

Article

A Multi-Source Circular Geodesic Voting Model for Image Segmentation

Shuwang Zhou ^{1,2}, Minglei Shu ²  and Chong Di ^{2,*}
¹ College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China; ftdy123@outlook.com

² Shandong Artificial Intelligence Institute, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250014, China; shuml@sdas.org

* Correspondence: dich@sdas.org

Abstract: Image segmentation is a crucial task in artificial intelligence fields such as computer vision and medical imaging. While convolutional neural networks (CNNs) have achieved notable success by learning representative features from large datasets, they often lack geometric priors and global object information, limiting their accuracy in complex scenarios. Variational methods like active contours provide geometric priors and theoretical interpretability but require manual initialization and are sensitive to hyper-parameters. To overcome these challenges, we propose a novel segmentation approach, named PolarVoting, which combines the minimal path encoding rich geometric features and CNNs which can provide efficient initialization. The introduced model involves two main steps: firstly, we leverage the PolarMask model to extract multiple source points for initialization, and secondly, we construct a voting score map which implicitly contains the segmentation mask via a modified circular geometric voting (CGV) scheme. This map embeds global geometric information for finding accurate segmentation. By integrating neural network representation with geometric priors, the PolarVoting model enhances segmentation accuracy and robustness. Extensive experiments on various datasets demonstrate that the proposed PolarVoting method outperforms both PolarMask and traditional single-source CGV models. It excels in challenging imaging scenarios characterized by intensity inhomogeneity, noise, and complex backgrounds, accurately delineating object boundaries and advancing the state of image segmentation.

Keywords: geodesic voting; image segmentation; multi-source; polar representation; geodesic model



Citation: Zhou, S.; Shu, M.; Di, C. A Multi-Source Circular Geodesic Voting Model for Image Segmentation. *Entropy* **2024**, *26*, 1123. <https://doi.org/10.3390/e26121123>

Academic Editor: Amelia Carolina Sparavigna

Received: 14 November 2024

Revised: 13 December 2024

Accepted: 19 December 2024

Published: 22 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Image segmentation is a fundamental task in computer vision, playing a crucial role in applications such as medical imaging [1,2], autonomous vehicles [3,4], and agriculture [5,6]. Given its wide-ranging applications, image segmentation has been extensively studied over the past few decades, resulting in the development of numerous methods aimed at solving this problem efficiently [7,8].

The advent of deep learning, particularly convolutional neural networks (CNNs), has led to significant advancements in image segmentation [9]. CNNs can automatically learn image features and contextual information from large-scale datasets, thereby markedly improving segmentation performance. Specific neural network architectures, such as U-Net [10], DeepLab [11], and PolarMask [12,13], are designed with multi-scale processing capabilities, enabling them to handle features and objects of varying scales simultaneously. However, existing neural network-based segmentation models primarily rely on local appearance and abstract features from convolutional layers based on regular image grids. These models often lack explicit consideration of geometric constraints and global object information, which could further enhance segmentation precision.

Meanwhile, researchers have also investigated energy functional-based active contour methods for image segmentation [14,15]. The core idea of these methods is to define an

energy functional that accounts for specific image attributes, such as grayscale values, color information, and texture features. By employing optimization algorithms to minimize this energy functional, an optimal image segmentation result can be obtained. A notable advantage of active contour methods is their ability to design an energy functional based on the characteristics of the segmentation target. This approach naturally integrates prior information about the target, thereby enhancing model interpretability. However, a major limitation is that these methods do not automatically extract features from training samples. In practical applications, they often require manual or semi-automatic contour initialization based on test images and are highly sensitive to hyperparameters.

Given the limitations of neural network-based and active contour-based image segmentation methods, researchers have started to integrate these approaches to develop segmentation techniques with enhanced performance across various applications [16]. In the remainder of this section, we provide a brief overview of recent image segmentation methods that combine neural networks with active contour detection, followed by a discussion of the motivation behind our work.

Integration of Neural Networks and Active Contour Models for Image Segmentation

In [17], the segmentation task is reframed as a contour detection problem, combining CNNs with active contour modeling to explicitly track contours without relying on dense intermediate predictions. Building on this integration, other studies have leveraged the synergy between CNNs and active contour methods for improved accuracy and robustness in diverse industrial and inspection tasks. For instance, the authors of [18] apply a lightweight depthwise CNN following an active contour-based segmentation stage to classify complex multi-defect wafer maps, achieving high accuracy under intricate manufacturing conditions. Similarly, the approach in [19] enhances deep learning-based contour detection by incorporating domain knowledge and specialized edge extraction to accurately delineate yarn boundaries, improving inspection performance in challenging textile environments.

Several methods have focused on refining local region modeling and parameter estimation within active contour frameworks. In [20], a local region-based active contour model employs Bayes' theorem and spatial regularization to achieve robustness against noise when segmenting images with intensity inhomogeneity. Additionally, the authors of [21] combine CNN-generated parameter maps with texture descriptors, improving both speed and reliability during contour evolution. These techniques have shown particular success in medical imaging applications. For example, the authors of [22] fuse a modified active shape model with CNNs to accurately extract brain boundaries in MRI scans. In another instance, the authors of [23] integrate fuzzy connectedness with an autoencoder CNN, followed by post-processing with active contours, to enhance breast tumor segmentation in ultrasound images.

However, most of the existing approaches that integrate neural networks with active contour modeling still rely heavily on localized cues and lack a more global geometric understanding. In addition, many of these methods focus on specific applications, tailoring their designs and parameter settings to particular imaging scenarios and object classes, which limits their generalizability. These observations motivate the development of a method that not only leverages the strengths of deep learning but also integrates global geometric constraints to provide a more robust and general segmentation solution.

This paper proposes a novel approach that combines geometric constraints with neural networks. Specifically, it introduces a multi-source point circular geometric voting (CGV) method, termed PolarVoting. The PolarMask deep learning model [12] is utilized to extract feature representations at different scales and to obtain the primary contour of the object. Based on the center point and the source contour points identified by PolarMask, a new variant of the CGV model is proposed, where the voting score map is constructed using multiple source points. By leveraging the representational power of neural networks in conjunction with the global geometric information inherent in the image, the proposed PolarVoting method achieves precise and robust segmentation. It is particularly effective in

complex imaging scenarios, accurately delineating object boundaries despite challenges such as intensity inhomogeneity, noise, and intricate background structures.

The contributions of this paper are summarized as follows:

1. We introduce a variational PolarVoting model for image segmentation that blends the enhancement from the geometric regularization of the CGV approach and the learning capabilities of the deep learning-based PolarMask model. This fusion leverages the classical PolarMask model to generate initialization information, thereby enhancing segmentation precision, especially in complicated scenarios.
2. The traditional CGV model utilizes a single user-provided source point to create the voting score map, which may suffer from unexpected difficulties for complex images. Our model advances the CGV model by incorporating multiple source points, an enhancement designed to achieve more robust segmentation, particularly for intricate object boundaries.
3. The efficacy of the proposed PolarVoting model is evaluated on both real-world natural and medical images. Experimental results demonstrate that the PolarVoting model indeed outperforms both PolarMask and the traditional single-source CGV model, underscoring its effectiveness across varied contexts.

2. Preliminaries

In this paper, we propose the integration of the PolarMask deep learning model [12,13] with the CGV method [24], thereby leveraging the representational capabilities of deep neural networks alongside the geometric regularization inherent in geodesic models. This section introduces the core principles underlying the PolarMask model and the CGV model, as well as a discussion of their respective limitations, which serve as the impetus for the development of the novel image segmentation model presented in this paper.

2.1. PolarMask Model

The PolarMask model is a widely recognized single-shot instance segmentation framework that predicts image masks in polar coordinates both effectively and efficiently. This approach has demonstrated promising performance in object image segmentation tasks. First, we present the overall framework of PolarMask, followed by a discussion of its limitations.

2.1.1. PolarMask Segmentation

As shown in Figure 1, the PolarMask model introduces an efficient way to represent the image segmentation mask using the polar coordinate system, where an instance mask is represented by a single center and n straight rays. For a given instance mask, PolarMask first samples a candidate center (x_c, y_c) of the instance and a set of boundary points (x_i, y_i) for $i = 1, 2, \dots, n$ located on the predicted contour. From the center point x_c, y_c , n rayline-like straight segments are emitted uniformly at equal angular intervals $\Delta\theta$, with their lengths determined by the distance from the center point to corresponding boundary point (x_i, y_i) . The angular resolution is predefined (e.g., $n = 8, \Delta\theta = 45^\circ$). The length r of each rayline-like segment is predicted through a learning process. As a consequence, the image segmentation consists of an instance center classification processing and a dense distance regression processing in polar coordinates. With the polar coordinate representation, the PolarMask mask incorporates the polar IoU loss and soft polar centerness for instance center classification and dense coordinate regression, which further enhances its performance. For more detailed information, we refer the reader to [12,13].

The PolarMask model integrates a backbone network, a modified feature pyramid network, and a task-specific head [25]. Given the center location (x_c, y_c) and the lengths of n rays, the position of each corresponding contour point (x_i, y_i) is calculated by:

$$x_i = \cos \theta_i \times d_i + x_c \quad (1)$$

and

$$y_i = \sin \theta_i \times d_i + y_c. \quad (2)$$

Starting from the angle 0° , the boundary points are sequentially connected, ultimately forming the complete contour and the corresponding mask.

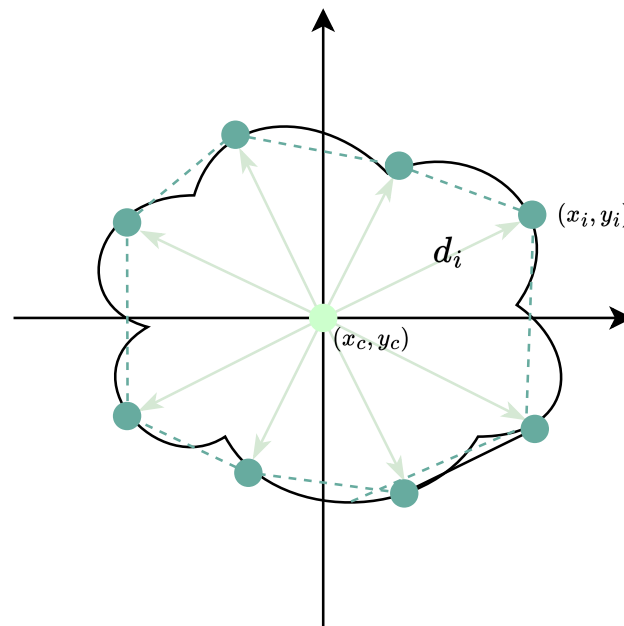


Figure 1. Illustration for a polar representation of a mask. The light green dot denotes the center point (x_c, y_c) and the dark green dots are the sampled boundary points (x_i, y_i) for $1 \leq i \leq n$.

