



OPEN

Automatic wavelet-based 3D nuclei segmentation and analysis for multicellular embryo quantification

Tzu-Ching Wu^{1,4}, Xu Wang^{1,2,3,4}, Linlin Li², Ye Bu² & David M. Umulis^{1,2}✉

Identification of individual cells in tissues, organs, and in various developing systems is a well-studied problem because it is an essential part of objectively analyzing quantitative images in numerous biological contexts. We developed a size-dependent wavelet-based segmentation method that provides robust segmentation without any preprocessing, filtering or fine-tuning steps, and is robust to the signal-to-noise ratio. The wavelet-based method achieves robust segmentation results with respect to True Positive rate, Precision, and segmentation accuracy compared with other commonly used methods. We applied the segmentation program to zebrafish embryonic development INTOTO for nuclei segmentation, image registration, and nuclei shape analysis. These new approaches to segmentation provide a means to carry out quantitative patterning analysis with single-cell precision throughout three dimensional tissues and embryos and they have a high tolerance for non-uniform and noisy image data sets.

In developmental biology, quantification of cell morphology, tissue patterning, and gene expression within a developing embryo at different stages provides detailed information that drives our understanding of differentiation and the function of signaling networks driving cell fate specification. Several nuclei segmentation methods have been developed to track nuclei position in fluorescent images and extract the spatiotemporal expression of each nucleus^{1–4}. Most of the nuclei segmentation algorithms need image pre-processing steps before segmentation, and the settings in pre-processing are highly dependent on the microscopy imaging environment⁵. Global and adaptive thresholding segmentation methods are the simplest and the most common nuclei segmentation methods that use single or multiple threshold values on an image histogram to distinguish between an object and background^{6,7}.

However, most real images lack significant thresholding points or low contrast foreground and background⁸. The Watershed method is another common segmentation method also capable of dividing overlapping nuclei^{9,10}. Nearby pixels with similar features are classified topologically and defined as a catchment basin that are separated by a watershed ridge line from the adjacent catchment basin. Using the watershed method can divide overlapping objects, however, it may cause an over-cutting problem. Other methods such as active contour models (ACMs)^{11,12}, level-set¹³ and graph cuts¹⁴ are also applied in the field of nuclei segmentation. Various level-set and model-based methods have been developed for accurate embryo nuclei segmentation^{15–17}, although they are computationally expensive or require pre-knowledge¹⁸. To segment and track chromosome dynamics in the early *Drosophila* embryo, a fast level-set method was proposed¹⁹. A 3D level-set method was also applied to detect nuclei and mitotic chromosomes in *Drosophila* embryos²⁰. A region-based active contour method was developed in 3D cell counting and segmentation of vertebrae in early embryogenesis²¹. We applied a number of these methods to quantify nuclei number and shape in zebrafish, however we encountered a number of problems associated with imaging large amounts of nuclei in a relatively large sample. Specifically, intensity attenuation effects due to absorption and scattering of fluorescent light in deep layer specimens can lead to intensity decay and underestimates in the segmentation results²². Moreover, densely packed nuclei are often present in an embryo which requires overlapping nuclei detection strategies²³. To develop the new approaches, we leverage wavelet transforms that are common in signal processing and are applied in many fields where the method is used to

¹Department of Agriculture and Biological Engineering, Purdue University, West Lafayette, IN 47907, USA. ²Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907, USA. ³Bioland Laboratory, Guangzhou Regenerative Medicine and Health Guangdong Laboratory, Guangzhou 510005, China. ⁴These authors contributed equally: Tzu-Ching Wu and Xu Wang. ✉email: dumulis@purdue.edu

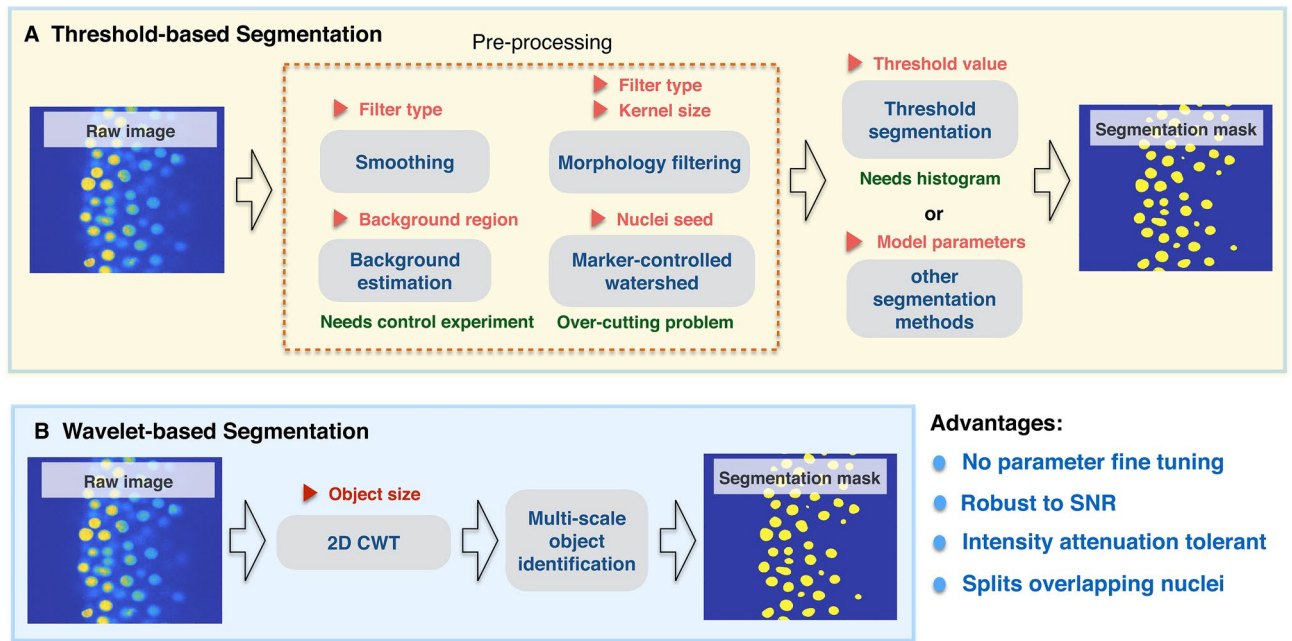


Figure 1. Comparison between threshold-based segmentation and wavelet-based segmentation methods. Preprocessing steps are crucial for most threshold-based segmentation methods and determine the quality of segmentation performance. Increased parameter numbers in preprocessing steps and segmentation method increases result complexity. Red triangle symbols indicate parameters required for each step and may need parameter fine tuning.

isolate the object from a noisy signal. The 2D WTMM (2D wavelet transform modulus maxima) method was first introduced for multiscale edge detection and segmentation of regions of interest in cell nuclei^{24,25} and CT detection²⁶.

Several image analysis software packages have been established, such as ImageJ²⁷, Fiji²⁸, and instrument-bundled packages: ZEN and cellSens. However, commercial bundled packages lack flexibility and may lack affordable options for on-the-fly segmentation, and open source tools are either limited in specific applications, not robust in segmentation results, or not convenient to use. For nuclei segmentation tools, widely used software CellProfiler^{29,30}, provides pre-processing, object identification, and cell counting modules, but it is limited to only 2D images. Additional related software includes Matlab toolbox OML that is capable of automatic cell segmentation. Spatzcells³¹, FISH-quant³² and MINS³³ are Matlab executable codes that can generate segmentation cell masks and spot recognition for fluorescent 2D or 3D images. LSDCAS (The Large-Scale Digital Cell Analysis System)^{34,35} provides the ability for live cell image segmentation and tracking of cell trajectories. RACE³⁶ was designed for automated 3D cell segmentation and cell reconstruction for embryos, based on a watershed segmentation approach that, in our hands is more sensitive to non-uniform or variable-quality images (Table S1).

In this paper we present a wavelet-based automatic 3D nuclei segmentation method that is object size-dependent, robust to noise and intensity attenuation, and can divide overlapping nuclei (Fig. S1–S3) (Table S3). The new segmentation method we introduced in this article can overcome challenges in other contexts for analysis including the need to generate training data for AI or ML methods and it also does not require thresholding values based on intensity. Lastly, the wavelet-based segmentation integrates the segmentation and the smoothing step and there is no need to apply additional pre-processing steps. Overall, the process is composed of six steps, including a 2D continuous wavelet transform, multi-scale object identification, 3D alignment, first and second nuclei division steps and outlier removal. Unlike many segmentation methods, the wavelet-based method has a strong ability to segment objects in noisy images without any preprocessing steps. The method herein combines a 2D continuous wavelet transform (Fig. 1), and the only parameter needed for object determination is the wavelet scale factor determined by target object size, and no other preprocessing steps are needed. Through comparison between several wavelet coefficient maps, wavelet-based segmentation can easily distinguish the object region from the background in images and it can do this without any preprocessing or thresholding steps. We tested the model against manually curated 3D embryonic data and found that it has robust and stable segmentation performance with respect to True Positive Rate (TP rate), precision, and segmentation accuracy when compared to other methods. We integrated the wavelet-based nuclei segmentation, image registration, topology and geometry packages into an automatic interactive analysis platform named WaveletSEG for embryological developmental studies. Additionally, we developed a 3D ground truth annotation tool, synthetic image generator, and a segmented training dataset export tool and data visualization interface for additional data analysis and data validation in WaveletSEG. We quantified 3D nuclei and intensity spatial distributions, topology features and shape classification results at 4.7, 5.7, and 6.5 h post fertilization (hpf) zebrafish using WaveletSEG.

