

Circle Representation for Medical Instance Object Segmentation

Juming **Xiong**¹, Ethan H. **Nguyen**², Yilin **Liu**², Ruining **Deng**², Regina N **Tyree**³, Hernan **Correa**⁴, Girish **Hiremath**³, Yaohong **Wang**⁴, Haichun **Yang**⁴, Agnes B. **Fogo**⁴, Yuankai **Huo**^{1,2,4}

¹ Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, USA

² Department of Computer Science, Vanderbilt University, Nashville, TN, USA

³ Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Vanderbilt University Medical Center, Nashville, TN, USA

⁴ Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA

Abstract

Recently, circle representation has been introduced for medical imaging, designed specifically to enhance the detection of instance objects that are spherically shaped (e.g., cells, glomeruli, and nuclei). Given its outstanding effectiveness in instance detection, it is compelling to consider the application of circle representation for segmenting instance medical objects. In this study, we introduce CircleSnake, a simple end-to-end segmentation approach that utilizes circle contour deformation for segmenting ball-shaped medical objects at the instance level. The innovation of CircleSnake lies in these three areas: (1) It substitutes the complex bounding box-to-octagon contour transformation with a more consistent and rotation-invariant bounding circle-to-circle contour adaptation. This adaptation specifically targets ball-shaped medical objects. (2) The circle representation employed in CircleSnake significantly reduces the degrees of freedom to two, compared to eight in the octagon representation. This reduction enhances both the robustness of the segmentation performance and the rotational consistency of the method. (3) CircleSnake is the first end-to-end deep instance segmentation pipeline to incorporate circle representation, encompassing consistent circle detection, circle contour proposal, and circular convolution in a unified framework. This integration is achieved through the novel application of circular convolution within the context of circle detection and instance segmentation. In practical applications, such as the detection of glomeruli, nuclei, and eosinophils in pathological images, CircleSnake has demonstrated superior performance and greater rotation invariance when compared to benchmarks. The code has been made publicly available at: <https://github.com/hrlblab/CircleSnake>.

Keywords

Contour-based, CircleSnake, Detection, Segmentation, Image Analysis, Pathology

Article informations

<https://doi.org/10.59275/j.melba.2025-8bad> ©2025 Juming Xiong, Ethan H. Nguyen, Yilin Liu, Ruining Deng, Regina N Tyree, Hernan Correa, Girish Hiremath, Yaohong Wang, Haichun Yang, Agnes B. Fogo, Yuankai Huo. License: CC-BY 4.0

Volume 3, Received: 07/2024, Published 10/2025

Corresponding author: yuankai.huo@vanderbilt.edu



1. Introduction

The deep learning model for natural images is frequently applied in the context of medical imaging to identify and highlight regions of interest within medical images (Lo et al., 2018; Liang et al., 2022; Xiong et al., 2023). However, medical images often present unique challenges compared to natural images. Unlike natural images which are typically displayed in a fixed orientation, medical images can be viewed at various angles (Yang et al., 2020; Nguyen et al., 2021), making traditional rectangu-

lar bounding boxes less consistent across different angles (Fig. 1). In such cases, circles may offer more flexible and consistent representations of delineating objects or areas of interest. This representation allows for a more consistent and orientation-independent identification of objects, which is crucial in biomedical tissue quantification across different acquisition scenarios. Moreover, the detection method can yield less rotational variances often seen in medical imagery, ensuring that key features are accurately captured regardless of which angle the image is acquired, especially when

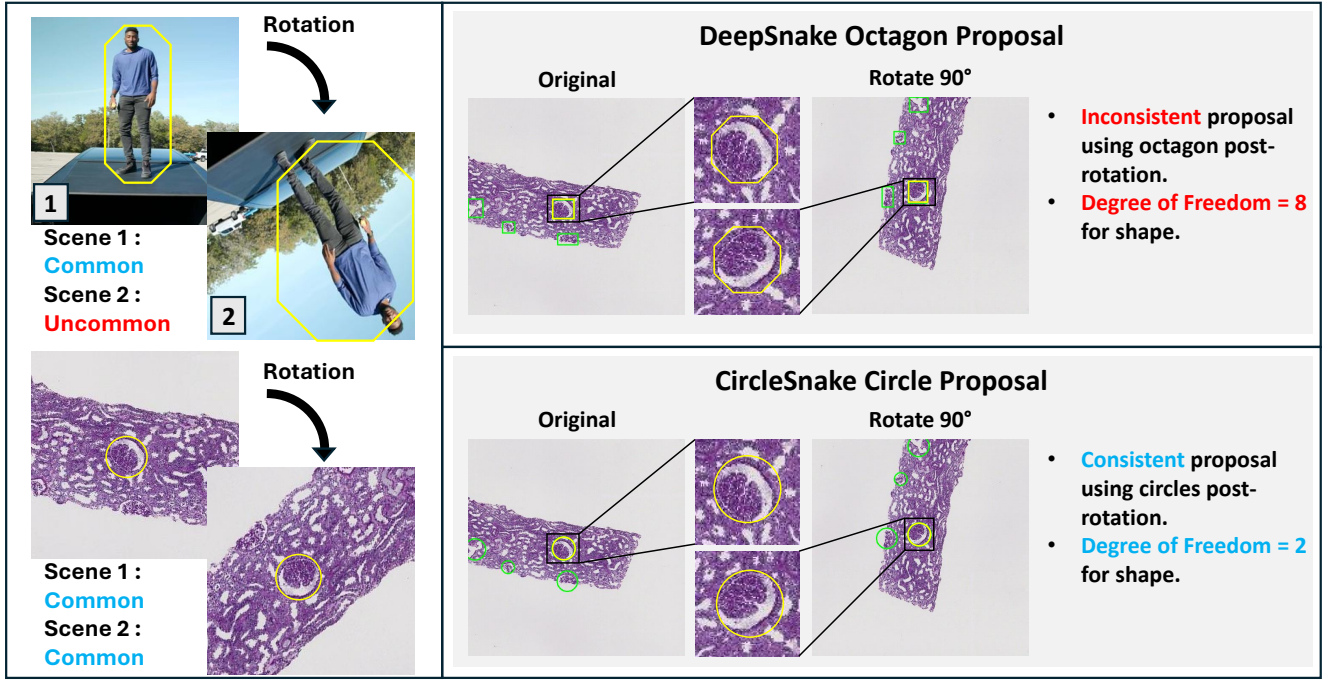


Figure 1: **Comparison of polygon representation and the proposed circle representation.** The left panel shows that samples of glomeruli can be scanned at any angle of rotation. The right panel highlights the difference between the octagon proposal and the proposed circle proposal on ball-shaped objects. The proposed CircleSnake yields a more rotation consistent representation while using fewer DoF.

detecting ball-shaped objects (Nguyen et al., 2022).

Recently, the concept of circle representation has been introduced as a representation optimized for medical imaging, particularly for ball-shaped object detection, such as glomeruli, nuclei, inflammatory cells such as eosinophils, and tumors (Nguyen et al., 2022; Yang et al., 2020; Nguyen et al., 2021; Luo et al., 2021; Liu et al., 2023; Zhang et al., 2023). The exceptional effectiveness of circle representation makes it an attractive candidate for adaptation in the field of instance object segmentation.

In this paper, we propose a simple contour-based end-to-end instance segmentation method that utilizes the circle representation, called CircleSnake, for the robust segmentation of ball-shaped medical objects. The “bounding circle” is introduced for both detection and initial contour representation on the ball-shaped objects. Once the center location of the lesion is obtained, only degrees of freedom (DoF) = 2 is required to form the bounding circle, while DoF = 8 is required for the bounding octagon. Briefly, the contribution of this study is threefold:

- **Circle Representation:** Our proposal introduces a unified circle representation pipeline for the segmentation of ball-shaped biomedical objects. This pipeline includes three integrated components: (1) circle detection, (2) circle contour proposal, and (3) circular convolution. It achieves superior segmentation performance while requiring a

reduced DoF for fitting.

- **Optimized Biomedical Object Segmentation:** To the best of our knowledge, CircleSnake, our proposed method, represents the first instance of a contour-based end-to-end segmentation approach that is optimized for ball-shaped biomedical objects.