2.1.2. Limitations of the PolarMask Model

The PolarMask model offers a novel approach to single-shot image segmentation, significantly improving performance and simplifying the training process. However, PolarMask still faces significant challenges when addressing image segmentation tasks, particularly in the following three scenarios:

1. The PolarMask model constructs the final contour by directly connecting the predicted points on the instance contour. This means that to achieve a more detailed representation of the contour, a greater number of polar rays must be predicted. For instance, as shown in Figure 2a, constructing the contour of an instance using only 36 points would lead to poor segmentation results. Furthermore, as previously mentioned, the mask regression branch has a shape of $H \times W \times n$ to predict the length of each polar ray. When the number of rays is large, the branch structure becomes more complex, making training more challenging. Therefore, selecting an appropriate value for the hyper-parameter n is a delicate task for PolarMask.
2. In the PolarMask model, only the point with the greatest length for each ray is preserved. However, if there are multiple points along the ground truth contour, as in the case of the example shown in Figure 2b, PolarMask struggles to accurately capture the true contour of the instance, even when the number of rays is large. As a result, PolarMask encounters significant challenges in segmenting U-shaped instances.
3. The PolarMask model does not incorporate contour regularization. In scenarios where image quality is low or distorted by noise, such as in medical CT images, the absence of regularization can lead to unexpected and irregular segmentation results, highlighting significant room for improvement.



(a) Result of insufficient number of rays.

(b) Result for a U-shaped instance.

Figure 2. Result in challenging scenarios.

2.2. The Circular Geodesic Voting Model

The circular geodesic voting model (CGV) [24] is a novel variational model designed to address the interactive image segmentation problem, particularly effective for complex tasks such as medical image segmentation. This section introduces the key concepts underlying the method and discusses the limitations that this paper aims to address.

2.2.1. Geodesic Minimal Path Model with Asymmetric Quadratic Metrics

The general geodesic minimal path problem is defined on a bounded domain $\Omega \subset \mathbb{R}^n$ equipped with a metric $\mathcal{F}(\mathbf{x}, \mathbf{u})$ that depends on positions $\mathbf{x} \in \Omega$ and orientations $\mathbf{u} \in \mathbb{R}^n$. This metric \mathcal{F} defines a norm at each point \mathbf{x} as:

$$\mathcal{F}_{\mathbf{x}}(\mathbf{u}) := \mathcal{F}(\mathbf{x}, \mathbf{u}), \quad (3)$$

where the norms $\mathcal{F}_{\mathbf{x}}$ must be positive $\mathcal{F}_{\mathbf{x}}(\mathbf{u}) > 0$ whenever $\mathbf{u} \neq \mathbf{0}$, 1-homogeneous, and satisfy the triangle inequality. However, in general, these norms are allowed to be asymmetric:

$$\mathcal{F}_{\mathbf{x}}(\mathbf{u}) \neq \mathcal{F}_{\mathbf{x}}(-\mathbf{u}). \quad (4)$$

The length of a regular curve γ with respect to the metric \mathcal{F} is measured as:

$$\mathcal{L}_{\mathcal{F}}(\gamma) = \int_0^1 \mathcal{F}(\gamma(t), \gamma'(t)) dt. \quad (5)$$

The minimal action map $U(\mathbf{x})$ is defined as:

$$U(\mathbf{x}) := \inf\{\mathcal{L}_{\mathcal{F}}(\gamma); \gamma \in \mathcal{A}_{s,x}\}, \quad (6)$$

where $\mathcal{A}_{s,x}$ represents the collection of all Lipschitz paths $\gamma: [0, 1] \rightarrow \Omega$ such that $\gamma(0) = s$ and $\gamma(1) = x$. This map is the unique viscosity solution to an Eikonal partial differential equation (PDE) [26,27]:

$$\begin{cases} \mathcal{F}_{\mathbf{x}}^*(\nabla U(\mathbf{x})) = 1, & \forall \mathbf{x} \in \Omega \setminus \{s\}, \\ U(s) = 0, \end{cases} \quad (7)$$

where $\mathcal{F}_{\mathbf{x}}^*$ is the dual norm of $\mathcal{F}_{\mathbf{x}}$, defined for all $\mathbf{u} \in \mathbb{R}^n$ by:

$$\mathcal{F}_{\mathbf{x}}^*(\mathbf{u}) := \sup_{\mathbf{v} \neq \mathbf{0}} \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\mathcal{F}_{\mathbf{x}}(\mathbf{v})}. \quad (8)$$

Based on the definition of the dual norm in (8), the corresponding optimal direction map Ψ is obtained by:

$$\Psi(\mathbf{x}, \mathbf{u}) := \arg \max_{\mathbf{v} \neq \mathbf{0}} \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\mathcal{F}_{\mathbf{x}}(\mathbf{v})}, \quad \forall \mathbf{x} \in \Omega, \quad \forall \mathbf{u} \in \mathbb{R}^n. \quad (9)$$

The geodesic $C_{s,x}$ is obtained by reversing the geodesic $\hat{C}_{x,s}$ with $C_{s,x}(0) = s$ and $C_{s,x}(1) = x$, where $\hat{C}_{x,s}$ is tracked through the following ordinary differential equation (ODE) involving the minimal action map U and the optimal direction map Ψ :

$$\begin{cases} \hat{C}'_{x,s}(t) \propto -\Psi(\hat{C}_{x,s}(t), \nabla U(\hat{C}_{x,s}(t))), \\ \hat{C}_{x,s}(0) = x. \end{cases} \quad (10)$$

Numerically, the ODE expressed in Equation (10) is solved using methods such as Heun's or Runge–Kutta's, or more robustly with the numerical method proposed by [28].

The asymmetric quadratic metrics \mathcal{F}_{AQ} used in the geodesic minimal path model within CGV are defined by:

$$\mathcal{F}_x := \mathcal{F}_{AQ}(\mathbf{x}, \mathbf{u}) = \sqrt{\langle \mathbf{u}, \mathcal{M}(\mathbf{x})\mathbf{u} \rangle + \langle \omega(\mathbf{x}), \mathbf{u} \rangle_+^2}, \quad (11)$$

where $\mathcal{M} : \mathcal{M} \rightarrow \mathcal{S}_{2+}^+$ is a tensor field, $\omega : \mathcal{M} \rightarrow \mathbb{R}^2$ is a vector field, and $\langle \mathbf{u}, \mathbf{v} \rangle_+^2 = (\max\{0, \langle \mathbf{u}, \mathbf{v} \rangle\})^2$ with $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$.

To leverage edge asymmetry features, two anisotropic and asymmetric minimal paths can be traced from a source point to an endpoint on the target edge along two opposite edge tangent directions based on the asymmetric quadratic metric. A closed contour is then formed by combining these two minimal paths, which share the same source and target points, to delineate the target boundary. For this purpose, the asymmetric quadratic metrics, constructed with different rotation matrices \mathcal{M}_i ($i \in \{0, 1\}$) and vector fields ω_i in the second part of the metric, are used to track minimal paths in counter-clockwise and clockwise directions. The corresponding asymmetric quadratic metrics are defined as:

$$\mathcal{F}_{AQ}^0(\mathbf{x}, \mathbf{u}) = \sqrt{\langle \mathbf{u}, \mathcal{M}(\mathbf{x})\mathbf{u} \rangle + \langle \omega_0(\mathbf{x}), \mathbf{u} \rangle_+^2}, \quad (12)$$

$$\mathcal{F}_{AQ}^1(\mathbf{x}, \mathbf{u}) = \sqrt{\langle \mathbf{u}, \mathcal{M}(\mathbf{x})\mathbf{u} \rangle + \langle \omega_1(\mathbf{x}), \mathbf{u} \rangle_+^2}. \quad (13)$$

Unlike the isotropic metric, which is independent of path tangents, the asymmetric quadratic metrics can generate different minimal paths with the same source point and endpoint by reversing the vector field component.

2.2.2. Overall Pipeline

The overall pipeline of the CGV model includes the following four components: computing the adaptive cut, building the set of endpoints, performing geodesic voting, and final segmentation.

Computing the Adaptive Cut

The adaptive cut [15] is integrated into the segmentation framework in conjunction with the minimal path model associated with the asymmetric quadratic metric. The use of the adaptive cut helps avoid the shortcut problem by preventing minimal paths from crossing the adaptive cut.

The adaptive cut is generated by computing a minimal path that connects a landmark point z inside the target region to a point x on the image boundary. It is constructed based on the Cohen–Kimmel minimal path model, where the metric in Equation (5) is simplified as $\mathcal{F}_{\text{Iso}}(\mathbf{x}, \mathbf{u}) = \psi(\mathbf{x})\|\mathbf{u}\|$. The potential function ψ is designed so that the resulting minimal path intersects the target boundary only once. The geodesic distance map $U_z : \mathcal{M} \rightarrow \mathbb{R}_0^+$ can be computed by solving the PDE (7) associated with the metric \mathcal{F}_{Iso} .

The endpoint $\mathbf{b} \in \partial\mathcal{M}$ of the target adaptive cut C_z can be identified by finding the point with the minimum distance value, i.e.,

$$\mathbf{b} := \arg \min_{\mathbf{x} \in \partial\mathcal{M}} U_z(\mathbf{x}). \quad (14)$$

Finally, the adaptive cut C_z can be generated by solving the gradient descent ODE (10) on U_z such that $C_z(0) = \mathbf{p}$ and $C_z(1) = \mathbf{b}$. Since the adaptive cut intersects the target boundary only once, it allows for tracking the minimal path from one side of the cut to the other.

Building the Set of Endpoints

The farthest point sampling (FPS) scheme [29] is employed to generate a set of points $A = \{\mathbf{p}_k \in \mathcal{M}; 1 \leq k \leq N\}$ with $N \geq 3$, which is then used to construct the set of endpoints for computing anisotropic and asymmetric minimal paths.

All points in the adaptive cut C_z are utilized as the initial model to search for the farthest points. The set of sampled farthest points is denoted as \mathcal{P} . In the first iteration, \mathcal{P} is set as $\mathcal{P}^{(0)} = C_z$, where \mathcal{P}^k represents the updated set at the k -th iteration. The first farthest point \mathbf{p}_1 is detected via the corresponding geodesic distance map U_{C_z} with $\mathcal{P} = \mathcal{P}^{(0)}$, as follows:

$$\mathbf{p}_1 = \arg \max_{\mathbf{x} \in \mathcal{M}} U_{\mathcal{P}}(\mathbf{x}). \quad (15)$$

The set \mathcal{P} is then updated to $\mathcal{P}^{(1)} = \mathcal{P}^{(0)} \cup \{\mathbf{p}_1\}$. In the k -th iteration (where $1 < k \leq N$), the k -th farthest point \mathbf{p}_k is detected by:

$$\mathbf{p}_k = \arg \max_{\mathbf{x} \in \mathcal{M}} U_{\mathcal{P}}(\mathbf{x}). \quad (16)$$

This process yields the updated set $\mathcal{P}^{(k)} = \mathcal{P}^{(k-1)} \cup \{\mathbf{p}_k\}$. Finally, the set of farthest points A is generated as:

$$A = \mathcal{P} \setminus C. \quad (17)$$

Performing Geodesic Voting

Using the sampled endpoints, a set of minimal paths for constructing the voting score map can be established by connecting each endpoint in the farthest point set A with asymmetric quadratic paths, based on the asymmetric quadratic metric. All minimal paths originate from the source point, which is the intersection of the target boundary and the adaptive cut. Let $\Phi_{AQ} = \{G_{i,j}\}_{1 \leq i \leq N}^{j \in \{0,1\}} \subset \text{Lip}([0,1], \mathcal{M})$ represent the set of voting paths, where N is the total number of endpoints. For each endpoint $\mathbf{p}_i \in A$, two minimal paths, $G_{i,0}$ and $G_{i,1}$, corresponding to the metrics \mathcal{F}_{AQ}^0 and \mathcal{F}_{AQ}^1 , are computed from the initial point s along both sides of the adaptive cut. The geodesic distance maps $U_{i,0}$ and $U_{i,1}$ are calculated using the adapted Hamiltonian fast marching (HFM) method in conjunction with the adaptive cut.

