ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

DITET

Lucía Pancorbo Fernández
23-946-437

Automatic Segmentation of Ventricular Volumes from Cardiac Optoacoustic Data

Semester Project

Master's in Biomedical Engineering
Multiscale Functional and Molecular Imaging Lab
Swiss Federal Institute of Technology (ETH) Zurich

Supervision

Dr. Çağla Özsoy, Dr. Luis Deán-Ben & Prof. Daniel Razansky
May 21, 2024

Abstract

In this semester project, an automatic segmentation algorithm is proposed and evaluated on cardiac optoacoustic volumes for the segmentation of the left ventricle. The method, Restricted Propagation, involves the segmentation of the central slice and its propagation to the rest of the volume with an algorithm based on active contours. A constraint on the maximum diameter of the masks is set to prevent them from overflowing the cavity and ensuring consistency between contiguous slices.

The data consists of two Langendorff-isolated murine hearts, a fibrotic transgenic model and the wild-type control. The left ventricle is segmented for volume computation at each time point. A good perfoming segmentation method should capture the contraction dynamics in the temporal profile of the volumes. The algorithm is tested against the foundation segmentation model MedSAM.

The proposed model succeeded in segmenting the left ventricle in most cases. However, the masks were not accurate enough to capture a periodic pattern of the volume, indicating that the method needs further improvement to be able to extract quantitative information. The MedSAM model presented a superior performance in that aspect. Nonetheless, the Restricted Propagation masks were better at delineating the ventricle in some slices, showing potential for fully automatic and reliable segmentation. Code is available in `https://github.com/lpancorbo/L_Pancorbo_Automatic_Segmentation_of_Ventricular_Volumes_from_Cardiac_Optoacoustic_Data`

Contents

1   Introduction

1.1   Optoacoustic Imaging

Optoacoustic Imaging (OA) leverages the photoacoustic effect to obtain optical absorption contrast at high penetration depth. The tissue is probed with a short-pulsed laser, producing a local thermoelastic expansion. This causes a pressure wave that is detected with an ultrasound transducer to create a map of the molecular absorbers distribution in the tissue (1). The penetration barrier imposed by light diffusion is overcome while still achieving optical absorption contrast (2). This allows the non-invasive acquisition of images at centimetre-scale depths and $\sim$100 μm resolution (3). State-of-the-art OA tomography systems achieve fast volumetric imaging speeds determined by the pulse repetition frequency of the laser (100 Hz). Spectral specificity is attained by imaging at different optical wavelengths, a technique known as multi-spectral optoacoustic tomography (MSOT) (4).

Langendorff hearts are common in cardiac studies. This technique maintains isolated-perfused mammalian hearts beating, which enables the imaging of cardiac dynamics without the ribs, skin and blood (5; 6). OA imaging has unlocked high spatiotemporal resolution at the centimetre-scale field of view allowing the 4D imaging of the entire isolated murine beating heart (6). Ultrafast OA (laser pulse of $>$ 1 KHz) goes one step further and can even capture the propagation of mechanical waves across the cardiac walls (7).

1.2   Segmentation of Medical Images

Image segmentation plays a crucial role in the medical field by enabling the precise delineation of anatomical structures and the extraction of quantitative information. Manual segmentation is the gold standard for medical images but is time-consuming, has high variability between users, and lacks scalability. Automatic segmentation methods can be broadly categorized into non-trainable/traditional and trainable approaches (8).

Simple traditional methods include intensity thresholding and region growing (9), while more advanced techniques are watershed segmentation and active contours (10; 11). The latter is also known as snakes. It consists of initializing a contour and minimizing an energy functional to guide it towards image features such as edges. In the traditional snake model, the energy functional contains an external energy term defined by the image characteristics, and an internal energy

that controls the deformation of the countour. Some variants of the snake model define the external energy differently to solve convergence and computational problems. Examples are the gradient vector flow (GVF) model, the balloon model, and the geodesic active contours model (12; 13; 14). Overall, active contours is a well-established segmentation method that has been applied widely in the medical field (15; 16; 17).

On the other hand, trainable methods use machine learning to learn features from data and perform specific segmentation tasks. These methods include the implementation of common deep learning models like the U-Net architecture, which was proposed for biomedical image segmentation (18). Nonetheless, these are task-specific and require training with a large dataset of manually segmented masks, which are often not available. This drawback explains the growing interest in task-agnostic or foundation models. These models are trained on large datasets with high variability and can perform well on downstream tasks without re-training (19). An example is the Segment Anything Model (SAM), which was trained on the largest segmentation dataset containing a billion masks of natural images (20). Its adaptation to medical images (MedSAM) aims to achieve universal promptable segmentation in the medical field (21). Its training dataset contained 10 imaging modalities such as MR, CT, Endoscopy, Ultrasound, and X-Ray.

1.3   Problem Statement

Manual segmentation is time-consuming, especially when dealing with the temporal evolution of large volumes. For this reason, only selected frames are segmented, and the measurements and conclusions are then generalised to the whole set. This restriction of the amount of data available for analysis motivates the search for an automatic solution that can segment and extract information from every volume and time point.

In this project, an automatic segmentation algorithm for the left ventricle (LV) is proposed. To evaluate its performance, we have optoacoustic images from Langendorff-isolated mice hearts. The dataset comes from a previous study and contains a transgenic mouse model for fibrosis (TG) and a wild-type control (WT) (22). Our starting hypothesis is that the algorithm will be able to segment the LV at every time point, being able to capture the heart rhythm in the temporal profile of the resulting volume. The method is also evaluated against the foundation model for medical imaging segmentation, MedSAM.

2 Methods

The proposed algorithm, Restricted Propagation, was evaluated using cardiac optoa-coustic data for left ventricle (LV) segmentation. This data comes from a prior study that implements volumetric multi-spectral optoacoustic tomography (vMSOT) for beat-to-beat characterization of Langendorff-perfused mice hearts (22). The dataset consists of 2 hearts: a transgenic $Fosl - 2^{tg}$ mouse model (TG) to assess myocardial fibrosis and the wild-type control $Fosl - 2^{wt}$ (WT). $128 \times 128 \times 128$ voxels were acquired over a field-of-view of $11 \times 11 \times 11$ mm, resulting in a spatial resolution of $\sim 86\mu$m. Volumes were acquired at a sampling frequency of 100 Hz and we used single-wavelength vMSOT signals at 700 nm.

