

Article

MC-DBRG: 3D Point Cloud Instance Segmentation for Automatic Seismic Fault Interpretation via Multi-Scale Contour Constraints Delaunay-Based Region Growing

Lu Huang , Chunxia Zhang , Jiangshe Zhang *, Kai Sun and Liming Han

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China;
lchl199544@stu.xjtu.edu.cn (L.H.); cxzhang@mail.xjtu.edu.cn (C.Z.); kaisun@mail.xjtu.edu.cn (K.S.);
3125307027@stu.xjtu.edu.cn (L.H.)

* Correspondence: jszhang@mail.xjtu.edu.cn

Abstract

Automatic fault interpretation is essential for seismic geological modeling; however, fault intersections, noise interference, and topological complexity frequently lead to over-segmentation or under-segmentation. To address these challenges, we propose multi-scale contour constraints Delaunay-based region growing (MC-DBRG), an instance segmentation method for three-dimensional (3D) fault point clouds. This approach effectively extracts individual fault instances from fault probability volumes. The pipeline begins with voxel downsampling and multi-scale feature extraction applied to the fault point cloud. A Gaussian Mixture Model then performs binary classification to delineate explicit and closed fault contour walls. These walls function as physical barriers during the subsequent Delaunay triangulation-based region growing, which initializes from non-contour smooth points. Implementing these constraints guides the growth process and prevents topological adhesion among intersecting faults. To correct over-segmentation caused by data discontinuities in the preliminary results, we introduce a global manifold secondary geometric merging strategy. A macroscopic planar trend constraint is subsequently applied to accurately reassign and backfill contour points and noise. Experimental results indicate that MC-DBRG resolves boundary ambiguity and spatial adhesion within complex intersecting faults. The generated fault models preserve high topological integrity and precise instance labels, establishing a reliable foundation for automated seismic geological modeling.

Keywords: automatic fault interpretation; 3D point cloud; instance segmentation; delaunay triangulation; region growing; contour constraints



Academic Editor: Atsushi Mase

Received: 1 July 2026

Revised: 16 July 2026

Accepted: 20 July 2026

Published: 30 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The automated interpretation of three-dimensional (3D) seismic data and fault modeling are technically challenging and time-consuming components of underground oil and gas exploration and geological carbon sequestration assessment [1,2]. Fault systems, as the primary manifestation of subsurface structural discontinuities, directly control structural formation, trap development, and fluid migration pathways. Traditional techniques rely heavily on manually designed seismic geometric attributes such as coherence and variance [3–5]. These methods are sensitive to sedimentary unconformities and chaotic seismic facies [6], resulting in low robustness in environments with poor signal-to-noise ratios. Recently, deep learning methods, particularly convolutional neural networks [7–9],

have enabled the direct generation of high-precision fault likelihood volumes from raw seismic data by extracting complex non-linear features.

However, voxel-level probability prediction represents only the preliminary stage of automated fault interpretation. Existing end-to-end fault instance recognition algorithms struggle with generalization and boundary delineation when applied to real fault zones characterized by complex spatial structures and multi-scale features. Practical geological engineering and numerical simulations require 3D discrete fault surface models with independent instance labels, clear geometric morphology, and complete topological structures [10,11]. Furthermore, fine-grained hydrocarbon exploration requires the quantitative evaluation of geometric attributes such as fault strike, dip, and extension length [12]. Consequently, transforming unordered, multi-scale fault probability volumes containing residual noise into discrete fault instances with accurate geometric contours and clear topological relationships remains a critical bottleneck in intelligent 3D geological modeling [13].

Converting seismic fault probability volumes into 3D point clouds for spatial geometric segmentation [14–16] has become the standard paradigm for addressing instance segmentation in 3D fault data. While methods such as Random Sample Consensus (RANSAC) and Moving Least Squares (MLS) perform well in extracting single regular planes [17,18], RANSAC-based methods struggle to fit faults with high curvature, whereas Moving Least Squares (MLS) techniques are prone to erroneously blending nearby independent clusters due to search neighborhood overlaps. In contrast, the Delaunay-based Region Growing (DBRG) [19] algorithm demonstrates excellent mesh connectivity in general 3D surface reconstruction [20,21]. The application of this algorithm to complex seismic fault data, however, presents two main challenges. First, subsurface faults frequently exhibit complex spatial configurations, such as Y-shaped bifurcations or X-shaped intersections, where the spatial distances between fault planes are extremely small and are accompanied by sharp changes in normal vectors. Because region growing algorithms [22] rely heavily on the local smoothness assumption, the absence of explicit global boundary constraints allows the growth process to penetrate weak physical boundaries, incorrectly stitching multiple independent fault planes into a single volume. This topological leakage compromises instance independence and introduces severe errors into subsequent fault sealing evaluations [23]. Second, existing constrained region growing algorithms are highly sensitive to hyperparameters [24,25]. Relying solely on local normal vector similarity for growth blocking leads to unstable segmentation results that depend heavily on the neighborhood size [26]. An excessively small neighborhood amplifies local geological noise, generating numerous pseudo-fault fragments, whereas an excessively large neighborhood smooths out critical fault intersection boundaries. This hyperparameter sensitivity requires tedious manual tuning across datasets, restricting the practical applicability of the algorithm.

To overcome these limitations, we propose MC-DBRG, an unsupervised 3D fault point cloud instance segmentation framework based on multi-scale contour constraints and Delaunay triangulation-based region growing. By operating entirely without supervised training, the proposed method completely circumvents the necessity for labor-intensive mask annotation. This training-free characteristic serves as a central contribution of the work, significantly enhancing the practical applicability of the framework in complex geological scenarios where obtaining large-scale annotated data is prohibitive. The proposed framework decouples the global multi-instance segmentation task into four stages: feature-driven contour recognition, multi-constraint controlled region growing, global manifold secondary merging, and trend-guided attribution backfilling. Initially, the contour recognition stage reduces data redundancy via voxel downsampling [27,28] and extracts multi-scale geometric features [29,30]. A Gaussian Mixture Model [31,32] then performs binary classification with soft-threshold confidence control to isolate complex fault intersec-

tions. A Delaunay topological neighborhood [21] voting mechanism repairs undetected holes, constructing explicit and closed contour walls. This stage is highly robust to local hyperparameter choices due to the use of multi-scale feature fusion and probability density estimation. Next, the region growing process [22] initiates from non-contour smooth points acting as seeds. The application of multiple constraints, including physical wall blocking, local density consistency, topological area mutation, and normal vector similarity, confines the growth process to independent fault cores, preventing topological adhesion and leakage. To address fragmented clusters resulting from the preliminary growth, a secondary merging stage utilizes Principal Component Analysis [33,34] to approximate micro-fault manifolds. These manifolds are then projected and merged on a macroscopic plane, eliminating over-segmentation caused by data discontinuities. Finally, an attribution backfilling stage assigns labels to contour points and noise from the outside in, guided by global macroscopic planar trends. Experimental results on synthetic and real datasets demonstrate that MC-DBRG effectively resolves boundary ambiguity and spatial adhesion in complex intersecting faults. The resulting fault models maintain high topological integrity and precise instance labels, establishing a reliable foundation for automated seismic geological modeling [35–38].

The remainder of this paper is organized as follows. Section 2 details the proposed MC-DBRG framework. Section 3 evaluates the method on synthetic and real seismic data, and Section 4 concludes the paper.

