3D Nuclei Segmentation for Multi-cellular Quantification of Zebrafish Embryos Using NISNet3D

Linlin Li, Liming Wu, Alain Chen, Edward J. Delp, David M. Umulis,; Purdue University ; West Lafayette, IN

Abstract

Advances in imaging of developing embryos in model organisms such as the fruitfly, zebrafish, and mouse are producing massive data sets that contain 3D images with every cell and readouts of signaling activity in every cell of an embryo. In Zebrafish embryos, determining the locations of nuclei is crucial for the study of the spatial-temporal behavior of these cells and the control of gene expression during the developmental process. Traditional image processing techniques suffer from bad generalizations, often relying on heuristic measurements that narrowly apply to specific data types, microscope settings, or other image characteristics. Machine learning techniques, and more specifically convolutional neural networks, have recently revolutionized image processing and computer vision. A well-known challenge in developing these algorithms is the lack of curated training data. We developed a new, manually-curated nuclei segmentation data set for four complete zebrafish embryos containing over 8,000 cells each. The whole-mount zebrafish embryos at different development stages were hand-labeled with 3D volumetric segmentation of nuclei. Two full embryo data sets were used for training the 3D nuclei instance segmentation network NISNet3D, and the other two embryos were used to validate the training results. We provide both qualitative and quantitative evaluation results for each of the volumes using multiple evaluation metrics. We also provide fully curated and manually segmented embryo data sets, along with raw images, for the image processing community.

Introduction

Three-dimensional whole-mount immunofluorescent imaging approaches have contributed broadly to the study of embryonic development typically by imaging via confocal or light-sheet microscopy [3]. Computational imaging methods are applied to analyze and quantify the embryo's nuclei over space and time during crucial periods of early development. Segmentation and localization techniques, widely employed in microscopic imaging, allow for this analytical characterization. Several nuclei segmentation methods are available and utilized widely by the imaging community.

Zebrafish (*Danio rerio*) is a popular model organism to study developmental biology. Its physiological and genetic similarities with humans make its study relevant for areas such as drug discovery, molecular genetics, and different diseases such as cancer. Moreover, in developmental biology, some of the characteristics displayed by Zebrafish make them an ideal model organism. The transparent embryo's development outside the mother and rapid development allows for convenient study and observation. Thus, zebrafish embryos has be wildly used to study morphogenetic events in developmental biology. High-quality 3D or volumetric *in situ* and *in vivo* data with nuclei position along with protein/mRNA distribution provide comprehensive information for understanding biological events during development, such as signaling and gene regulatory networks, cell movement with fate mapping, and cell lineage tracing. Such data require a high-resolution image with accurate segmentation to identify nuclei position and quantify the fluorescence intensity of the protein and mRNA targets of interest. For instance, our previous study focused on quantifying the expression and distribution of proteins and mRNAs in relation to bone morphogenetic protein(BMP) signaling components in zebrafish embryo[26, 12, 16]. During embryonic development, morphogen gradients guide individual cell fate decisions, specifically in the zebrafish embryo, BMP concentration gradients determine the positional information for dorsal-ventral body axis formation[14]. Phosphorylated-Smad (P-Smad) accumulated in the nucleus is a direct readout for the BMP signaling activities. Quantified measuring of the fluorescent-antibody stained intensity of P-Smad5 over the embryo relies upon an accurate segmentation of the nuclear channel.

Traditional nuclear segmentation techniques rely on the contrast between foreground and background, or distinctive pixel features[5, 6, 15]. Preprocessing and calibration steps are also commonly applied to account for imaging environments. The variability and characteristics of the nuclei from real-life microscopic zebrafish images can prove challenging to the previously mentioned methods. Furthermore, images have a high density of overlapping nuclei. Techniques such as the watershed algorithm attempt to segment in these conditions, but over-cutting problems may occur[19, 4]. Other methods striving for accuracy are computationally expensive and may require previous knowledge[1, 8, 10]. Previously we developed different methods including threshold method[26] and wavelet-based segmentation method on zebrafish embryo imaging data addressing some of these obstacles[22]. In the Wavelet-based segmentation method developed by Wu[22] some of the challenges associated with preprocessing and calibrations steps of threshold-based segmentation were alleviated. The Wavelet method achieves a higher True Positive (TP) rate and Precision than Derivative Sum (DS), point-wise and other methods. However, with a larger set of test images, the performance decreased and the time to carry out the segmentation increased. Overlapping nuclei and areas with a high density of nuclei proved especially difficult for the wavelet method, requiring massive amounts of computational time for embryos with more than 8000 nuclei.

Recently, deep learning (DL) methods has been applied to investigate biomedical microscopy images analysis[23]. Several deep learning models have been developed specifically for image segmentation especially using neural nets to predict the masks of cells[2, 9, 17] and usually providing accurate and robust results across a range of image types. However, the major bottleneck in

using deep learning methods in analyzing microscopy images is obtaining the large amounts of manually annotated (ground truth) data needed for training, validation, and testing. Manually annotating 3D whole-mount image data is a labor-intensive and time-consuming process and requires expert verification. In this paper, we manually hand-labeled the nuclei mask of four whole-mount zebrafish embryos collected at different stages during the early development as the ground truth data. The data was used to train a pre-developed network NISNet3D, a true 3D segmentation methods based on a 3D Convolutional Neural Network (CNN)[20]. The method presented herein started by training a NISNet3D model that accelerates our process time on large embryo data sets and has improved the accuracy of 3D segmentation output. Notably, we started with a pre-trained NISNet3D that was already trained on other types of synthetic microscopy volumes generated from SpCycleGAN [7]. The pre-trained NISNet3D is described and can be obtained from[20].