The segmented masks are compared with results using the foundation model MedSAM for comparison (21). Besides visual inspection of the results, temporal profiles of the LV volumes are computed. Finally, the Intersection over Union (IoU) metric is used, getting values between 0 and 1 according to how well-aligned the mask from the Restricted Propagation method is to the MedSAM mask.

2.1 Restricted Propagation

The proposed automatic segmentation algorithm has two main steps: automatic segmentation of the central slice and mask propagation to the rest of the volume. The first step involves blob detection and active contours, while the second implements active contours with a restriction on the maximum diameter. This pipeline is based on the assumption that the central slice corresponds to the maximum area of the ventricle and that the spatial distance between slices (spatial resolution) is small enough to ensure consistent mask propagation.

2.1.1 AutoSegment Pipeline

To automatically segment the central slice, the AutoSegment pipeline is implemented. An overview is shown in Figure 1. This method involves several steps, starting with the L1/TV (image denoising) model (23), which is based on the minimization of the data-fidelity term (L1-norm) and the Total-Variation regularization term. This pre-processing of the image creates piece-wise smooth regions suitable for segmentation.

Next, blob detection based on the Laplacian of Gaussian (LoG) method is used to identify a unique blob within the slice, which is assumed to fall within the left ventricle (24). The LoG operation consists of a Gaussian smoothing followed by a

Laplacian operator, that computes the second derivative of the image at each point. Local maxima of the resulting image are marked as blobs, defined as a circular area with different intensity values than the surroundings. We apply blob detection with increasing blob sizes, by increasing the standard deviation values of the Gaussian kernel, until we either detect a single blob or reach the maximum blob size that can fit in the image. If a single blob is found, indicating the presence of the LV, we proceed to the next step. Otherwise, if multiple blobs are detected or no blobs are found, one can conclude that ventricles are not present in the slice, and an empty mask is generated.

If a single blob is found, it applies active contours (morphological Chan-Vese method) to segment the ventricles (25). Also known as Active Contours without Edges (ACWE), the Chan-Vese method varies from traditional active contours in not using the gradient of the image as a stopping term, which improves performance in objects with smooth or discontinuous boundaries (26). Morphological ACWE is preferred for this problem as it improves stability and speed. Instead of directly solving the partial differential equation used in traditional active contours, the algorithm approximates its numerical solution by applying morphological operators successively. Starting from a circle at the found blob coordinates and with the same radius as the Gaussian kernel, the contour evolves iteratively until it has segmented the ventricle. The main assumption of ACWE is that the foreground intensities are different on average from the background intensities.
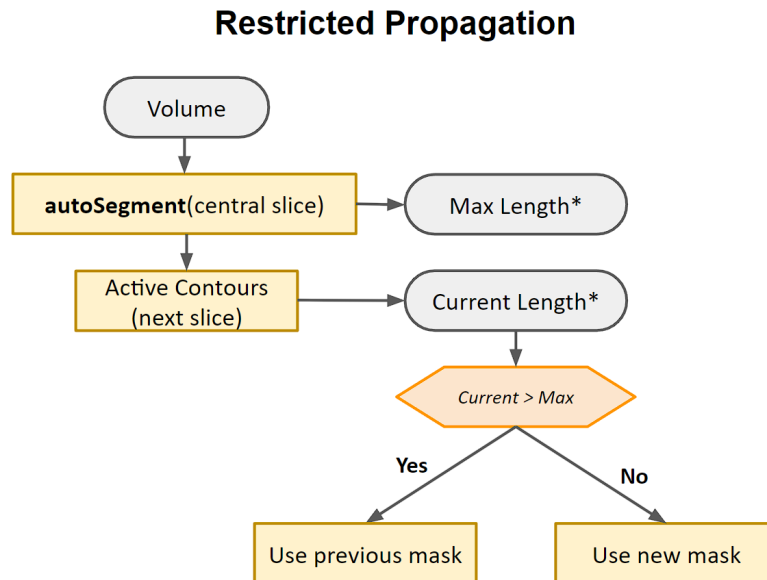
2.1.2   Mask Propagation

Once we have a successful segmentation of the ventricle in the central slice, the algorithm iteratively extends the mask to adjacent slices in the volume, ensuring consistent segmentation throughout (see flowchart in Figure 2). First, the whole volume is pre-processed with the L1/TV denoising model. The central slice is segmented with the autoSegment algorithm and the maximum Feret diameter, computed as the largest Euclidean distance among all pixels (27), is set as an upper limit.

The mask in the central slice is used to initialize morphological active contours on the next contiguous slice. To prevent the over-expansion of segmentation boundaries, the maximum Feret diameter of the propagated contour is restricted to the upper limit. This means that after performing active contours, the resulting mask is checked for this restriction. If the condition is not met, the mask from the previous slice is applied to the current one without any change. This process is iterated both upwards and downwards from the central slice.
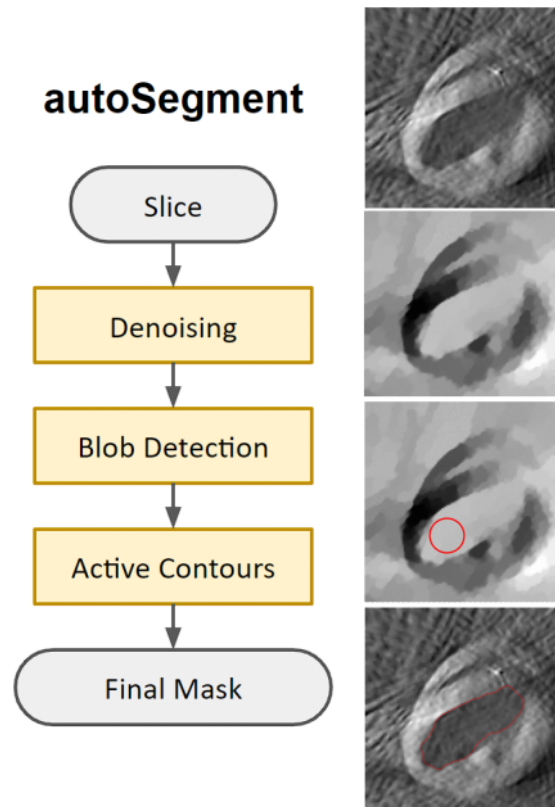
Figure 1: AutoSegment Pipeline with an example slice illustrating each step. The image corresponds to the central sagittal slice of the TG heart.


Figure 2: Restricted Propagation flowchart. The mask from the central slice is propagated to the rest of the volume.