2. Methods

2.1. Overview of the Method

While seismic fault probability volumes extracted via deep learning accurately depict the spatial distribution of fault systems, they fundamentally remain voxel-level binary probability representations. To transition from continuous probability volumes to discrete 3D fault instances, we convert high-confidence fault voxels into an initial 3D fault point cloud. To resolve topological leakage and over-segmentation in complex intersection regions, we abandon traditional global homogeneous growth strategies and propose MC-DBRG, a progressive multi-surface instance segmentation framework for fault point clouds. Figure 1 illustrates the algorithmic architecture and data flow.

The framework divides the segmentation task into four interconnected phases. We begin with feature-driven contour recognition and wall reinforcement, constructing explicit geometric blocking boundaries for the ensuing spatial topological growth. Guided by these boundaries, a multi-constraint controlled region growing phase extracts the core skeleton manifolds of independent faults. Because data discontinuities often introduce over-segmentation artifacts, we apply a global manifold secondary geometric merging step to eliminate them. The pipeline concludes with a trend-guided attribution backfilling stage that ensures precise label assignment across the entire point cloud.

By applying a dynamic decoupling strategy based on divide-and-conquer principles and manifold reconstruction, MC-DBRG resolves spatial adhesion and topological leakage at complex intersections. Combining efficient voxel downsampling preprocessing with the merging and backfilling mechanisms preserves geometric fidelity and computational efficiency, even when processing industrial point clouds containing tens of millions of points. This pipeline establishes a reliable data foundation for subsequent automated geological modeling.

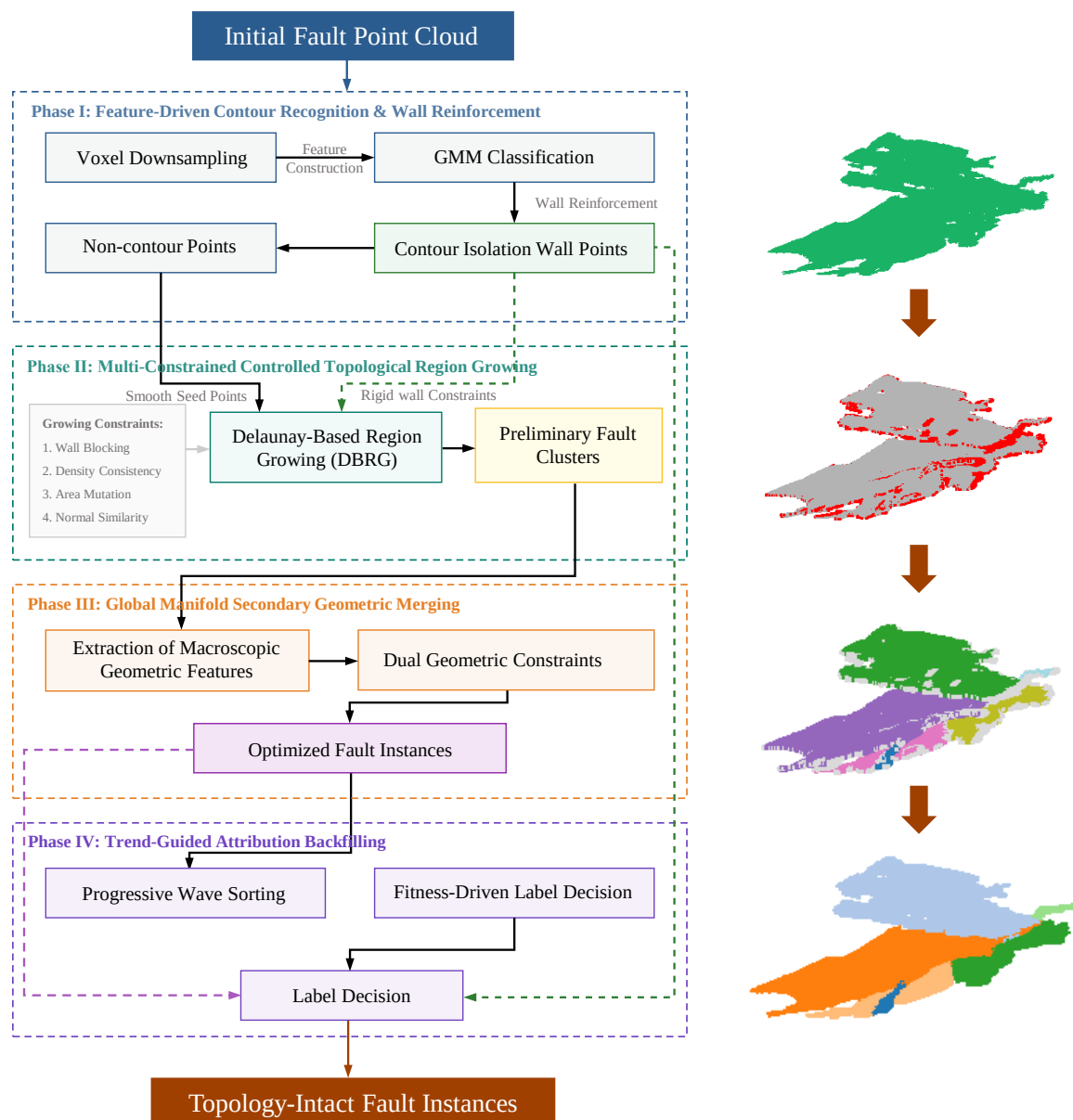


Figure 1. Algorithmic Architecture and Data Flow of the MC-DBRG Framework for Progressive Multi-Surface Instance Segmentation of Fault Point Clouds. Left: Different colored background regions represent distinct stages of the pipeline. Right: Different colors of the fault point cloud indicate separate segmented instance categories.

2.2. Feature-Driven Contour Recognition and Wall Reinforcement

2.2.1. Motivation for Contour Recognition

Fault pinch-outs and multi-surface intersection lines, such as Y-shaped and X-shaped configurations, exhibit severe geometric discontinuities. Constrained by resolution limits, fault probability volumes generated by deep learning models often suffer from spatial ambiguity and high noise at these boundaries. Consequently, traditional region growing algorithms are highly susceptible to topological leakage. We address this by explicitly decoupling contour boundary points from the remaining points to construct closed isolation walls at the intersections. This approach partitions the complex global topology graph into independent, locally smooth subgraphs, physically preventing the region growing process from propagating erroneously (details are shown in the Figure 2).