Method

Four whole-mount zebrafish embryos were fixed in four different development stages (4.7hpf, 5.7hpf, 6.3phf, and 8hpf) and imaged with various acquisition methods for nuclei staining and imaged through different Microscope settings and excitation wavelengths. We wanted to provide robust and diverse input data so that we could account for expected variability between different labs imaging the embryos. 3D medical image segmentation software application ITK-SNAP has was to generate the ground-truth masks [24]. Students and lab technicians with biomedical engineering and medical background contributed in preparation and hand labeling. Following the labeling, we used use the convolutional neural network, NISNet3D ([20]) for 3D nuclei instance segmentation in our zebrafish embryo volumes. The architecture of NISNet3D is shown in Figure 1. NISNet3D is a modified 3D U-Net model with residual blocks, attention gates, and shortcut feature concatenations.

We used Embryo 1 and Embryo 2 to train NISNet3D and used Embryo 3 and Embryo 4 for testing. Since Embryo 1 and 2 are large and most regions within these two volumes do not contain any nuclei, we cropped them into 571 subvolumes each of size 64×64×64 and containing at least one nucleus. The NISNet3D was trained with the Adam optimizer [11], with a constant learning rate of 0.001 for 200 epochs. The weight coefficients for loss functions are set to the default, which is $\lambda_1 = 1$, and $\lambda_2 = \lambda_3 = 10$. The parameter T_m used for nuclei separation is set to the default, which is 0. NISNet3D is capable of inferencing on a very large volume of any size by using a divide-and-conquer inference strategy. Specifically, it splits the given volume into many fixed-sized subvolumes where each subvolume overlaps with adjacent subvolumes. Then it inferences each of the subvolumes, fuses the overlapping regions to avoid errors and concatenates these subvolumes back to a large segmentation mask.

Dataset Acquisition

Embryo 1: The embryo was fixed at 4.7 hpf with 4% paraformaldehyde at 4°C overnight. Nuclei were stained by SYTOX Orange (Invitrogen, S11368) along with p-Smad. The embryo was mounted in BABB reagent (Benzyl alcohol (Sigma-Aldrich, B1042) and Benzyl benzoate (Sigma Aldrich, B6630), 1:2 volume ratio) and scanned using a Zeiss LSM 800 upright

confocal microscope with a 20x water immersion lens. The image was collected with a total pixel size of 1946x1945, 348 z slices and 0.446 um scales to x and y pixel size, and 1 um voxel depth of z step size.

Embryo 2: We used the protocol shown in our published paper [25]. DAPI was used to stain the nuclei along with RNAscope Fluorescent multiplex detection reagents (ACD, #320851) of *noggin* and *chordin* mRNA. The Embryo was fixed in 8hpf. Embryos were mounted in 1% low melting agarose on 35 mm glass bottom microwell dishes (Matek, P35G-1.5-10-C). Whole-mount embryos were fixed at 8hpf and imaged with a 20 ×/1.0 Plan-Apochromat water immersion lens (D = 0.17 M27 75 mm). The image was collected with total pixel size 2019x1981, 188 z slices, 0.52 um scales to x and y pixel size, and 2 um voxel depth of z step size.

Embryo 3: This data set was adopted from our previous published paper [26]. The image was collected with LD LCI Plan-Achromat 25x/0.8 Imm Corr DIC M27 multi-immersion lens. The embryo was fixed at 5.7hpf, applied with clearing method BABB and the yolk was removed and imaged with the orientation from the animal closest to the coverslip. The image was collected with a total pixel size of 1024x1024, 122 z slices, 0.5535106 um scales to x and y pixel size, and 2.2 um voxel depth of z step size.

Embryo 4: The imaging acquisition setting for Embryo 4 was the same as Embryo 2. The embryo was fixed at 6.3hpf. DAPI was used to stain the nuclei along with *bmp2b* and *chordin* mRNA and total pixel size 1945x1945, 282 z slices.

All four embryo image data sets were hand labeled as outlined above by students and lab technicians in our lab who have biomedical engineering and medical background. The ground truth was generated through the 3D medical images segmentation software application ITK-SNAP [24]. Over 30,000 individual nuclei were labeled in a 3D mask through this process.

Neural Network Model

As shown in Figure 1, NISNet3D is a modified 3D U-Net model with residual blocks, attention gates, and shortcut feature concatenations. It also replaced all MaxPooling layers with 3D Convolution blocks of stride size 2. Since encoder-decoder-based networks use bottom-up approaches for image segmentation, they first generate semantic segmentation by assigning each pixel/voxel a class label and then split the objects based on the semantic segmentation masks. Using this approach, post-processing steps such as the watershed method are typically necessary to separate touching objects. In NISNet3D, the instance segmentation is an extension of 3D CentroidNet[21], which is achieved by simultaneously learning a 3D vector field volume that contains the centroid and boundary information for each nucleus. The loss function of NISNet3D is shown in Equation 1.

$$\mathscr{L}(S,\hat{S},V,\hat{V}) = \lambda_1 \mathscr{L}_{TL}(S,\hat{S}) + \lambda_2 \mathscr{L}_{FL}(S,\hat{S}) + \lambda_3 \mathscr{L}_{MSE}(V,\hat{V})$$

$$(1)$$

where $S, \hat{S}, V$ and $\hat{V}$ represent the ground truth binary masks, the output binary segmentation masks from NISNet3D, the ground truth 3D vector field volumes, and the output 3D vector field volumes from NISNet3D, respectively. In addition, λ_1, λ_2, and λ_3 are the weight coefficients for the Tversky Loss L_{TL} [18],
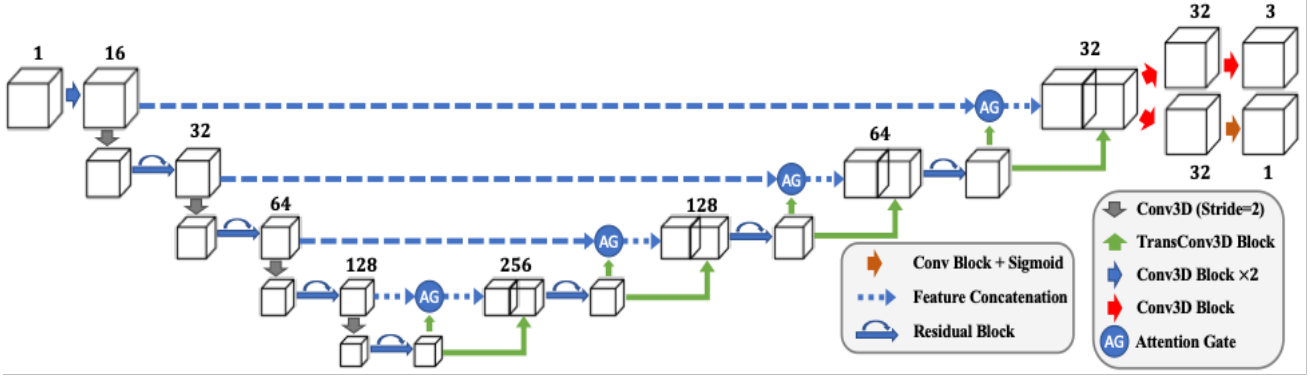
Figure 1. *The architecture of NISNet3D.*