- **Rotation Consistency:** The proposed circle representation results in less DoF of fitting, improved segmentation efficiency, and enhanced rotation consistency. As shown in (Fig. 1), tissue samples can be viewed at various angles. Consequently, improved rotational consistency could enhance the robustness in detecting identical objects within the same tissue, potentially leading to increased reproducibility in image analysis.

This study expands upon our prior conference paper (Nguyen et al., 2022) by introducing fundamental enhancements in this manuscript. (1) More Comprehensive Experiments: We conduct more rigorous and comprehensive experiments using three public and in-house datasets, featuring various ball-shaped objects such as glomeruli, nuclei, and cells. (2) Methodological Depth: The methodology is presented with greater depth, including additional benchmarks, comprehensive mathematical derivations, and an elaborate experimental design. (3) Rotation Invariance Analysis: We provide deeper analyses of rotation invariance to better evaluate performance under varying conditions.

(4) Reproducible Open-source Research: The open-source implementation has been updated to support both instance object detection and instance segmentation in a unified codebase <https://github.com/hrlblab/CircleSnake>.

2. Related Works

2.1 Instance Segmentation

Mask-based instance segmentation can be further divided into two-stage (He et al., 2018; Cai and Vasconcelos, 2021) and one-stage methods (Bolya et al., 2019; Ying et al., 2019).

Two-stage methods usually perform detection in two steps: (i) region proposal and (ii) object classification and segmentation mask regression. Mask-RCNN (He et al., 2018) establishes the foundation for two-stage, binary mask based detection systems. Mask-RCNN (He et al., 2018) consists of a region proposal network (RPN) and a Region-based CNN (R-CNN) approach (Girshick et al., 2014; Girshick, 2015) that is capable of pixel-level object instance segmentation. To achieve pixel-level precision, these binary masks classify pixels as part of the object (foreground) or not (background), simplifying the computational process and allowing for fast and accurate delineation of object boundaries. Following the introduction of Mask-RCNN, numerous algorithms have been developed to enhance its effectiveness, including different architectures (Chen et al., 2018, 2019), different backbone (Fang et al., 2023), transfer learning strategy (Zhang et al., 2020), different training strategy (Wang et al., 2021), and feature aggregation mechanisms (Liu et al., 2018). While these two-stage detectors can achieve state-of-the-art results, they are often structurally complex and need more time to infer.

One-stage methods eliminate the region proposal step and encapsulate all computations in a single network. With the introduction of YOLACT (Liu et al., 2016), these types of methods have attracted academic attention for their high computational efficiency. YOLACT employs separate prediction heads that simultaneously identify object classes, bounding boxes, and mask coefficients for each detected object. The final instance masks are dynamically assembled by combining these prototype masks with the specific mask coefficients. After YOLACT, There were many methods to improve its performance by using different feature wrapping module (Liu et al., 2021), multi-scale feature extraction (Zeng et al., 2022), and adjusting model structure (Bolya et al., 2022). Now, the performance of one-stage object detectors that use binary masks is nearly equivalent to that of two-stage detectors with anchors, yet they offer quicker inference times.

2.2 Medical Object Segmentation

2.2.1 Binary Mask-Based Methods

In the existing literature, several works have focused on pixel-based instance [segmentation](#) methodologies within regional proposals at the pixel level. Notable examples include the works of Gadermayr et al. (2017); Bueno et al. (2020); Kannan et al. (2019); Ginley et al. (2019), which tackle the challenges of feature extraction and precise pixel-level prediction through varied strategies. For instance, Gadermayr et al. (2017) refines convolutional architectures to capture intricate details, while Bueno et al. (2020); Kannan et al. (2019); Ginley et al. (2019) address segmentation challenges in specific pathologies, such as glomeruli, using specialized convolutional networks. Moreover, the Mask R-CNN framework introduced by He et al. (2018) represents a significant advancement by unifying detection and segmentation into a single model. This integration not only underscores its impact on instance segmentation but also reinforces its relevance in the broader context of medical image analysis. Despite these promising developments, the structural complexity and high computational cost of these methods continue to pose challenges for real-time applications.

2.2.2 Contour-based Methods

Contour-based methods such as DeepSnake (Peng et al., 2020) have the potential to be faster and simpler. Huang et al. (2022) adapts ideas from U-Net (Ronneberger et al., 2015) and DeepSnake (Peng et al., 2020) to right ventricular segmentation. However, this methodology exhibits suboptimal performance when applied to circular objects in pathology images, and the computational simplicity is not adequately achieved (Fig. 2).

3. Methods

The proposed CircleSnake method, shown in Fig. 3, is an instance segmentation method inspired by the DeepSnake method (Peng et al., 2020). DeepSnake is a deep learning-based method for object contour detection and instance segmentation that integrates active contour models, known as "snakes," with convolutional neural networks to accurately delineate object boundaries by iteratively refining polygonal contours initialized from bounding boxes. CircleSnake, as described, modifies this approach by replacing bounding boxes with bounding circles for the initial contour proposals, utilizing circle detection algorithms to directly generate contours that are better suited for circular objects. The primary difference between DeepSnake and CircleSnake lies in their initialization strategies: while DeepSnake uses bounding boxes to create initial polygonal contours suitable for various object shapes, CircleSnake employs bounding circles to provide a closer initial approximation for circular objects,

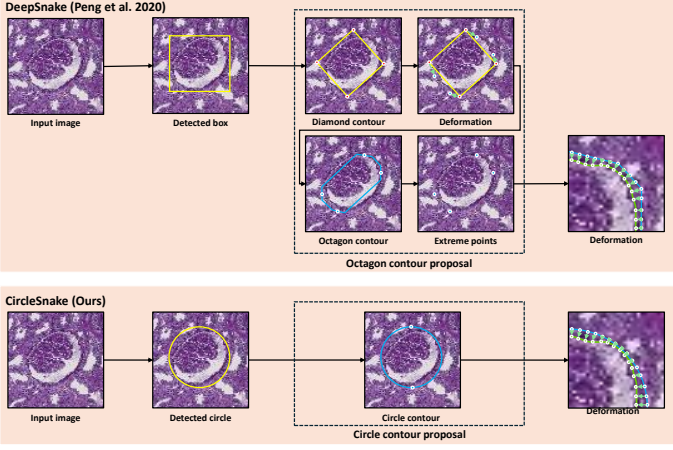


Figure 2: **Circle contour proposal.** This figure presents the differences between the “bounding box to octagon contour” representation and the proposed “bounding circle to circle contour” representation. Our circle contour proposal avoids relying on complex extreme points and deformation-based contour generation by introducing a straightforward circle proposal. This approach seamlessly connects circle detection and deformation-based segmentation in an end-to-end manner without incurring additional computational overhead. In other words, the circle detection step itself can be directly employed as a circle contour proposal.

potentially enhancing efficiency and accuracy by reducing the number of refinement iterations needed. Both methods use neural networks to adjust and refine the contours, but CircleSnake’s approach may offer advantages in applications involving predominantly circular objects by offering a better starting point for contour refinement compared to the more general-purpose initialization used in DeepSnake. CircleSnake aggregates the circle detection (Yang et al., 2020; Nguyen et al., 2021, 2022) and circle contour proposals for initial object segmentation (Fig. 1). This method involves starting with a basic circular contour, which is then refined through a process of circular convolution. This process adjusts the vertices of the contour to align more closely with the actual boundary of the object (Fig. 2). This adjustment is repeated in several iterations to accurately capture the final shape of the object for segmentation purposes. In contrast to the DeepSnake method, which depends on bounding boxes and polygonal contours, the proposed strategy is more straightforward and consistent.