For an endpoint \mathbf{p}_i , the minimal paths $G_{i,0}$ and $G_{i,1}$ are obtained with respect to the geodesic distance maps $U_{i,0}$ and $U_{i,1}$. The minimal path $G_{i,j}$ with the lower weighted curve length is selected for voting score construction, as it has a higher likelihood of accurately depicting the target boundary. In this context, the weighted lengths of $G_{i,0}$ and $G_{i,1}$ are $U_{i,0}(\mathbf{p}_i)$ and $U_{i,1}(\mathbf{p}_i)$, respectively. The path with the lower length between $G_{i,0}$ and $G_{i,1}$ is denoted by \hat{G}_i , where $\hat{G}_i = G_{i,0}$ if $U_{i,0}(\mathbf{p}_i) < U_{i,1}(\mathbf{p}_i)$, and $\hat{G}_i = G_{i,1}$ otherwise. For convenience, define:

$$\Phi_D = \{\hat{G}_i\}_{1 \leq i \leq N}. \quad (18)$$

The voting score map $\mathcal{V}^D : \mathcal{M} \rightarrow [0, \infty)$, associated with the initial point s , is defined as:

$$\mathcal{V}^D(\mathbf{x}) = \sum_{G \in \Phi_D} \chi_X(G), \quad (19)$$

where $\chi_X : \text{Lip}([0,1], \Omega) \rightarrow \{0,1\}$ is a path detector defined by:

$$\chi_X(\gamma) = \begin{cases} 1, & \exists t \in [0,1] \text{ such that } \gamma(t) = \mathbf{x}, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

The value $\chi_X(\gamma)$ equals 1 if the curve γ passes through the point \mathbf{x} , and 0 otherwise.

For each pair of minimal paths $G_{i,0}$ and $G_{i,1}$ leading to $\mathbf{p}_i \in A$, these paths are considered to form a closed contour that accurately depicts the target if the geodesic flows at the endpoint are inverse to each other. The Euclidean scalar product of the vector fields at point \mathbf{p}_i , derived from the geodesic flow maps $\mathbf{V}_{i,0}$ and $\mathbf{V}_{i,1}$, is expressed as:

$$\mathcal{T}(\mathbf{p}_i) = \frac{\langle \mathbf{V}_{i,0}(\mathbf{p}_i), \mathbf{V}_{i,1}(\mathbf{p}_i) \rangle}{\|\mathbf{V}_{i,0}(\mathbf{p}_i)\| \|\mathbf{V}_{i,1}(\mathbf{p}_i)\|}, \quad (21)$$

with $G_{i,0}(1) = G_{i,1}(1) = \mathbf{p}_i$. Given a threshold $\zeta \in [-1, 0)$, the minimal paths $G_{i,0}$ and $G_{i,1}$ linked to \mathbf{p}_i are regarded as forming a valid closed path if $\mathcal{T}(\mathbf{p}_i) < \zeta$. Let $\Phi_{\mathcal{F}} = \{G_{n,j}\}_{1 \leq n \leq M, 0 \leq j \leq 1}$ denote the set of all minimal paths from Φ_{AQ} that satisfy this condition, where $M \in (0, N]$ is a positive integer representing the number of selected minimal paths. The second voting score map $\mathcal{V}^{\mathcal{F}} : \mathcal{M} \rightarrow [0, \infty)$ is then established as:

$$\mathcal{V}^{\mathcal{F}}(\mathbf{x}) = \sum_{G \in \Phi_{\mathcal{F}}} \chi_X(G). \quad (22)$$

The final voting score map $\mathcal{V} : \mathcal{M} \rightarrow [0, \infty)$, associated with the initial point s , is defined as:

$$\mathcal{V}(\mathbf{x}) = \alpha \mathcal{V}^{\mathcal{F}}(\mathbf{x}) + \beta \mathcal{V}^{\mathcal{D}}(\mathbf{x}), \quad (23)$$

where the weight parameters α and β control the relative importance of the two voting score maps.

Final Segmentation

In the voting score map, a high value of $\mathcal{V}(\mathbf{x})$ indicates a strong likelihood that \mathbf{x} belongs to the target boundary. A thresholding procedure is then applied to the voting score map \mathcal{V} to obtain the initial segmentation results. Additionally, mathematical morphological operators are employed to refine these results, ensuring that the final contours have a width of a single grid point.

2.2.3. Limitations

The CGV model demonstrates significant advantages by integrating the disconnection constraint within the image domain with the anisotropic and asymmetric properties of image edges. This approach allows for the detection of boundary contours of complex objects using only a randomly placed landmark point within the target region. Despite these strengths, the model has certain limitations that warrant further consideration:

1. The CGV model is inherently designed for interactive image segmentation, as it relies on the manual input of a landmark point. Consequently, it is unsuitable for scenarios requiring fully automatic segmentation.
2. In constructing the score map, the model utilizes a single point on the instance contour as the source point. While effective for simpler contours, this approach may struggle to accurately capture the complexity of highly intricate instance contours.

3. A Multi-Source Circular Geodesic Voting Model

In this section, we present the core contribution of this work: integrating the geometric properties of the CGV approach with the learning capabilities of the deep learning-based PolarMask model. This integration enhances the traditional CGV model by incorporating multiple source points to improve the computation of the voting score map.

An overview of the proposed PolarVoting framework is shown in Figure 3. The method involves two primary steps: (i) updating the vertices, which are derived from the PolarMask model and located near the target boundary, and (ii) computing a series of continuous curves, each of which passes through a vertex and connects the center point to the image domain boundary.

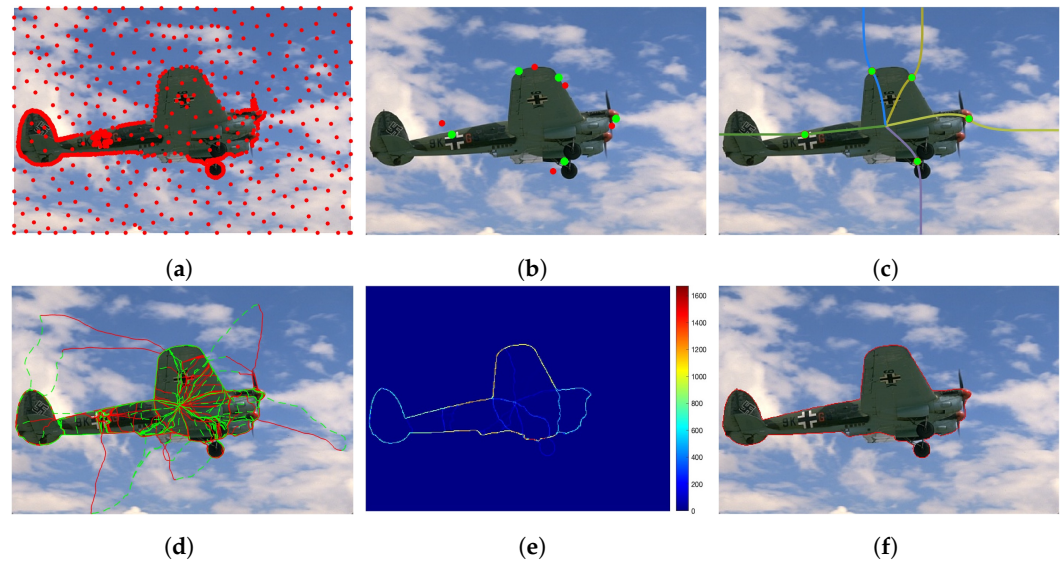


Figure 3. Overview of the proposed PolarVoting framework. (a) The original image, where the red dots represent the sampled farthest points. (b) Updating of vertices: the red dots denote contour points obtained from PolarMask, while the green dots represent the updated points. (c) Construction of multiple adaptive cuts. (d) Visualization of minimal paths using different metrics. (e) Visualization of the voting score map. (f) The final segmentation contour, represented by the red line.

3.1. Update the Vertices

Suppose that the output from the classical PolarMask model is a set of landmark points $\{\tilde{\mathbf{p}}_j\}_{1 \leq j \leq J}$, where J is a positive integer. We further suppose that these points are distributed along the target boundary in a clockwise order. However, these points sometimes are not exactly located in the target boundary. In order to solve this problem, we move these points towards the close edge points identified using the image gradient magnitude. Specifically, let $\mathbf{I} = (I_1, I_2, I_3) : \Omega \rightarrow \mathbb{R}^3$ be a vector-valued image (color image), where I_k for $k = 1, 2, 3$ represent the respective channel of the color image. As introduced in [30], the gradients of the color image \mathbf{I} , denoted by $\nabla \mathbf{I}$, are defined in its Jacobian matrix, that is

$$\nabla \mathbf{I}(\mathbf{x}) = [\nabla G_\sigma * I_1, \nabla G_\sigma * I_2, \nabla G_\sigma * I_3], \quad (24)$$

where $\nabla G_\sigma = (\partial_a G_\sigma, \partial_b G_\sigma)$ are the standard euclidean gradients of the Gaussian kernel with deviation σ and $\partial_a G_\sigma$ (resp. $\partial_b G_\sigma$) is the partial differential of G_σ along a -axis (resp. b -axis). Then, we construct a symmetric quadratic definite matrix field

$$W(\mathbf{x}) = \nabla \mathbf{I}(\mathbf{x}) \nabla \mathbf{I}(\mathbf{x})^\top \quad (25)$$

whose eigenvalues are $\lambda_1(\mathbf{x})$ and $\lambda_2(\mathbf{x})$ with $\lambda_1(\mathbf{x}) \leq \lambda_2(\mathbf{x})$, such that we define $f : \Omega \rightarrow \mathbb{R}_0^+$ as the image edge indicator such that

$$f(\mathbf{x}) = \lambda_1(\mathbf{x}) + \lambda_2(\mathbf{x}). \quad (26)$$

In essence, high values of the function $f(\cdot)$ usually imply strong appearance of image edges. As a consequence, we consider rectifying the landmark points $\{\tilde{\mathbf{p}}_j\}_{1 \leq j \leq J}$ by detecting an optimal point, denoted by \mathbf{p}_j , in a neighborhood $\mathbb{T}_\xi(\tilde{\mathbf{p}}_j)$ which surrounds the original landmark point $\tilde{\mathbf{p}}_j$, where $\xi \in \mathbb{R}^+$ is a positive scalar value defining the size of $\mathbb{T}_\xi(\tilde{\mathbf{p}}_j)$.