Focal Loss L_{FL} [13], and Mean Squared Error Loss L_{MSE}, respectively. Figure 2 shows the block diagram of nuclei instance segmentation steps for NISNet3D.

The estimated 3D vector field volume, denoted as $\hat{I}^{vec}$, is a 3-channel volume of size 3×X×Y×Z where each voxel is an estimated 3D vector starting from current voxel and ending at the nearest nucleus centroid. $\hat{I}^{vec(x)}, \hat{I}^{vec(y)}$, and $\hat{I}^{vec(z)}$ are the x-, y-, and z-channel of $\hat{I}^{vec}$, respectively, representing spatial shifts of current voxels toward the nearest nuclei centroids in the x-, y-, and z- directions. The 3D vectors on the boundaries of touching nuclei point to very different directions and can have large gradients. The gradient of 3D vector field volume can be obtained using Equation 2.

$$\nabla \hat{I}^{vec} = [\nabla \hat{I}^{vec(x)}, \nabla \hat{I}^{vec(y)}, \nabla \hat{I}^{vec(z)}]^T$$
$$= [\frac{\partial \hat{I}^{vec(x)}}{\partial x}, \frac{\partial \hat{I}^{vec(y)}}{\partial y}, \frac{\partial \hat{I}^{vec(z)}}{\partial z}]^T \quad (2)$$
$$= [S_x * \hat{I}^{vec(x)}, S_y * \hat{I}^{vec(y)}, S_z * \hat{I}^{vec(z)}]^T$$

where $S_x, S_y, and S_z$ represent 3D Sobel filters in the x-, y-, and z-directions, and * is the convolution operation. The 3D vector field volume can be further decoded into a 3D gradient map denoted as I^{grad} using Equation 3.

$$I^{grad} = max(\partial \hat{I}^{vec(x)}, \partial \hat{I}^{vec(y)}, \partial \hat{I}^{vec(z)}) \quad (3)$$

The 3D gradient map is a gray-scale volume where nuclei boundaries have high intensities. Since NISNet3D uses 3D marker-controlled watershed segmentation to separate touching nuclei, it needs to generate accurate markers for each nucleus using the 3D gradient map. Equation 4 shows the marker generation steps.

$$I^{blob} = \sigma(I^{mask} - \tau(I^{grad}, T_m)) \quad (4)$$

$\tau(x,y)$ is a thresholding function that sets the value of x to 1 if x is greater than a threshold y. $\tau(I^{grad}, T_m)$ is used to convert the gray-scale gradient map to a binary volume where nuclei boundaries are highlighted. By subtracting $\tau(I^{grad}, T_m)$ from the binary segmentation mask I^{mask}, the interior regions of nuclei are obtained. $\sigma(x)$ is a function that sets x to 1 if x is negative. I^{blob}

denotes the initial markers which need to be refined using conditional erosion described in Equation 5.

$$I^{mark} = \delta_{t_f}(\delta_{t_c}(I^{blob}, B_c), B_f) \quad (5)$$

where $\delta_{t_c}(I^{blob}, B_c)$ denotes the morphological erosion operation on I^{blob} using coarse structure element B_c until each object size is smaller than t_c voxels. Similarly, $\delta_{t_f}(x, B_f)$ is the morphological erosion operation on x using fine structure element B_f until each object size is smaller than t_f voxels. B_c and B_f are shown in Figure 3, and $T_m, t_c, and t_f$ are set as default values.