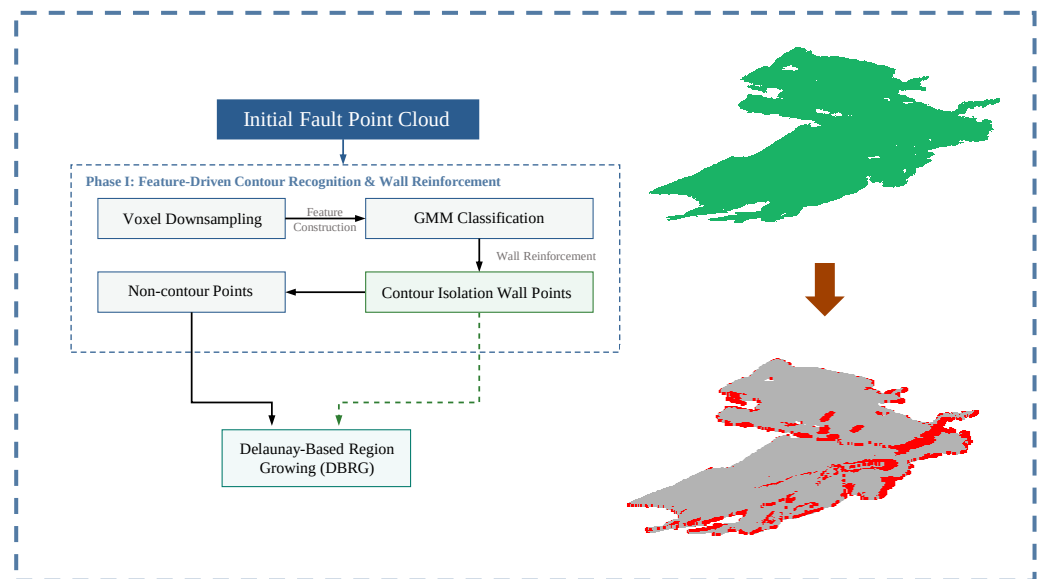


Figure 2. Workflow of the feature-driven contour recognition and wall reinforcement phase. In the fault point cloud on the right, the red points represent the identified contour points, while the gray points denote the non-contour points.

2.2.2. Voxel Downsampling Preprocessing

Three-dimensional (3D) seismic fault probability point clouds are typically massive and exhibit uneven density. To balance computational efficiency with topological consistency, we apply a voxel downsampling strategy. This method divides the 3D space into regular cubic grids with an edge length of v_s and replaces the points within each voxel with their geometric centroid. Running in $\mathcal{O}(N)$ time, this process homogenizes the global spatial density. It effectively eliminates local sampling artifacts and provides an unbiased data foundation for subsequent geometric feature calculations.

2.2.3. Dynamic Neighborhood Mechanism and Multi-Scale Feature Construction

A significant spatial scale disparity exists between fault intersection boundaries and single smooth surfaces. Relying on a fixed neighborhood size often blurs boundaries when the neighborhood is too large, or amplifies noise when it is too small. We therefore employ a multi-scale dynamic neighborhood mechanism to adaptively construct a local covariance matrix C_k corresponding to each spatial radius r_k within the set $R_{\text{scale}} = \{r_1, r_2, \dots, r_m\}$. For a specific scale r_k , the matrix is defined as:

$$C_k = \frac{1}{|N_k(p_i)|} \sum_{q \in N_k(p_i)} (q - \bar{p}_k)(q - \bar{p}_k)^T, \quad (1)$$

where $N_k(p_i)$ denotes the local neighborhood of point p_i constrained by radius r_k , and $\bar{p}_k$ is the center of this neighborhood. It is important to clarify that the covariance matrix is computed separately for each of the m spatial scales. Applying Principal Component Analysis to each C_k independently yields a set of scale-specific eigenvalues $\lambda_{1,k} \geq \lambda_{2,k} \geq \lambda_{3,k} \geq 0$. These scale-indexed eigenvalues are then utilized to compute the high-dimensional composite feature vector $[C_{p,k}, S_{p,k}, N_{p,k}, \theta_{\max,k}, D_{\min,k}]$ for each respective scale.

The 3D geometric features are defined as follows:

- Surface variation (C_p) measures the degree to which the manifold deviates from the tangent plane:

$$C_{p,k} = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}. \quad (2)$$

- Sphericity (S_p) quantifies the local scatter:

$$S_{p,k} = \frac{\lambda_3}{\lambda_1}. \quad (3)$$

- Non-planarity (N_p) evaluates the deviation from a planar structure:

$$N_{p,k} = 1 - \frac{\lambda_2 - \lambda_3}{\lambda_1}. \quad (4)$$

To extract 2D projection features, we project the neighborhood points onto the principal component tangent plane. We then calculate the maximum angle $\theta_{\max,k}$ between the connected projected points and the minimum distance $D_{\min,k}$ to the projection center. If a target point lies on an intersecting ridge or a physical boundary, its projection falls on the edge of the local convex hull, causing a significant increase in both $\theta_{\max,k}$ and $D_{\min,k}$ (see Figure 3). This multi-scale feature fusion mechanism bypasses manual parameter tuning and fundamentally improves the algorithm's robustness against complex geological noise.

Case A: Smooth Surface
(Target $\approx$ Interior Point)

Case B: Intersecting Ridge / Boundary
(Target falls on Hull Edge)

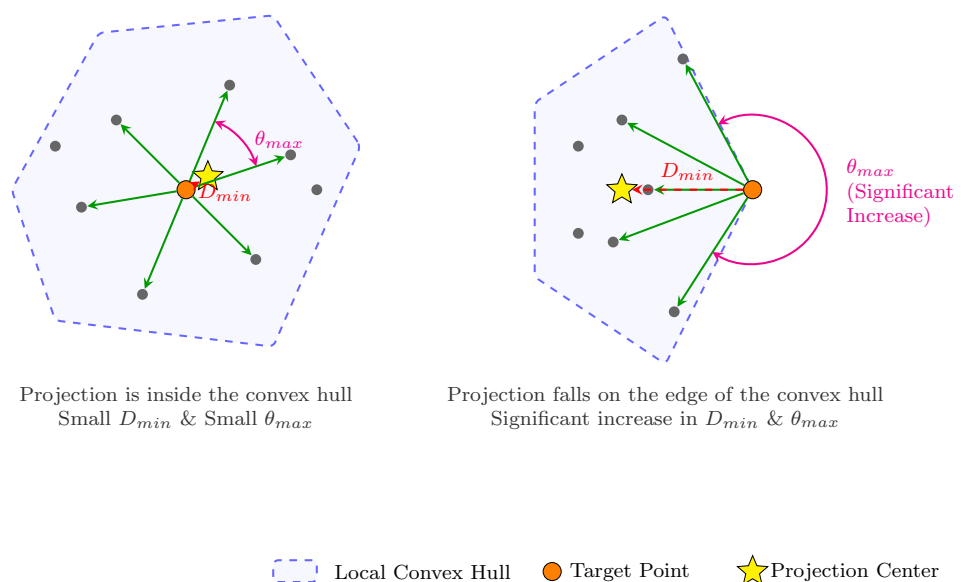


Figure 3. Schematic of 2D projection feature extraction on the principal component tangent plane. The gray points represent the neighbors of the target point. (Case A) For an interior point on a smooth surface, the neighborhood projection is distributed within the local convex hull. (Case B) For a point on an intersecting ridge or physical boundary, the target falls on the edge of the convex hull.

2.2.4. Feature Decoupling and Confidence Control Based on Gaussian Mixture Models

To enhance the algorithm's generalization capability across heterogeneous geological regions, we introduce a Gaussian Mixture Model (GMM) to perform unsupervised binary classification within the feature space. The GMM adaptively fits a two-component multi-variate Gaussian distribution and outputs the posterior probability $P(\text{contour}|p_i)$ that a sampled point belongs to the contour class. To ensure robust edge blocking, we implement a soft-threshold control mechanism based on this probability:

$$\text{Label}(p_i) = \begin{cases} \text{Contour}, & P(\text{contour}|p_i) \geq \tau \\ \text{Smooth}, & P(\text{contour}|p_i) < \tau. \end{cases} \quad (5)$$

Setting an appropriate confidence threshold τ incorporates fuzzy transition zone points into the contour wall candidates. This strategy guarantees that ambiguous boundary points capable of causing spatial adhesion are retained, thereby constructing a rigid safety barrier against topological leakage. Geologically, this barrier directly corresponds to the highly deformed damage zones at fault intersections, which exhibit severe geometric distortions compared to smooth fault planes. By effectively capturing the underlying bimodal statistical distribution of these distortions, the GMM fundamentally translates complex geological variations into reliable topological isolation walls.