3.1 Circle Object Detection

The method for detecting circular objects in this approach is based on the CircleNet framework (Nguyen et al., 2022). The $I \in \mathbb{R}^{W \times H \times 3}$ is an input image with height H and width W . The Center Point Localization network produces the center localization of each object within a heatmap $\hat{Y} \in$

$[0, 1]^{\frac{W}{R} \times \frac{H}{R} \times C}$ where C is the number of candidate classes and R is a downsampling factor. Within the heatmap, $\hat{Y}_{xyc} = 1$ is the center of the lesion and $\hat{Y}_{xyc} = 0$ is the background. All these terminologies are defined by Zhou et al. (2019). Following convention, Law and Deng (2018); Zhou et al. (2019), the target center point is splat on a heatmap as a 2D Gaussian kernel:

$$Y_{xyc} = \exp \left(-\frac{(x - \tilde{p}_x)^2 + (y - \tilde{p}_y)^2}{2\sigma_p^2} \right) \quad (1)$$

where the x and y are the center point of the ground truth, $\tilde{p}_x$ and $\tilde{p}_y$ are the downsampled ground truth center point, and σ_p is the kernel standard deviation. The training loss is L_k penalty-reduced pixel-wise logistic regression with focal loss (Lin et al., 2017):

$$L_k = \frac{-1}{N} \sum_{xyc} \begin{cases} (1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\ (1 - Y_{xyc})^\beta (\hat{Y}_{xyc})^\alpha \log(1 - \hat{Y}_{xyc}) & \text{otherwise} \end{cases} \quad (2)$$

where the hyper-parameters α and β are set 2 and 4 to kept the same as Lin et al. (2017). To further refine the prediction location, the ℓ_1 -norm offset prediction loss L_{off} is used.

To ascertain the central point, we propose identifying the top n peaks, where n is set to 100. Each peak’s value is either greater than or equal to its 8-connected neighbors. These n selected center points are denoted as $\hat{\mathcal{P}} = (\hat{x}_i, \hat{y}_i)_{i=1}^n$. For each object, its center point is represented by an integer coordinate (x_i, y_i) , which is derived from $\hat{Y}_{x_i y_i c}$ and L_k . Additionally, the offset $(\delta \hat{x}_i, \delta \hat{y}_i)$ is calculated from L_{off} . The formulation of the bounding circle incorporates both the center point $\hat{p}$ and the radius $\hat{r}$, thereby creating a comprehensive representation of the spatial domain:

$$\hat{p} = (\hat{x}_i + \delta \hat{x}_i, \hat{y}_i + \delta \hat{y}_i). \quad \hat{r} = \hat{R}_{\hat{x}_i, \hat{y}_i}. \quad (3)$$

where $\hat{R} \in \mathbb{R}^{\frac{W}{R} \times \frac{H}{R} \times 1}$ contains the radius prediction for each pixel, optimized by

$$L_{radius} = \frac{1}{N} \sum_{k=1}^N |\hat{R}_{p_k} - r_k|. \quad (4)$$

where r_k is the ground truth radius for each object k . Finally, the overall objective is

$$L_{det} = L_k + \lambda_{radius} L_{radius} + \lambda_{off} L_{off}. \quad (5)$$

Following Zhou et al. (2019), we set $\lambda_{radius} = 0.1$ and $\lambda_{off} = 1$.

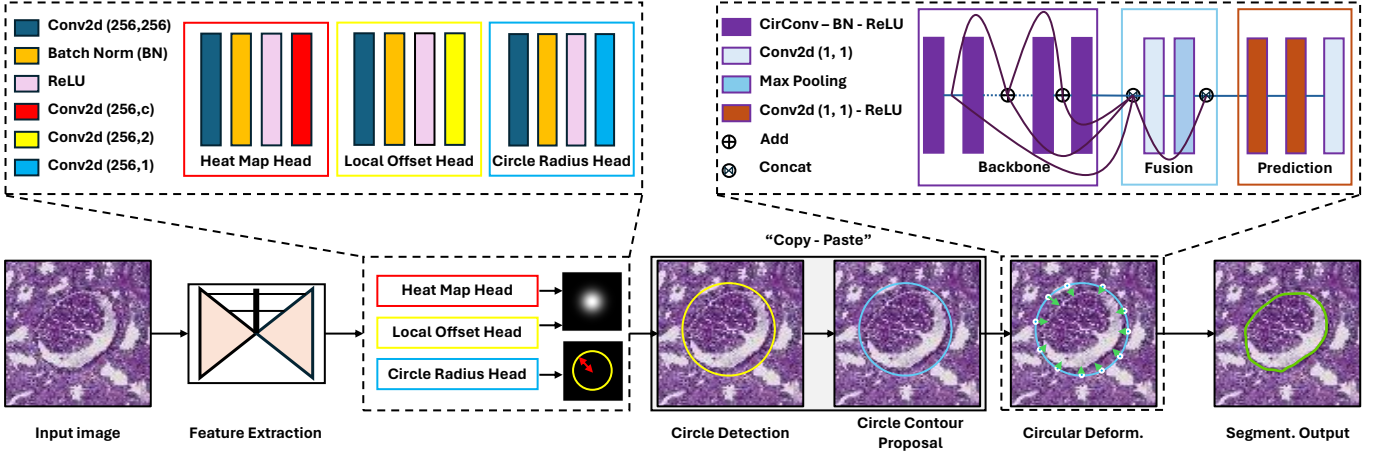


Figure 3: **Overview of CircleSnake.** A backbone network functions as a feature extractor for the subsequent three head networks. The head networks for heatmap and local offset identify the circle’s center point, whereas the head network for circle radius calculates the circle’s radius to achieve bounding circle detection. This bounding circle establishes the initial contour. Subsequently, this contour deforms to the final contour through the use of a circular convolutional network

3.2 Circle Contour Proposal and Deformation

Bounding circle-based detection for each target object is accomplished using CircleNet-based object detection, as described in Nguyen et al. (2021). Directly derived from the circle representation shown in Fig. 2, the initial circle contour proposal marks a departure from the previously complex deformation process and the use of extreme point-based octagon contour proposals. This octagon is formed by four extreme points in DeepSnake: the topmost, leftmost, bottommost, and rightmost pixels of an object, respectively. This innovation simplifies and standardizes the contour proposal approach. Defined by its radius and center point, the circle proposal is established by uniformly sampling N initial points

$\mathbf{x}_i^{circle} | i = 1, 2, \dots, N$ from the circle contour, beginning at the top-most point $\mathbf{x}_1^{circle}$. Correspondingly, the ground truth contour is created by clockwise sampling of N vertices along the boundary of the object. Adhering to the guidelines in Peng et al. (2020), the value of N is fixed at 128, facilitating a consistent and streamlined process.

For a contour with N vertices denoted as $\mathbf{x}_i^{circle} | i = 1, \dots, N$, we initiate by creating feature vectors for each vertex. The feature maps are extracted by applying a CNN backbone to the input image. This CNN backbone is shared with the detector in our instance segmentation pipeline. The image feature is computed using bilinear interpolation at the vertex coordinate. The feature f_i^{circle} corresponding to a vertex $\mathbf{x}_i^{circle}$ comprises a combination of the learning-based features and the vertex’s coordinates, represented as $[F(\mathbf{x}_i^{circle}); \mathbf{x}_i^{circle}]$. F signifies the feature maps. These input features are then treated as a one-dimensional discrete

signal $f : \mathbb{Z} \rightarrow \mathbb{R}^D$ on the contour of the circle. In line with the approach detailed in Peng et al. (2020), circular convolution is employed for the learning of features. The circular convolution is defined as:

$$(f_N^{circle} * k)_i = \sum_{j=-r}^r (f_N^{circle})_{i+j} k_j, \quad (6)$$

where $k : [-r, r] \rightarrow \mathbb{R}^D$ is a learnable kernel function, while the operator $*$ is the standard convolution. Following Peng et al. (2020), the kernel size of the circular convolution is fixed to be nine.

The CircleSnake model adopts the architecture of DeepSnake (Peng et al., 2020) for its convolutional implementation, structured into three main components: backbone, fusion, and prediction. The backbone consists of eight layers, each composed of "Conv-Bn-ReLU" sequences. This segment incorporates residual skip connections and utilizes standard convolutional layers for feature extraction. The fusion block is designed to integrate information across contour points at multiple scales. In this process, features from all layers in the backbone are combined and passed through a 1×1 convolutional layer followed by a max pooling layer. The prediction head, serving as the final component of the network, is equipped with three 1×1 convolutional layers. Its primary function is to output vertex-wise offsets. Regarding the loss function, it is specifically designed to accommodate the iterative contour deformation process within the CircleSnake framework:

$$L_{iter} = \frac{1}{N} \sum_{i=1}^N l_1(\tilde{x}_i^{circle} - x_i^{gt}). \quad (7)$$

where x_i^{gt} is the ground truth boundary point and $\tilde{x}_i^{circle}$ is the deformed contour point. Following Peng et al. (2020), we regress the N offsets in 3 iterations.