Experimental Setup

In our experiments, we use a pre-trained NISNet3D that was already trained on other types of synthetic microscopy volumes generated from SpCycleGAN [7]. The pre-trained NISNet3D is described and can be obtained from [20]. We used Embryo 1 and Embryo 2 to continue training and used Embryo 3 and Embryo 4 for testing the trained NISNet3D on Embryo 1 and Embryo 2. Since Embryo 1 and 2 are large and most regions within these two volumes do not contain any nuclei, we cropped them into 571 subvolumes each of size $64 \times 64 \times 64$ and contain at least one nucleus. The NISNet3D was trained with the Adam optimizer [11] with constant learning rate 0.001 for 200 epochs. The weight coefficients for loss functions are set to the default, which are $\lambda_1 = 1$, and $\lambda_2 = \lambda_3 = 10$. The parameter T_m used for nuclei separation is set to the default, which is 0. NISNet3D is capable of inferencing on a very large volume of any size by using a divide-and-conquer inference strategy. Specifically, it splits the given volume into many fixed-sized subvolumes where each subvolume overlaps with adjacent subvolumes. Then it inferences on each of the subvolumes, fuses the overlapping regions to avoid errors, and concatenates these subvolumes back to a large segmentation mask. During inference, Embryo 3 is up-sampled on the z-direction by a factor of 2 using bilinear interpolation to make the nuclei look more like a sphere or ellipsoid. Both the microscopy volume and the ground truth annotation for Embryo 3 is of size $1024 \times 1024 \times 244$. The segmentation results of Embryo 3 and Embryo 4 will be discussed in the next sections.
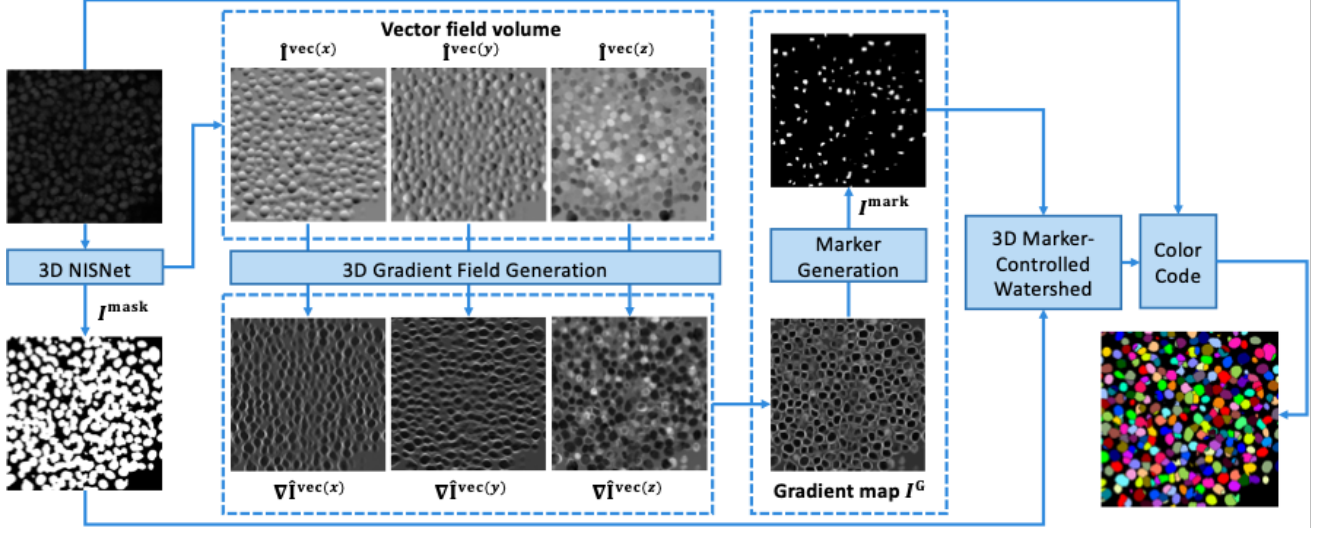
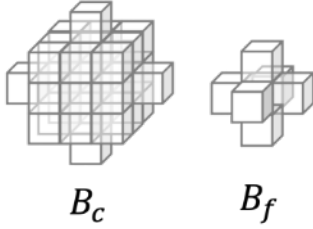**Figure 2.** *Block diagram of 3D nuclei instance segmentation using NISNet3D*

Figure 3. *Coarse structuring element (left) and fine structuring element (right) used for conditional morphological erosion*

Results

Segmentation Results

We demonstrated the segmentation results using NISNet3D on Embryo 3 through the z planes in Figure 4. The first row shows the different XY focal planes of the original embryo volume from the confocal image data. The second row is the corresponding instance segmentation results obtained using NISNet3D. The third row compared the overlapping region through the slices. Figure 5 shows the zoom-in results of Embryo 4 in both 2D and 3D volumes. The results are color-coded, and different color represents different nuclei instances. The third row shows the overlay of the second row on the first row. The results show a highly consistent matching in both 2D and 3D fashion. To quantitatively evaluate the accuracy of the segmentation result, we applied multiple evaluation metrics in the following section for the results obtained from the evaluation data set of Embryo 3 and Embryo 4.

Quantitative Evaluation

We define N_v^{gt} as the total number of annotated nuclei in v-th embryo volume. Similarly, N_v^{det} denotes the number of detected or segmented nuclei for the v-th embryo volume. We use the Mean Absolute Percentage Error (MAPE), which is shown in Equation 6, to measure the nuclei counting accuracy of our neural network model,

$$MAPE = \frac{100\%}{N} \sum_{v=1}^{N} \left| \frac{N_v^{det} - N_v^{gt}}{N_v^{gt}} \right| \quad (6)$$

where N is the total number of embryo volumes for evaluation, and $|\cdot|$ denotes the absolute value. The evaluation results for volume Embryo 3 and Embryo 4 are shown in Table 1.

Table 1: Nuclei counting accuracy evaluation using Mean Absolute Percentage Error (MAPE)

Volume Name	Number of Ground Truth Nuclei	Number of Detected Nuclei	MAPE (%)
Embryo 3	9464	9479	0.16
Embryo 4	7646	7767	1.58

Voxel-based Segmentation Accuracy

Voxel-based segmentation metrics evaluate how accurate the methods segment the nuclei from the background structures. Note that Voxel-based segmentation metrics do not distinguish between different nuclei instances and only evaluate whether a voxel is correctly segmented or not. We first define $n_{tp}, n_{tn}, n_{fp}, n_{fn}, n_{total}$ as the number of True Positives (voxels segmented as nuclei correctly), number of True Negatives (voxels segmented as background correctly), number of False Positives (background voxels falsely segmented as nuclei), number of False Negatives (nuclei voxels falsely segmented as background), and the total number of voxels in an embryo volume.