2.2.5. Wall Reinforcement via Delaunay Topological Neighborhood Voting

The initially classified contour walls may still contain small gaps. To prevent the region growing process from penetrating these holes and erroneously connecting distinct fault instances, we design a morphological closing algorithm based on Delaunay topological neighborhood voting. The algorithm begins by applying local Delaunay triangulation to the point cloud to establish an explicit topological network. For any smooth point p_j , the nodes within its first-order Delaunay neighborhood $N_{\text{Delaunay}}(p_j)$ vote to reinforce the center point:

$$\text{If } \frac{\sum_{k \in N_{\text{Delaunay}}(p_j)} \mathbb{I}(\text{Label}(p_k) = \text{Contour})}{|N_{\text{Delaunay}}(p_j)|} \geq \alpha_{\text{vote}}, \quad \text{then Label}(p_j) \leftarrow \text{Contour}. \quad (6)$$

Here, $\mathbb{I}(\cdot)$ denotes the indicator function and α_{vote} is the voting threshold. Operating on the topological manifold of the point cloud, this procedure is equivalent to a morphological closing operation. It tightly stitches boundary gaps and forms a spatially closed, continuous physical isolation wall at multi-surface intersections and pinch-outs, providing a secure barrier for the subsequent region growing algorithm.

The complete technical flow of the feature-driven contour recognition and topological wall reinforcement, transitioning from raw point clouds to discrete explicit boundaries, is detailed and formalized in Algorithm 1.

Algorithm 1 MC-DBRG Phase 1: Feature-Driven Contour Recognition and Wall Reinforcement

Require: Initial 3D fault point cloud P_{input} ; Voxel grid edge length v_s ; Multi-scale spatial radius set $R_{\text{scale}} = \{r_1, r_2, \dots, r_m\}$; GMM confidence threshold τ ; Delaunay neighborhood voting threshold α_{vote} .

Ensure: Downsampled point cloud P ; Global Delaunay network $\mathcal{G}_{\text{Del}}$; Label array Label for all points in P (Label(p) $\in$ {Contour, Smooth}).

```

1:  $P \leftarrow \text{VoxelDownsample}(P_{\text{input}}, v_s)$  ▷ Homogenize global spatial density
2: for each point  $p_i \in P$  do
3:   for each radius  $r_k \in R_{\text{scale}}$  do
4:     Construct local covariance matrix  $C_k$  within constrained neighborhood  $N_k(p_i)$ 
5:     Compute scale-specific feature  $\mathbf{f}_{i,k} \leftarrow [C_{p,k}, S_{p,k}, N_{p,k}, \theta_{\max,k}, D_{\min,k}]$ 
6:   end for
7:    $\mathbf{f}_i \leftarrow [\mathbf{f}_{i,1}, \mathbf{f}_{i,2}, \dots, \mathbf{f}_{i,m}]$  ▷ Concatenate into an aggregated composite vector
8: end for
9:  $\mathcal{M}_{\text{GMM}} \leftarrow \text{GMM}(\{\mathbf{f}_i\})$  ▷ Unsupervised binary classification
10: for  $p_i \in P$  do
11:   Evaluate  $P(\text{contour} | p_i)$  via  $\mathcal{M}_{\text{GMM}}$ 
12:   if  $P(\text{contour} | p_i) \geq \tau$  then
13:     Label( $p_i$ )  $\leftarrow$  Contour
14:   else
15:     Label( $p_i$ )  $\leftarrow$  Smooth
16:   end if
17: end for

```

Algorithm 1 *Cont.*

```

18:  $\mathcal{G}_{\text{Del}} \leftarrow \text{Delaunay}(P)$  ▷ Establish explicit topology
19: for  $p_j \in P$  where  $\text{Label}(p_j) == \text{Smooth}$  do
20:    $N_{\text{Del}}(p_j) \leftarrow \text{Delaunay}(\mathcal{G}_{\text{Del}}, p_j)$ 
21:    $\gamma \leftarrow \frac{1}{|N_{\text{Del}}(p_j)|} \sum_{q \in N_{\text{Del}}(p_j)} \mathbb{I}(\text{Label}(q) == \text{Contour})$  ▷  $\mathbb{I}(\cdot)$  is the indicator function
22:   if  $\gamma \geq \alpha_{\text{vote}}$  then
23:      $\text{Label}(p_j) \leftarrow \text{Contour}$  ▷ Morphological closing via topological voting
24:   end if
25: end for
26: return  $P, \mathcal{G}_{\text{Del}}, \text{Label}$ 

```

2.3. Multi-Constraint Controlled Topological Region Growing**2.3.1. Geological Applicability and Topological Advantages of DBRG**

In 3D point cloud segmentation, Random Sample Consensus tends to fit regular geometric planes, making it prone to model mismatch when processing curved or densely parallel subsurface fault clusters. Moving Least Squares, a local smooth approximation algorithm, often blurs fault pinch-outs or intersection edges due to over-smoothing. In contrast, region growing algorithms rely on local continuity propagation, which aligns with the physical characteristics of faults that exhibit local smoothness and global continuity. We replace traditional K-nearest neighbor (KNN) searches with Delaunay-based Region Growing (DBRG). While traditional methods execute time-consuming neighborhood retrievals during the growth phase, DBRG requires constructing a global Delaunay triangulation network only once. This reduces the query complexity for adjacency relationships to $\mathcal{O}(1)$, significantly improving computational throughput for large-scale point clouds. Furthermore, traditional KNN searches often terminate abruptly in sparse regions. In contrast, a Delaunay network inherently preserves topological connectivity, adaptively bridging areas with heterogeneous densities (details are shown in the Figure 4).