The restricted propagation method had to be adapted to the WT dataset. In this case, the denoising is performed by a median filter and morphological Geodesic Active Contours (GAC) is used instead of ACWE (14). This method differs from the traditional active contours in the formulation of the energy functional and can be used to segment objects with discernible yet noisy or broken borders.

2.2   Foundation Segmentation Model: MedSAM

The MedSAM model is a task-agnostic image segmentation model pre-trained on a large dataset with multiple image modalities. Figure 3 illustrates the network architecture. An image encoder based on a vision transformer (ViT) extracts features from the image, while a prompt encoder incorporates the bounding box input by the user. Finally, a mask decoder generates the segmentation masks.
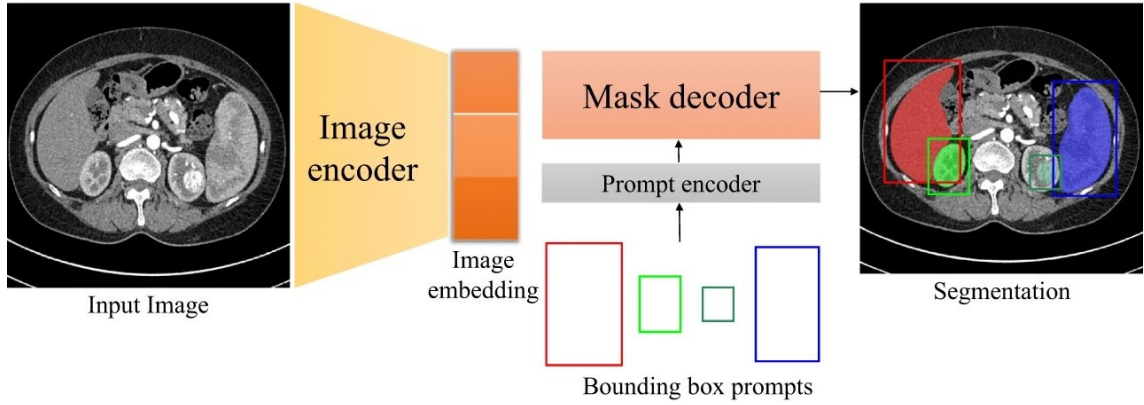
Figure 3: MedSAM network architecture. Figure adapted from (21).

The available trained model and the inference code were used to segment the optoacoustic volumes slice-by-slice. For this, images are first denoised with the L1/TV algorithm and resized into the model's input requirements $(1024 \times 1024 \times 3)$ by interpolation and duplication of the values in the color dimension. Then, 3 ranges of slice indices were defined and 3 bounding boxes were visually selected. The size of the ventricle is similar in all the slices belonging to the same range, so that a single bounding box could fit them all. Figure 4 shows the input prompts used for bottom, central, and top example slices for the TG and WT data set.
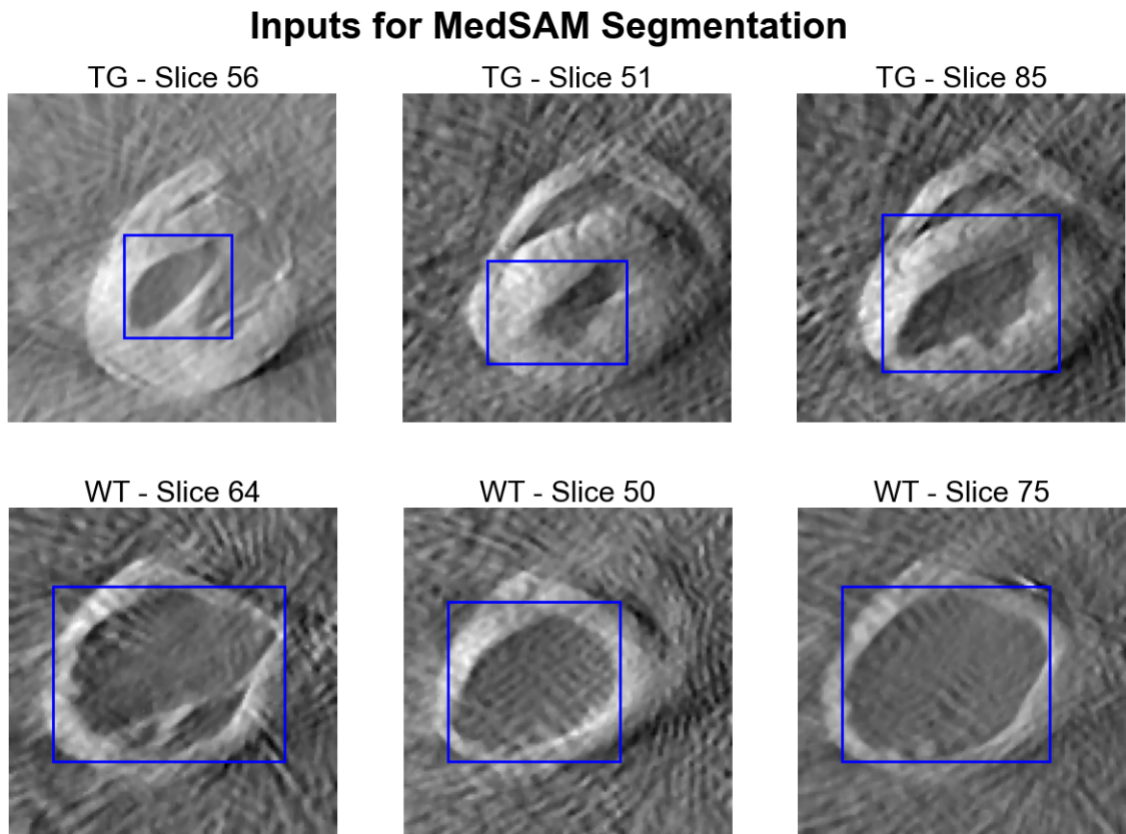
Figure 4: Bounding boxes used as prompt for the pre-trained MedSAM model. A total of 3 boxes are defined for each data set, ensuring that it marks the LV in all the slices and time points.