In this work, this procedure involves two main steps, where the first one is the construction of the neighbourhood $\mathbb{T}_\xi(\tilde{\mathbf{p}}_j)$. In order to take into account the image data, we apply the minimal weighted curve length associated to each point $\tilde{\mathbf{p}}_j$ to recover the

small region $\mathbb{T}_\xi(\tilde{\mathbf{p}}_j)$. Specifically, this is implemented by estimating the following map of minimal weighted curve length

$$\mathcal{D}_{\tilde{\mathbf{p}}_j}(\mathbf{x}) = \min_{\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}} \int_0^1 \mathcal{R}(\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}(u), \gamma'_{\tilde{\mathbf{p}}_j, \mathbf{x}}(u)) du \quad (27)$$

where $\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}} : [0, 1] \rightarrow \Omega$ is a curve of Lipschitz continuity, subject to $\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}(0) = \tilde{\mathbf{p}}_j$ and $\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}(1) = \mathbf{x}$. Let λ_i ($i = 1, 2$) be the eigenvalues of the matrix field W as defined in Equation (25). In addition, let $\mathbf{r} : \Omega \rightarrow \mathbb{R}^2$ be a vector field such that the direction $\mathbf{r}(\mathbf{x})$ points from \mathbf{x} to the closest image edge. For this purpose, we apply the method [31] to construct such a vector field \mathbf{r} and we further suppose that $\mathbf{r}_\perp(\mathbf{x})$ is a vector perpendicular to $\mathbf{r}(\mathbf{x})$. The function $\mathcal{R} : \Omega \times \mathbb{R}^2 \rightarrow [0, \infty]$ is a standard Riemannian metric expressed as

$$\mathcal{R}(\mathbf{x}, \mathbf{u}) = \mathbf{r}_\perp(\mathbf{x})\mathbf{r}_\perp(\mathbf{x})^\top + \exp(\mu f(\mathbf{x}))\mathbf{r}(\mathbf{x})\mathbf{r}(\mathbf{x})^\top \quad (28)$$

where $\mu > 0$ is a positive constant. The computation of the minimal weighted curve length $\mathcal{D}_{\tilde{\mathbf{p}}_j}$ can be efficiently estimated using the Hamiltonian fast marching method [28,32]. Moreover, each point \mathbf{x} leads to a minimal path $\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}^*$, whose Euclidean length is $L(\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}^*)$. Then, we define the neighbourhood $\mathbb{T}_\xi(\tilde{\mathbf{p}}_j)$ as

$$\mathbb{T}_\xi(\tilde{\mathbf{p}}_j) = \left\{ \forall \mathbf{x} \in \Omega, L(\gamma_{\tilde{\mathbf{p}}_j, \mathbf{x}}^*) \leq \xi \right\}. \quad (29)$$

Eventually, the rectified points \mathbf{p}_j can be obtained by

$$\mathbf{p}_j = \arg \max_{\mathbf{x} \in \mathbb{T}_\xi(\tilde{\mathbf{p}}_j)} f(\mathbf{x}). \quad (30)$$

An example of updating five PolarMask contour points is illustrated in Figure 3c.

3.2. Construct Multiple Adaptive Cuts

Now, we have a set of rectified landmark points $\{\mathbf{p}_j\}_j$ distributed along the boundary of the target and a center point \mathbf{c} inside the target. In contrast to the adaptive circular geodesic model [15] whose adaptive cut only connects the center point to a point at the image domain boundary $\partial\Omega$, the introduced method is able to simultaneously pass through a specified point such as a landmark point \mathbf{p}_j , allowing us to blend the benefits from the landmark points and the circular minimal paths. We denote by \mathcal{C}_j as the adaptive cut such that $\mathcal{C}_j(0) = \mathbf{c}$, $\mathcal{C}_j(1) \in \partial\Omega$ and $\mathcal{C}_j(u) = \mathbf{p}_j$, for all $1 \leq j \leq J$.

Basically, an adaptive cut \mathcal{C}_j is a continuous curve, and in our work this curve is generated as the concatenation of two minimal paths. In other words, we suppose that

$$\mathcal{C}_j = \mathcal{C}_{1,j} \cup \mathcal{C}_{2,j} \quad (31)$$

where $\mathcal{C}_{1,j}, \mathcal{C}_{2,j} : [0, 1] \rightarrow \Omega$ are two open curves. The operator \cup is used for curve concatenation [33–35] such that

$$(\gamma_1 \cup \gamma_2)(u) = \begin{cases} \gamma_1(2u), & \text{if } 0 \leq u \leq \frac{1}{2}, \\ \gamma_2(2u - 1), & \text{if } \frac{1}{2} \leq u \leq 1. \end{cases} \quad (32)$$

where $\gamma_1, \gamma_2 : [0, 1] \rightarrow \Omega$ are two curves.

Now, we describe how to compute the curves $\mathcal{C}_{1,j}$ and $\mathcal{C}_{2,j}$. In our work, we compute both paths using the Euler–Mumford elastica model [36,37], such that the path curvature can be taken into account. Basically, the energy of this elastica model can be formulated as:

$$\text{Length}(\gamma) := \int_0^1 \phi(\gamma(u), \gamma'(u) / \|\gamma'(u)\|) (1 + \tau^2 \kappa(u)^2) du \quad (33)$$

where $\phi : \Omega \times \mathbb{S}^1 \rightarrow \mathbb{R}^+$ is a cost function with $\mathbb{S}^1 = [0, 2\pi]$ being an interval of periodic boundary condition, $\kappa : [0, 1] \rightarrow \mathbb{R}$ is the curvature of the regular path $\gamma : [0, 1] \rightarrow \Omega$, and $\tau > 0$ is a parameter that controls the relative importance of the path curvature κ . The cost function ϕ is dependent on the direction $\mathbf{n}_\theta = (-\sin \theta, \cos \theta)$ for any angle $\theta \in \mathbb{S}^1$. In this work, we compute the cost ϕ using the image gradient-related matrix W defined in Equation (25). Let $\mu > 0$ be a weight parameter and the cost ϕ be estimated by

$$\phi(\mathbf{x}, \theta) = \exp(-\mu g(\mathbf{x}, \theta) / \|g(\mathbf{x}, \theta)\|_\infty) \quad (34)$$

where the function g is defined as $g(\mathbf{x}, \theta) = \langle \mathbf{n}_\theta, W(\mathbf{x}) \mathbf{n}_\theta \rangle$.

This is implemented in two steps. The first step is to compute $\mathcal{C}_{1,j}$, which is treated as a curvature-penalized minimal path that globally minimizes the elastica energy under a specified condition

$$\inf_{\gamma} \text{Length}(\gamma), \text{ subject to } \begin{cases} \gamma(0) = \mathbf{c}, \\ \gamma(1) = \mathbf{p}_j. \end{cases} \quad (35)$$

We adopt the method of orientation lifting, as presented in [36,37], to solve the minimization of the length (33), yielding the minimal path $\mathcal{C}_{1,j}$ such that $\mathcal{C}_{1,j}(0) = \mathbf{c}$ and $\mathcal{C}_{1,j}(1) = \mathbf{p}_j$.

Secondly, we apply the similar procedure to generate the minimal path $\mathcal{C}_{1,j}$ obeying that $\mathcal{C}_{2,j}(0) = \mathbf{p}_j$ and $\mathcal{C}_{2,j}(1) \in \partial\Omega$. This is performing by solving

$$\min_{\mathbf{x} \in \partial\Omega} \inf_{\gamma} \text{Length}(\gamma), \text{ subject to } \begin{cases} \gamma(0) = \mathbf{p}_j, \\ \gamma(1) = \mathbf{x}. \end{cases} \quad (36)$$

The target adaptive cut \mathcal{C}_j can be generated using equation (31). Repeating this procedure, we eventually obtain a series of adaptive cuts $\{\mathcal{C}_j\}$, $1 \leq j \leq J$.

An example of the construction of multiple adaptive cuts, with the number of vertices set to five, is illustrated in Figure 3b.

3.3. Circular Geodesic Voting with Distance Competition and Multiple Cuts

First of all, we construct a set of farthest points

$$Q = \{\mathbf{q}_k \in \Omega \mid 1 \leq k \leq K\}$$

which are usually taken as the endpoints for computing the voting paths. Note that in the circular geodesic voting method, the user is required to provide the contrast prior to the image gray levels. More precisely, the gray levels at most of the points inside the target region are supposed to be locally higher than the background. In order to remove this constraint and to facilitate the segmentation, we introduce a geodesic distance competition procedure which is able to identify locally the contrast of image gray levels, providing that the target boundaries are unknown.

Recall that each landmark point \mathbf{p}_j is assigned with an adaptive cut \mathcal{C}_j whose unit normal vectors are denoted by $\mathcal{N}_j(u) \propto \mathbf{M}(\pi/2) \mathcal{C}'_j(u)$, where $\mathbf{M}(\pi/2)$ is the clockwise rotation matrix with angle $\pi/2$ and \propto is the positively proportional operator. In this case, each landmark point \mathbf{p}_j corresponds to two offset points $\mathbf{p}_j^l = \mathbf{p}_j - \mathcal{N}_j(u)\epsilon$ and $\mathbf{p}_j^r = \mathbf{p}_j + \mathcal{N}_j(u)\epsilon$, where $\epsilon > 0$ is a sufficiently small constant. As a consequence, the point \mathbf{p}_j^l (resp. \mathbf{p}_j^r) is placed at the left (resp. right) side of the adaptive cut \mathcal{C}_j .