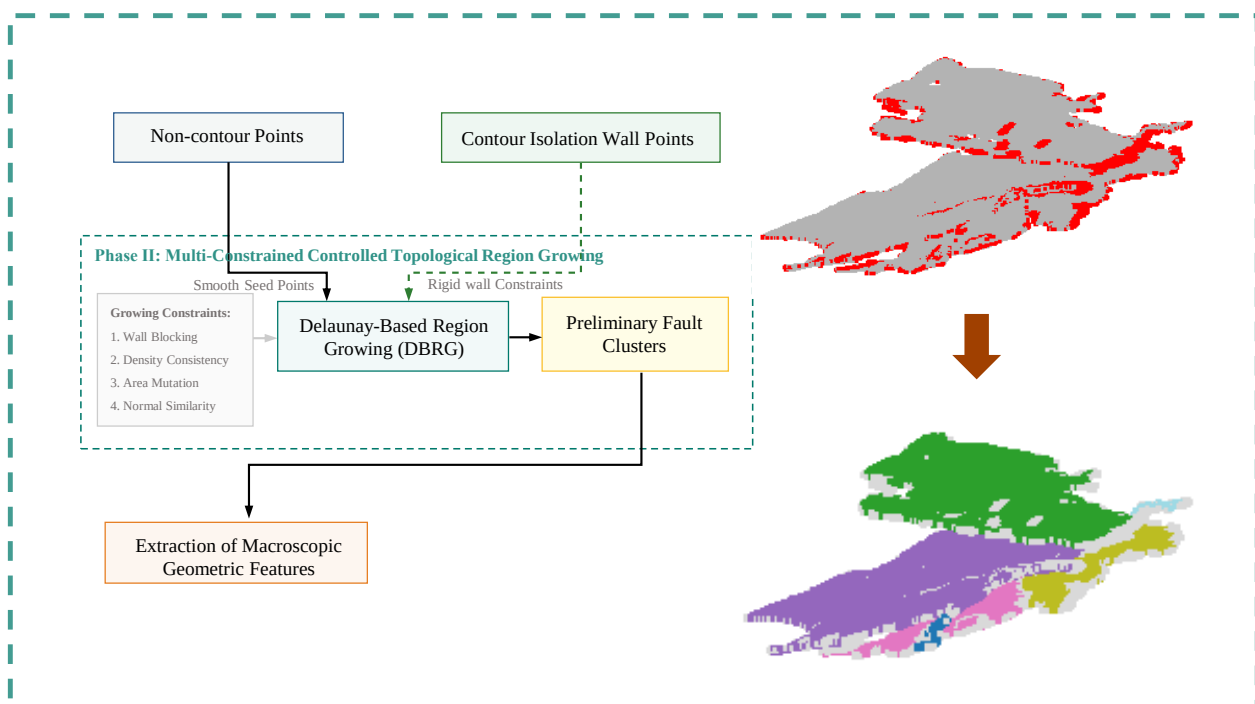


Figure 4. Workflow of the multi-constraint controlled topological region growing phase.

2.3.2. Four Controlled Spatial Topological Constraints

Although DBRG offers strong connectivity, its reliance on the local smoothness assumption can cause spatial adhesion between adjacent fault instances if left unregulated. To mitigate this, we strictly limit the initial seed points to the smooth non-contour point set extracted in the first stage. During the region-growing traversal, candidate adjacent points are selected from the entire unvisited local Delaunay neighborhood, rather than just the smooth set. We design one hard physical blocking constraint and three soft manifold constraints to adaptively monitor and guide the controlled growth process. It is crucial to emphasize that for a candidate point to be successfully merged into the current fault instance, it must simultaneously satisfy all three soft constraints (a logical AND relationship) without triggering the hard barrier:

- **Hard Physical Blocking (Contour Walls):** We treat the contour point set identified in the first stage as a rigid boundary. During traversal, if an unvisited candidate point is labeled as a contour wall node, the algorithm halts expansion in that direction and bypasses the point. This physically severs paths that lead to cross-fault adhesion.
- **Soft Constraint 1: Local Density Consistency.** To prevent growth into sparse noise regions, the algorithm monitors density fluctuations using the edge lengths of the topological network. Specifically, the spatial distance $\text{dist}(p_c, p_n)$ from the current node p_c to a candidate p_n must not exceed τ_d times the median edge length within its first-order neighborhood:

$$\text{dist}(p_c, p_n) \leq \tau_d \cdot \text{median}(\{\text{dist}(p_c, p_k) \mid p_k \in N_{\text{Delaunay}}(p_c)\}). \quad (7)$$

- **Soft Constraint 2: Topological Area Mutation.** This constraint captures discontinuities at the manifold scale. We define a local topological feature area using the squared edge length, $\Delta S_{\text{new}} = \text{dist}^2(p_c, p_n)$. Once the growth history exceeds a threshold, the area of a newly expanded patch must satisfy:

$$\Delta S_{\text{new}} \leq \tau_a \cdot \frac{1}{M} \sum_{m=1}^M \Delta S_{t-m}. \quad (8)$$

where the right side represents the average area of historical patches within a sliding window of size M . Exceeding the mutation threshold τ_a indicates a cross-fault misconnection, prompting the system to reject the extension.

- **Soft Constraint 3: Directional Similarity.** To ensure a smooth transition of geological strike and dip within a single fault instance, we require the absolute dot product between the normal vector n_c of the current node and n_n of the candidate to stay within a predefined strike tolerance θ_{thresh} :

$$|\langle n_c, n_n \rangle| \geq \cos(\theta_{\text{thresh}}). \quad (9)$$

This guarantees smooth coplanar extension.

Following this rigorously constrained DBRG process, the fault main bodies acquire precise discrete instance labels. The scattered points blocked by these constraints, alongside the contour wall points from the first stage, are deferred to subsequent stages for manifold merging and attribution backfilling.

The detailed execution logic of the aforementioned topological propagation, strictly regulated by the physical barriers and manifold conditions, is summarized in Algorithm 2.

Algorithm 2 MC-DBRG Phase 2: Multi-Constraint Controlled Topological Region Growing

Require: Downsampled point cloud P ; Global Delaunay network $\mathcal{G}_{\text{Del}}$ from Phase 1; Label array Label from Phase 1; Local density mutation threshold τ_d ; Topological area mutation threshold τ_a ; Directional strike tolerance angle θ_{thresh} ; Sliding history window size M .

Ensure: Set of preliminary fault core clusters $\mathcal{P}_{\text{cores}} = \{P_1, P_2, \dots, P_K\}$.

```

1:  $\mathcal{P}_{\text{unvisited}} \leftarrow P$  ▷ Initialize unvisited pool with all points
2:  $\mathcal{P}_{\text{smooth}} \leftarrow \{p \in P \mid \text{Label}(p) == \text{Smooth}\}$ 
3:  $\mathcal{P}_{\text{cores}} \leftarrow \emptyset$ 
4: while there exists an unvisited seed  $p_{\text{seed}} \in \mathcal{P}_{\text{smooth}} \cap \mathcal{P}_{\text{unvisited}}$  do
5:    $P_c \leftarrow \{p_{\text{seed}}\}, Q \leftarrow [p_{\text{seed}}]$ 
6:    $\mathcal{P}_{\text{unvisited}} \leftarrow \mathcal{P}_{\text{unvisited}} \setminus \{p_{\text{seed}}\}$ 
7:   Initialize sliding window  $W_{\text{history}}$  for historical patch areas
8:   while  $Q$  is not empty do
9:      $p_c \leftarrow \text{Dequeue}(Q)$ 
10:    for  $p_n \in \text{Delaunay}(\mathcal{G}_{\text{Del}}, p_c)$  where  $p_n \in \mathcal{P}_{\text{unvisited}}$  do
11:      if  $\text{Label}(p_n) == \text{Contour}$  then
12:        continue ▷ Hard physical wall barrier blocking
13:      end if
14:      ▷ Evaluate the three soft manifold constraints:
15:       $C_1 \leftarrow \text{dist}(p_c, p_n) \leq \tau_d \cdot \text{median}(\{\text{dist}(p_c, p_k) \mid p_k \in N_{\text{Del}}(p_c)\})$ 
16:       $C_2 \leftarrow \text{dist}^2(p_c, p_n) \leq \tau_a \cdot \text{average}(W_{\text{history}})$ 
17:       $C_3 \leftarrow |\langle n_c, n_n \rangle| \geq \cos(\theta_{\text{thresh}})$ 
18:      if  $C_1$  and  $C_2$  and  $C_3$  then
19:         $P_c \leftarrow P_c \cup \{p_n\}$ 
20:         $\mathcal{P}_{\text{unvisited}} \leftarrow \mathcal{P}_{\text{unvisited}} \setminus \{p_n\}$ 
21:         $\text{Enqueue}(Q, p_n)$ 
22:        Update  $W_{\text{history}}$  with new area  $\text{dist}^2(p_c, p_n)$ 
23:      end if
24:    end for
25:  end while
26:   $\mathcal{P}_{\text{cores}} \leftarrow \mathcal{P}_{\text{cores}} \cup \{P_c\}$ 
27: end while
28: return  $\mathcal{P}_{\text{cores}}$ 