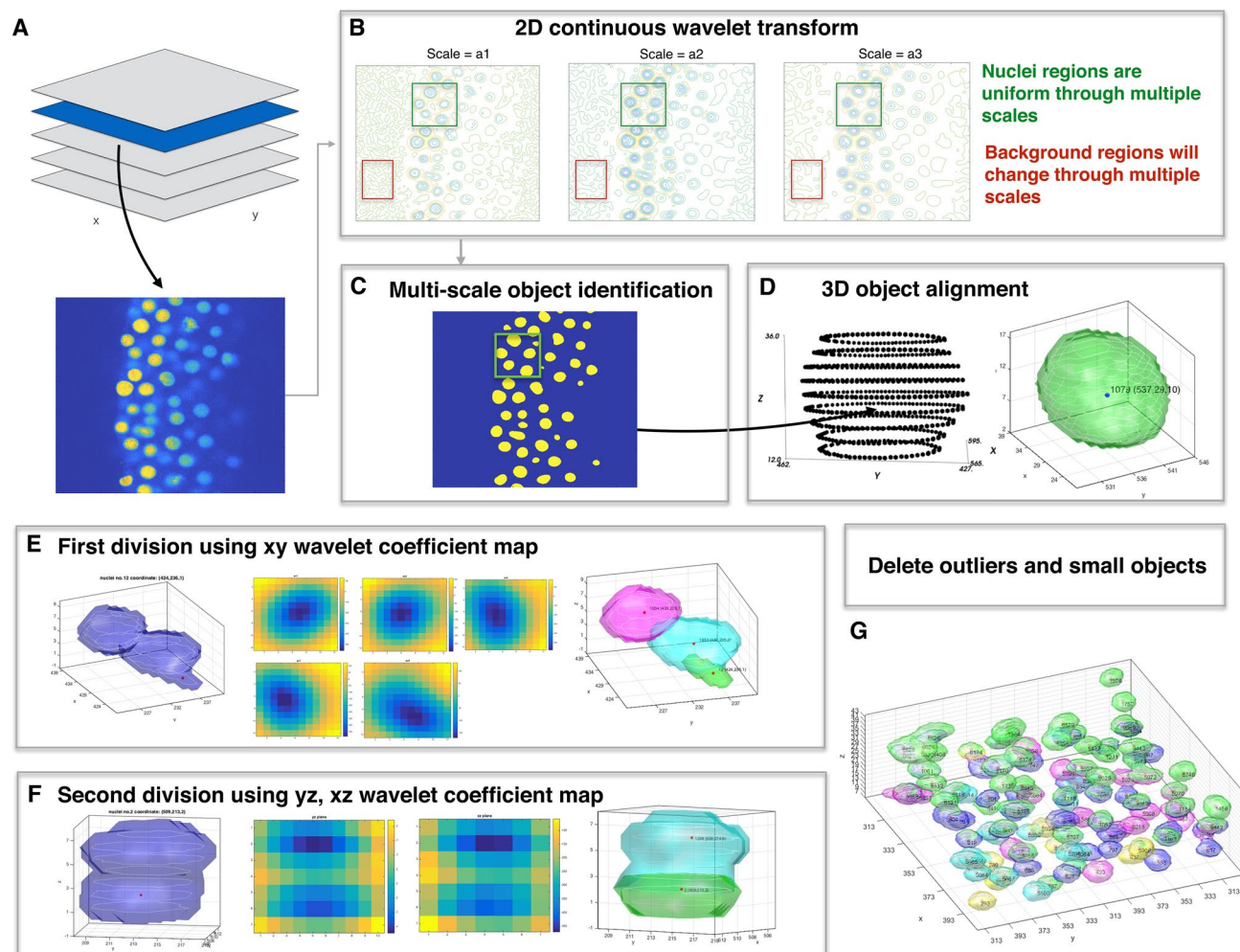


Figure 2. Overview of wavelet-based nuclei segmentation method. **(A)** Import 3D raw image files. **(B)** After importing 3D images, 2D continuous wavelet transform with appropriate wavelet scale factors are applying to each 2D z-stack. **(C)** 2D segmentation masks are obtained by applying multi-scale object identification to wavelet coefficient maps from previous step. **(D)** In 3D object alignment step, connected identified 2D masks from neighboring slices forms a 3D foreground object. **(E)** First division step if more than one center or center shift on neighboring z slice wavelet coefficient maps. **(F)** Second division step in z direction if multiple centers found on yz or xz wavelet coefficient maps. **(G)** Delete 3D nuclei voxels smaller than nuclei size $\times 0.2$.

Results

Underlying methods of the wavelet-based segmentation method. The six main steps of the wavelet-based nuclei segmentation method (Fig. 2) are: (1) application of a 2D continuous wavelet transform (2D CWT), (2) multi-scale object identification, (3) 3D object alignment, (4) first round of nuclei division based on xy plane wavelet coefficients, (5) second division of nuclei based on z direction, and (6) deletion of outliers and small objects. In step one, two-dimensional continuous wavelet transforms with a range of wavelet scale factors are applied to each Z-stack slice of 3D sections common as output from confocal imaging (see Online Methods and Fig. 2A,B). The wavelet scale factor range is based on the object size we want to identify. In line with this, we designed an efficient multi-scale search algorithm that compares different scales of the CWT coefficients and groups shape-similar regions such as the nuclei masks based on contour line similarities in the CWT coefficients diagram. Through the comparison between wavelet coefficient matrices over different scales, regions that meet the similarity index criteria are masked and labeled as nuclei in each stack (see Online Methods, Fig. 2C and Fig. S4).

Following 2D identification, nuclei are reconstructed by 3D object alignment that aligns 2D nuclei cross-sections to recreate 3D nuclei shapes by comparing consecutive nuclei segmentation results from the preceding step. The method computes the intersection of the mask coverage area of the neighboring z-plane. If the intersection area among adjacent z-planes is greater than 80% and if the max intensity is not lower than 20% nuclei center intensity, it is inferred that the adjacent mask coverage area should belong to the same 3D nuclei object (see Online Methods and Fig. 2D).

In threshold and other segmentation approaches, a major challenge is errors in nuclei identification due to imaging conditions that lead to irregular or poor contrast boundaries or intensity attenuation that systematically

distort measured fluorescence intensity. These factors compound in 3D and additional steps are needed for identification. We designed a two-step 3D nuclei division method by comparing the calculated wavelet coefficient planes in both vertical and horizontal directions to overcome this overlapping issue. The local minimum on each Z-slice wavelet coefficient plane is determined. If there is a shift in the wavelet coefficient *local minimum* on two consecutive z-planes and the shift distance is larger than 1/3 nuclei radius, this is indicative of two different nuclei objects, and it will be marked as likely belonging to different nuclei (see Online Methods, Fig. 2E, and Fig. S5). For identification of potential separate overlapping nuclei in the YZ-plane and XZ-plane, a similar center point detection is used (Fig. 2F, and Fig. S6). If multiple wavelet coefficient local minimum points are found on both the YZ-plane and the XZ-plane, it is likely they come from two nearby nuclei (see Online Methods). Lastly, noise or false positive nuclei after the additional segmentation are removed based on object size filtering (Fig. 2G).

Evaluation of segmentation methods with synthetic and real images. To benchmark the wavelet and other methods that are coded into the WaveletSEG program, we developed tunable test data as well as a ground-truth data to manually segment 3D images for validation. In addition to the wavelet method, the following methods are included in the program: point wise method³⁷, Otsu method⁶, and Derivatives Sum (DS) method³⁸. The validation datasets we used to evaluate segmentation performance included synthetic image datasets that were created by using a synthetic data generator (see Online Methods and Fig. S7), publicly open image datasets³⁹, and our ground truth image dataset determined by multiple rounds of manual segmentation. We also generated synthetic nuclei overlapping datasets to test the ability to divide overlapping nuclei.