3 Results and Discussion
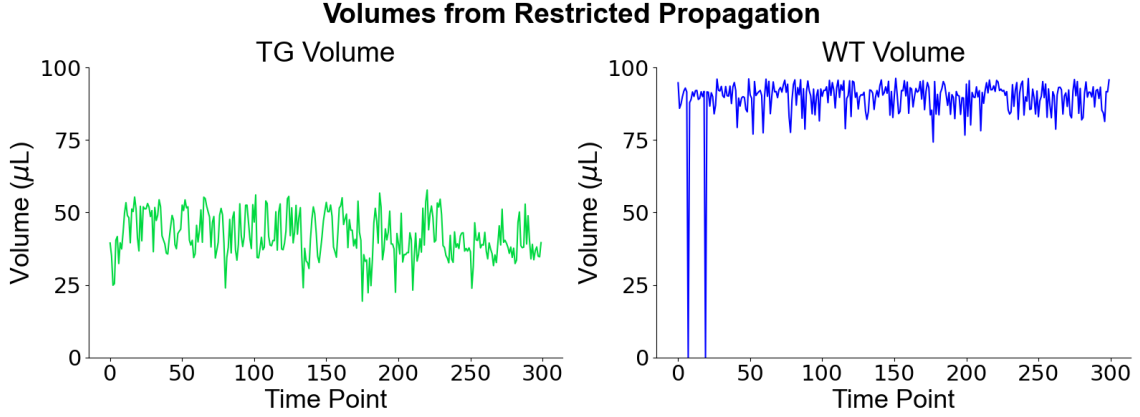
3.1 Restricted Propagation

The visual inspection of the segmentation results indicates overall effectiveness, as observed in the consistent delineation of ventricular walls across most slices. Figure 5 shows examples from each data set. The method slightly underestimates the areas for the TG, while it overestimates them for the WT. The central slices exhibit particularly reliable segmentation results, while some inaccuracies are noted at the extreme slices. The restriction on the maximum diameter occasionally leads to unchanged masks across consecutive slices. However, due to the small spatial resolution, the differences between contiguous slices are minimal, ensuring that the same mask fits neighbouring slices with relatively good accuracy.
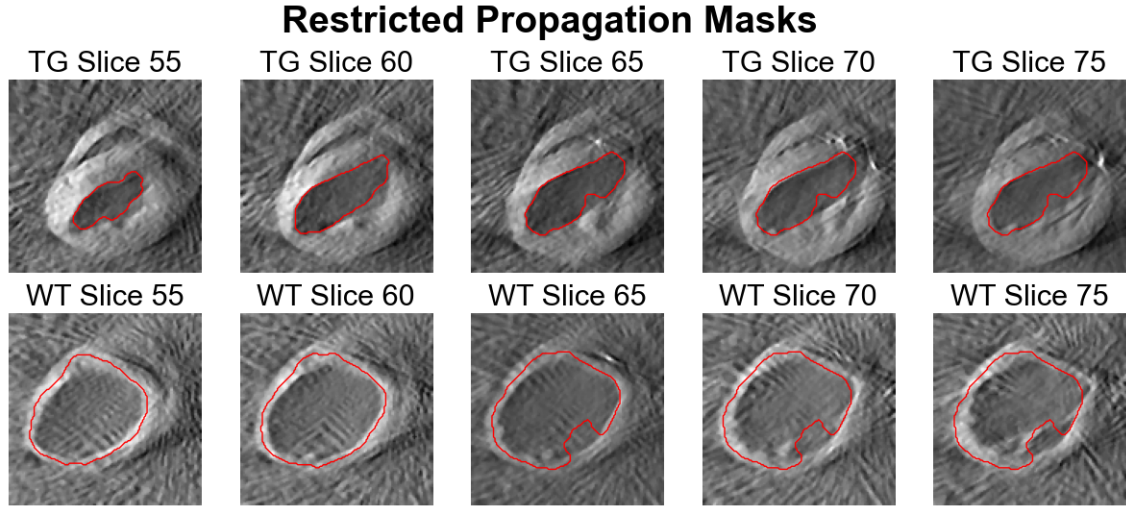
Figure 5: Masks propagated from the central slice following the restricted propagation pipeline. **Top:** Transgenic mouse time point 170. **Bottom:** Wild-Type mouse time point 190.

Volumes are computed by segmenting all slices for a given time point and multiplying the total number of foreground voxels by the voxel size. Repeating this for all time points results in the temporal profiles of ventricular volumes (See Figure 6). There is a large difference between the volume values obtained for TG and WT. This difference, which is noticeable in the slice images, is probably accentuated by the respective underestimation and overestimation done by the segmentation model. The TG's LV has a mean volume of 42.05 μL, while the WT's ventricle has an average volume of 89.93 μL.

While ventricular contraction is visually imperceptible in the WT dataset, it can

Figure 6: Calculated left ventricular volumes for 3 seconds using the Restricted Propagation algorithm. **Left:** Transgenic mouse **Right:** Wild-Type mouse.

be appreciated in the TG dataset. However, both profiles show noisy and aperiodic patterns lacking any contraction rhythm. This suggests that the method needs to be tuned, via adjusting parameters, before enabling the quantitative analysis of cardiac dynamics. Additionally, it is worth noting that there are 2 cases where no mask is generated for the WT dataset. In those cases, the blob detection algorithm failed to identify the LV.

While the quantitative results of the proposed segmentation method are not what we expected, the masks visually align with ventricular boundaries suggesting some degree of success. Achieving a more accurate delineation of the walls is needed to capture the heart rhythm effectively. Note that the method relies on multiple parameters, including thresholds in blob detection, parameters of active contour and denoising algorithms, that need to be adjusted for each dataset. This highlights the need for methods that generalize to different contrasts, resolutions, and other data-specific features to ensure robust and reliable segmentation.

3.2   MedSAM Segmentation

Figure 7 shows examples of the MedSAM results. The sharp segmentation masks prove the foundation nature of the method, outlining the LV's walls in a well-defined manner. This method often obtains small foreground classifications outside of the main contour (See TG Slice 60). While this could be understood as noise, a closer inspection reveals that these parts of the images are of the same intensity as the inside of the ventricle and are probably inner cavities. In other cases like in WT Slice 75, there are background classifications inside the main mask that match the intensity of the walls, probably indicating some part of the cardiac wall. The ability