We use the Type-I and Type-II errors to estimate the rate of falsely segmenting a background voxel as a nuclei voxel, and the rate of falsely segmenting a nuclei voxel as a background voxel. We use Accuracy to estimate the rate of correctly segmenting nuclei voxels and rejecting back. The Type-I error, Type-II error,
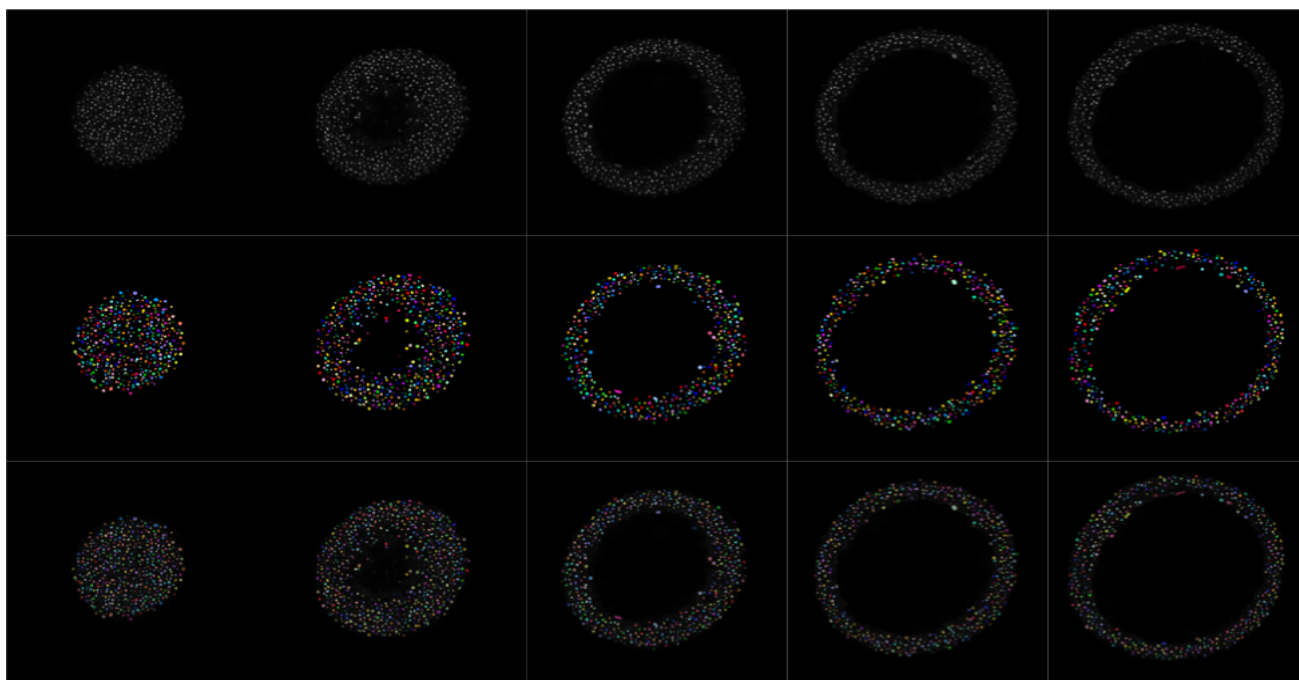
Figure 4. Segmentation for the entire volume (Embryo 3). First row: original embryo volume. Second row: color-coded segmentation and different color represents different nuclei instances. Third row: the overlay of the color-coded segmentation masks on the original embryo volume. Each column represents a X-Y focal plane
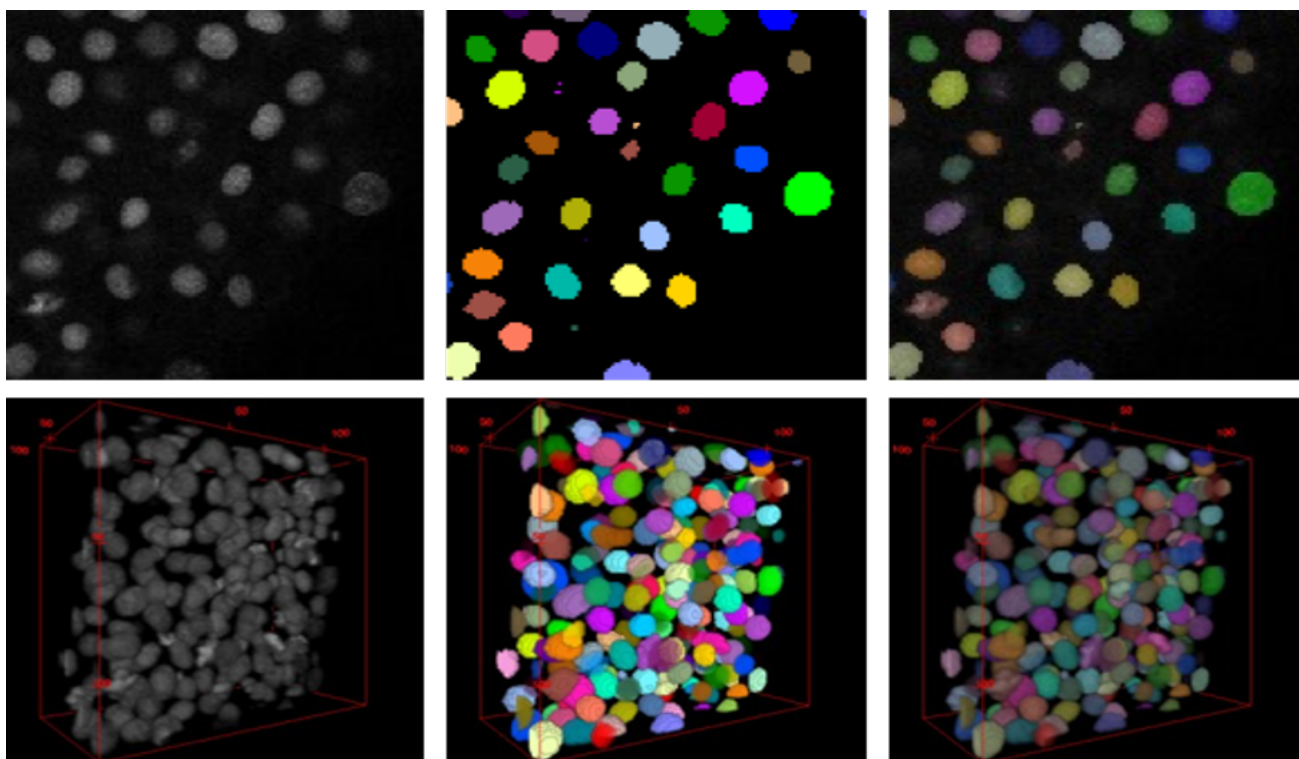