4. Experimental Design

4.1 Data

4.1.1 Glomeruli Dataset

To construct a dataset for glomeruli analysis, renal biopsy whole slide images (WSIs) were obtained and annotated. Kidney tissues underwent routine processing, embedding in paraffin, and were sectioned to a thickness of $3\text{ }\mu\text{m}$, followed by staining with hematoxylin and eosin (HE), periodic acid–Schiff (PAS), or Jones. The samples were anonymized, with studies receiving Institutional Review Board (IRB) approval by Vanderbilt University Medical Center (VUMC) under protocol number 202459 on December 28, 2020. The dataset comprised 704 glomeruli from 42 biopsies for training, 98 from 7 for validation, and 147 from 7 for testing. Given the relative size of a glomerulus (Puelles et al., 2011), the original high-resolution images ($0.25\text{ }\mu\text{m}/\text{pixel}$) were downsampled to $4\text{ }\mu\text{m}/\text{pixel}$. Image patches of 512×512 pixels containing at least one glomerulus were randomly sampled, resulting in a comprehensive dataset with 7,040 training, 980 validation, and 1,470 testing images.

4.1.2 Nuclei Dataset

The 2018 Multi-Organ Nuclei Segmentation challenge dataset includes 30 tissue images, each measuring 1000×1000 pixels, and contains 21,623 manually annotated nuclear boundaries. These images were extracted from various organs sourced from *The Cancer Genomic Atlas* (TCGA) using H&E-stained WSIs at $40\times$ magnification. Additionally, a separate testing dataset with 14 images of the same dimensions was created, including lung and brain tissues exclusively for testing. To create training and validation datasets, we randomly selected 10 patches sized at 512×512 pixels from each of the 30 training/validation images, resulting in a total of 300 images. Similarly, 140 images were obtained from the 14 testing images. This dataset comprised 200 training, 100 validation, and 140 testing images. It's worth noting that the 2018 Multi-Organ Nuclei Segmentation challenge dataset contains more objects (21,623 nuclei) than the glomeruli dataset (802 glomeruli) before any data augmentation. The 2018 Multi-Organ Nuclei Segmentation challenge dataset is publicly accessible for research purposes.

4.1.3 Eosinophils Dataset

All image patches were cut from 50 WSIs from the Pediatric Eosinophils dataset with a $40\times$ objective lens (Liu

et al., 2023). The Eosinophils dataset contains over 12,000 annotations in 50 annotated WSIs. All human subjects research got IRB approval by VUMC under protocol number 230223 on March 2, 2023. To enhance the model's discernment of eosinophils characteristics, we established four distinct classes which include eosinophils (eos), papilla eos, red blood cells (RBC), and RBC clusters. To evaluate the performance of models, the dataset was divided into three subsets: train, val (validation), and test, with a split ratio of 7:1:2. Specifically, there are 4,842 annotations for "eos", 2,789 annotations for "papillae eos", 426 annotations for "RBC", and 524 annotations for "RBC Cluster" in the training dataset. 690 annotations for "eos", 430 annotations for "papillae eos", 40 annotations for "RBC", and 12 annotations for "RBC Cluster" are in the validation dataset. 1,389 annotations for "eos", 813 annotations for "papillae eos", 100 annotations for "RBC", and 151 annotations for "RBC Cluster" are in the validation dataset.

4.2 Experimental Design

The implementation of CircleSnake's object instance segmentation and backbone networks followed DeepSnake's official PyTorch implementations. The Common Objects in Context (COCO) pre-trained model (Lin et al., 2014) was used to initialize all models. All experiments were conducted on the same workstation with a 24 GB Nvidia RTX A5000. For both the glomeruli and Eos experiments, the hyperparameters were set to maximum epoch = 50, learning rate = $5e-4$, batch size = 16, and optimizer = Adam. For the nuclei experiment, the batch size was set to 4 due to memory constraints.

As baseline methods, Faster-RCNN (Ren et al., 2015), Mask-RCNN (He et al., 2018), CenterNet (Zhou et al., 2019), DeepSnake (Peng et al., 2020), were chosen for their superior object detection and object instance segmentation performance. ResNet-50 (He et al., 2016) and deep layer aggregation (DLA) network (Yu et al., 2018) were used as backbone networks for these different methods. For CircleSnake, we followed the original implementation (Nguyen et al., 2021) and used DLA for the backbone networks.

To ensure the experiment's fairness, no data augmentation was performed on any model during training, and no hyperparameter optimization was performed for any specific model. During validation, all models were tested with the best epoch.

4.3 Evaluation Metrics

In the domain of computer vision, the COCO evaluation metrics have gained prominence as a means to assess the performance of object detection and segmentation algorithms. These metrics were developed as an integral component of the COCO dataset and challenge, offering a standardized

framework for quantifying the accuracy and efficacy of such algorithms. In the evaluation process, we employed several key metrics, including Average Precision (AP), AP_{50} (indicating Intersection over Union, or IOU, threshold at 0.5), AP_{75} (IOU threshold at 0.75), AP_S (pertaining to objects at a small scale with area less than 32^2), and AP_M (associated with objects at a medium scale with area greater than 32^2 and less than 96^2). In the context of segmentation, the Dice score was also utilized as part of the evaluation criteria. During the evaluation process, all evaluations were conducted using consistent threshold settings.

4.4 Rotation Consistency Score

To rigorously evaluate the resilience and precision of our model across varying angles, we have adopted the Rotation Consistency Score as a key metric (Nguyen et al., 2021). The rotation consistency score is calculated by using average Dice score (before and after rotation), where 1 means all boxes/circles overlapped while 0 means no boxes/circles overlapped. A higher rotation consistency score indicates that the predicted rotated test image, when returned to its original orientation, aligns more accurately with the unrotated test image. This results in a greater overlap during visualization, demonstrating better consistency in the model's predictions. This decision is rooted in the necessity to rotate the original test images by 90 degrees, a choice specifically made to mitigate the effects of intensity interpolation that might arise with arbitrary angles. By employing this structured rotation, we can effectively transform the segmentation masks from the rotated images back to their original orientation. This allows for direct comparison and calculation of the Rotation Consistency Score against the manually segmented results. We conducted two phases of experiments, where the first rotation process was only implemented during the test phase and was intentionally not used as a data augmentation strategy during the training of all methods. The second rotation process was also implemented during the test phase, but 50% of the data was rotated during the training of all methods. This approach ensures a focused assessment of the model's performance in handling rotational variations, providing a clear insight into its robustness and accuracy under controlled conditions.

5. Results

5.1 Glomeruli Detection and Segmentation Performance

As seen in Table 1, the proposed CircleSnake method using the DLA as a backbone achieves 0.623 segmentation AP, 0.894 segmentation AP_{50} , 0.762 AP_{75} , 0.543 detection AP_S , 0.719 segmentation AP_M . A qualitative comparison between DeepSnake and CircleSnake can be seen in (Fig. 4) part (a). The arrows are displayed only in areas

where the predictions differ from the manual segmentation. Yellow arrows indicate discrepancies in predictions without rotation compared to the manual segmentation, while blue arrows highlight discrepancies in predictions after a 90-degree rotation compared to the manual segmentation.

5.2 Nuclei Detection and Segmentation Performance

The dataset from the 2018 Multi-Organ Nuclei Segmentation Challenge (Kumar et al., 2017, 2019) was also applied in the experiment. As seen in Table 2, the proposed CircleSnake-DLA method reach 0.485 detection AP, 0.845 detection AP_{50} , 0.518 AP_{75} , 0.495 detection AP_S , 0.418 segmentation AP_M . A qualitative comparison between Mask-RCNN, DeepSnake, and CircleSnake can be seen in Fig. 4 section (b).

5.3 Eosinophils Detection and Segmentation Performance

In pursuit of confirming the model's broad applicability and generalizability, we decided to introduce the Eosinophils dataset (Liu et al., 2023) into our study. As indicated in Table 3, the performance of the CircleSnake achieve 0.344 detection AP, 0.727 detection AP_{50} , 0.298 AP_{75} , 0.340 detection AP_S , 0.403 segmentation AP_M . A qualitative comparison between Mask-RCNN, DeepSnake, and CircleSnake can be seen in Fig. 4 part (c).