Three evaluation criteria were applied to compare the accuracy of four segmentation methods. If a representative point was nearest-neighbor of a point in ground truth dataset and visa-versa, the object was regarded as a true positive (TP). If only the former condition was met, the segmented result is considered a false negative (FN). If only the latter condition was met, the object was regarded as a false positive (FP). Then the true positive rate (TP rate) is equal to the true positive number divided by the ground truth number, and Precision is defined as the TP number divided by the sum of TP number with FP number. TP rate defined in Eq. (1) represents the ratio of successful nuclei segmentation within all ground truth nuclei data, and Precision is defined in Eq. (2). The false negative rate is the false negative number divided by the ground truth number. We also use segmentation accuracy to estimate pixel-wise segmentation performance, which is defined as the overlapping region between ground truth segmentation region and segmentation region in pixel level.

$$TP\ rate = TP/P = TP/(TP + FN) \quad (1)$$

$$Precision = TP/(TP + FP) \quad (2)$$

We use the Jaccard similarity index J (Eq. (3)) to define segmentation accuracy³⁹ which means the amount of overlap between the segmentation results S and ground truth annotation results R.

$$J(R, S) = \frac{|R \cap S|}{|R \cup S|} \quad (3)$$

In test data we added multiple different types of image noise that would normally be encountered in an imaging environment. For Gaussian white noise (Fig. 3A), the wavelet-based segmentation method has the best TP rate and Precision followed by DS method. Otsu method obtained the lowest Precision and the point-wise method performed the poorest in TP rate. In salt and pepper noise cases (Fig. 3B), both TP rate and Precision were equal to one when segmented by the wavelet-based method, and all other method results are lower than 0.9 in both TP rate and Precision. In images with Intensity attenuation (Fig. 3C), the point-wise method has very high false positive rates and Otsu method performed the poorest in TP rate. In this test, the wavelet-based segmentation method had the highest TP rate and Precision. In the next step we examined the ability to divide nuclei in overlapping images for four segmentation methods (Fig. 3D). Compared to three other segmentation methods, the wavelet-based method has the best ability to divide overlapping nuclei in both TP rate and Precision.

Next, we compared the pixel-level segmentation accuracy on publicly open image datasets using the wavelet-based method and Otsu's method. Figure 3E shows that segmentation accuracy is above 0.85 for the wavelet-based method, and smaller than 0.4 for Otsu's method. To further test the segmentation ability in real image datasets, we chose five zebrafish embryo images and selected sub-regions that we manually curated for ground truth data and labeled the nuclei positions using our ground truth annotation tool (Fig. S8). Both the TP rate and Precision of the wavelet method had the highest scores (Fig. 3F).

Outline of main structure and interface of the WaveletSEG matlab program. WaveletSEG is an open-source Matlab-based imaging research program we developed that runs on Windows, Mac and Linux systems with complete Graphical User Interface (GUI) and it is outlined in Fig. 4. Users can directly import either raw image files or data files with calculated results which helps them to check results or rerun processing with a different setting. 3D segmentation and quantification or other results such as embryo topology features can be easily displayed or saved in the WaveletSEG data visualization system. Users can also create scatter plots by selecting menu options to explore the relationship between them. For the examples used herein, we show the fluorescent distributions of the BMP signaling transducer phosphorylated SMAD (PSMAD) along the embryo dorsal/ventral axis, or the nuclei size spatial distribution from the outer embryo layer inwards.

The WaveletSEG program consists of five main functions that run independently: 1. Nuclei identification, 2. Embryo orientation, 3. Shape classification, 4. Profile extraction, and 5. Time lapse steps (Fig. S9). In the nuclei identification step, users can do intensity calibration by using nearby nuclei intensity or do 3D rotation to adjust

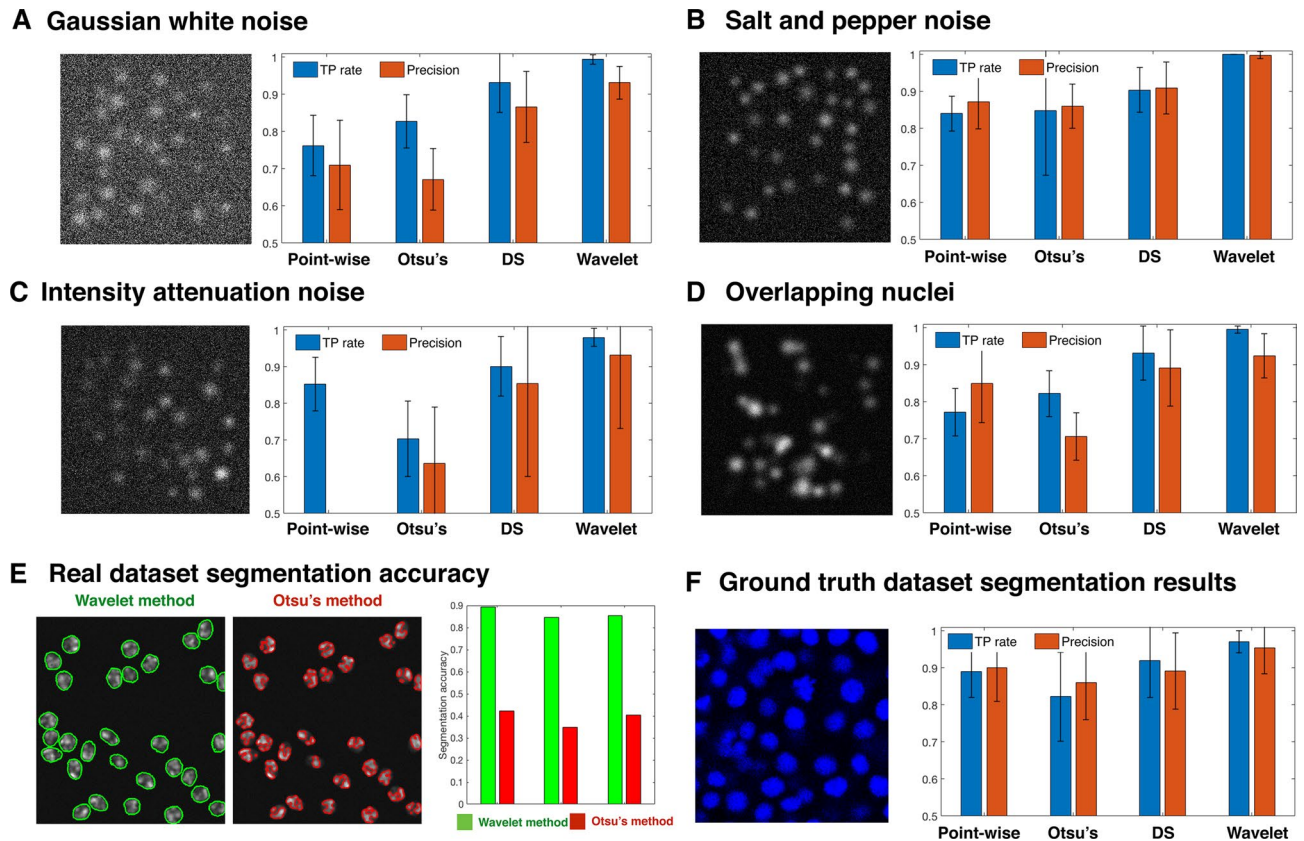


Figure 3. Performance criteria and segmentation accuracy of four segmentation methods on synthetic images and ground truth image datasets. (A–D) True positive (TP) rate and Precision on four kinds of noisy synthetic images using three segmentation methods and the wavelet method. Error bars indicate standard deviation. (A) Samples of synthetic images (N = 20) Gaussian white noise and TP rate and Precision when applying point-wise method, Otsu method, DS method and wavelet-based segmentation method. (B) Samples of salt and pepper noise synthetic images (N = 20) and TP rate and Precision for four segmentation methods. (C) Intensity attenuation noisy synthetic image samples (N = 20) and TP rate and Precision for four segmentation methods. (D) Samples of overlapping synthetic image samples (N = 20) and TP rate and Precision for four segmentation methods. (E) Segmentation accuracy when applying Otsu method and wavelet-based segmentation method on three Fluo-N2DH-GOWT1 sample images. (F) True positive rate and Precision on five ground truth images when using point-wise method, Otsu method, DS method and wavelet-based segmentation method.

the embryo DV and AP axis direction after wavelet-based segmentation. After the nuclei segmentation step, the embryo coordinate system is automatically created and each nucleus is given a coordinate list value (see Methods), and the 3D topology feature for each nucleus is also calculated. In step 3, we can apply nuclei shape classifications to analyze the 3D nuclei shape type and features which resulted from step 1 and 2. In the profile extraction step, we project the nuclei quantification and features into one averaged distributed sphere plane (see Online Methods). If an imported image is a 4D (3D + time) image, the user can run time-lapse analysis in step 5. In step 5, nuclei tracking is performed by segmenting all the nuclei in a time-lapse image sequence using wavelet-based segmentation. This is achieved one step at a time, and the spatiotemporal overlap between corresponding segmented regions in consecutive time steps is compared, providing information about coordinate values, topology features and shape classification.