Figure 5. Segmentation for a $128 \times 128 \times 64$ subvolume (Embryo 3). First column: original embryo image/volume. Second column: color-coded segmentation image/volume. Third row: the overlay of the color-coded segmentation masks on the original embryo image/volume. First row: a image slice of the subvolume. Second row: the 3D visualization of the subvolume.

and Accuracy is shown in Equation eq:acc.

$$Type-I\ error = \frac{n_{fp}}{n_{total}},$$
$$Type-II\ error = \frac{n_{fn}}{n_{total}}, \tag{7}$$
$$Accuracy = \frac{n_{tp}+n_{tn}}{n_{total}}$$

We further define the Sensitivity and Specificity in Equation 8.

$$Sensitivity = \frac{n_{tp}}{n_{tp}+n_{fn}},$$
$$Specificity = \frac{n_{tn}}{n_{tn}+n_{fp}} \tag{8}$$

Sensitivity, also known as the True Positive Rate, is used to estimate the rate of correctly segmented nuclei voxels among all ground truth nuclei voxels. Similarly, we use Specificity, also known as the True Negative Rate, to estimate the rate of correctly rejected voxels among all ground truth background voxels. To measure how well the method segments the foreground nuclei, we use the Dice coefficient and Intersection-over-Union (IoU), which are shown in Equation 9, to measure the overlap between all ground truth nuclei voxels and segmented nuclei voxels.

$$Dice = \frac{2n_{tp}}{2n_{tp}+n_{fp}+n_{fn}},$$
$$IoU = \frac{n_{tp}}{n_{tp}+n_{fp}+n_{fn}} \tag{9}$$

We also compared the performance of NISNet3D to a 3D watershed [4] method on the Otsu [15]-thresholded volume. Ostu's method is first used to determine a threshold to separate the foreground (nuclei) and background. 3D watershed is then used to separate individual instances of nuclei.The evaluation results of voxel-based segmentation accuracy for volumes Embryo 3 and Embryo 4 are shown in Table 2.

The results show high accuracy of voxel-based evaluation results for both Embryo 3 and Embryo 4. The average precision of NISNet3D results is over 98%, both Type-I and Type-II errors are equal or less than 1%, Sensitivity is over 70%, and also with high specificity, dice, and IoU. Notably, we also get a high precision result for both of the embryo data by using the Otsu and watershed method in most of the voxel-based evaluation except Sensitivity. To evaluate the segmentation results in individual nuclei we further included the object-based segmentation accuracy evaluation as described in the following section.

Object-based Segmentation Accuracy

We also use the object-based nuclei instance segmentation metrics to evaluate how accurately the method segments individual nuclei. Suppose in the segmentation volume, different nuclei are marked with different intensities. We first define $t = \frac{|G \cap S|}{|G \cup S|}$ as the Intersection-over-Union threshold between a ground truth nucleus G and a segmented nucleus S. If the IoU between G and S is greater than t then we say G and S are matched. Then we define $N_{tp}^t, N_{fp}^t, N_{fn}^t$ as the number of ground truth nuclei with "matched" detections, the number of detected nuclei without "matched" ground truth nuclei, and the number of remaining

ground truth nuclei without any detected nuclei "matched" under the IoU threshold t. Note that the $N_{tp}^t, N_{fp}^t, N_{fn}^t$ defined here is different from the n_{tp}, n_{fp}, n_{fn} defined in last section. To avoid evaluation bias, we obtain the metrics under different IoU thresholds $t \in T_{ious} = 0.25, 0.3, \ldots, 0.45$ and provide the mean of them. As shown in Equation 10, we define the mean Precision (mP), mean Recall (mR), and mean F_1 score as the mean of precision, Recall, and F_1 score evaluated under all IoU thresholds T_{ious}.

$$mP = \frac{1}{|T_{ious}|} \sum_{t \in T_{ious}} \frac{N_{tp}^t}{N_{tp}^t + N_{fp}^t},$$
$$mR = \frac{1}{|T_{ious}|} \sum_{t \in T_{ious}} \frac{N_{tp}^t}{N_{tp}^t + N_{fn}^t}, \tag{10}$$
$$F_1 = \frac{1}{|T_{ious}|} \sum_{t \in T_{ious}} \frac{2N_{tp}^t}{2N_{tp}^t + N_{fp}^t + N_{fn}^t}$$

In addition, we adopt a commonly used object detection evaluation metric in computer vision benchmarks, known as Average Precision (*AP*), which can be obtained by estimating the area under the accumulated Precision-recall Curve. We define AP_t as the *AP* evaluated using the IoU threshold t. We further provide the mean Average Precision (*mAP*) shown in Equation 11.

$$mAP = \frac{1}{|T_{ious}|} \sum_{t \in T_{ious}} AP_t \tag{11}$$

The evaluation results using mP, mR, mF_1, AP_t, and mAP are shown in Table 3. Compared to the voxel-based segmentation accuracy evaluation results, the NISNet3D results show a much higher mean precision(*mP*), mean Recall(*mR*), mean F_1 score(*mF_1*), average precision(*AP*), and mean average precision(*mAP*) compare to the Otsu and Watershed method. These results indicate that the NISNet3D results remain at high accuracy in the segmentation of 3D nuclei and highly match the hand-labeled segmentation data compared to the traditional segmentation method.

Figure 6 demonstrate the Precision, Recall, F_1 *score*, and Average Precision evaluated under different IoU thresholds from 0.25 to 0.45 for Embryo 3 (Left) and Embryo 4 (Right). The results match our expectation with the performance of NISNet3D, since the higher threshold requires a higher match with semantic segmentation results through voxel-wised evaluation.