5.4 Semantic Segmentation Results and Rotation Consistency

For the semantic segmentation results and rotation consistency comparison. As shown in Table 4 and Fig. 5 part (a). Our method, CircleSnake can achieve a Dice score of 0.828 with 0.796 rotation consistency score and the Dice score result has a significant difference from the other two methods in the Glomeruli dataset. For the Nuclei dataset Table 4 and Fig. 5 part (b), CircleSnake achieves 0.8 for the Dice score which also has a significant difference from the other two methods and 0.799 in the rotation consistency score. In the Eosinophils dataset Table 4 and Fig. 5 part (c), CircleSnake achieved 0.743 for the Dice score which has a significant difference from the DeepSnake method and 0.752 in the rotation consistency score.

6. Ablation Study

As seen in Table 6, the StarDist (Weigert and Schmidt, 2022) U-Net-based segmentation method, which is a widely used approach designed for instance segmentation of star-convex objects such as nuclei, achieves a Dice Score of 0.618 and a Rotation Consistency Score of 0.621 on the Nuclei dataset. StarDist uses a U-Net backbone to predict distances to object boundaries and classifies object centers,

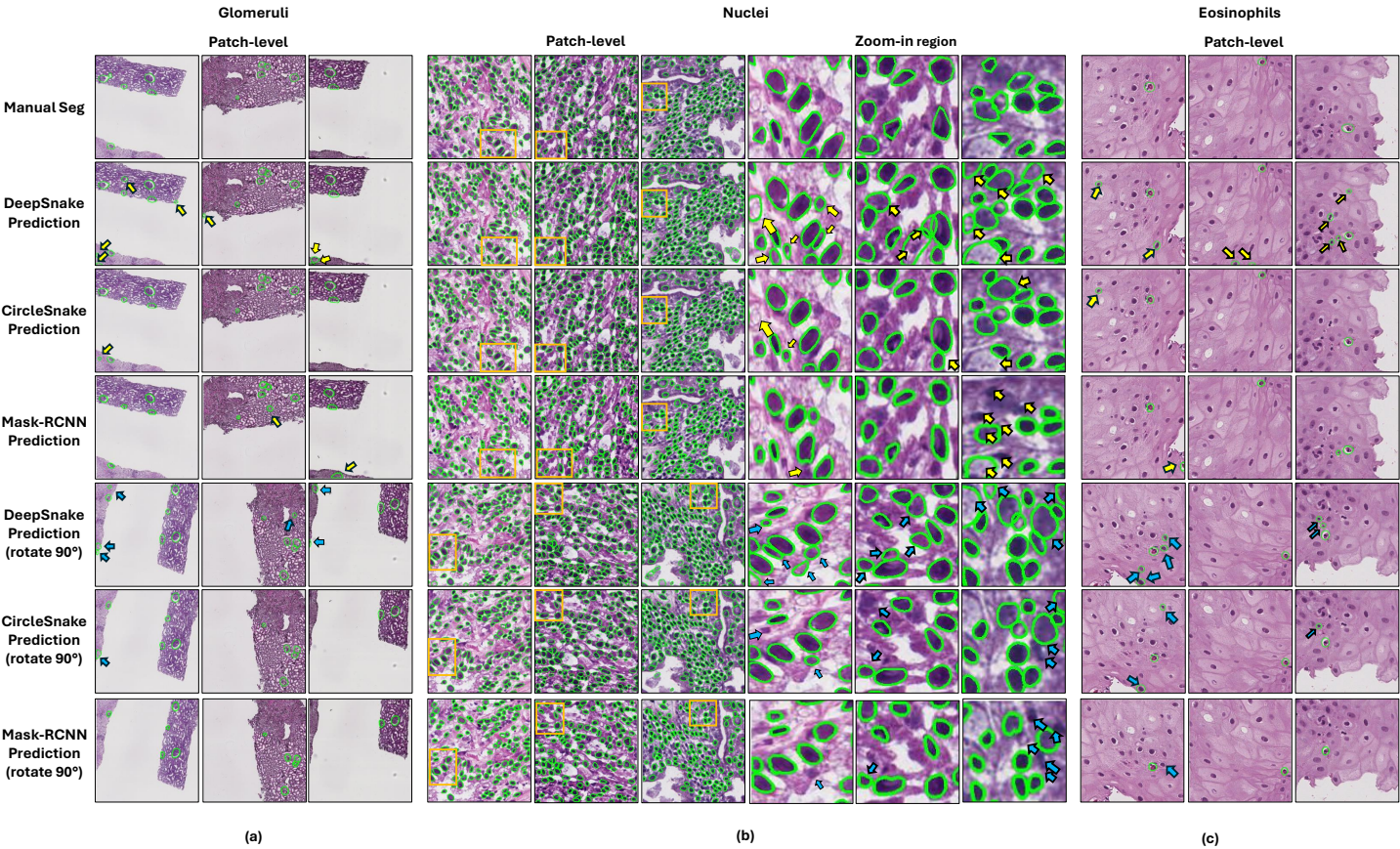


Figure 4: **Qualitative Comparison.** This figure shows the Qualitative comparison between deepsnake segmentation results and circlesnake segmentation results. Each orange box indicates the location of each selection. Each yellow arrow shows the differences between the manual segmentation results and prediction results. Each blue arrow shows the differences only between the manual segmentation results and the results in rotated 90-degree prediction. (a) is for the Glomeruli dataset, (b) is for the Nuclei dataset, and (c) is for the Eosinophils dataset.

Table 1: Detection and Instance Segmentation Performance on Glomeruli.

Methods	Backbone	AP	$AP_{(50)}$	$AP_{(75)}$	$AP_{(S)}$	$AP_{(M)}$
CenterNet (Law and Deng, 2018)	DLA-34	0.547	0.872	0.643	0.432	0.640
CircleNet (Nguyen et al., 2022)	DLA-34	0.570	0.857	0.691	0.466	0.647
Faster-RCNN (Ren et al., 2015)	ResNet-50	0.578	0.878	0.691	0.466	0.670
Mask-RCNN (Detection) (He et al., 2018)	ResNet-50	0.576	0.877	0.675	0.470	0.660
Mask-RCNN (Segmentation) (He et al., 2018)	ResNet-50	0.590	0.878	0.705	0.453	0.685
DeepSnake (Detection) (Peng et al., 2020)	DLA-34	0.527	0.881	0.609	0.424	0.610
DeepSnake (Segmentation) (Peng et al., 2020)	DLA-34	0.548	0.877	0.669	0.411	0.642
CircleSnake (Detection) (Ours)	DLA-34	0.603	0.877	0.734	0.534	0.670
CircleSnake (Segmentation) (Ours)	DLA-34	0.623	0.894	0.762	0.488	0.719

Table 2: Detection and Instance Segmentation Performance on Nuclei.

Methods	Backbone	AP	$AP_{(50)}$	$AP_{(75)}$	$AP_{(S)}$	$AP_{(M)}$
CenterNet (Law and Deng, 2018)	DLA-34	0.404	0.818	0.342	0.440	0.290
CircleNet (Nguyen et al., 2022)	DLA-34	0.461	0.843	0.462	0.473	0.332
Faster-RCNN (Ren et al., 2015)	ResNet-50	0.364	0.679	0.349	0.365	0.336
Mask-RCNN (Detection) (He et al., 2018)	ResNet-50	0.369	0.679	0.359	0.367	0.348
Mask-RCNN (Segmentation) (He et al., 2018)	ResNet-50	0.366	0.678	0.369	0.365	0.446
DeepSnake (Detection) (Peng et al., 2020)	DLA-34	0.401	0.827	0.315	0.408	0.345
DeepSnake (Segmentation) (Peng et al., 2020)	DLA-34	0.425	0.827	0.394	0.425	0.454
CircleSnake (Detection) (Ours)	DLA-34	0.485	0.845	0.518	0.495	0.217
CircleSnake (Segmentation) (Ours)	DLA-34	0.436	0.837	0.419	0.436	0.418

Table 3: Detection and Instance Segmentation Performance on Eosinophils.