```

2.4. Global Manifold Secondary Geometric Merging

2.4.1. Causes of Over-Segmentation and Motivation for Merging

The region growing process in the second phase is strictly confined to the non-contour smooth point set. The edge isolation walls constructed in the first phase act as physical barriers at complex intersections or transition zones, segregating smooth regions that belong to the same primary fault manifold. Because the region growing wavefront cannot cross these walls, it generates spatially isolated discrete sub-clusters within a single fault, resulting in over-segmentation. This enforced over-segmentation is a deliberate “break-then-build” strategy designed to prevent irreversible erroneous cross-connections (under-segmentation) at chaotic intersections. By initially sacrificing large-scale connectivity, we ensure that every generated sub-cluster remains geometrically pure. To ultimately restore the topological connectivity of the main fault from a macroscopic perspective, we introduce a global manifold secondary geometric merging strategy based on a disjoint-set data structure. This approach fuses spatially scattered point cloud clusters that share consistent macroscopic features (details are shown in the Figure 5).

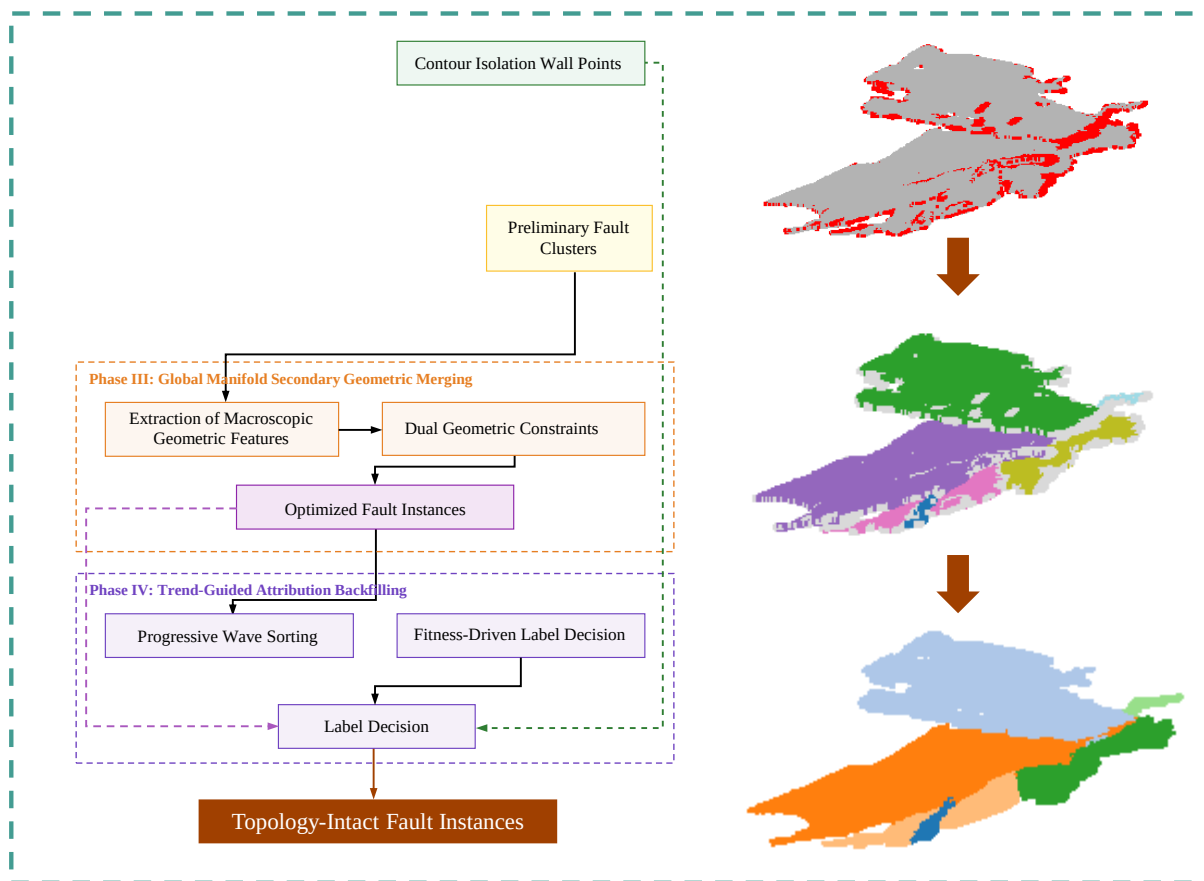


Figure 5. Workflow of the global manifold secondary geometric merging and trend-guided attribution backfilling phase. Distinct colors in the fault point cloud on the right indicate that they have been segmented into different instance classes.

2.4.2. Extraction of Macroscopic Geometric Features

For any valid fault point cloud cluster P_i segmented in the second phase, we compute its macroscopic geometric manifold features using Principal Component Analysis. We define the macroscopic geometric centroid as the arithmetic mean of the coordinates of all points in the cluster. This representation is robust to local geological noise and is formalized as

$$\mathbf{c}_i = \frac{1}{|P_i|} \sum_{\mathbf{p} \in P_i} \mathbf{p}. \quad (10)$$

To obtain the macroscopic normal vector, we construct the global covariance matrix of the cluster and perform eigenvalue decomposition. The eigenvector corresponding to the minimum eigenvalue serves as the global principal normal vector $\mathbf{n}_i$ of the fault plane. Jointly fitting all points in the cluster avoids the ambiguity associated with microscopic single-point normals.

2.4.3. Dual Geometric Constraints

We initialize a disjoint-set data structure to maintain the topological relationships among multiple fault clusters, evaluating the relative geometry between any two clusters P_i and P_j using two constraints. The first constraint ensures that their spatial intersection angle falls within a defined strike tolerance θ_{merge} . Taking the absolute dot product of their macroscopic normal vectors accounts for opposite normal orientations in 3D space, accurately identifying parallel candidate planes:

$$|\langle \mathbf{n}_i, \mathbf{n}_j \rangle| \geq \cos(\theta_{\text{merge}}). \quad (11)$$

The second constraint defines a coplanar projection error d_{proj} by taking the maximum projection magnitude of the centroid connection vector $\mathbf{v}_{ij} = \mathbf{c}_j - \mathbf{c}_i$ onto both normal vectors:

$$d_{\text{proj}} = \max(|\langle \mathbf{v}_{ij}, \mathbf{n}_i \rangle|, |\langle \mathbf{v}_{ij}, \mathbf{n}_j \rangle|). \quad (12)$$

If this error satisfies $d_{\text{proj}} \leq d_{\text{max}}$, P_j is considered to lie on the virtual plane extended from P_i . When both constraints are met, the clusters are merged within the disjoint-set, achieving global manifold reconstruction from the discrete fragments.