3.3.1. The Construction of Voting Paths

In the CGV model, two asymmetric quadratic metrics \mathcal{F}_{AQ}^0 (12) and \mathcal{F}_{AQ}^1 (13) are adopted to track minimal paths in counter-clockwise and clockwise directions. This paper considers two sets, Π_j^r and Π_j^l , as the search spaces for minimal paths with respect to \mathbf{p}_j :

$$\Pi_j^r := \{\gamma \in \text{Lip}([0, 1], \Omega) \mid \gamma(0) = \mathbf{p}_j^r, \forall u \in [0, 1], \gamma(u) \notin \mathcal{C}_j\},$$

$$\Pi_j^l := \{\gamma \in \text{Lip}([0, 1], \Omega) \mid \gamma(0) = \mathbf{p}_j^l, \forall u \in [0, 1], \gamma(u) \notin \mathcal{C}_j\}.$$

In other words, the set Π_j^r involves all the paths that emanate from \mathbf{p}_j^r . Given an asymmetric quadratic metric, the geodesic distance between two points is the minimal length of an admissible curve γ . By this definition, we construct two geodesic distance maps $\mathcal{D}_j^{0,r}$ and $\mathcal{D}_j^{1,r}$:

$$\mathcal{D}_j^{0,r}(\mathbf{x}) = \inf \left\{ \int_0^1 \mathcal{F}_{AQ}^0(\gamma(u), \gamma'(u)) du \mid \gamma \in \Pi_j^r, \gamma(1) = \mathbf{x} \right\} \quad (37)$$

$$\mathcal{D}_j^{1,r}(\mathbf{x}) = \inf \left\{ \int_0^1 \mathcal{F}_{AQ}^1(\gamma(u), \gamma'(u)) du \mid \gamma \in \Pi_j^r, \gamma(1) = \mathbf{x} \right\}. \quad (38)$$

In the eikonal PDE framework [32,38,39], the distances maps $\mathcal{D}_j^{0,r}(\mathbf{x})$ and $\mathcal{D}_j^{1,r}(\mathbf{x})$ are the viscosity solutions to the following PDEs

$$\begin{cases} \max_{\dot{\mathbf{x}} \neq 0} \frac{\langle \nabla \mathcal{D}_j^{0,r}(\mathbf{x}), \dot{\mathbf{x}} \rangle}{\mathcal{F}_{AQ}^0(\mathbf{x}, \dot{\mathbf{x}})} = 1, & \forall \mathbf{x} \neq \mathbf{p}_j^r \\ \mathcal{D}_j^{0,r}(\mathbf{p}_j^r) = 0. \end{cases} \text{ and } \begin{cases} \max_{\dot{\mathbf{x}} \neq 0} \frac{\langle \nabla \mathcal{D}_j^{1,r}(\mathbf{x}), \dot{\mathbf{x}} \rangle}{\mathcal{F}_{AQ}^1(\mathbf{x}, \dot{\mathbf{x}})} = 1, & \forall \mathbf{x} \neq \mathbf{p}_j^r \\ \mathcal{D}_j^{1,r}(\mathbf{p}_j^r) = 0. \end{cases} \quad (39)$$

The geodesic distance map $\mathcal{D}_j^{0,r}$ (resp. $\mathcal{D}_j^{0,l}$) corresponds to a set of minimal paths $\{\mathcal{G}_k^{0,r}\}_{1 \leq k \leq K}$ (resp. $\{\mathcal{G}_k^{0,l}\}_{1 \leq k \leq K}$), satisfying that $\mathcal{G}_k^{0,r}(1) = \mathbf{q}_k$ (resp. $\mathcal{G}_k^{0,l}(1) = \mathbf{q}_k$) where $\mathbf{q}_k \in Q$ is a sampled point. In other words, each sampled point \mathbf{q}_k leads to two minimal paths, and our goal is to choose a proper one so as to compute the voting scores. For this purpose, we estimate a binary-valued map $\mathbf{b}_r : \Omega \rightarrow \{0, 1\}$ such that

$$\mathbf{b}_r(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathcal{D}_j^{1,r}(\mathbf{x}) < \mathcal{D}_j^{0,r}(\mathbf{x}), \\ 0, & \text{otherwise.} \end{cases} \quad (40)$$

Let \mathcal{G}_k^r be the voting path such that $\mathcal{G}_k^r(0) = \mathbf{p}_j^r$ and $\mathcal{G}_k^r(1) = \mathbf{q}_k$. Using the map \mathbf{b}_r , one can identify \mathcal{G}_k^r as follows:

$$\mathcal{G}_k^r = \begin{cases} \mathcal{G}_k^{1,r}, & \text{if } \mathbf{b}_r(\mathbf{x}) = 1, \\ \mathcal{G}_k^{0,r}, & \text{otherwise.} \end{cases} \quad (41)$$

Similar to the construction of \mathcal{G}_k^r , one can obtain the voting paths \mathcal{G}_k^l obeying that $\mathcal{G}_k^l(0) = \mathbf{p}_j^l$ and $\mathcal{G}_k^l(1) = \mathbf{q}_k$. Note that the binary map \mathbf{b}_l for identifying the voting paths can be simply set to $\mathbf{b}_l = \mathbf{b}_r$ or can be computed by Equation (40) using the geodesic distance maps $\mathcal{D}_j^{0,l}$ and $\mathcal{D}_j^{1,l}$. Finally, we can also obtain K voting paths \mathcal{G}_k^l . In Figure 3d, we illustrate the voting paths originating from the farthest points for illustration.

3.3.2. Voting Score Map with Multiple Cuts

For each pair of points \mathbf{p}_j^r and \mathbf{p}_j^l , one has two sets of voting paths $\{\mathcal{G}_k^r\}_{1 \leq k \leq K}$ and $\{\mathcal{G}_k^l\}_{1 \leq k \leq K}$. Using the geodesic voting method introduced in the section on “Section Performing Geodesic Voting”, we can compute a voting score map $\mathcal{V}_j : \Omega \rightarrow \mathbb{R}^+$ using (23) by setting the parameter $\alpha = 0$. This means that we only use the circular path voting part.

As illustrated in the example in Figure 3e, the final voting score map \mathcal{V} can be simply set as the summation of all the \mathcal{V}_j :

$$\mathcal{V}(\mathbf{x}) = \sum_{j=1}^J \mathcal{V}_j(\mathbf{x}). \quad (42)$$

4. Experiments

In this section, we conduct experiments to compare the proposed multi-source circular geodesic voting model with both deep learning and geodesic voting methods through qualitative and quantitative analyses.

4.1. Experimental Configuration

4.1.1. Datasets

We employ both natural and medical images in our experiments to assess the performance of the methods.

- *Natural Images:* We utilize the Berkeley Segmentation Dataset 500 (BSDS500) [40] as the dataset for natural images, which serves as a standard benchmark for contour detection. This dataset is specifically designed to evaluate natural edge detection, encompassing not only object contours but also interior and background boundaries.
- *Medical Images:* For medical images, we select X-ray computed tomography (CT) scans from [41]. This dataset is commonly used in medical image segmentation evaluations, with image resolution ranging from 100×100 to 150×150 pixels, depending on the specific context of each test.

For both datasets, we select 80% of the images as the training set for deep learning-based methods, with the remaining 20% used as the test set for performance evaluation.

4.1.2. Baseline Methods

The baseline methods include the deep learning-based PolarMask [13] and the geodesic voting-based CGV [24]. For PolarMask, the training process follows that in [13], where ResNet101 and ResNeXt101 serve as the backbone networks, and the number of rays is set to 36. For CGV, the parameter settings align with those in [24]. Specifically, the number of endpoints N is set to 1000, and the parameters α and β used to compute the final voting score in Equation (23) are both set to 1. The threshold value for the final segmentation result is set to 100.

4.2. Qualitative Comparison Analysis

Figure 4 presents qualitative comparison results on natural images. The green lines denote the segmentation contours. In the original images (Column 1), target objects are surrounded by complex backgrounds, have intricate shapes, and exhibit intensity inhomogeneity. Additionally, some target boundaries are not clearly defined, making the segmentation task challenging. The PolarMask method (Column 2) struggles to capture the detailed contours of objects when boundaries are complex, as seen in the images in the second and fourth rows. It may also miss essential object components due to the limited number of points on the contour, as observed in the images in the first and fifth rows. The circular geodesic voting method (Column 3) encounters the “short-cut” problem when the object contains intricate textures, as evident in the images in the first, second, third, and fifth rows. This method is also influenced by strong background shapes because it uses only a single source point to compute the geodesic distance map. In contrast, the proposed multi-source circular geodesic voting method (Column 4) overcomes these limitations, accurately detecting target boundaries even in complex scenarios.

Figure 5 illustrates qualitative comparison results on medical CT images. In the first column, the original images highlight the challenges of segmenting the liver contour due to its low contrast with adjacent tissues, which complicates boundary differentiation. The liver’s irregular shape and texture, along with the partial volume effect, result in blurred

edges that further complicate segmentation. Given the difficulty of this task, the PolarMask method (Column 2) fails to capture a complete set of contour points around the liver. The circular geodesic voting method (Column 3) is affected by inherent noise along the liver boundary and artifacts from nearby structures, such as the ribs or diaphragm, leading to incorrect areas being included in the contour. By contrast, the proposed multi-source circular geodesic voting method (Column 4) demonstrates robustness against noise and accurately delineates the target boundaries.

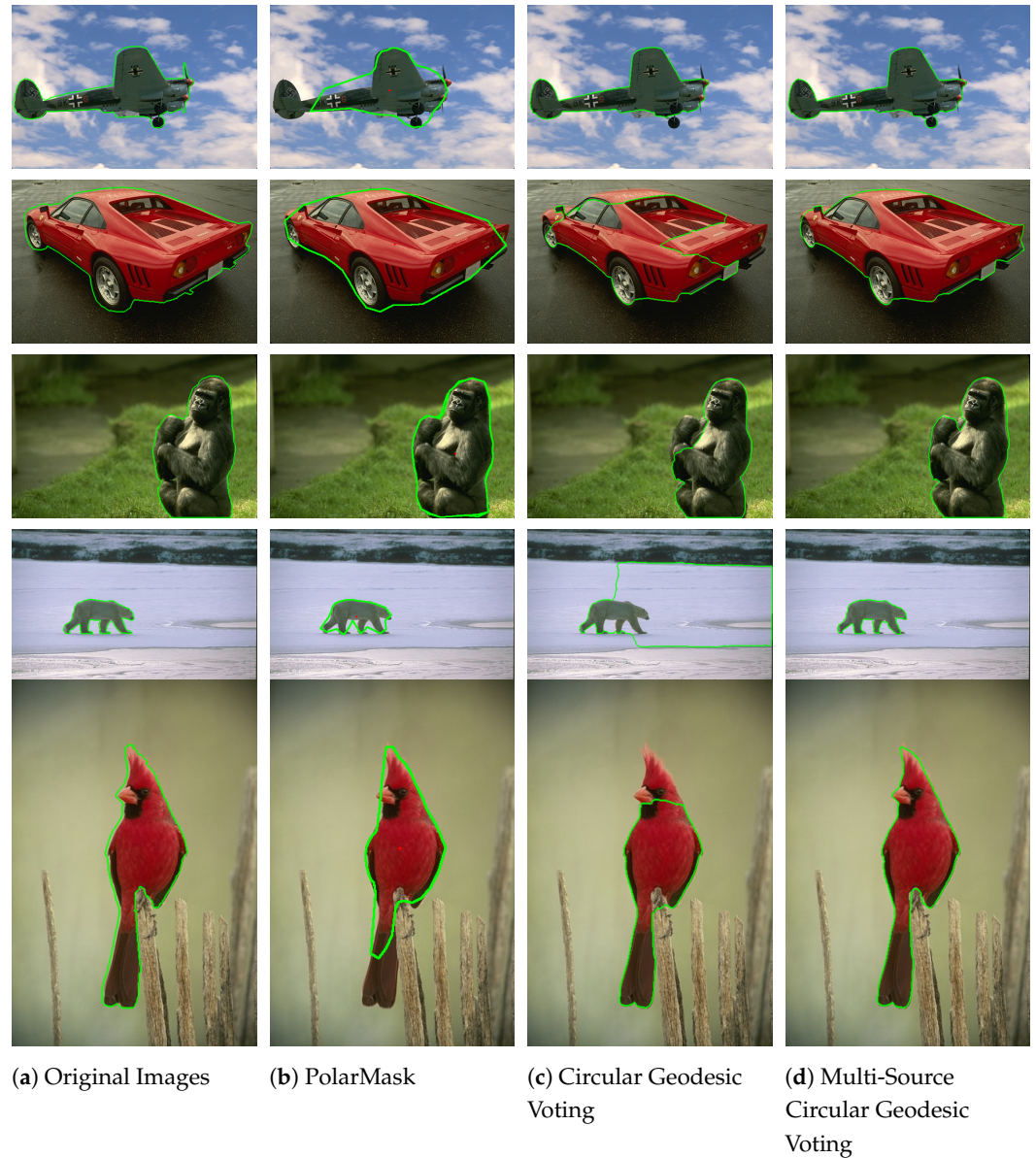


Figure 4. Qualitative comparison results on nature images, where the green lines denote the segmentation contours. Column 1 displays the original images with ground truth segmentation contours indicated by green lines. Columns 2–4 show the segmented results produced by the PolarMask, circular geodesic voting, and multi-source circular geodesic voting methods, respectively.