Methods	Backbone	AP	$AP_{(50)}$	$AP_{(75)}$	$AP_{(S)}$	$AP_{(M)}$
CenterNet (Law and Deng, 2018)	DLA-34	0.327	0.729	0.230	0.351	0.318
CircleNet (Nguyen et al., 2022)	DLA-34	0.340	0.706	0.292	0.368	0.332
Faster-RCNN (Ren et al., 2015)	ResNet-50	0.331	0.696	0.237	0.332	0.343
Mask-RCNN (Detection) (He et al., 2018)	ResNet-50	0.335	0.700	0.246	0.336	0.354
Mask-RCNN (Segmentation) (He et al., 2018)	ResNet-50	0.330	0.690	0.243	0.282	0.391
DeepSnake (Detection) (Peng et al., 2020)	DLA-34	0.300	0.720	0.149	0.301	0.323
DeepSnake (Segmentation) (Peng et al., 2020)	DLA-34	0.305	0.712	0.169	0.255	0.360
CircleSnake (Detection) (Ours)	DLA-34	0.344	0.727	0.298	0.340	0.358
CircleSnake (Segmentation) (Ours)	DLA-34	0.332	0.723	0.263	0.282	0.403

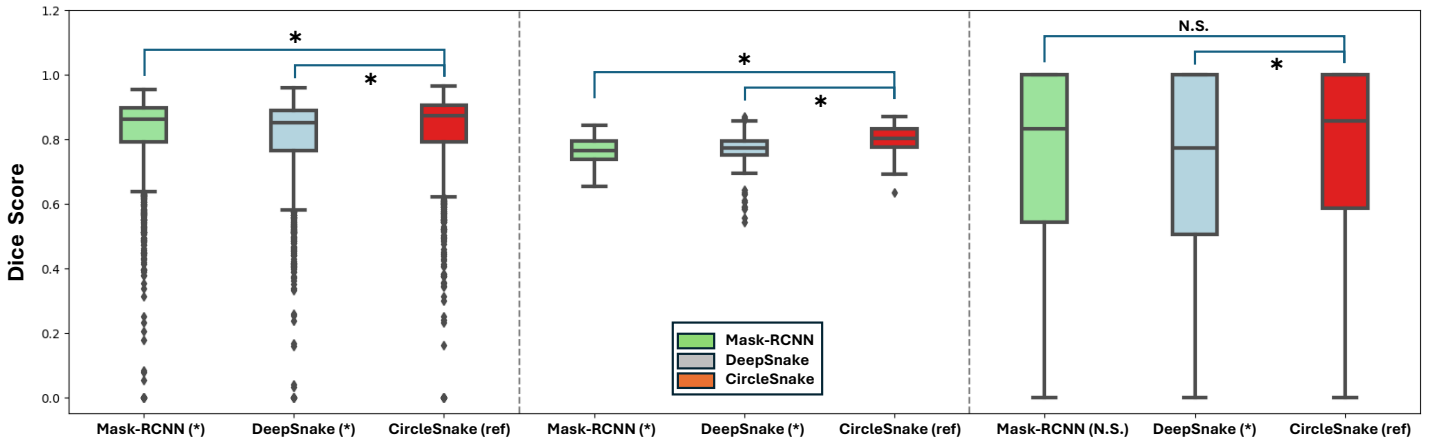


Figure 5: **Statistical Analysis for Dice Score** The figure shows the boxplots of dice score for Mask-RCNN, DeepSnake, and CircleSnake on Glomeruli, Nuclei, and Eosinophils test datasets. The Wilcoxon signed-rank test is performed with CircleSnake as the reference (“re”) method, to compare with other methods. “*” represents the significant ($p < 0.05$) differences, while “N.S.” means the difference is not significant. (a) is for the Glomeruli dataset, (b) is for the Nuclei dataset, and (c) is for the Eosinophils dataset.

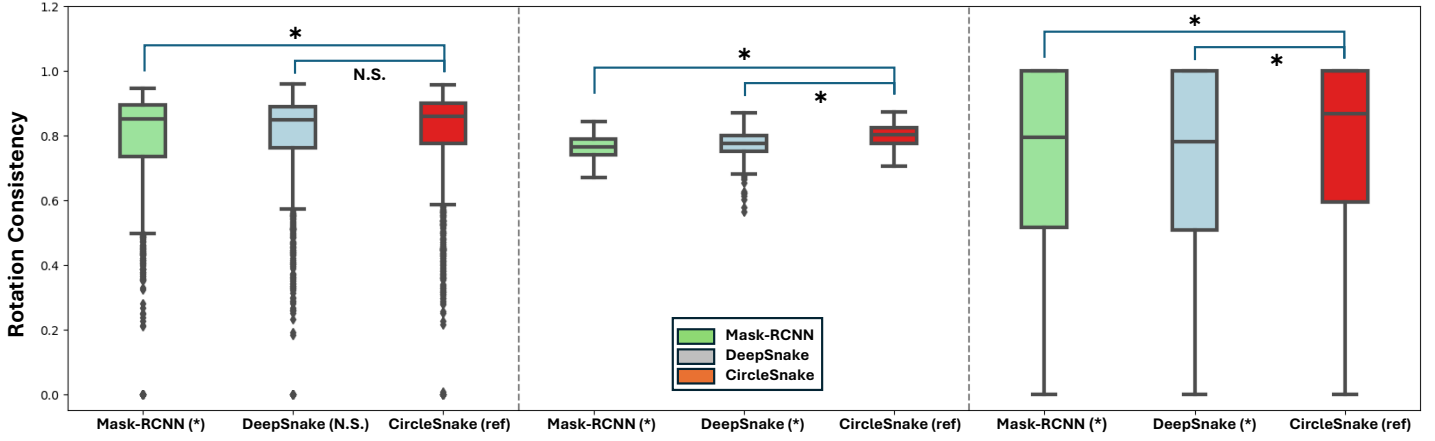


Figure 6: **Statistic Analysis for Rotation Consistency Score** The figure shows the boxplots of rotation consistency score for Mask-RCNN, DeepSnake, and CircleSnake on Glomeruli, Nuclei, and Eosinophils test datasets. The Wilcoxon signed-rank test is performed with CircleSnake as the reference (“re”) method, to compare with other methods. “*” represents the significant ($p < 0.05$) differences, while “N.S.” means the difference is not significant. (a) is for the Glomeruli dataset, (b) is for the Nuclei dataset, and (c) is for the Eosinophils dataset.

Table 4: Semantic Segmentation Results and Rotation Consistency.

Dataset	Methods	Backbone	Dice	Rotation Consistency
Glomeruli	Mask-RCNN (He et al., 2018)	ResNet	0.813	0.779
Glomeruli	DeepSnake-DLA (Zhou et al., 2019)	DLA	0.801	0.793
Glomeruli	CircleSnake (Ours)	DLA	0.828	0.796
Nuclei	Mask-RCNN (He et al., 2018)	ResNet	0.763	0.763
Nuclei	DeepSnake-DLA (Zhou et al., 2019)	DLA	0.765	0.767
Nuclei	StarDist (Weigert and Schmidt, 2022)	U-Net	0.618	0.621
Nuclei	CircleSnake (Ours)	DLA	0.800	0.799
Eosinophils	Mask-RCNN (He et al., 2018)	ResNet	0.710	0.691
Eosinophils	DeepSnake-DLA (Zhou et al., 2019)	DLA	0.684	0.690
Eosinophils	CircleSnake (Ours)	DLA	0.743	0.752

Table 5: Semantic Segmentation Results and Rotation Consistency. (training with rotation augmentation)

Dataset	Methods	Backbone	Dice	Rotation Consistency
Glomeruli	Mask-RCNN (He et al., 2018)	ResNet	0.829	0.839
Glomeruli	DeepSnake-DLA (Zhou et al., 2019)	DLA	0.819	0.830
Glomeruli	CircleSnake (Ours)	DLA	0.834	0.847
Nuclei	Mask-RCNN (He et al., 2018)	ResNet	0.768	0.763
Nuclei	DeepSnake-DLA (Zhou et al., 2019)	DLA	0.797	0.787
Nuclei	CircleSnake (Ours)	DLA	0.804	0.805
Eosinophils	Mask-RCNN (He et al., 2018)	ResNet	0.747	0.746
Eosinophils	DeepSnake-DLA (Zhou et al., 2019)	DLA	0.732	0.740
Eosinophils	CircleSnake (Ours)	DLA	0.751	0.750

making it particularly effective for segmenting objects with distinct shapes. In comparison, our method, CircleSnake, outperforms StarDist on both metrics, demonstrating superior segmentation accuracy and robustness to rotational variations. This highlights the effectiveness of our approach in handling challenging datasets with high object density and complex object boundaries.