Discussion

One of the key advantages of using deep learning for nuclei segmentation is the ability to learn complex, hierarchical features from the input images. This allows the model to capture subtle differences in the shape and appearance of nuclei, which is important for accurately segmenting individual nuclei in crowded or overlapping regions. Another advantage is the ability to handle large amounts of data, which is important for nuclei segmentation as it often requires processing large microscopy images with thousands of nuclei. Deep learning models can also be trained on a diverse set of images, which can help improve the generalization of the model and make it more robust to different types of nuclei. In this paper, we applied a recently released deep learning-based method known as NISNet3D in 3D nuclei segmentation of

Table 2. Nuclei semantic segmentation evaluation using Precision, Type-I Error, Type-II Error, Sensitivity, Specificity, Dice coefficient and Intersection-over-Union (IoU)

Volume + Method	Precision (%)	Type-I Error (%)	Type-II Error (%)	Sensitivity(%)	Specificity(%)	Dice(%)	IoU(%)
Embryo 3, NISNet3D	98.75	0.23	1.01	71.39	99.75	80.25	67.06
Embryo 3, Otsu and Watershed	98.36	0.08	1.55	56.36	99.91	71.07	55.12
Embryo 4, NISNet3D	99.18	0.54	0.27	80.11	99.44	72.82	57.26
Embryo 4, Otsu and Watershed	99.17	0.41	0.41	69.77	99.58	69.86	53.68

Table 3. Nuclei object-based instance segmentation evaluation using mean Precision (mP), mean Recall (mR), mean F1 score (mF1), Average Precision (AP), and mean Average Precision (mAP)

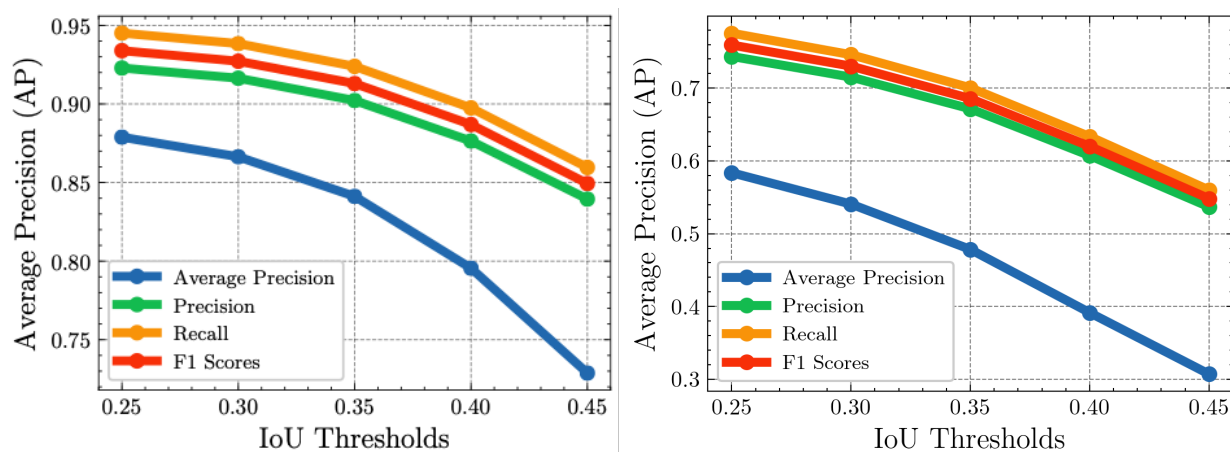| Volume + Method	mP	mR	mF_1	$AP_{0.25}$	$AP_{0.45}$	mAP
Embryo 3, NISNet3D	89.15	91.29	90.21	86.41	71.38	82.22
Embryo 3, Otsu + Watershed	8.81	59.61	15.35	9.56	2.92	5.94
Embryo 4, NISNet3D	65.49	68.31	66.87	58.38	30.71	46.03
Embryo 4, Otsu + Watershed	0.005	5.20	0.009	0.004	0.000	0.0008

Figure 6. *Precision, Recall, F1 scores, and Average Precision using different IoU thresholds for Embryo 3 (left) and Embryo 4 (right) using NISNet3D*

zebrafish embryos. We introduced a large-scale zebrafish embryo dataset which includes four whole-mount zebrafish embryos that were fixed in four different development stages. The embryos are imaged with various acquisition methods for nuclei staining and through different microscopes and excitation wavelengths. The images are also manually annotated and verified by experts. To demonstrate the usefulness of this dataset, we benchmarked the performance of NISNet3D for both qualitative and quantitative evaluation. Our results indicated that NISNet3D has shown promise for 3D nuclei segmentation in zebrafish embryos and has achieved good performance. The dataset can be used for 3D nuclei counting and instance segmentation and further be helpful for quantitative analysis of zebrafish embryos in biological applications.

Data availability

The original image and label are available at https://doi.org/10.5281/zenodo.7671626. The results generated by NISNet3D for the dataset are available upon email request to the author Linlin Li. The NISNet3D code is available in the cited article. [20]

Acknowledgement

References

[1] Mark E Ambühl, Charles Brepsant, J-J Meister, Alexander B Verkhovsky, and Ivo F Sbalzarini. High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *Journal of microscopy*, 245(2):161–170, 2012.

[2] Flávio HD Araújo, Romuere RV Silva, Daniela M Ushizima, Mariana T Rezende, Cláudia M Carneiro, Andrea G Campos Bianchi, and Fátima NS Medeiros. Deep learning for cell image segmentation and ranking. *Computerized Medical Imaging and Graphics*, 72:13–21, 2019.

[3] Evan Bardot, Nikos Tzavaras, Deanna L Benson, and Nicole C Dubois. Quantitative whole-mount immunofluorescence analysis of cardiac progenitor populations in mouse embryos. *JoVE (Journal of Visualized Experiments)*, (128):e56446, 2017.

[4] Serge Beucher. The watershed transformation applied to image segmentation. *Scanning Microscopy*, 1992(6):28, 1992.

[5] Hongmin Cai, Zhong Yang, Xinhua Cao, Weiming Xia, and Xiaoyin Xu. A new iterative triclass thresholding technique in image segmentation. *IEEE transactions on image processing*, 23(3):1038–1046, 2014.