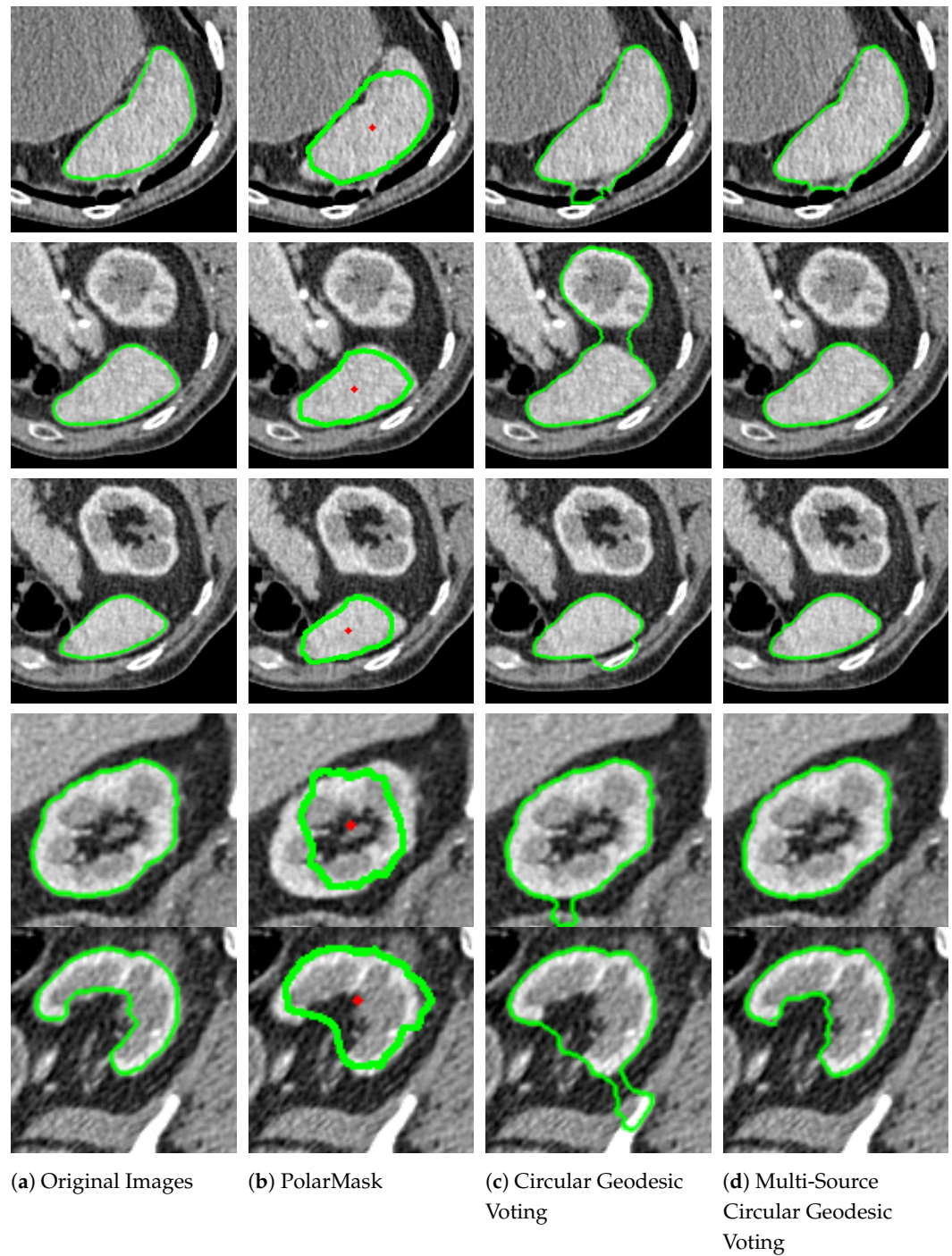


Figure 5. Qualitative comparison results on CT images. The green lines denote the segmentation contours. Column 1 displays the original images. Columns 2–4 show the segmented results produced by the PolarMask, circular geodesic voting, and multi-source circular geodesic voting methods, respectively.

4.3. Quantitative Performance Analysis

To compare the methods rigorously and convincingly, we conduct quantitative comparisons using the Dice score \mathcal{J} , chosen for its effectiveness in measuring the overlap between predicted paths and ground truth regions [15,24]. The accuracy of the tested methods is measured by the Dice score, defined as follows:

$$\mathcal{J} = \frac{\#|S \cap G|}{\#|S|}, \quad (43)$$

where S represents the set of grid points traversed by the evaluated paths, G denotes the region corresponding to the ground truth, and $\#|S|$ indicates the number of elements in the set S . The accuracy score \mathcal{J} ranges within the interval $[0, 1]$, with higher values of \mathcal{J} indicating better performance.

To quantitatively evaluate the segmentation accuracy of the proposed model, we perform a comparative analysis of different methods on the test set of the CT image dataset. Table 1 presents the quantitative comparison results for the PolarMask, CGV, and PolarVoting methods on the test set. These metrics report the mean accuracy and standard deviation of segmentation performance across the dataset.

Table 1. The quantitative comparison results of the PolarMask, circular geodesic voting, and multi-source circular geodesic voting methods on the test set of CT images.

PolarMask		CGV		PolarVoting	
Mean	Std	Mean	Std	Mean	Std
0.2928	0.1841	0.6639	0.3344	0.8561	0.1541

From the results in Table 1, it is evident that the PolarVoting method achieves a significantly higher mean accuracy (0.8561) compared to both PolarMask (0.2928) and single-source CGV (0.6639), with a notably lower standard deviation (0.1541) than that of the single-source CGV method (0.3344). This suggests that the PolarVoting approach provides more consistent segmentation outcomes.

In contrast, PolarMask demonstrates the lowest performance in both mean accuracy and standard deviation, indicating insufficient robustness and precision for CT image segmentation in this context. While the single-source CGV method outperforms PolarMask, it remains inferior to the proposed PolarVoting approach in terms of both accuracy and consistency. The box plots of the Dice scores in Figure 6 further support these findings.

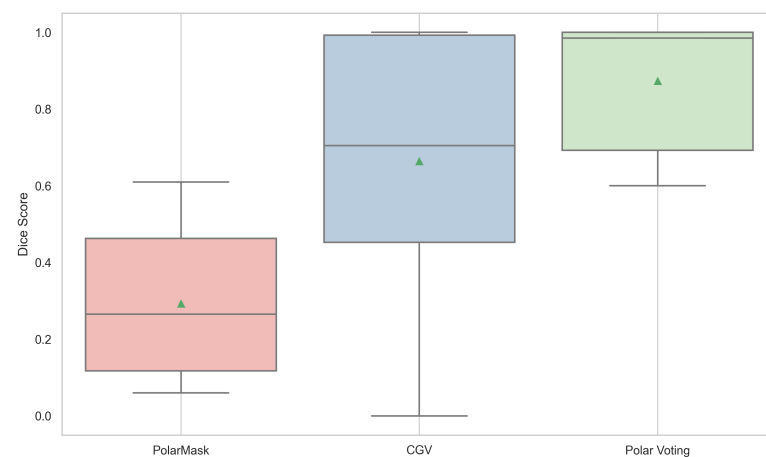


Figure 6. Box plots of the Dice scores for the PolarMask, circular geodesic voting, and multi-source circular geodesic voting methods on the test set of CT images. The green triangles represent the mean Dice score.

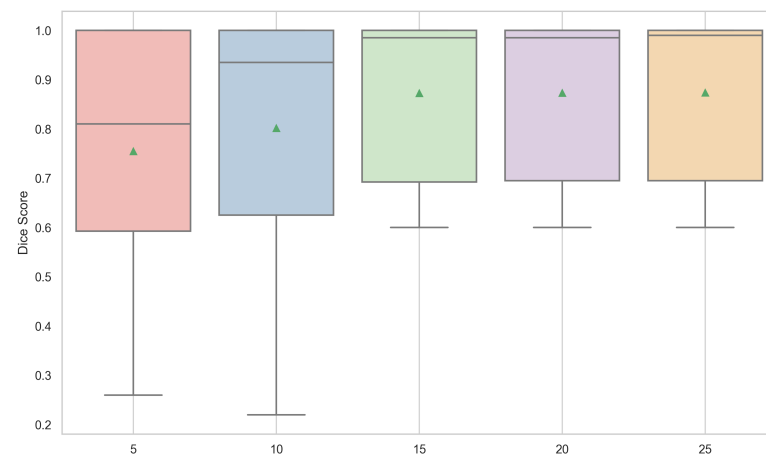
4.4. Influence of Source Points on Performance

4.4.1. Effect of the Number of Source Points

Table 2 and Figure 7 present the mean and standard deviation of the Dice scores, along with bar plots showing the Dice scores for different numbers of source points in the multi-source circular geodesic voting method on the test set of CT images.

Table 2. The quantitative comparison results of different numbers of source points for multi-source circular geodesic voting methods on the test set of CT images.

Number of Source Points									
5		10		15		20		25	
Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
0.7550	0.2432	0.8017	0.2436	0.8728	0.1576	0.8733	0.1572	0.8739	0.1576

**Figure 7.** Box plots of the Dice scores for different numbers of source points for multi-source circular geodesic voting methods on the test set of CT images. The green triangles represent the mean Dice score.

From these results, we observe that the Dice score improves as the number of source points increases from 5 to 25. Specifically, the mean Dice score rises from 0.7550 with 5 source points to 0.8739 with 25 points. This trend suggests that a higher number of source points enhances segmentation performance in the multi-source circular geodesic voting approach, likely due to the additional spatial information provided by multiple source points, allowing for more accurate alignment of geodesic votes within the CT image domain.

Furthermore, the standard deviation of the Dice scores decreases slightly as the number of source points exceeds 10. For example, with 5 and 10 source points, the Std is relatively high, around 0.243, indicating greater variability in performance across different samples in the test set. However, from 15 source points onward, the Std reduces to approximately 0.157, indicating that the method's performance becomes more consistent as the number of source points increases. This increased stability can be attributed to a more robust spatial representation achieved with a higher number of source points, which helps to minimize performance fluctuations.

The box plots in Figure 7 further illustrate the distribution of Dice scores across different configurations. With five source points, the Dice score distribution shows greater variability, including a wider range with lower minimum values. In contrast, the distributions become more compact and centered around higher median values as the number of source points increases, particularly from 15 to 25. This indicates that the voting-based method performs more reliably and effectively with an increased number of source points, leading to improved segmentation accuracy.