WaveletSEG also includes a synthetic data generator, a 3D ground truth labeling system, 2D and 3D segmentation viewer sub-GUI, and segmentation method comparison sub-GUI, (see Online Methods and Fig. S7–S11). Using the CompareSEG sub-GUI, users can also easily validate segmentation results by utilizing different methods and display TP, FN, or FP nuclei directly. In WaveletSEG, users can also export 2D or 3D segmentation masks as the machine learning training datasets for further object detection or semantic segmentation studies.

Zebrafish embryo quantification. Development of an animal embryo involves molecular signaling to coordinate cellular rearrangement and proliferation. To quantify these processes in zebrafish, we quantified zebrafish BMP signaling along the AP and DV axes and calculate embryo properties including thickness, and cell division state. These properties provide a complete embryo picture (Fig. 5A) that includes the 3D embryo geometry, quantitative signaling information and other properties^{40,41}. After segmentation, 3D rotation and deletion of the EVL cell later, (see Online Methods and Fig. S12–S13), we generated 2D hemispherical surfaces and projected every nuclei onto the 2D surface to obtain intensity distributions of nuclear properties (Fig. 5B,C).

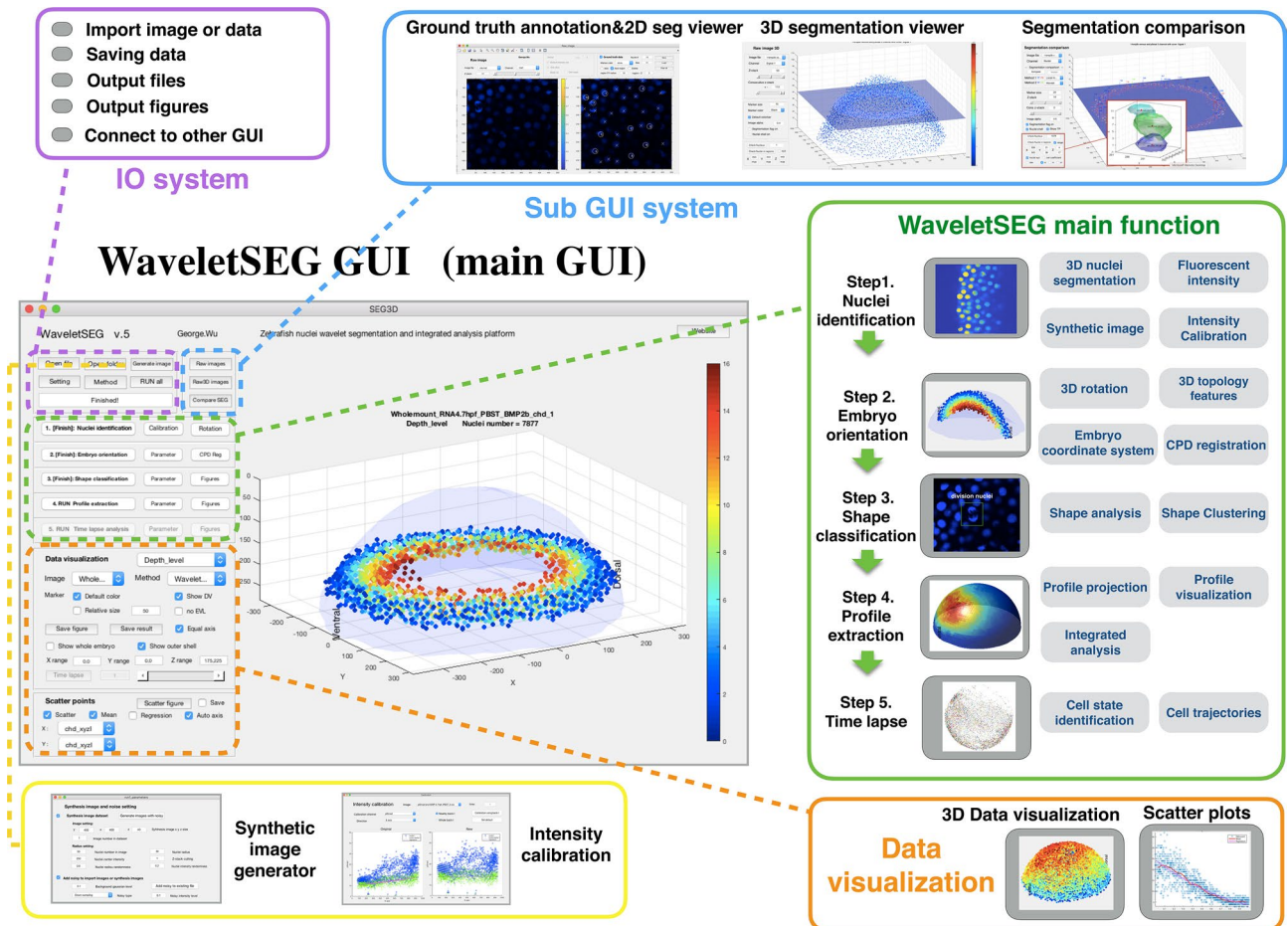


Figure 4. WaveletSEG software main function structure and GUI. WaveletSEG is the image processing analysis platform which integrates all main and extension functions including IO system, extension GUI and data visualization system in one GUI. The WaveletSEG main function block consists of five main steps that can run independently including 1. Nuclei identification, 2. Embryo orientation, 3. Shape classification, 4. Profile extraction, and 5. Time lapse steps. In the nuclei identification step, users can do intensity calibration by using nuclei intensity or after wavelet-based segmentation. In IO system block, users can directly import microscope image files or intermediate data files into WaveletSEG, and save or output data results or figures directly from the GUI. We also developed a set of segmentation validation tools in WaveletSEG including synthetic data generator, 3D ground truth labeling system, 2D and 3D segmentation viewer sub-GUI, and segmentation method comparison extension GUI. In WaveletSEG data visualization block, 3D segmentation and quantification or other results such as embryo topology features can be easily displayed or saved in the WaveletSEG data visualization system. User can also create scatter plots by selecting menu options to explore the relationship between them.

Other outputs include whole embryo, sub region quantification results (Fig. 5D), DV axis pSMAD intensity distributions (Fig. 5G), or population average properties after the coherent point drift (CPD) registration (Fig. 5E) (see Online Methods).

In early embryo development of the zebrafish, cell rearrangements guide the establishment of embryonic axes and layers^{42,43} and embryonic cells spread over the yolk mass while the blastoderm thins. The radial intercalation movement between deep cell layer (DCL) and enveloping layer (EVL) is regarded as the main driving force for epiboly, and the cell density increase near the EVL plays an important role in this process⁴⁴. Figure 5H, J shows the Depth level spatial distribution for nuclei in the DV direction cross section of 4.7 hpf (N = 9, left), 5.7 hpf (N = 10, center) and 6.5 hpf Zebrafish embryo (N = 8, right) (Fig. 5F). They have significant differences not only in nuclei number, but also in the nuclei spatial distribution, density and thickness. In Fig. 5H (right), we observed the cell accumulation (Hypoblast) and involution movement near the blastoderm margin.

Emboly is another principle coordinated cell movement to form hypoblast which contains the involution and intercalation movements in the early gastrulation stage. Involuting cells move and accumulate in the blastoderm margin to narrow and elongate the embryonic axis, eventually forming endoderm and mesoderm that characterizes the gastrulation stage in Zebrafish^{45,46}. Convergence and extension of epiblast deep cells result from intercellular space decrease and cell density contributions to the internalization of hypoblast cells from the margin. For further investigation of spatial and temporal analysis of embryonic structure distribution, we