The systematic procedure for fusing over-segmented fragments and extracting robust macroscopic properties through these dual geometric constraints is presented in Algorithm 3.

Algorithm 3 MC-DBRG Phase 3: Global Manifold Secondary Geometric Merging

Require: Set of preliminary fault core clusters $\mathcal{P}_{\text{cores}}$ from Phase 2; Macroscopic merge strike tolerance θ_{merge} ; Maximum coplanar projection error d_{max} .

Ensure: Set of merged fault clusters $\mathcal{P}_{\text{merged}}$; Macroscopic properties for each merged cluster l : global centroids $\mathbf{c}_l$ and principal normal vectors $\mathbf{n}_l$.

```

1: Initialize Disjoint-Set Data Structure  $\mathcal{DS}$  with all clusters in  $\mathcal{P}_{\text{cores}}$ 
2: for  $P_i, P_j \in \mathcal{P}_{\text{cores}}$  ( $i \neq j$ ) do
3:   Compute macroscopic centroids  $\mathbf{c}_i, \mathbf{c}_j$  via arithmetic coordinate averaging
4:   Compute global principal normals  $\mathbf{n}_i, \mathbf{n}_j$  using standard PCA
5:   if  $|\langle \mathbf{n}_i, \mathbf{n}_j \rangle| \geq \cos(\theta_{\text{merge}})$  then
6:      $\mathbf{v}_{ij} \leftarrow \mathbf{c}_j - \mathbf{c}_i$ 
7:      $d_{\text{proj}} \leftarrow \max(|\langle \mathbf{v}_{ij}, \mathbf{n}_i \rangle|, |\langle \mathbf{v}_{ij}, \mathbf{n}_j \rangle|)$ 
8:     if  $d_{\text{proj}} \leq d_{\text{max}}$  then
9:        $\mathcal{DS}.\text{Union}(P_i, P_j)$   $\triangleright$  Merge geologically continuous isolated sub-clusters
10:    end if
11:  end if
12: end for
13:  $\mathcal{P}_{\text{merged}} \leftarrow \mathcal{DS}.\text{Clusters}()$ 
14: for  $P_l \in \mathcal{P}_{\text{merged}}$  do
15:   Compute final global centroid  $\mathbf{c}_l$  and principal normal  $\mathbf{n}_l$ 
16: end for
17: return  $\mathcal{P}_{\text{merged}}, \{\mathbf{c}_l\}, \{\mathbf{n}_l\}$ 

```

2.5. Trend-Guided Attribution Backfilling

2.5.1. Challenges of Backfilling at Thick Intersecting Walls

Following topological growth and secondary merging, the primary fault manifolds are established. However, many unclassified points remain, consisting primarily of the edge isolation wall points identified in the first phase and scattered points that triggered soft constraints. Traditional nearest-neighbor attribution methods, such as K-nearest neighbor voting, lack macroscopic trend guidance, frequently causing label confusion and mismatching in spatially ambiguous boundary zones. To resolve this, we introduce an outside-in progressive wave stripping mechanism constrained by global macroscopic planar trends to achieve precise label backfilling (details are shown in the Figure 5).

2.5.2. Wave Propagation Sorting Mechanism

To propagate backfilling decisions smoothly, we construct a global KD-Tree from the established fault body point sets. The algorithm calculates the nearest Euclidean distance $d_{\text{init}}(\mathbf{p}_u)$ from all unprocessed contour points $\mathbf{p}_u$ to the surfaces of the formed fault bodies. It then sorts the unprocessed points by ascending distance to form a wave-like evolution queue:

$$\text{Sort}(\{\mathbf{p}_u\}) \quad \text{s.t.} \quad d_{\text{init}}(\mathbf{p}_1) \leq d_{\text{init}}(\mathbf{p}_2) \leq \dots \leq d_{\text{init}}(\mathbf{p}_M). \quad (13)$$

This queue prioritizes easily separable points closest to the fault plane extensions. As the queue advances, label assignment encroaches on the walls from the outside in. Delaying the assignment of structurally complex multi-surface intersection centers until the final stage minimizes ambiguity during label propagation.

2.5.3. Label Decision Strategy

For a target point $\mathbf{p}_{\text{target}}$ in the queue, the algorithm retrieves its K nearest classified neighbors via the KD-Tree to identify the set of candidate fault instances L_{can} within the local space. For each candidate label $l \in L_{\text{can}}$, we define a fitness function:

$$\text{Fitness}(l) = \frac{1}{d_{\min,l} + w \cdot d_{\text{macro_plane},l} + \epsilon}. \quad (14)$$

Here, $d_{\min,l}$ is the minimum Euclidean distance from the target point to the points of class l in the local neighborhood, and ϵ is a small positive constant to prevent division by zero. The term $d_{\text{macro_plane},l}$ represents the vertical projection distance from the target point to the fitted plane of the candidate fault, evaluating its alignment with the extrapolated trend line of the fault manifold:

$$d_{\text{macro_plane},l} = |\langle \mathbf{p}_{\text{target}} - \mathbf{c}_l, \mathbf{n}_l \rangle|. \quad (15)$$

The variables $\mathbf{c}_l$ and $\mathbf{n}_l$ are the global centroid and principal normal of the candidate fault instance, respectively. In ambiguous thick-wall regions with intersecting surfaces, we assign a dominant weight w to the global trend term. The target point is assigned to the fault instance with the highest fitness score:

$$\text{Label}(\mathbf{p}_{\text{target}}) = \arg \max_{l \in L_{\text{can}}} \text{Fitness}(l). \quad (16)$$

This multi-dimensional cost function accurately backfills complex boundary points into their respective faults, yielding a final 3D fault instance segmentation model with complete topology and precise labels.

The complete pipeline of this trend-guided attribution backfilling strategy, which assigns the final instance labels to the remaining point clouds from the outside in, is formalized in Algorithm 4.

Algorithm 4 MC-DBRG Phase 4: Trend-Guided Attribution Backfilling

Require: Downsampled point cloud P from Phase 1; Set of merged fault clusters $\mathcal{P}_{\text{merged}}$ from Phase 3; Global centroids $\{\mathbf{c}_l\}$ and principal normal vectors $\{\mathbf{n}_l\}$ from Phase 3; Backfilling retrieval neighborhood size K ; Global trend dominant weight w ; Small positive constant ϵ .

Ensure: Final 3D fault point cloud instance labels L for all points in P .