MedSAM Masks
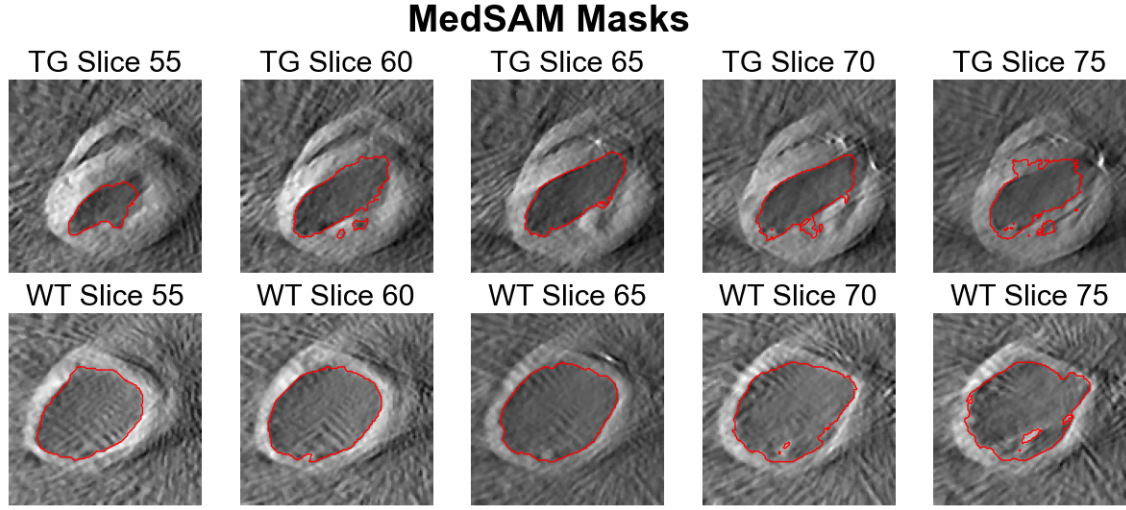
Figure 7: Segmentation of different slices using the MedSAM model. **Top:** Transgenic mouse time point 170. **Bottom:** Wild-Type mouse time point 190.

to make a distinction between walls and cavities illustrates the accuracy of the MedSAM model and confirms that we can use it as a baseline to quantify the accuracy of the proposed method.

When visualizing the temporal profiles of the volumes, we can see that the MedSAM model can effectively capture the heart rhythm of the TG heart, showing a periodic pattern (Figure 8). This is not the case for the WT profile. However, this motion was not perceptible in the images indicating that the image quality was not suitable for this task. The mean volumes are 38.35 μL and 74.21 μL for the TG and WT's LV respectively. Additionally, the TG temporal profile results in a well-defined spectrum in the frequency domain, with a clear fundamental peak at 6.67 Hz - 400.2 bpm (See Figure 9, Left). The primary peak is accompanied by secondary peaks of decreasing magnitude, corresponding to higher harmonics. The WT profile does not result in a clean spectrum, but we can still differentiate a fundamental peak at 8.67 Hz - 520.2 bpm (Figure 9, Right). This matches the normal resting state heart rate in mice, which is around 500-700bpm (28).

3.3   Performance Comparison

Results have proven the superior performance of the MedSAM model, exhibiting a robust segmentation of all the slices and being able to capture the contraction and relaxation dynamics. Nonetheless, the proposed Restricted Propagation method is also capable of finding and delineating the ventricle in most slices. The proposed method provides multiple advantages including the definition of a maximum diame-

Volumes from MedSAM Segmentation
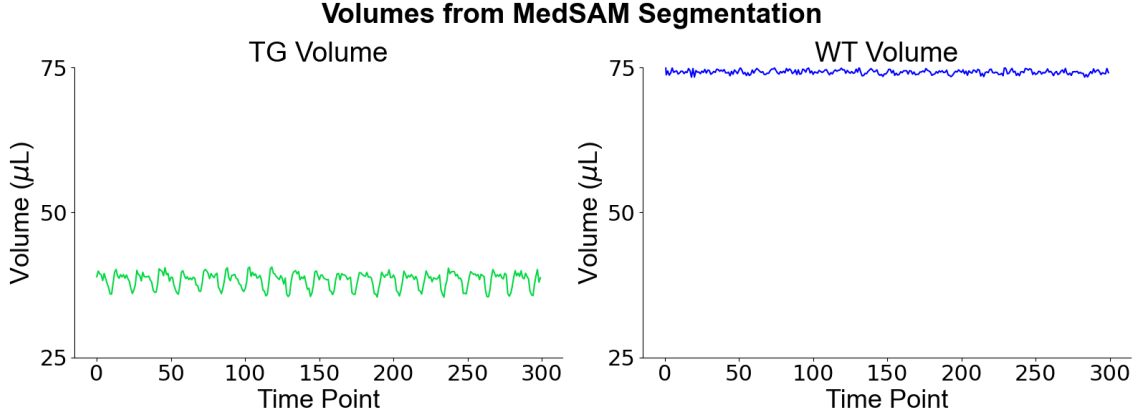
Figure 8: Calculated left ventricular volumes for 3 seconds using MedSAM segmentation model. **Left:** Transgenic mouse **Right:** Wild-Type mouse.

Fourier Transforms of MedSAM Volumes
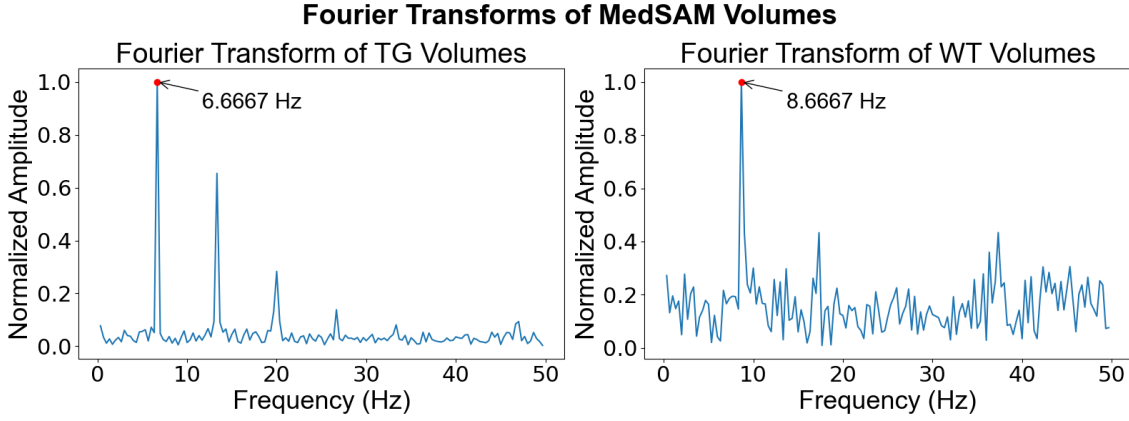
Figure 9: Fourier Transforms of the volumes computed from MedSAM segmentations. **Left:** Transgenic mouse **Right:** Wild-Type mouse.

ter to prevent the mask from overflowing and the lack of user input prompts, which makes the method fully automatic. Moreover, the implementation of active contours makes the segmentation decisions more dependent on the cavity structure than on the intensity values. This can be observed when comparing WT Slice 75 in Figures 5 and 7. The MedSAM model prioritizes the cohesiveness of intensity values, segmenting all the dark area except the wall. On the other hand, the Restricted Propagation method stops the segmentation at the wall, delineating only the left ventricle and excluding the right ventricle.