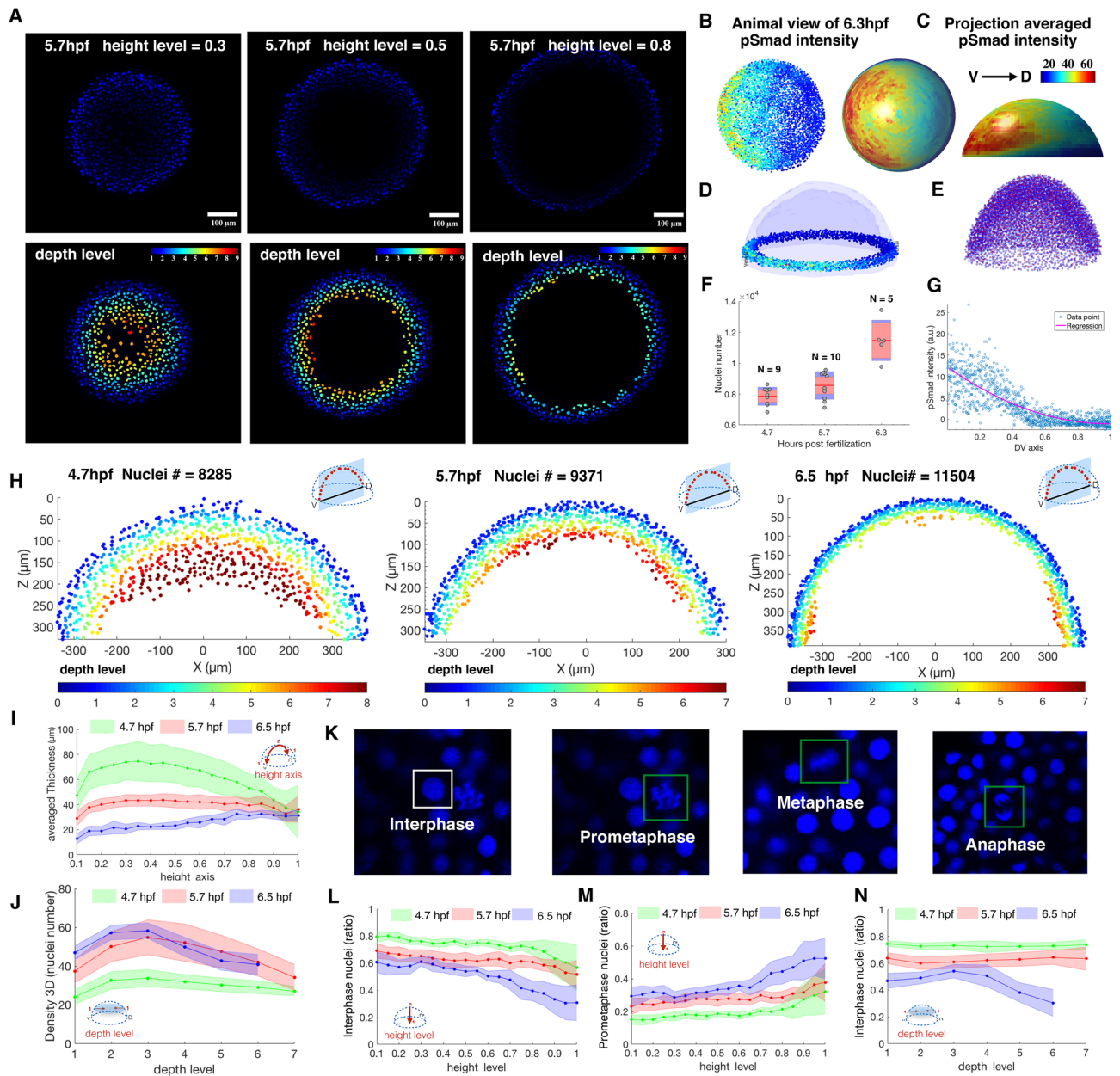


Figure 5. Example quantification of nuclear pSMAD gradient in Zebrafish embryo and spatial analysis of embryo coordinate system, 3D topology features and cell cycle phase patterns between 4.7, 5.7, 6.5 hpf Zebrafish embryo. (A) Z-slice DAPI nuclear stain (upper) and depth level for each nucleus on height level = 0.3, 0.5, 0.8 in 5.7 hpf Zebrafish embryo (lower). (B, C) Animal view of nuclear pSMAD intensity of all nuclei from the 6.3 hpf embryo (left), and animal and lateral view of averaged pSMAD intensity by projecting to uniform distributed semisphere surface using WaveletSEG (center and right). (D) Data visualization of nuclear pSMAD intensity for 6.5 hpf Zebrafish embryo in marginal region using WaveletSEG. Light blue and dark blue transparent shells indicate inner and outer boundary surfaces of all nuclei in embryo. (E) Embryo alignment using coherent point drift (CPD) method in WaveletSEG. (F) Nuclei number counts of embryo vs different developmental time (4.7 hpf: N = 9, 5.7 hpf: N = 10, 6.5 hpf: N = 5) which also applied in (i–j, l–n). Gray dots are individual embryo nuclei number. Red line red box region, and blue region show the mean nuclei number, 95% confidence region and region between one standard deviation. (G) Distribution of nuclear pSMAD intensity vs DV axis in embryo margin region in (C). Blue dot and red line indicate individual nuclei intensity and averaged intensity along DV axis. (H) Depth level spatial distribution for nuclei in DV direction cross section plane $\pm 30 \mu\text{m}$ of 4.7 hpf (left), 5.7 hpf (center) and 6.5 hpf Zebrafish embryo (right). (I) Averaged nuclear layer thickness through embryo height axis in 4.7, 5.7 and 6.5 hpf embryo developmental time. (J) Averaged nuclear density through embryo height axis in 4.7, 5.7 and 6.5 hpf embryo developmental time. Green, red and blue dark line indicate mean value, and light green, light red and light blue regions indicate one standard deviation region. (K) Examples of cell cycle phase in mitosis based on nuclei shape classification results. (L) Interphase nuclei count ratios through embryo height level in 4.7, 5.7 and 6.5 embryo developmental time. (M) Prometaphase nuclei count ratios through embryo height level in 4.7, 5.7 and 6.5 embryo developmental time. (N) Interphase nuclei count ratios through embryo depth level in 4.7, 5.7 and 6.5 embryo developmental time. (L–N) Green, red and blue dark lines indicate mean value, and light green, light red and light blue regions indicate one standard deviation region.

introduced eight 3D topology features including Size, Thickness, Density2D, Density3D, Neighbor2D, Neighbor3D, H sorting, EVL (see Online Methods and Fig. S14) which can be directly calculated in WaveletSEG. For example, Fig. 5I can be used to describe the averaged thickness in epiboly process in 4.7, 5.7 and 6.5 embryo developmental time. We also found that the thickness increases near the margin position in 6.5hpf, which is the signal for emboly process. The gradual separation of the EVL lineage correlates with the flattening of the blastoderm on the yolk cell and is accompanied by both an increase in tension and cell shape changes within the EVL⁴⁷. It has been suggested that the separation of the EVL as a lineage might be the consequence of increased tension that causes EVL cell divisions to occur preferentially within the plane of the EVL. At the end of gastrula stage, the cell density is decreased at the ventral side of embryo.

Lastly, we used WaveletSEG to monitor the cell proliferation and division patterns during morphogenesis to regulate the embryo shape and growth. Nuclei segmentation and shape classification results have been used to estimate cell proliferation and cell cycle phases⁴⁸ because nuclei are spherical during early prophase, irregular shape with nuclear envelope break down, and become ellipsoidal shape in metaphase⁴⁹(Fig. 5K). Here we created the cell-cycle phases pattern in zebrafish embryo based on nuclei shape features (see Online Methods, Fig. 5K,L,M,N and Fig. S15).

Figure 5L,M show the interphase and prometaphase nuclei count ratios at 4.7, 5.7, and 6.5hpf. We found that in all time stages, the ratio of interphase nuclei decreases from the top of the embryo to the bottom, and the prometaphase nuclei ratio increases from between 0.2 and 0.3 at 4.7 hpf to between 0.3 and 0.5 at 6.5 hpf. It showed that cell proliferation rates are low in the animal region of the embryo and higher in the zebrafish margin. In Fig. 5N, interphase nuclei ratio in 6.5 hpf is heavily decreased from 0.5 to 0.25 after the depth level is more than 3, which means the proliferation rate is higher in deep cells in the shield stage. These results are consistent with previous data⁵⁰ that shows abundant mitotic cells at 50% epiboly and especially during the shield stage.

Discussion

As illustrated above, we have developed a novel wavelet-based image segmentation algorithm for robust 3D nuclei segmentation that demonstrates good segmentation performance in many image datasets. Recent technological advances in biological imaging are focused on developing fully automated and large-scale imaging algorithms, and wavelet-based segmentation will be a good fit because of its reliability in the face of noise and intensity attenuation and the only necessary parameter is the wavelet scale factor which is chosen based on property object size and no preprocessing or fine-tuning steps are needed. We also presented a new way to divide overlapping nuclei according to the self-similarity between multiple wavelet coefficient maps.

We developed WaveletSEG to determine embryonic patterning information for molecular signaling, gene regulation, and 3D positional data for nuclei at various stages of the cell cycle. This integrated image analysis platform helps define the embryo coordinate systems to provide embryo axes for geometry features and quantification. After 3D nuclei segmentation for whole embryos, the post processing in WaveletSEG provides convenient tools to rotate the embryo, remove EVL layer nuclei and select subregions for additional quantification. Overall, this quantification and post-processing program is designed to support converting imaging data into information that can be used to infer biological mechanism, gain quantitative insight into development and develop quantitative data sets of pattern formation and signal regulation by morphogens.