Figure 8 illustrates the execution time of the PolarVoting model as a function of the number of source points. The results indicate that the execution time increases linearly with the number of source points. This linear relationship suggests that the computational complexity of the PolarVoting model is directly influenced by the number of source points used in the segmentation process. While a higher number of source points may improve segmentation accuracy by better approximating the target boundary, it also increases the

computational burden. This trade-off between execution time and accuracy highlights the importance of selecting an appropriate number of source points based on application-specific requirements.

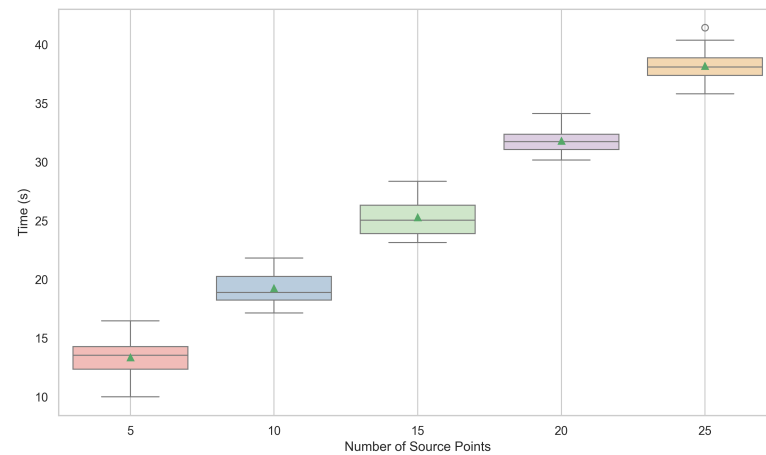


Figure 8. Box plots of the execution time for different numbers of source points for multi-source circular geodesic voting methods on the test set of CT Images. The green triangles represent the mean execution time.

4.4.2. Effect of the Placement of Source Points

In the PolarVoting model, the initial placement of source points is determined using PolarMask, and their locations are subsequently updated to align with the nearby target boundary. To evaluate the impact of the initial placement of source points, a PolarMask model is trained to generate 36 source points. A fixed number of source points, specifically 5, 10, and 15, are then sampled using different random seeds. This procedure is repeated 10 times for each configuration.

For each image in the test set of the CT image dataset, the Dice score is computed, and the variance across the 10 iterations is calculated. The distribution of these variances is depicted in Figure 9. The results indicate that the variance distributions for different numbers of source points are similar and remain small. This demonstrates that the location update strategy employed by the PolarVoting model ensures stable performance, rendering it insensitive to the initial placement of source points.

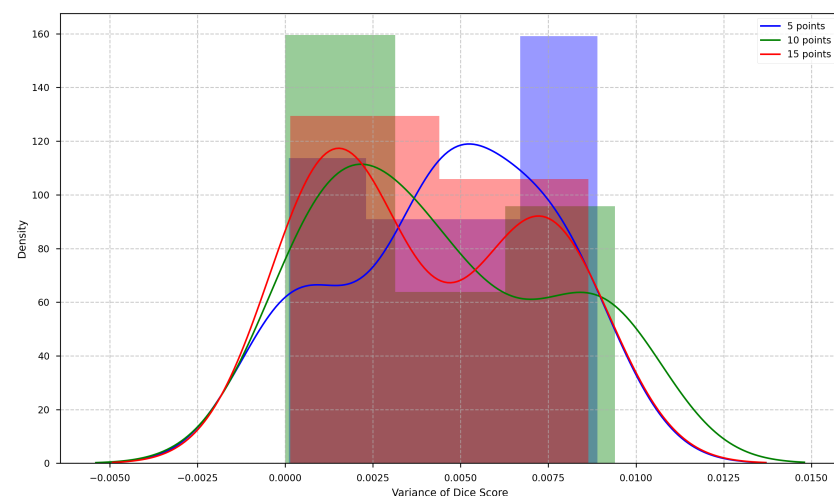


Figure 9. Variance distribution of Dice scores for different initial placements of source points in the PolarVoting model on the test set of CT images.

4.5. Performance Under Adverse Conditions

In this section, the performance of the proposed multi-source CGV model is evaluated under adverse conditions, including images affected by noise, blurring, and poor brightness.

Figures 10–12 present the segmentation results of PolarVoting under various adverse conditions. For noise conditions, Gaussian noise with variances of 0.01, 0.02, and 0.05 is added to the images. For blur conditions, the images are blurred using a rotationally symmetric Gaussian low-pass filter with σ values of 1, 2, and 3. For brightness variations, the brightness of the images is adjusted by factors of 0.2, 0.5, and 1.3, where values less than 1 darken the image and values greater than 1 brighten it. The segmentation results demonstrate that the PolarVoting model is robust to noise, blur, and brightness variations. It effectively tracks the correct boundaries even in the presence of significant noise, severe blurring, or extremely low brightness.

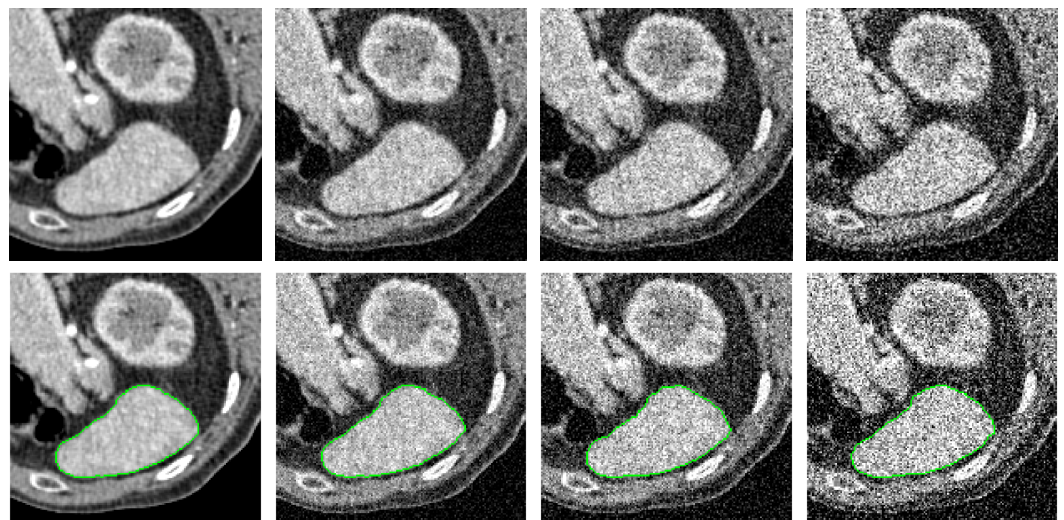


Figure 10. Performance of the PolarVoting model under noise conditions on CT images. The green lines represent the segmentation contours. Column 1 shows the original image alongside its segmentation result. Columns 2–4 present images affected by Gaussian noise with variances of 0.01, 0.02, and 0.05, respectively, along with the corresponding segmentation results.

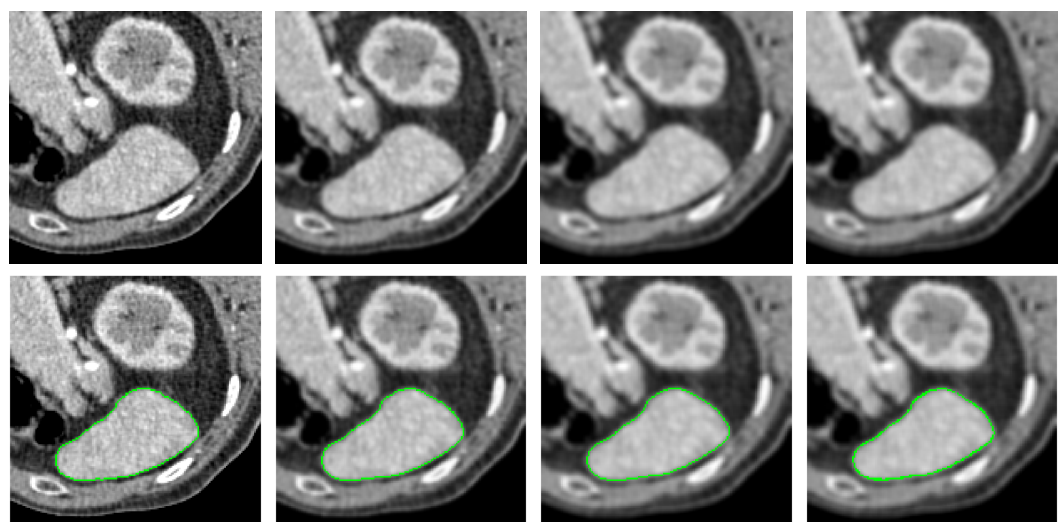


Figure 11. Performance of the PolarVoting model under blur conditions on CT images. The green lines represent the segmentation contours. Column 1 shows the original image alongside its segmentation result. Columns 2–4 display images with Gaussian blur levels of $\sigma = 1$, $\sigma = 2$, and $\sigma = 3$, respectively, along with the corresponding segmentation results.

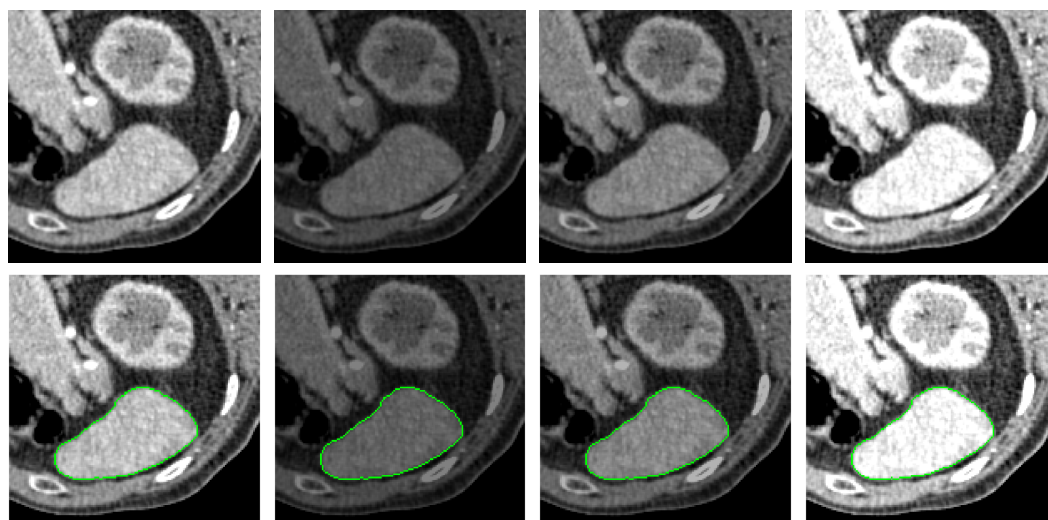


Figure 12. Performance of the PolarVoting model under brightness variations on CT images. The green lines represent the segmentation contours. Column 1 shows the original image alongside its segmentation result. Columns 2–4 display images with brightness adjustment factors of 0.5, 0.7, and 1.3, respectively, along with the corresponding segmentation results.