Table 1 displays the computed volumes averaged across systolic and diastolic phases for the three different methods. Results from manual segmentation were obtained in a previous study from two time points (systole and diastole) and are considered the ground-truth values (22). Thus, the Restricted Propagation method achieves a mean absolute error (MAE) of 2.37 μL and 20.68 μL for the TG and WT

Table 1: Mean Left Ventricular Volumes

Heart	Method		
	Manual[*]	Restricted Propagation	MedSAM
Transgenic	39.67 μL	42.05 μL	38.35 μL
Wild-Type	69.25 μL	89.93 μL	74.21 μL

[*]Values from previous study (22).

cases, respectively. The MedSAM model proved a superior performance, getting an MAE of 1.32 μL 4.96 μL for TG and WT. It is worth mentioning that there is only one heart per dataset, and that the quality of the WT recordings was suboptimal. Consequently, we have mostly relied on the TG results when evaluating the performance of the models.

To provide a quantitative analysis of the proposed method's performance, we have compared it to the baseline: the MedSAM masks. For this purpose, we have used the Intersection-over-Union (IoU) metric, which measures the overlap between predicted and ground-truth foreground regions divided by the union providing a value between 0 and 1 (29). This metric is often used in object segmentation tasks and is preferred to pixel-wise classification accuracy due to class imbalance (more background than foreground pixels).

After noting a distinction in accuracy at the extreme slices of the volume related to the central slices, it was decided to compute the IoU values for each slice index. Figure 10 shows the resulting IoU for all the 300 time points grouped in box plots, having one box for each slice index. For the TG data, Figure 10 Top, the mentioned increase in accuracy for the central slices is observed. This indicates there is more similarity between the Restricted Propagation results and the MedSAM results in this range. Conversely, the IoU values for the WT data remain constant across slices (Figure 10 Bottom).

Figure 11 illustrates the similarities between the Restricted Propagation (in red) and the MedSAM masks (in blue) for the TG heart. The intersection between the masks used for IoU computation is highlighted in yellow. Three examples with good scores are shown in the top row. In these cases, both masks are very well aligned and delineate the LV walls. In the bottom row, there are three examples with bad scores. As expected, they correspond to slices at the extremes of the volume. Surprisingly, for the first and last example, it is the Restricted Propagation mask the one that best segments the ventricle. This shows its potential for reliable segmentation.
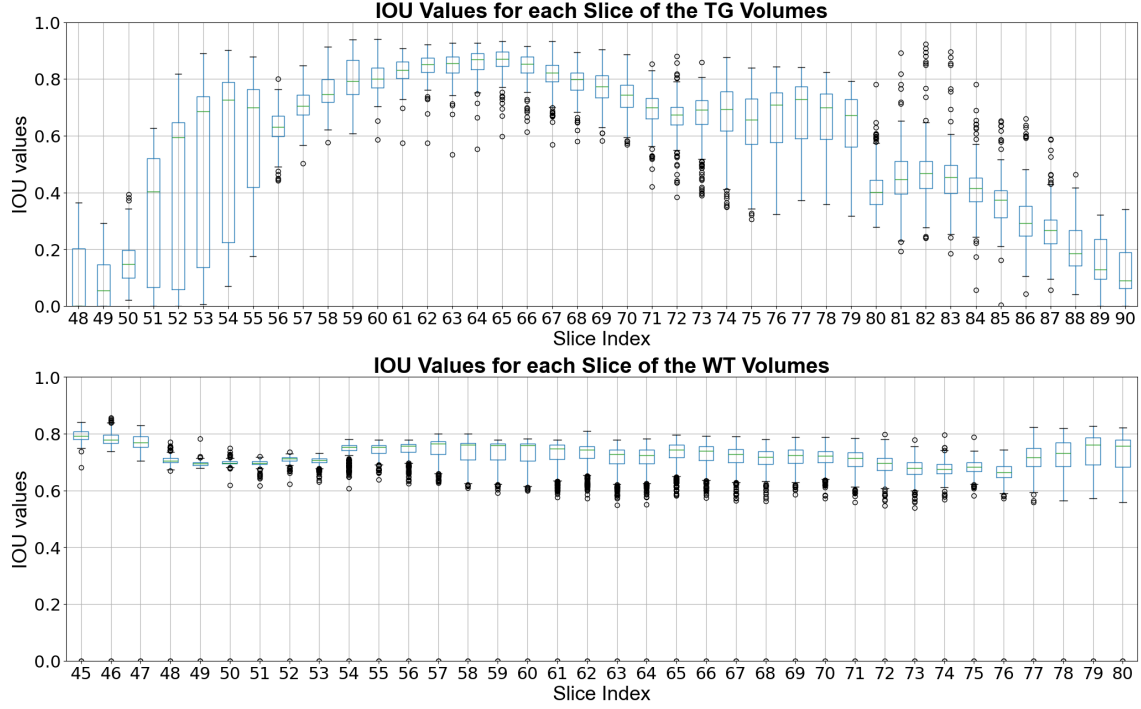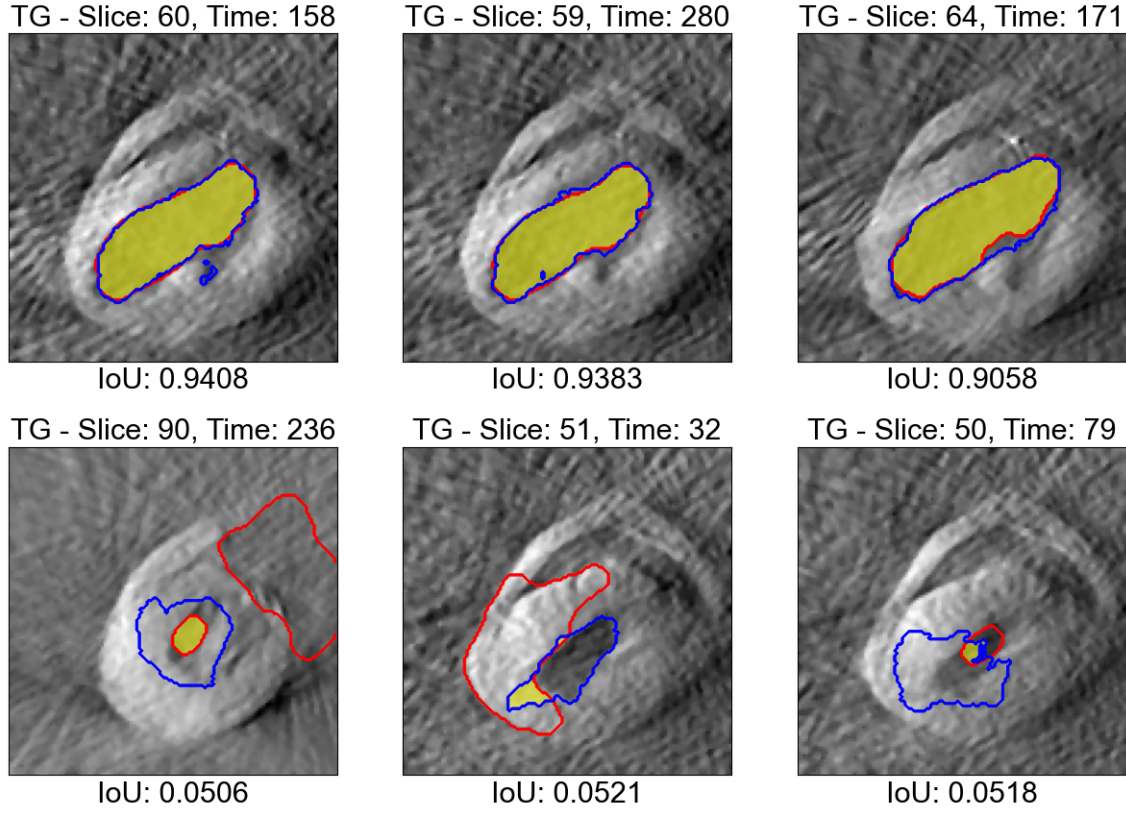
Figure 10: Intersection over Union (IoU) values between the Restricted Propagation and the MedSAM masks across slices for 300 time points. **Top:** Transgenic mouse **Bottom:** Wild-Type mouse.