Materials and methods

2D wavelet transform and multi-scale object identification. 1D discrete wavelet transforms (1D DWTs) and 1D continuous wavelet transforms (1D CWTs) are the most commonly applied wavelet tools in 1D signal processing and time series analysis, and are used in signal decomposition in both time domain and frequency domain^{51,52}. We extend the 1D CWT to the two-dimensional continuous wavelet transform (2D CWT) and apply it to the spatial domain of the image⁵³. Supposing that $f(x, y)$ is continuous and differentiable 2D image data, we choose the 2D Mexican hat function as the wavelet mother function and perform 2D CWT with the translation factors a and b and the wavelet scale factor s . We compare the nuclei signal shape with three typical mother wavelet functions including Morlet wavelet, Mexican Hat wavelet, and Spline wavelet and found Mexican Hat is the most similar mother wavelet function. $\psi(x, y)$ is the 2D Mexican hat function and $K_{2D}(s, a, b)$ is the wavelet coefficient matrix (Eq. (4), (5)):

$$K_{2D}(s, a, b) = \frac{1}{\sqrt{s}} \iint_{-\infty}^{\infty} f(x, y) \psi\left(\frac{x-a}{s}, \frac{y-b}{s}\right) dx dy \quad (4)$$

$$\psi(x, y) = \frac{1}{\pi\sigma^2} \left(1 - \frac{1}{2} \left(\frac{x^2 + y^2}{\sigma^2}\right)\right) e^{-\frac{x^2 + y^2}{\sigma^2}} \quad (5)$$

Figure 2 shows the nuclei fluorescence microscopy raw image (Fig. 2A) and we applied 2D CWT with a range of wavelet scale factors based on nuclei size to each z-plane on a 3D raw image (Fig. 2B). For example, we chose wavelet scale factor 9, 10, and 11 on zebrafish embryo imaging because its nuclei radius is about 9 μ m. We designed a search algorithm to obtain a 2D segmentation mask by comparing the multi-scale wavelet coefficient matrix(s), which resulted from 2D CWT using different scale factors (Fig. 2C). Local minima were first determined on every wavelet coefficient matrix, and if local minima appeared in the same positions on all wavelet coefficient matrixes, they were identified as potential nuclei peak locations. In the next step, we calculated all zero-value cross sections in potential nuclei peak locations on multi-scale wavelet coefficient matrixes. If the difference in cross section areas was less than 10%, they were averaged and identified as nuclei 2D masks.

In 3D microscope imaging, the Z axial resolution may be different from the lateral (XY) resolution, so it's difficult to use the same wavelet scaling factor in both the Z axial direction and lateral direction. This is the reason to use the 2D wavelet transform in each image rather than applying 3D wavelet transform.

3D object alignment. The following step is a 3D object alignment that aligns 2D nuclei cross-sections to recreate 3D nuclei shapes by comparing consecutive nuclei segmentation results from the preceding step (Fig. 3D). The method computes the intersection of the mask coverage area of the neighboring z-plane. If the intersection area among adjacent z-planes is greater than 80% and if the max intensity is not lower than 20% nuclei center intensity, it is inferred that the adjacent mask coverage area should belong to the same 3D nuclei object. In some cases, there will be overlapping problems on the image, or sometimes, when the z-slice cutting distance is large, the two nuclei will be treated as one, and subsequent division steps will then be followed.

First and second division on XY wavelet coefficient plane, XZ and YZ plane and delete small object. For this purpose, we designed division steps in two stages to divide overlapping nuclei or overlapping 2D masks using wavelet coefficient maps. For each 3D nuclei object that we obtained from the previous step, we compared Z-slice wavelet coefficient planes and found the center of each plane that is at a local minimum (Fig. 2E). If there is more than one wavelet coefficient local minimum or the displacement of center is bigger than the nuclei radius on neighboring Z slices wavelet coefficients, we assigned a new nuclei object by cutting the median line between two centers, or a new nuclei in the case of center displacement. Using the center of Z-plane wavelet coefficients, most y or z direction overlapping nuclei can be divided.

In the second-stage division step, we divided z-direction overlapping nuclei by checking both the xz plane and the yz plane wavelet coefficient (Fig. 2F). If there were more than local minimum points present in both xz and yz plane, a wavelet coefficient division plane was set in the center plane between the two centers, and a new 3D nuclei object is created. In the last step of the wavelet-based segmentation protocol, we removed the 3D nuclei voxels where the size was smaller than nuclei size $\times 0.2$ (Fig. 2G).

Synthetic image testing. To provide a dataset of synthetic images in validating segmentation methods, we developed a convenient interface to generate a synthetic image. Users can define synthetic image's 3D size and the number of image and synthetic nuclei signals based on the nuclei number, nuclei radius and intensity, and the randomness of the nuclei radius and intensity between 0 to 1. If the radius randomness is 0.2, this means the generated nuclei radius = defined nuclei radius $r \pm \text{uniform}(0, 1) \text{ random variable} \times 0.2 \times r$. Nuclei are generated in the following order: Synthetic nuclei with random 3D positions are generated inside the image. If the distance of this nuclei relative to the previously generated nuclei, or the distance to the image boundary is smaller than nuclei $r \times 1.5$, we regenerate the synthetic nuclei. A 3D Gaussian function with nuclei intensity and standard deviation equal to nuclei radius $r \times 0.4$ is added to the synthetic image. We also created a synthetic overlapping nuclei dataset to evaluate the segmentation method's ability to isolate overlapping nuclei. Subsequently, half of the nuclei were randomly created, and additional nuclei were iteratively generated and added to the dataset once its distance from the previous nuclei set was smaller than a nuclei radius $r \times 1.5$. After initial nuclei are distributed throughout the image, we add both white noise and user-defined noise types with user-defined noise level noise density (ND) onto the synthetic image that we generated in the previous step. White noise is white Gaussian noise with user-defined variance. User-defined noise types include Gaussian white noise and salt-and-pepper noise with ND and intensity attenuations with image gradient ND, in which pixel intensity will increase continuously from right to left with intensity ND times on the left side of the image.

Point-wise method, Otsu method, DS method and parameter screening using GA. To evaluate the segmentation performance of the wavelet-based segmentation method, we compared it with three commonly used segmentation methods, including the point-wise method, the Otsu method and the DS algorithm. The point-wise method is the simplest method to estimate the intensity of nuclei center points i . Raw images were firstly smoothed using $9 \times 9 \times 3$ kernel, followed by h-maxima and h-minima transform to suppress all background signals. After combining nearby local maxima closer than 6 pixels, the remaining local maxima were assumed to be nuclei center points and intensity was calculated after applying $6 \times 6 \times 3$ spherical kernel.

The Otsu method is the most popular threshold-based segmentation method to separate foreground and background pixels by finding the optimum threshold to satisfy minimum intra-class variance. We applied a symmetric Gaussian low-pass filter and sharpened the filter with specific radius and amount on the raw image. We then used Wiener filter or pixel-wise adaptive low-pass Wiener filter to deblur the image and remove noise. To this end, we used the multilevel thresholding Otsu method to decide threshold values.

Derivatives Sum algorithm (DS algorithm) started with applying denoising filters such as a Gaussian filter or a non-linear isotropic diffusion filter on the image, and the 2D spatial derivatives were computed to get the image Gauss gradient, Laplacian determinant, and Hessian determinant. A mask function F was calculated by combining the first and second spatial derivatives with weight parameters, and the Otsu method was applied here to obtain binary image slices. 3D nuclei segmentation was obtained by connecting the neighboring 2D masks, and we used a genetic algorithm for the parameter screening in this method⁵⁴.

Wavelet segmentation time complexity and computational time. The time complexity of the wavelet-based 3D nuclei segmentation method can be estimated by calculating the Big O complexity in each step. In the first step, a two-dimensional continuous wavelet transform is applied to each Z-stack slice of 3D image with scale factor s in range N_s . The Big O complexities of 2D continuous wavelet transform and H-minima transform are $O(N^2)$ so the overall complexity of this step will be $O(N^2 N_z N_s) \sim O(N^3)$ because N_s is much

smaller than N . In multi-scale object identification step, intersection areas among adjacent z planes are calculated $Nz-1$ times so the complexity will be $O(N^2(Nz - 1)) \sim O(N^3)$. First division and second division steps are applied to M nuclei segmented in previous steps. The complexity of first division step can be estimated by comparison between adjacent z planes for M segmented nuclei with size $n_2 \times n_z$. The Big O complexities of first division step is $O(n^2 n_z M) \sim O(M)$ because M is much larger than n and n_z . The second division and delete outliers steps should have the same complexities as the first division step. The overall expected time complexity of the wavelet segmentation method will be $O(N^3 + M)$.