[6] Yan Xue Dong. Review of otsu segmentation algorithm. In *Advanced Materials Research*, volume 989, pages 1959–1961. Trans Tech Publ, 2014.

[7] Chichen Fu, Soonam Lee, David Joon Ho, Shuo Han, Paul Salama, Kenneth W Dunn, and Edward J Delp. Three dimensional fluorescence microscopy image synthesis and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 2221–2229, 2018.

[8] Prabhakar R Gudla, K Nandy, J Collins, KJ Meaburn, T Misteli, and SJ Lockett. A high-throughput system for segmenting nuclei using multiscale techniques. *Cytometry Part A: The Journal of the International Society for Analytical Cytology*, 73(5):451–466, 2008.

[9] Carlos X Hernández, Mohammad M Sultan, and Vijay S Pande. Using deep learning for segmentation and counting within microscopy data. *arXiv preprint arXiv:1802.10548*, 2018.

[10] Humayun Irshad, Antoine Veillard, Ludovic Roux, and Daniel Racoceanu. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. *IEEE reviews in biomedical engineering*, 7:97–114, 2013.

[11] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[12] Linlin Li, Xu Wang, Junyi Chai, Xiaoqian Wang, Adrian Buganza-Tepole, and David M. Umulis. Determining the role of advection in patterning by bone morphogenetic proteins through neural network model-based acceleration of a 3d finite element model of the zebrafish embryo. *Frontiers in system biology*, 2022.

[13] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.

[14] Aasakiran Madamanchi, Mary C Mullins, and David M Umulis. Diversity and robustness of bone morphogenetic protein pattern formation. *Development*, 148(7):dev192344, 2021.

[15] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

[16] Aishwarya Pawar, Linlin Li, Arun K Gosain, David M Umulis, and Adrian Buganza Tepole. Pde-constrained shape registration to characterize biological growth and morphogenesis from imaging data. *Engineering with Computers*, pages 1–16, 2022.

[17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[18] Seyed Sadegh Mohseni Salehi, Deniz Erdogmus, and Ali Gholipour. Tversky loss function for image segmentation using 3d fully convolutional deep networks. In *International workshop on machine learning in medical imaging*, pages 379–387. Springer, 2017.

[19] Luc Vincent and Pierre Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(06):583–598, 1991.

[20] Liming Wu, Alain Chen, Paul Salama, Kenneth Dunn, and Edward Delp. Nisnet3d: Three-dimensional nuclear synthesis and instance segmentation for fluorescence microscopy images. *bioRxiv*, 2022.

[21] Liming Wu, Alain Chen, Paul Salama, Kenneth W Dunn, and Edward J Delp. 3d centroidnet: nuclei centroid detection with vector flow voting. In *2022 IEEE International Conference on Image Processing (ICIP)*, pages 651–655. IEEE, 2022.

[22] Tzu-Ching Wu, Xu Wang, Linlin Li, Ye Bu, and David M Umulis. Automatic wavelet-based 3d nuclei segmentation and analysis for multicellular embryo quantification. *Scientific reports*, 11(1):1–13, 2021.

[23] Fuyong Xing, Yuanpu Xie, Hai Su, Fujun Liu, and Lin Yang. Deep learning in microscopy image analysis: A survey. *IEEE transactions on neural networks and learning systems*, 29(10):4550–4568, 2017.

[24] Paul A. Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C. Gee, and Guido Gerig. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.

[25] JM Zinski, W Dou, Y Huang, D Umulis, and MC Mullins. Dynamics and shaping of the bmp signaling gradient by the bmp antagonists during dv axial patterning. 25, 2014.

[26] Joseph Zinski, Ye Bu, Xu Wang, Wei Dou, David Umulis, and Mary C Mullins. Systems biology derived source-sink mechanism of bmp gradient formation. *Elife*, 6:e22199, 2017.

Author Biography

Please submit a brief biographical sketch of no more than 75 words. Include relevant professional and educational information as shown in the example below.

Linlin Li received her BS in instrument science from Jilin University, China (2008), MS in Mechanical and Aerospace Engineering from Seoul National University, South Korea (2010).She received her Ph.D. in Biomedical Engineering from Purdue University (2021). She has been working as a Research Scientist at Purdue University since then. Her work has focused on system biology and using multi-dimensional models to study gene expression responses to protein signaling networks. Her research also includes artificial intelligence in biology to .

Liming Wu received his B. S. degree in Computer Science and Technology from Anhui University and Technology, China (2017).He received his M. S. degree in Electrical and Computer Engineering from Purdue University Northwest, USA (2019).He joined the Ph.D. program at the School of Electrical and Computer Engineering at Purdue University in August 2019 and worked at the Video and Image Processing Laboratory (VIPER) as a Research Assistant under the supervision of Professor Edward J. Delp since January 2020. His research focus on three-dimensional fluorescence microscopy image synthesis and analysis using machine learning.

Alain Chen is a Ph.D. student at the School of Electrical and Computer Engineering at Purdue University in West Lafayette, Indiana. He is a research assistant at the Video and Image Processing Laboratory under the supervision of Professor Edward Delp since May 2019. His research focus is applying deep learning and image processing to microscopy image analysis.Alain received his B. S. in Computer Science at University of Missouri in Columbia, Missouri in 2018.

Edward J. Delp is the The Charles William Harrison Distinguished Professor of Electrical and Computer Engineering and Professor of Biomedical Engineering at Purdue University in West Lafayette, Indiana. His research interests include image and video processing, image analysis, computer vision, machine learning, image and video compression, multimedia security, medical imaging, multimedia systems, communication and information theory.

Dr. David M. Umulis is the Dane A. Miller Head and Professor of the Weldon School of Biomedical Engineering at Purdue University and Director of the NSF-funded EMBRIO Institute. Umulis leads the Weldon School to advance discovery and translation research in biomedical imaging, neuroengineering, instrumentation and devices, computational biomedicine, and engineered biomaterials. His research interests at the interface of developmental biology, engineering, and medicine and uses simulation and AI combined with cell biology to discover mechanisms of wound repair and cellular defense.