Figure 12 shows the same comparison for the WT data. In the examples with the best IoU scores, the masks are not as well-aligned as for the TG data. Additionally, the MedSAM masks also fail to correctly delineate the LV walls, overflowing the ventricle in all cases. For the examples with bad scores, the Restricted Propagation masks fail to cover half of the ventricle, while the baseline method delineates the walls accurately in this case.

Classic segmentation methods depend on manually extracted features and heuristics. Active contours have predefined energy functions and optimization techniques, and their performance might be sensitive to initialization and parameter selection (26; 14). Foundation models, such as the Segment Anything Model SAM or its adaptation for medical images MedSAM (20; 21), are pre-trained on large datasets to automatically extract meaningful features and adapt to various imaging modalities and tasks. We had to perform some data-specific adjustments to the Restricted Propagation pipeline, while the MedSAM was directly applied to TG and WT data without any modification. It is worth noting that the MedSAM model was not trained on optoacoustic images, but could still generalized to predict accurate segmentation masks. Moreover, its performance could be improved if one bounding

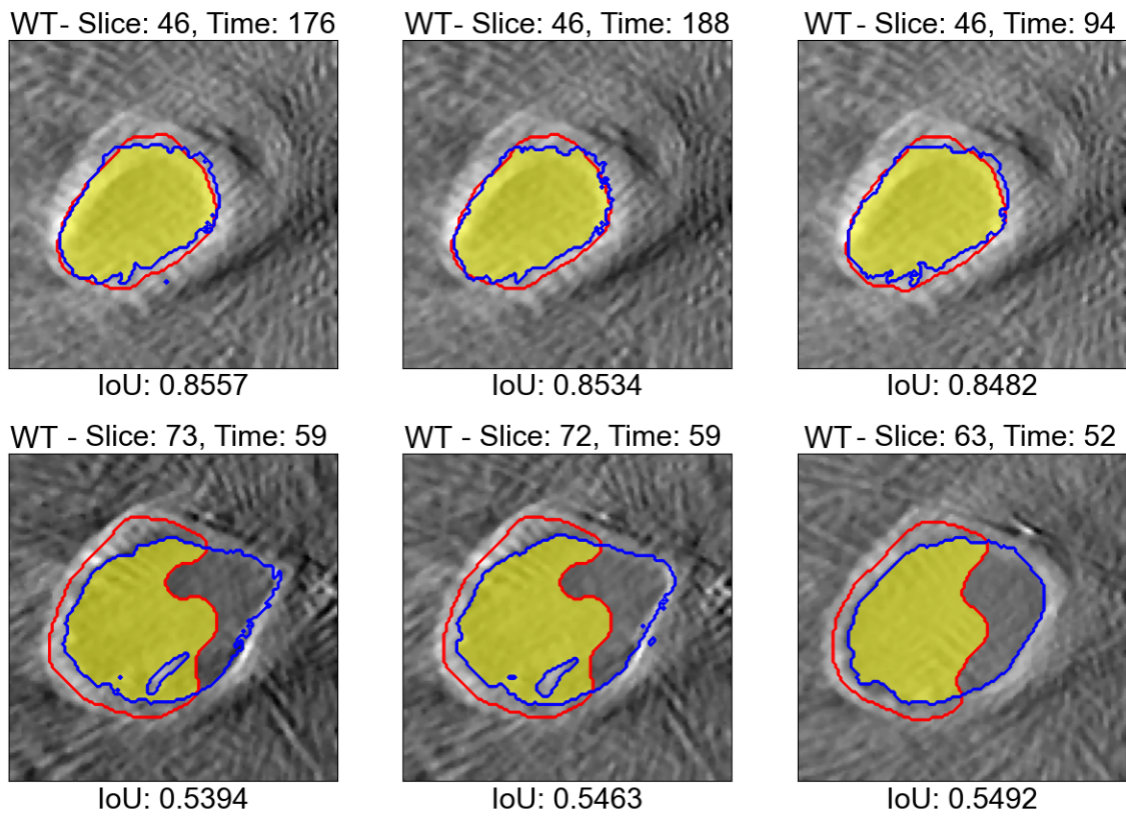**Restricted Propagation and MedSAM Comparison for TG**

Figure 11: Comparison of the restricted propagation (red) and MedSAM masks (blue) for TG with the intersection area highlighted in yellow. The Intersection over Union (IoU) values are annotated at the bottom of each plot.

box was defined per slice instead of three bounding boxes for the whole volume.