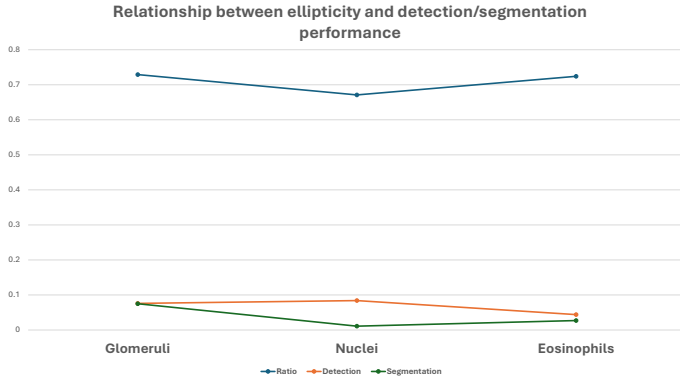


Figure 7: **Relationship between ellipticity and detection/segmentation performance.** The blue line shows the average ellipticity ratio for all objects on three different datasets. The orange line shows the detection performance on three different datasets. The green line shows the segmentation performance on three different datasets.

7. Discussion

In this study, we propose a contour-based method, CircleSnake, optimized for the segmentation of biomedical ball-shaped objects. Instead of using an octagon as the initial contour, CircleSnake directly uses the detection circle as the initial contour which has a simpler structure and is shown to offer superior detection or segmentation performance and rotation consistency. Specifically, CircleSnake outperforms the other two baseline methods in all evaluation metrics in the Glomeruli dataset in Table 1. It also has better performance except AP_M in the Nuclei dataset in Table 2 and also better results in AP , AP_{75} , and AP_M in Table 3. From Table 3, it can be observed that all models show a noticeable drop in performance on the eosinophil dataset. We believe this is primarily due to the less distinct characteristics of eosinophil cells. As shown in Figure 4, glomeruli exhibit the most prominent features, with well-defined boundaries that clearly separate the object from the background. Similarly, nuclei are relatively easier to distinguish due to their strong color contrast with the background. In contrast, eosinophil cells are characterized by intracellular inflammation, often appearing partially red, which makes their features more ambiguous and challenging for all models to detect effectively. This highlights the

difficulty of accurately segmenting objects with subtle or less distinct features.

As shown in Tables 4 and 5, adopting a circular deformation for the segmentation contour leads to better rotation consistency, regardless of whether rotation data augmentation is applied. This improved performance likely stems from the fact that while length and width measurements are sensitive to orientation, radii are inherently more rotation-invariant. To confirm this, we fit the segmentation annotations to ellipses and compute the mean ratio of their minor to major axes (where a ratio closer to 1 indicates a more rounded shape). We then examine how object roundness influences the performance gap in detection and segmentation between CircleSnake and DeepSnake. Our results demonstrate that while CircleSnake’s segmentation performance improves as objects become more circular, no similar trend is observed for detection performance, as illustrated in Figure 7.

Turning to the comparisons in Tables 1, 2, and 3, we observe that the Nuclei dataset—despite having the highest object density—does not yield the poorest performance. This finding suggests that object density alone does not dictate model success. Instead, COCO AP’s density-aware evaluation ensures that models capable of handling dense object distributions, such as those found in the Nuclei dataset, are fairly assessed for their ability to separate, identify, and accurately segment individual instances. Through this analysis, several factors contributing to errors emerge. For false positives, complex or irregular backgrounds may mimic object features, and overlapping objects in dense scenes can cause over-segmentation of non-existent instances. For false negatives, faint objects or those that are low-contrast relative to the background may be missed, while extreme size variation—particularly very small objects—increases the likelihood of detection failures.

The utilization of a circular contour as the foundational shape for deformation processes may not be ideally suited for objects of varying geometrical forms, such as ellipses or elongated shapes. This assertion is substantiated by the data presented in Table 2 and Table 3, where a discernible discrepancy is observed between the AP of object detection and that of segmentation. The higher AP in detection compared to segmentation might be attributable to the inherent limitations of a circular contour in accommodating the deformation needs of non-circular objects, necessitating a more extensive modification to fit diverse shapes. Furthermore, the current methodology, being predicated on a single-scale feature map, exhibits a marginal reduction in efficacy when dealing with objects of disparate scales, as evidenced by the results in Table 3. Consequently, it is posited that transitioning from a single-scale to a multi-scale feature map could potentially enhance the overall performance of the model, providing a more nuanced and

Table 6: Semantic Segmentation Results and Rotation Consistency for StarDist.

Dataset	Methods	Backbone	Dice	Rotation Consistency
Nuclei	StarDist (Weigert and Schmidt, 2022)	U-Net	0.618	0.621

adaptable approach to object detection and segmentation.

For the dataset utilization, these three datasets employed in this work are all within the domain of digital pathology, leveraging whole slide imaging. In the future, it would be beneficial to extend the proposed circle representation-based segmentation approach to other fields, such as lung nodule detection (CT slices), liver tumor detection (CT slices), cyst detection (MRI scans), and lesion detection in fundus images. Notably, the circle representation has already been applied in other imaging modalities for object detection, including aerial images (Huang et al., 2023) and radiological MR images (Camarasa et al., 2023), demonstrating its potential versatility across diverse applications.

8. Conclusion

In this paper, we propose CircleSnake, a simple contour-based end-to-end deep learning algorithm with simple architecture, that achieves (1) circle detection, (2) circle contour proposal, and (3) circular convolution. The CircleSnake method is optimized for ball-shaped biomedical objects, offering superior glomeruli, nuclei, and eosinophils instance segmentation performance and rotation consistency. The experimental results demonstrate that, in contrast to conventional bounding box and octagonal representations, the reduction of DoF has no adverse impact on the model's accuracy. Furthermore, it effectively sustains detection efficiency across diverse viewing angles.

Acknowledgments

This research was supported by NIH R01DK135597(Huo), DoD HT9425-23-1-0003(HCY), NIH NIDDK DK056942-(ABF). This work was also supported by Vanderbilt Seed Success Grant, Vanderbilt Discovery Grant, and VISE Seed Grant. This project was supported by The Leona M. and Harry B. Helmsley Charitable Trust grants G-1903-03793 and G-2103-05128. This research was also supported by NIH grants R01EB033385, R01DK132338, REB017230, R01MH125931, and NSF 2040462. We extend gratitude to NVIDIA for their support by means of the NVIDIA hardware grant.

Ethical Standards

The work follows appropriate ethical standards in conducting research and writing the manuscript, following all applicable laws and regulations regarding treatment of animals or human subjects. Specifically, it received IRB approval from Vanderbilt University Medical Center.

Conflicts of Interest

We declare we don't have conflicts of interest.

The 2018 Multi-Organ Nuclei Segmentation challenge dataset can be found at <https://monuseg.grand-challenge.org/Data/>. For the Glomeruli dataset and Eosinophils dataset, both are proprietary datasets developed in-house. We currently are unable to share the data due to stringent privacy and confidentiality agreements safeguarding the personal and sensitive information of the individuals involved.

