Deep learning, 3d ultrastructural analysis reveals quantitative differences in platelet and organelle packing in covid-19/sarscov2 patient-derived platelets, Platelets **34**, 2264978 (2023).

[5] P. Li, J. Funke, Ariadne, C. Ordish, S. Saalfeld, and J. Funke, Miccai challenge on circuit reconstruction from electron microscopy images (2017).

[6] H. Kaiming, Z. Xiangyu, R. Shaoqing, and S. Jian, Deep residual learning for image recognition, in *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)* (2016).

[7] H. Spiers, H. Songhurst, L. Nightingale, J. de Folterand Christopher J Peddie, A. Weston, A. Strange, S. Hindmarsh, C. Lintott, L. M. Collinson, and M. L. Jones, Deep learning for automatic segmentation of the nuclear envelope in electron microscopy data, trained with volunteer segmentations traffic, John Wiley Online Library 10.1111/tra.12789 (2021).

[8] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, Cellpose: a generalist algorithm for cellular segmentation, Nature Methods https://doi.org/10.1101/2020.02.02.931238 (2021).

[9] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, *U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications* (IEEE, 2021).

[10] O. Ronneberger, P. Fischer, and T. Brox, U-net: Convolutional networks for biomedical image segmentation, Springer Cham https://doi.org/10.1007/978-3-319-24574-4 (2015).

[11] National Institutes of Health, Biowulf high-performance computing, `https://hpc.nih.gov` (2024), accessed: 2024.

[12] Thermo Fisher Scientific, Amira-Avizo Software for 3D Visualization and Analysis (2024), [Software]. Available from: `https://www.thermofisher.com`.

[13] B. Serge and L. Christian, Use of watersheds in contour detection, In International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation (1979).

[14] A. Kornilov, I. Safonov, and I. Yakimchuk, A review of watershed implementations for segmentation of volumetric images, Journal of Imaging https://doi.org/10.3390/jimaging8050127 (2022).

[15] I. Pokrovskaya, M. Tobin, R. Desai, M. A. Aronova, J. A. Kamykowski, G. Zhang, S. Joshi, S. W. Whiteheart, R. D. Leapman, and B. Storrie, Structural analysis of resting mouse platelets by 3d-em reveals an unexpected variation in alpha-granule shape, Platelets 10.1080/09537104.2020.1799970 (2021).