To evaluate the computational time of wavelet-based segmentation method, we compared it with three commonly used segmentation methods, including the point-wise method, the Otsu method and the DS algorithm. Four different sizes of 3D images are tested and repeated three times, and the results are listed in Table S4. The result shows that the segmentation running time of wavelet segmentation method is similar to DS method and are two or three times larger than the point-wise method and Otsu's method in most sizes 3D images. Although the threshold based segmentation methods have less running time, additional time is required for pre-processing steps and parameter tuning steps in each image and each Z -slice. There is no pre-processing step required for our wavelet based segmentation method and the only required parameter is nuclei size so no parameter tuning time is needed. The hardware platform for testing was a 3.70 GHz Intel Core Xeon CPU and 32 GB RAM.

Ground truth labeling GUI. Creating embryonic 3D nuclei segmentation ground truth data is challenging and time-consuming. Here we introduced Raw_image sub-GUI, a 3D nuclei annotation tool that provides a convenient and efficient way to label 3D nuclei, which can also be used to evaluate nuclei segmentation results.

In Raw_image the user can directly label embryo nuclei by simple mouse clicking. There are two main types of labeling: nuclei center labeling that shows the nuclei index, and non-center nuclei labeling that is marked as a cross symbol. Non-center nuclei labeling can help to prevent double counting. After clicking on the screen, one sub-GUI will jump up with the specific region and a range of Z -slack planes (Fig. S8B). The user can directly assign the labeling by clicking. Raw_image can also be used to check segmentation results by either comparing segmentation masks with raw images or checking single nuclei masks.

CPD registration method. The coherent point drift (CPD) algorithm is a probabilistic-based point set registration method widely used in the field of pattern recognition⁵⁵. The main goal of point registration is to merge two or more points sets into one representative set. CPD treats the registration process as a probability density estimation problem and fits the Gaussian mixture model (GMM) of one-point set to a reference point set using maximum likelihood method, and the expectation maximization algorithm is used to make the most effective use of the optimization function. We use CPD to merge embryos in the same stage into a representative embryo one by one. The final GMM centroids are the representative embryo nuclei sets.

3D topology features. We designed eight 3D topological features to describe the topology and geometric structural changes for embryological development studies in Zebrafish. Nuclei size is the total pixel numbers for every 3D segmented nuclei. We used four indexes to describe the spatial nuclei density distribution for the embryo, including density 3D, density 2D, neighbor 3D and neighbor 2D distance. Density 3D is the total nuclei number with distance smaller than 30 pixels around one nucleus. The problem of using density 3D is the underestimation near the outermost or innermost layer. In this instance, we define density 2D as the total nuclei number closer within 30 pixels distance and also in the same layer (depth level). Neighbor distance is the average distance of the nuclei to the closest two nuclei which is suitable to find the outlier point in point set. The difference between neighbor 2D and neighbor only considers nuclei in the same layer.

Another important topological feature of the embryo is the thickness. We define thickness here as the vertical distance from the innermost layer of embryo. The innermost shell was divided into 360 sub-regions, and the vertical distance was calculated from nucleus to its corresponding plane. The outer epithelial monolayer of a zebrafish embryo is called the enveloping layer (EVL). Here we defined the inner EVL and outlier EVL as the layer with depth level equal to one or the maximum values. $H_sorting$ is the index to describe the dynamic process of nuclei movement by dividing every 500 nuclei from the top of the embryo to the embryo bottom.

Cell cycle phase pattern based on nuclei shape classification. To study the spatial and temporal shape distribution of embryo nuclei, we defined some rules relative to dividing them into spherical nuclei, irregular shape nuclei, elliptical nuclei and dividing nuclei. We calculated the sphericity value using $4 \times \pi \times \text{area}$ divided by squared perimeter in three planes (xy , yz and xz planes), and averaged them as the averaged sphericity value. The average aspect ratio is the mean of aspect ratio xy , yz , and xz . Aspect ratio xy is the proportional relationship between x axis range divided by y axis range.

Here we defined spherical nuclei if the mean sphericity value is larger than 0.9 which means the shape of nuclei is similar to a circle with the same area. If the mean sphericity value is smaller than 0.9 and the mean aspect ratio is larger than 0.7, we identified it as an irregular shape of nuclei. A nucleus is determined as either elliptical or dividing if the mean sphericity value is smaller than 0.9 and the mean aspect ratio is also smaller than 0.7. In addition, if the radius of this nuclei is smaller than the embryo mean radius and they are paired, they are considered dividing nuclei.

Embryo profile projection and data visualization. To indicate a better concentration or feature distribution on a three-dimensional multi-layer embryo, and also allow for the integration of similar stage embryos, we projected features of each nucleus into one reference two-dimensional surface. We generated a hemispherical surface with averaged distributed reference point which we created in using regular placement method. We

projected the nuclei into reference points with the closest spherical coordinate system angle (θ, ϕ). Thereafter, we summarized the concentration of nuclei in this reference point, as well as average topological features such as nuclei size, nuclei density or thickness.

We designed the interactive graphical user interface to visualize segmentation and analyze the results of the embryo directly. Users can also choose a particular region such as the embryo margin region by choosing z region range, or DV axis concentration distribution. In addition, users can remove EVL nuclei directly by clicking the no EVL icon on the data visualization board. In a scatter plot board, users can choose one coordinate system index in the first pop-up menu and one feature in second pop-up menu to analyze the spatial distribution on the embryo. Alternatively, users can also plot the relationship between two features using the scatterplot board.

Zebrafish lines. All the procedures done on zebrafish adults and embryos were approved by the Purdue University Institutional Animal Care and Use Committee (IACUC: 1501001180A004). All experiments were performed in accordance with relevant guidelines and regulations. Zebrafish were raised and maintained by the Pentair Aquatic Eco-Systems. Wide-type (TL) embryos were collected in E3 water when a one male-and-female pair was crossed for 30 min. Subsequently, embryos were left to develop to the desired stage at 28 °C.

Immunostaining. Embryos were penetrated in 0.1% Triton X-100 in PBS for 1 h at RT. Embryos were blocked in blocking buffer (4% BSA (Millipore Sigma, #126,615), 1% DMSO, 0.1% Triton X-100 in PBS) overnight at 4 °C, and then stained with anti-phosphoSmad1/5/8 antibody (Cell Signaling Technology, #9511) at 1:100 diluted in blocking buffer overnight at 4 °C. Then embryos were detected by goat anti-rabbit cross-adsorbed Alexa Fluor 647-conjugated antibody at 1:500 dilution with DAPI overnight at 4 °C (Thermo Fisher Scientific, #A21244).

Data availability

Raw data and segmentation result files that support the findings of the study are available at <https://github.com/George-wu509/WaveletSEG>.

Code availability

WaveletSEG is freely available for academic use as Supplementary Software or (including updated versions) at <https://github.com/George-wu509/WaveletSEG>. Source code is available upon signing of a material transfer agreement.

Received: 1 September 2020; Accepted: 9 April 2021

Published online: 10 May 2021

References

- McMahon, A., Supatto, W., Fraser, S. E. & Stathopoulos, A. Dynamic analyses of *Drosophila* gastrulation provide insights into collective cell migration. *Science* (80) **322**, 1546–1550 (2008).
- Keller, P. J., Schmidt, A. D., Wittbrodt, J. & Stelzer, E. H. K. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* (80) **322**, 1065–1069 (2008).
- Stegmaier, J. *et al.* Fast segmentation of stained nuclei in terabyte-scale, time resolved 3D microscopy image stacks. *PLoS ONE* **9**, 1–11 (2014).
- Meilhac, S. M. *et al.* Active cell movements coupled to positional induction are involved in lineage segregation in the mouse blastocyst. *Dev. Biol.* **331**, 210–221 (2009).
- Win, K. Y., Choomchuay, S., Hamamoto, K. & Raveesunthornkiat, M. Comparative study on automated cell nuclei segmentation methods for cytology pleural effusion images. *J. Healthc. Eng.* **2018**, 315–396 (2018).
- Otsu, N. Threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).
- Cai, H., Yang, Z., Cao, X., Xia, W. & Xu, X. A new iterative triclass thresholding technique in image segmentation. *IEEE Trans. Image Process.* **23**, 1038–1046 (2014).
- Vala, M. & Baxi, A. A review on Otsu image segmentation algorithm. *Int. J. Adv. Res. Comput. Eng. Technol.* **2**, 387–389 (2013).
- Vincent, L., Vincent, L. & Soille, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 583–598 (1991).
- Beucher, S. The Watershed Transformation Applied to Image Segmentation. in *Proceedings of the 10th Pfefferkorn Conference on Signal and Image Processing in Microscopy and Microanalysis* 299–314 (1992). doi:[10.1.1.24.5229](https://doi.org/10.1.1.24.5229).
- Dufour, A. *et al.* Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *IEEE Trans. Image Process.* **14**, 1396–1410 (2005).
- Zhang, B., Zimmer, C. & Olivo-Marin, J. C. Tracking fluorescent cells with coupled geometric active contours. in *2004 2nd IEEE International Symposium on Biomedical Imaging: Macro to Nano* vol. 1 476–479 (2004).
- Dzyubachyk, O., Van Cappellen, W. A., Essers, J., Niessen, W. J. & Meijering, E. Advanced level-set-based cell tracking in time-lapse fluorescence microscopy. *IEEE Trans. Med. Imaging* **29**, 852–867 (2010).
- Al-Kofahi, Y., Lassoued, W., Lee, W. & Roysam, B. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Trans. Biomed. Eng.* **57**, 841–852 (2010).
- Santella, A., Du, Z., Nowotschin, S., Hadjantonakis, A. K. & Bao, Z. A hybrid blob-slice model for accurate and efficient detection of fluorescence labeled nuclei in 3D. *BMC Bioinf.* **11**, 19–33 (2010).
- Gudla, P. R. *et al.* A high-throughput system for segmenting nuclei using multiscale techniques. *Cytom. Part A* **73**, 451–466 (2008).
- Ambühl, M. E., Brepsant, C., Meister, J. J., Verkhovskiy, A. B. & Sbalzarini, I. F. High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *J. Microsci.* **245**, 161–170 (2012).
- Irshad, H., Veillard, A., Roux, L. & Racoceanu, D. Methods for nuclei detection, segmentation, and classification in digital histopathology: a review-current status and future potential. *IEEE Rev. Biomed. Eng.* **7**, 97–114 (2014).
- Yau, C. & Wakefield, J. Quantitative image analysis of chromosome dynamics in early *Drosophila* embryos. in *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro-Proceedings* 264–267 (2007). doi:<https://doi.org/10.1109/ISBI.2007.356839>.
- Chinta, R. & Wasser, M. Three-dimensional segmentation of nuclei and mitotic chromosomes for the study of cell divisions in live *Drosophila* embryos. *Cytom. Part A* **81**, 52–64 (2012).