References

- Daniel Bolya, Chong Zhou, Fanyi Xiao, and Yong Jae Lee. Yolact: Real-time instance segmentation, 2019.
- Daniel Bolya, Chong Zhou, Fanyi Xiao, and Yong Jae Lee. Yolact++ better real-time instance segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(2):1108–1121, February 2022. ISSN 1939-3539. .
- Gloria Bueno, M Milagro Fernandez-Carrobles, Lucia Gonzalez-Lopez, and Oscar Deniz. Glomerulosclerosis identification in whole slide images using semantic segmentation. *Computer methods and programs in biomedicine*, 184:105273, 2020.
- Zhaowei Cai and Nuno Vasconcelos. Cascade r-cnn: High quality object detection and instance segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(5):1483–1498, 2021. .
- Robin Camarasa, Hoel Kervadec, M. Eline Kooi, Jeroen Hendrikse, Paul J. Nederkoorn, Daniel Bos, and Marleen de Bruijne. Nested star-shaped objects segmentation using diameter annotations. *Medical Image Analysis*, 90: 102934, 2023. ISSN 1361-8415. .
- Kai Chen, Jiangmiao Pang, Jiaqi Wang, Yu Xiong, Xiaoxiao Li, Shuyang Sun, Wansen Feng, Ziwei Liu, Jianping Shi,

- Wanli Ouyang, Chen Change Loy, and Dahua Lin. Hybrid task cascade for instance segmentation, 2019.
- Liang-Chieh Chen, Alexander Hermans, George Papandreou, Florian Schroff, Peng Wang, and Hartwig Adam. Masklab: Instance segmentation by refining object detection with semantic and direction features. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4013–4022, 2018. .
- Shuqi Fang, Bin Zhang, and Jingyu Hu. Improved mask r-cnn multi-target detection and segmentation for autonomous driving in complex scenes. *Sensors*, 23(8), 2023. ISSN 1424-8220.
- Michael Gadermayr, Ann-Kathrin Dombrowski, Barbara Mara Klinkhammer, Peter Boor, and Dorit Merhof. Cnn cascades for segmenting whole slide images of the kidney. *arXiv preprint arXiv:1708.00251*, 2017.
- Brandon Ginley, Brendon Lutnick, Kuang-Yu Jen, Agnes B Fogo, Sanjay Jain, Avi Rosenberg, Vighnesh Walavalkar, Gregory Wilding, John E Tomaszewski, Rabi Yacoub, et al. Computational segmentation and classification of diabetic glomerulosclerosis. *Journal of the American Society of Nephrology*, 30(10):1953–1967, 2019.
- Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
- Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn, 2018.
- Kai Huang, Lei Xu, Yingliang Zhu, and Penghui Meng. A u-snake based deep learning network for right ventricle segmentation. *Medical physics*, 2022.
- Renjie Huang, Yangguang Shi, Yuting He, Yongqiang Zheng, Guoqiang Xiao, and Ziruo Liu. Semantic circle detection and circle-inner segmentation for tree-wise citrus summer shoot management in aerial images. In *2023 IEEE International Conference on Image Processing (ICIP)*, pages 1090–1094. IEEE, 2023.
- Shruti Kannan, Laura A Morgan, Benjamin Liang, McKenzie G Cheung, Christopher Q Lin, Dan Mun, Ralph G Nader, Mostafa E Belghasem, Joel M Henderson, Jean M Francis, et al. Segmentation of glomeruli within trichrome images using deep learning. *Kidney international reports*, 4(7):955–962, 2019.
- Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE transactions on medical imaging*, 36(7):1550–1560, 2017.
- Neeraj Kumar, Ruchika Verma, Deepak Anand, Yanning Zhou, Omer Fahri Onder, Efstratios Tsougenis, Hao Chen, Pheng-Ann Heng, Jiahui Li, Zhiqiang Hu, et al. A multi-organ nucleus segmentation challenge. *IEEE transactions on medical imaging*, 39(5):1380–1391, 2019.
- Hei Law and Jia Deng. Cornernet: Detecting objects as paired keypoints. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 734–750, 2018.
- Hao Liang, Zhiming Cheng, Haiqin Zhong, Aiping Qu, and Lingna Chen. A region-based convolutional network for nuclei detection and segmentation in microscopy images. *Biomedical Signal Processing and Control*, 71:103276, 2022. ISSN 1746-8094. .
- Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.
- Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.
- Haotian Liu, Rafael A. Rivera Soto, Fanyi Xiao, and Yong Jae Lee. Yolactedge: Real-time instance segmentation on the edge, 2021.
- Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia. Path aggregation network for instance segmentation, 2018.
- Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.
- Yilin Liu, Ruining Deng, Juming Xiong, Regina N Tyree, Hernan Correa, Girish Hiremath, Yaohong Wang, and

- Yuankai Huo. Eosinophils instance object segmentation on whole slide imaging using multi-label circle representation, 2023.
- Ying-Chih Lo, Chia-Feng Juang, I-Fang Chung, Shin-Ning Guo, Man-Ling Huang, Mei-Chin Wen, Cheng-Jian Lin, and Hsueh-Yi Lin. Glomerulus detection on light microscopic images of renal pathology with the faster r-cnn. In *International Conference on Neural Information Processing*, pages 369–377. Springer, 2018.
- Xiangde Luo, Tao Song, Guotai Wang, Jieneng Chen, Yinan Chen, Kang Li, Dimitris Metaxas, and Shaoting Zhang. Scpm-net: An anchor-free 3d lung nodule detection network using sphere representation and center points matching. *arXiv preprint arXiv:2104.05215*, 2021.
- Ethan H Nguyen, Haichun Yang, Ruining Deng, Yuzhe Lu, Zheyu Zhu, Joseph T Roland, Le Lu, Bennett A Landman, Agnes B Fogo, and Yuankai Huo. Circle representation for medical object detection. *IEEE transactions on medical imaging*, 41(3):746–754, 2021.
- Ethan H Nguyen, Haichun Yang, Zuhayr Asad, Ruining Deng, Agnes B Fogo, and Yuankai Huo. Circlesnake: Instance segmentation with circle representation. In *International Workshop on Machine Learning in Medical Imaging*, pages 298–306. Springer, 2022.
- Sida Peng, Wen Jiang, Huaijin Pi, Xiuli Li, Hujun Bao, and Xiaowei Zhou. Deep snake for real-time instance segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8533–8542, 2020.
- Victor G Puelles, Wendy E Hoy, Michael D Hughson, Boucar Diouf, Rebecca N Douglas-Denton, and John F Bertram. Glomerular number and size variability and risk for kidney disease. *Current opinion in nephrology and hypertension*, 20(1):7–15, 2011.
- Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
- Shijie Wang, Guiling Sun, Bowen Zheng, and Yawen Du. A crop image segmentation and extraction algorithm based on mask rcnn. *Entropy*, 23(9), 2021. ISSN 1099-4300.
- Martin Weigert and Uwe Schmidt. Nuclei instance segmentation and classification in histopathology images with stardist. In *The IEEE International Symposium on Biomedical Imaging Challenges (ISBIC)*, 2022. .
- Juming Xiong, Yilin Liu, Ruining Deng, Regina N Tyree, Hernan Correa, Girish Hiremath, Yaohong Wang, and Yuankai Huo. Deep learning-based open source toolkit for eosinophil detection in pediatric eosinophilic esophagitis, 2023.
- Haichun Yang, Ruining Deng, Yuzhe Lu, Zheyu Zhu, Ye Chen, Joseph T Roland, Le Lu, Bennett A Landman, Agnes B Fogo, and Yuankai Huo. Circlenet: Anchor-free glomerulus detection with circle representation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 35–44. Springer, 2020.
- Hui Ying, Zhaojin Huang, Shu Liu, Tianjia Shao, and Kun Zhou. Embedmask: Embedding coupling for one-stage instance segmentation, 2019.
- Fisher Yu, Dequan Wang, Evan Shelhamer, and Trevor Darrell. Deep layer aggregation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2403–2412, 2018.
- Jiexian Zeng, Huan Ouyang, Min Liu, LU Leng, and Xiang Fu. Multi-scale yolact for instance segmentation. *Journal of King Saud University - Computer and Information Sciences*, 34(10, Part B):9419–9427, 2022. ISSN 1319-1578. .
- Hengxu Zhang, Pengpeng Liang, Zhiyong Sun, Bo Song, and Erkang Cheng. Circleformer: Circular nuclei detection in whole slide images with circle queries and attention. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 493–502. Springer, 2023.
- Qinghui Zhang, Xianing Chang, and Shanfeng Bian. Vehicle-damage-detection segmentation algorithm based on improved mask rcnn. *IEEE Access*, 8:6997–7004, 2020. .
- Xingyi Zhou, Dequan Wang, and Philipp Krähenbühl. Objects as points. *arXiv preprint arXiv:1904.07850*, 2019.