5. Conclusions

In this paper, we introduce a novel multi-source circular geodesic voting model for image segmentation, which effectively integrates the geometric regularization of traditional voting models which can take advantage of the adaptive representation capabilities of deep learning, particularly when carried out by the classical PolarMask model. By employing multiple source points for constructing the voting score map, the proposed model overcomes the limitations of single-source CGV approaches, especially when dealing with complex and noisy images in many segmentation tasks. Experimental results on both natural and medical image datasets demonstrate that the multi-source CGV model achieves superior accuracy and robustness compared to existing methods, such as PolarMask and single-source CGV, validating its potential for diverse image segmentation applications. Future work will be devoted to automated selection of source points and further optimization of the model for large-scale segmentation tasks, aiming to enhance both efficiency and accuracy across broader application domains.

Author Contributions: Conceptualization, S.Z., M.S. and C.D.; methodology, S.Z.; software, S.Z. and C.D.; validation, S.Z., M.S. and C.D.; formal analysis, S.Z.; investigation, S.Z.; resources, M.S.; data curation, S.Z. and C.D.; writing—original draft preparation, S.Z., M.S. and C.D.; writing—review and editing, S.Z. and C.D.; visualization, C.D.; supervision, C.D.; project administration, S.Z.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Shandong Province Key Research and Development Program (Major Scientific and Technological Innovation Project): 2023CXGC010113, 2020CXGC010501, 2020CXGC010903; Young Talent of Lifting Engineering for Science and Technology in Shandong, China (SDAST2024QTB014); and Qilu University of Technology (Shandong Academy of Sciences) Science and Education-Industry Projects: 2024ZDZX11, 2023RCKY247.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Luo, X.; Wang, G.; Liao, W.; Chen, J.; Song, T.; Chen, Y.; Zhang, S.; Metaxas, D.N.; Zhang, S. Semi-supervised medical image segmentation via uncertainty rectified pyramid consistency. *Med. Image Anal.* **2022**, *80*, 102517. [\[CrossRef\]](#)
- Ning, Z.; Zhong, S.; Feng, Q.; Chen, W.; Zhang, Y. SMU-Net: Saliency-guided morphology-aware U-Net for breast lesion segmentation in ultrasound image. *IEEE Trans. Med. Imaging* **2021**, *41*, 476–490. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhou, W.; Dong, S.; Lei, J.; Yu, L. MTANet: Multitask-aware network with hierarchical multimodal fusion for RGB-T urban scene understanding. *IEEE Trans. Intell. Veh.* **2022**, *8*, 48–58. [\[CrossRef\]](#)
- Yao, S.; Guan, R.; Huang, X.; Li, Z.; Sha, X.; Yue, Y.; Lim, E.G.; Seo, H.; Man, K.L.; Zhu, X.; et al. Radar-camera fusion for object detection and semantic segmentation in autonomous driving: A comprehensive review. *IEEE Trans. Intell. Veh.* **2023**, *9*, 2094–2128. [\[CrossRef\]](#)
- Deng, Y.; Xi, H.; Zhou, G.; Chen, A.; Wang, Y.; Li, L.; Hu, Y. An effective image-based tomato leaf disease segmentation method using MC-UNet. *Plant Phenomics* **2023**, *5*, 0049. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wu, F.; Yang, Z.; Mo, X.; Wu, Z.; Tang, W.; Duan, J.; Zou, X. Detection and counting of banana bunches by integrating deep learning and classic image-processing algorithms. *Comput. Electron. Agric.* **2023**, *209*, 107827. [\[CrossRef\]](#)
- Minaee, S.; Boykov, Y.; Porikli, F.; Plaza, A.; Kehtarnavaz, N.; Terzopoulos, D. Image segmentation using deep learning: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3523–3542. [\[CrossRef\]](#)
- Wang, Z.; Wang, E.; Zhu, Y. Image segmentation evaluation: A survey of methods. *Artif. Intell. Rev.* **2020**, *53*, 5637–5674. [\[CrossRef\]](#)
- Sultana, F.; Sufian, A.; Dutta, P. Evolution of image segmentation using deep convolutional neural network: A survey. *Knowl.-Based Syst.* **2020**, *201*, 106062. [\[CrossRef\]](#)
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; proceedings, part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
- Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [\[CrossRef\]](#)
- Xie, E.; Sun, P.; Song, X.; Wang, W.; Liu, X.; Liang, D.; Shen, C.; Luo, P. Polarmask: Single shot instance segmentation with polar representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 12193–12202.
- Xie, E.; Wang, W.; Ding, M.; Zhang, R.; Luo, P. Polarmask++: Enhanced polar representation for single-shot instance segmentation and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 5385–5400. [\[CrossRef\]](#)
- Zou, L.; Song, L.T.; Weise, T.; Wang, X.F.; Huang, Q.J.; Deng, R.; Wu, Z.Z. A survey on regional level set image segmentation models based on the energy functional similarity measure. *Neurocomputing* **2021**, *452*, 606–622. [\[CrossRef\]](#)
- Liu, L.; Chen, D.; Shu, M.; Cohen, L.D. Grouping Boundary Proposals for Fast Interactive Image Segmentation. *IEEE Trans. Image Process.* **2024**, *33*, 793–808. [\[CrossRef\]](#)
- Yang, D.; Peng, B.; Al-Huda, Z.; Malik, A.; Zhai, D. An overview of edge and object contour detection. *Neurocomputing* **2022**, *488*, 470–493. [\[CrossRef\]](#)
- Heidler, K.; Mou, L.; Loebel, E.; Scheinert, M.; Lefèvre, S.; Zhu, X.X. A deep active contour model for delineating glacier calving fronts. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5615912. [\[CrossRef\]](#)
- Choi, J.; Suh, D. A depthwise convolutional neural network model based on active contour for multi-defect wafer map pattern classification. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109707. [\[CrossRef\]](#)
- Xu, C.; Wang, J.; Tao, J.; Zhang, J.; Zheng, P. A knowledge augmented deep learning method for vision-based yarn contour detection. *J. Manuf. Syst.* **2022**, *63*, 317–328. [\[CrossRef\]](#)
- Li, Y.; Cao, G.; Wang, T.; Cui, Q.; Wang, B. A novel local region-based active contour model for image segmentation using Bayes theorem. *Inf. Sci.* **2020**, *506*, 443–456. [\[CrossRef\]](#)
- Nouri, M.; Baleghi, Y. An active contour model reinforced by convolutional neural network and texture description. *Neurocomputing* **2023**, *528*, 125–135. [\[CrossRef\]](#)
- Nguyen, D.H.; Nguyen, D.M.; Mai, T.T.; Nguyen, T.; Tran, K.T.; Nguyen, A.T.; Pham, B.T.; Nguyen, B.T. ASMCNN: An efficient brain extraction using active shape model and convolutional neural networks. *Inf. Sci.* **2022**, *591*, 25–48. [\[CrossRef\]](#)
- Biesok, M.; Juszczak, J.; Badura, P. Breast tumor segmentation in ultrasound using distance-adapted fuzzy connectedness, convolutional neural network, and active contour. *Sci. Rep.* **2024**, *14*, 25859. [\[CrossRef\]](#)
- Zhou, S.; Liu, H.; Liu, L.; Shu, M.; Chen, D.; Cohen, L.D. A Generalized Geodesic Voting Framework for Interactive Image Segmentation. *IEEE Trans. Instrum. Meas.* **2024**, *74*, 5001814. [\[CrossRef\]](#)
- Tian, Z.; Chu, X.; Wang, X.; Wei, X.; Shen, C. Fully convolutional one-stage 3d object detection on lidar range images. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 34899–34911.
- Lions, P.L. *Generalized Solutions of Hamilton-Jacobi Equations*; Pitman Publishing: Lanham, MD, USA, 1982.
- Sethian, J.A.; Vladimirsky, A. Ordered upwind methods for static Hamilton–Jacobi equations: Theory and algorithms. *SIAM J. Numer. Anal.* **2003**, *41*, 325–363. [\[CrossRef\]](#)
- Mirebeau, J.M. Anisotropic fast-marching on cartesian grids using lattice basis reduction. *SIAM J. Numer. Anal.* **2014**, *52*, 1573–1599. [\[CrossRef\]](#)

29. Peyré, G.; Cohen, L.D. Geodesic remeshing using front propagation. *Int. J. Comput. Vis.* **2006**, *69*, 145–156. [[CrossRef](#)]
30. Sochen, N.; Kimmel, R.; Malladi, R. A general framework for low level vision. *IEEE Trans. Image Process.* **1998**, *7*, 310–318. [[CrossRef](#)] [[PubMed](#)]
31. Xu, C.; Prince, J.L. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Process.* **1998**, *7*, 359–369.
32. Mirebeau, J.M.; Portegies, J. Hamiltonian fast marching: A numerical solver for anisotropic and non-holonomic eikonal PDEs. *Image Process. Line* **2019**, *9*, 47–93. [[CrossRef](#)]
33. Mille, J.; Bogleux, S.; Cohen, L.D. Combination of piecewise-geodesic paths for interactive segmentation. *Int. J. Comput. Vis.* **2015**, *112*, 1–22. [[CrossRef](#)]
34. Chen, D.; Zhu, J.; Zhang, X.; Shu, M.; Cohen, L.D. Geodesic paths for image segmentation with implicit region-based homogeneity enhancement. *IEEE Trans. Image Process.* **2021**, *30*, 5138–5153. [[CrossRef](#)] [[PubMed](#)]
35. Chen, D.; Mirebeau, J.M.; Shu, H.; Cohen, L.D. A Region-based Randers geodesic approach for image segmentation. *Int. J. Comput. Vis.* **2024**, *132*, 349–391. [[CrossRef](#)]
36. Chen, D.; Mirebeau, J.M.; Cohen, L.D. Global minimum for a Finsler elastica minimal path approach. *Int. J. Comput. Vis.* **2017**, *122*, 458–483. [[CrossRef](#)]
37. Mirebeau, J.M. Fast-marching methods for curvature penalized shortest paths. *J. Math. Imag. Vis.* **2018**, *60*, 784–815. [[CrossRef](#)]
38. Chen, D.; Spencer, J.; Mirebeau, J.M.; Chen, K.; Cohen, L.D. Asymmetric geodesic distance propagation for active contours. In Proceedings of the 29th British Machine Vision Conference (BMVC'18), Newcastle, UK, 3–6 September 2018.
39. Cohen, L.D.; Kimmel, R. Global minimum for active contour models: A minimal path approach. *Int. J. Comput. Vis.* **1997**, *24*, 57–78. [[CrossRef](#)]
40. Martin, D.; Fowlkes, C.; Tal, D.; Malik, J. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In Proceedings of the Eighth IEEE International Conference on Computer Vision, Vancouver, BC, Canada, 7–14 July 2001; Volume 2, pp. 416–423.
41. Spencer, J.; Chen, K.; Duan, J. Parameter-free selective segmentation with convex variational methods. *IEEE Trans. Image Process.* **2018**, *28*, 2163–2172. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.