21. Rizzi, B. & Sarti, A. Region-based PDEs for cells counting and segmentation in 3D+time images of vertebrate early embryogenesis. *Int. J. Biomed. Imaging* **2009**, 61 (2009).
22. Xu, H., Sepúlveda, L. A., Figard, L., Sokac, A. M. & Golding, I. Combining protein and mRNA quantification to decipher transcriptional regulation. *Nat. Methods* **12**, 739–742 (2015).
23. Toyoshima, Y. *et al.* Accurate automatic detection of densely distributed cell nuclei in 3D space. *PLoS Comput. Biol.* **12**, 600–793 (2016).
24. Arnéodo, A., Decoster, N. & Roux, S. G. A wavelet-based method for multifractal image analysis. I. Methodology and test applications on isotropic and anisotropic random rough surfaces. *Eur. Phys. J. B* **15**, 567–600 (2000).
25. Arnéodo, A., Decoster, N., Kestener, P. & Roux, S. G. A wavelet-based method for multifractal image analysis: from theoretical concepts to experimental applications. *Adv. Imaging Electron Phys.* **126**, 1–92 (2003).
26. Branco, M. R. & Pombo, A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* **4**, 780–788 (2006).
27. Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with imageJ. *Biophoton. Int.* **11**, 36–41 (2004).
28. Schindelin, J. *et al.* Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
29. Carpenter, A. E. *et al.* Cell Profiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, 62 (2006).
30. McQuin, C. *et al.* Cell Profiler 3.0: next-generation image processing for biology. *PLoS Biol.* **16**, 11 (2018).
31. Skinner, S. O., Sepúlveda, L. A., Xu, H. & Golding, I. Measuring mRNA copy number in individual *Escherichia coli* cells using single-molecule fluorescent in situ hybridization. *Nat. Protoc.* **8**, 1100–1113 (2013).
32. Mueller, F. *et al.* FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat. Methods* **10**, 277–278 (2013).
33. Lou, X., Kang, M., Xenopoulos, P., Muñoz-Descalzo, S. & Hadjantonakis, A. K. A rapid and efficient 2D/3D nuclear segmentation method for analysis of early mouse embryo and stem cell image data. *Stem Cell Rep.* **2**, 382–397 (2014).
34. Ianzini, F. *et al.* Development of the large scale digital cell analysis system. *Radiat. Prot. Dosimetry* **99**, 289–293 (2002).
35. Davis, P. J., Kosmacek, E. A., Sun, Y., Ianzini, F. & Mackey, M. A. The large-scale digital cell analysis system: an open system for nonperturbing live cell imaging. *J. Microsci.* **228**, 296–308 (2007).
36. Stegmaier, J. *et al.* Real-time three-dimensional cell segmentation in large-scale microscopy data of developing embryos. *Dev. Cell* **36**, 225–240 (2016).
37. Zinski, J. *et al.* Systems biology derived source-sink mechanism of bmp gradient formation. *Elife* **6**, 14 (2017).
38. Rajasekaran, B., Uriu, K., Valentin, G., Tinevez, J. Y. & Oates, A. C. Object segmentation and ground truth in 3D embryonic imaging. *PLoS One* **11**, 94 (2016).
39. Ulman, V. *et al.* An objective comparison of cell-tracking algorithms. *Nat. Methods* **14**, 1141–1152 (2017).
40. Li, L., Wang, X., Mullins, M. C. & Umulis, D. M. Evaluation of BMP-mediated patterning in zebrafish embryos using a growing finite difference embryo model. *BioRxiv* **58**, 5471. <https://doi.org/10.1101/585471> (2019).
41. Huang, Y. & Umulis, D. M. Scale invariance of BMP signaling gradients in zebrafish. *Sci. Rep.* **9**, 167 (2019).
42. Lepage, S. E. & Bruce, A. E. Zebrafish epiboly: mechanics and mechanisms. *Int. J. Dev. Biol.* **54**, 1213–1228 (2010).
43. Morita, H. *et al.* The physical basis of coordinated tissue spreading in zebrafish gastrulation. *Dev. Cell* **40**, 354–366.e4 (2017).
44. Concha, M. L. & Adams, R. J. Oriented cell divisions and cellular morphogenesis in the zebrafish gastrula and neurula: a time-lapse analysis. *Development* **125**, 983–994 (1998).
45. Bensch, R., Song, S., Ronneberger, O. & Driever, W. Non-directional radial intercalation dominates deep cell behavior during zebrafish epiboly. *Biol. Open* **2**, 845–854 (2013).
46. Warga, R. M. & Kimmel, C. B. Cell movements during epiboly and gastrulation in zebrafish. *Development* **108**, 569–580 (1990).
47. Bruce, A. E. Zebrafish epiboly: Spreading thin over the yolk. *Dev. Dyn.* **245**, 244–258 (2016).
48. Gul-Mohammed, J., Arganda-Carreras, I., Andrey, P., Galy, V. & Boudier, T. A generic classification-based method for segmentation of nuclei in 3D images of early embryos. *BMC Bioinf.* **15**, 10000 (2014).
49. Chan, C. J., Heisenberg, C. P. & Hiiragi, T. Coordination of morphogenesis and cell-fate specification in development. *Curr. Biol.* **27**, R1024–R1035 (2017).
50. Mendieta-Serrano, M. A., Schnabel, D., Lomelí, H. & Salas-Vidal, E. Cell proliferation patterns in early zebrafish development. *Anat. Rec.* **296**, 759–773 (2013).
51. Liò, P. Wavelets in bioinformatics and computational biology: state of art and perspectives. *Bioinformatics* **19**, 2–9 (2003).
52. Du, P., Kibbe, W. A. & Lin, S. M. Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics* **22**, 2059–2065 (2006).
53. Antoine, J. P., Carrette, P., Murenzi, R. & Piette, B. Image analysis with two-dimensional continuous wavelet transform. *Signal Process.* **31**, 241–272 (1993).
54. Bäck, T. & Schwefel, H.-P. An overview of evolutionary algorithms for parameter optimization. *Evol. Comput.* **1**, 1–23 (1993).
55. Myronenko, A. & Song, X. Point set registration: coherent point drifts. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 2262–2275 (2010).

Acknowledgements

We thank the help of Matt Thompson and Michelle Ingle for generating the nucleus ground truth data; Madeline Ku for codes testing.

Author contributions

Conceptualization, T.C.W., X.W., and D.U.; Methodology, T.C.W., X.W., Y.B. and D.U.; Software, T.C.W., X.W., and D.U.; Validation, T.C.W., X.W., L.L.L. and D.U.; Investigation, T.C.W., X.W., and D.U.; Data Curation, T.C.W., X.W., and D.U.; Writing – Original Draft, T.C.W., X.W., and D.U.; Writing – Review & Editing, T.C.W., X.W., and D.U.; Visualization, T.C.W., X.W., and D.U.; Supervision, D.U.; Project administration, D.U.; Funding Acquisition, D.U.

Funding

This work was supported by National Institutes of Health [R01GM132501].

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-88966-2>.

Correspondence and requests for materials should be addressed to D.M.U.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021