The proposed Restricted Propagation shows great potential as a fully automatic segmentation pipeline. However, it has several disadvantages, such as its limited ability to adapt to different contrasts and the cascading effect, where errors in one slice are propagated to subsequent ones (refer to Slices 72 and 73 at time 59 in Figure 12). Nonetheless, this method was more accurate and efficient than directly applying the autoSegment function individually to each slice in the volume, while ensuring consistency across slices.

Figure 12: Comparison of the restricted propagation in red, and MedSAM masks in blue, for WT with the intersection area highlighted in yellow. The Intersection over Union (IoU) values are annotated at the bottom of each plot.

4   Conclusion

We aimed to establish a fully automated segmentation method for cardiac cavities in optoacoustic images. The pipeline consisted of the initial segmentation of the central slice and its subsequent propagation with morphological active contours to the rest of the volumes. The segmentation results are promising, succeeding in classifying most foreground pixels accurately. However, the LV areas were often underestimated or overestimated, preventing the quantitative analysis of the volumes. The proposed method was compared against MedSAM segmentation results, which provided more precise masks and allowed the characterization of the LV temporal dynamics.

Improvements to the Restricted Propagation model could include starting with a manual segmentation at the central slice, as well as defining more than one initial slice from which the masks are propagated. This is expected to solve the error-cascading issue. Additionally, the images could be scaled to a larger size as in the MedSAM pre-processing step to obtain sharper masks.

Future work should focus on refining the proposed pipeline to generalize to multiple contrasts and data-specific features. One step in this direction could be including histogram equalization as a pre-processing step and exploring different active contours methods and parameter configurations.

References

[1] L. V. Wang and J. Yao, "A practical guide to photoacoustic tomography in the life sciences," *Nature methods*, vol. 13, no. 8, pp. 627–638, 2016.

[2] X. Deán-Ben, S. Gottschalk, B. Mc Larney, S. Shoham, and D. Razansky, "Advanced optoacoustic methods for multiscale imaging of in vivo dynamics," *Chemical Society Reviews*, vol. 46, no. 8, pp. 2158–2198, 2017.

[3] J. Weber, P. C. Beard, and S. E. Bohndiek, "Contrast agents for molecular photoacoustic imaging," *Nature methods*, vol. 13, no. 8, pp. 639–650, 2016.

[4] X. Luís Deán-Ben and D. Razansky, "Adding fifth dimension to optoacoustic imaging: volumetric time-resolved spectrally enriched tomography," *Light: Science & Applications*, vol. 3, no. 1, pp. e137–e137, 2014.

[5] R. M. Bell, M. M. Mocanu, and D. M. Yellon, "Retrograde heart perfusion: the langendorff technique of isolated heart perfusion," *Journal of molecular and cellular cardiology*, vol. 50, no. 6, pp. 940–950, 2011.

[6] H.-C. A. Lin, X. L. Déan-Ben, M. Reiss, V. Schöttle, C. A. Wahl-Schott, I. R. Efimov, and D. Razansky, "Ultrafast volumetric optoacoustic imaging of whole isolated beating mouse heart," *Scientific reports*, vol. 8, no. 1, p. 14132, 2018.

[7] Ç. Özsoy, A. Özbek, M. Reiss, X. L. Deán-Ben, and D. Razansky, "Ultrafast four-dimensional imaging of cardiac mechanical wave propagation with sparse optoacoustic sensing," *Proceedings of the National Academy of Sciences*, vol. 118, no. 45, p. e2103979118, 2021.

[8] C. E. Cardenas, J. Yang, B. M. Anderson, L. E. Court, and K. B. Brock, "Advances in auto-segmentation," in *Seminars in radiation oncology*, vol. 29, no. 3.   Elsevier, 2019, pp. 185–197.

[9] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[10] A. P. Mangan and R. T. Whitaker, "Partitioning 3d surface meshes using watershed segmentation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 4, pp. 308–321, 1999.

[11] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.

[12] C. Xu and J. L. Prince, "Gradient vector flow: A new external force for snakes," in *Proceedings of IEEE computer society conference on computer vision and pattern recognition*.   IEEE, 1997, pp. 66–71.

[13] L. D. Cohen, "On active contour models and balloons," *CVGIP: Image understanding*, vol. 53, no. 2, pp. 211–218, 1991.

[14] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International journal of computer vision*, vol. 22, pp. 61–79, 1997.

[15] G. Sapiro, R. Kimmel, and V. Caselles, "Object detection and measurements in medical images via geodesic deformable contours," in *Vision Geometry IV*, vol. 2573. SPIE, 1995, pp. 366–378.

[16] R. Malladi, R. Kimmel, D. Adalsteinsson, G. Sapiro, V. Caselles, and J. A. Sethian, "A geometric approach to segmentation and analysis of 3d medical images," in *Proceedings of the workshop on mathematical methods in biomedical image analysis*. IEEE, 1996, pp. 244–252.

[17] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, "A geometric snake model for segmentation of medical imagery," *IEEE Transactions on medical imaging*, vol. 16, no. 2, pp. 199–209, 1997.

[18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[19] J. Ma and B. Wang, "Towards foundation models of biological image segmentation," *Nature Methods*, vol. 20, no. 7, pp. 953–955, 2023.

[20] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.

[21] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," *Nature Communications*, vol. 15, no. 1, p. 654, 2024.

[22] Ç. Özsoy *et al.*, "Five-dimensional optoacoustic characterization of cardiac fibrosis and arrhythmias driven by ap-1 transcription factor fosl-2," *Manuscript submitted for publication to Light: Science and Applications*, 2024.

[23] C. A. Micchelli, L. Shen, Y. Xu, and X. Zeng, "Proximity algorithms for the l1/tv image denoising model," *Advances in Computational Mathematics*, vol. 38, pp. 401–426, 2013.

[24] T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention," *International Journal of Computer Vision*, vol. 11, no. 3, pp. 283–318, 1993.

[25] P. Marquez-Neila, L. Baumela, and L. Alvarez, "A morphological approach to curvature-based evolution of curves and surfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 2–17, 2013.

[26] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266–277, 2001.

[27] H. G. Merkus, *Particle size measurements: fundamentals, practice, quality.* Springer Science & Business Media, 2009, vol. 17.

[28] P. M. Janssen, B. J. Biesiadecki, M. T. Ziolo, and J. P. Davis, "The need for speed: mice, men, and myocardial kinetic reserve," *Circulation research*, vol. 119, no. 3, pp. 418–421, 2016.

[29] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *International symposium on visual computing.* Springer, 2016, pp. 234–244.