- 1: Construct global spatial index $\mathcal{T}_{\text{KD}}$ using all points contained in $\mathcal{P}_{\text{merged}}$
 - 2: $\mathcal{P}_{\text{unlabeled}} \leftarrow \{p \in P \mid p \text{ has no assigned instance label}\}$
 - 3: Sort $\mathcal{P}_{\text{unlabeled}}$ into wave-propagation queue Q_{wave} ascending by Euclidean distance to $\mathcal{P}_{\text{merged}}$
 - 4: **while** Q_{wave} is not empty **do**
 - 5: $\mathbf{p}_{\text{target}} \leftarrow \text{Dequeue}(Q_{\text{wave}})$
 - 6: $L_{\text{can}} \leftarrow \text{Labels}(\mathcal{T}_{\text{KD}}, \mathbf{p}_{\text{target}}, K)$
-

Algorithm 4 *Cont.*

```

7:   for  $l \in L_{\text{can}}$  do
8:     Compute  $d_{\min,l} \leftarrow \min_{\mathbf{p} \in P_l} \|\mathbf{p}_{\text{target}} - \mathbf{p}\|$ 
9:     Compute  $d_{\text{macro\_plane},l} \leftarrow |\langle \mathbf{p}_{\text{target}} - \mathbf{c}_l, \mathbf{n}_l \rangle|$ 
10:     $\text{Fitness}(l) \leftarrow \frac{1}{d_{\min,l} + w \cdot d_{\text{macro\_plane},l} + \epsilon}$   $\triangleright$  Evaluate macro trend fitness penalty
11:  end for
12:   $\text{Label}(\mathbf{p}_{\text{target}}) \leftarrow \arg \max_{l \in L_{\text{can}}} \text{Fitness}(l)$ 
13:  Update  $\mathcal{T}_{\text{KD}}$  by inserting  $\mathbf{p}_{\text{target}}$  with its assigned label
14: end while
15:  $L \leftarrow \{\text{Label}(p) \mid p \in P\}$ 
16: return  $L$ 

```

3. Results and Discussion

To evaluate the segmentation accuracy and topological integrity of the MC-DBRG framework, we conduct comparative experiments on synthetic and real seismic fault datasets. All experiments and computations are performed on a workstation equipped with an NVIDIA GeForce RTX 5070Ti GPU (16 GB VRAM) (NVIDIA, Santa Clara, CA, USA) and an AMD Ryzen 7 9700X 8-Core CPU @ 3.80 GHz (Santa Clara, CA, USA).

3.1. Experimental Setup and Unsupervised Geometric Evaluation System

3.1.1. Synthetic Dataset and Environment Configuration

We validate our algorithm using a large-scale synthetic seismic fault dataset originally proposed by Wu et al. [38]. This dataset contains 200 independent fault volumes, which we crop into $128 \times 128 \times 128$ voxel grids. This specific grid size is deliberately chosen to strike an optimal balance between preserving the spatial continuity of geological fault structures and avoiding out-of-memory errors. The voxels store continuous fault probabilities generated by the front-end network, featuring complex Y- and X-shaped intersecting fault systems with varying degrees of spatial noise. It should be noted that while synthetic datasets theoretically allow for the retention of instance labels, this specific dataset was originally designed for binary segmentation. Consequently, it only provides 0-1 binary ground truth labels (fault versus background) and lacks absolute instance-level identifiers. During preprocessing, we convert high-confidence voxels (probability > 0.5) into an initial spatial point cloud P_{input} and apply voxel downsampling ($v_s = 3.0$) to reduce high-frequency sampling redundancy. All experiments run independently on the same computing platform.

3.1.2. Baseline Algorithms and Unsupervised Evaluation Metrics

We compare our method against three representative baseline algorithms. Mean-Shift clustering is a non-parametric density estimation method that adaptively partitions point cloud clusters using local density gradients. Multi-stage RANSAC provides an iterative extraction approach for regular planes, constructing global mathematical plane hypotheses and stripping inliers while scaling its distance threshold to the average point cloud spacing. Finally, Delaunay-based region growing utilizes the first-order adjacency of a global Delaunay triangulation and evolves the graph according to normal vector similarity.

Because subsurface faults lack absolute instance ground truth, we design an unsupervised geometric evaluation system based on intracluster manifold reconstruction. For each segmented fault cluster, we define a local neighborhood to compute the covariance matrix, extracting its eigenvalues and local normal vectors to calculate three manifold metrics:

- Intracluster Curvature Volatility (ICV) measures the spatial consistency of surface smoothness. Defining the local curvature of a sampled point as $\sigma = \lambda_3 / (\lambda_1 + \lambda_2 + \lambda_3)$, ICV represents the variance of these curvatures within the cluster:

$$\text{ICV} = \text{Var}(\{\sigma_k \mid p_k \in P_c\}). \quad (17)$$

A lower ICV indicates the absence of severe geometric mutations, confirming that the segmentation isolates boundaries without incorporating scattered noise at intersections.

- Intracluster Coplanar Consistency (ICC) evaluates the geometric alignment of microscopic directional manifolds. We compute the dominant normal vector n_{global} of the macroscopic plane via global eigenvalue decomposition. ICC is the arithmetic mean of the absolute cosine values between the local and global normal vectors:

$$\text{ICC} = \frac{1}{N_c} \sum_{k=1}^{N_c} |\langle n_{\text{local},k}, n_{\text{global}} \rangle|. \quad (18)$$

Values closer to 1.0 indicate better normal alignment and stronger manifold continuity.

- Geometrical Anisotropy Index (GAI) measures how closely the macroscopic manifold converges to a standard 2D geological plane. Using the non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ from the global covariance matrix of P_c , we define:

$$\text{GAI} = \frac{\lambda_2 - \lambda_3}{\lambda_1}. \quad (19)$$

While our pipeline utilizes PCA to guide local microscopic sub-clusters, GAI is applied strictly as a macroscopic global metric on the final merged fault instances. It independently evaluates whether the overall spatial merging strategy successfully prevented global under-segmentation or over-segmentation. A value approaching 1.0 indicates a highly continuous, regular surface, thereby quantifying the algorithm's ability to prevent fragmentation.

3.2. Performance of Different Algorithms on the Synthetic Dataset

Based on the unsupervised evaluation system, the Table 1 presents the global average quantitative metrics for the four algorithms across 200 synthetic seismic fault test samples, while the Figure 6 displays the corresponding segmentation results. For instance segmentation purity and manifold preservation, the proposed MC-DBRG method achieves the lowest curvature volatility with an ICV of 5.94×10^{-4} . In contrast, Region Growing and DBRG both yield an ICV of 1.53×10^{-3} . Traditional methods lack explicit boundary blocking, causing their growth wavefronts to suffer from topological leakage at Y-shaped and X-shaped intersection ridges. This leakage erroneously stitches intersecting planes with different orientations into a single instance, a limitation clearly visible in the segmentation results where general region growing algorithms fail to separate intersecting planes. By utilizing physical isolation walls reinforced through Gaussian Mixture Model soft-threshold confidence control and Delaunay topological closing operations, our method effectively locks intersection boundaries and maintains intra-instance manifold consistency.

In terms of strike smoothness and normal fidelity, the proposed method attains the highest coplanar consistency score of 0.986, surpassing the explicit plane-fitting approach of Multi-RANSAC at 0.984. When processing complex fault systems characterized by noise and local non-linear bending, Multi-RANSAC frequently suffers from outlier mismatches and fragmentation. Our global manifold secondary geometric merging strategy, combined with the trend-guided stripping and backfilling mechanism, successfully reconstructs the

over-segmented fragments caused by the isolation walls. This combination also ensures that boundary thick-wall points are accurately assigned around the macroscopic extrapolated trend line of the fault.

Table 1. Global average quantitative results on 200 synthetic seismic fault volume test samples using four segmentation methods are presented, where ICV denotes intracluster curvature volatility (lower is better), ICC denotes intracluster coplanar consistency (higher is better), and GAI denotes geometrical anisotropy index (higher is better).

Methods	ICV	ICC	GAI
Mean-Shift	1.59×10^{-3}	0.937	0.422
Multi-RANSAC	1.14×10^{-3}	0.984	0.686
DBRG	1.53×10^{-3}	0.940	0.231
MC-DBRG	5.94×10^{-4}	0.986	0.758

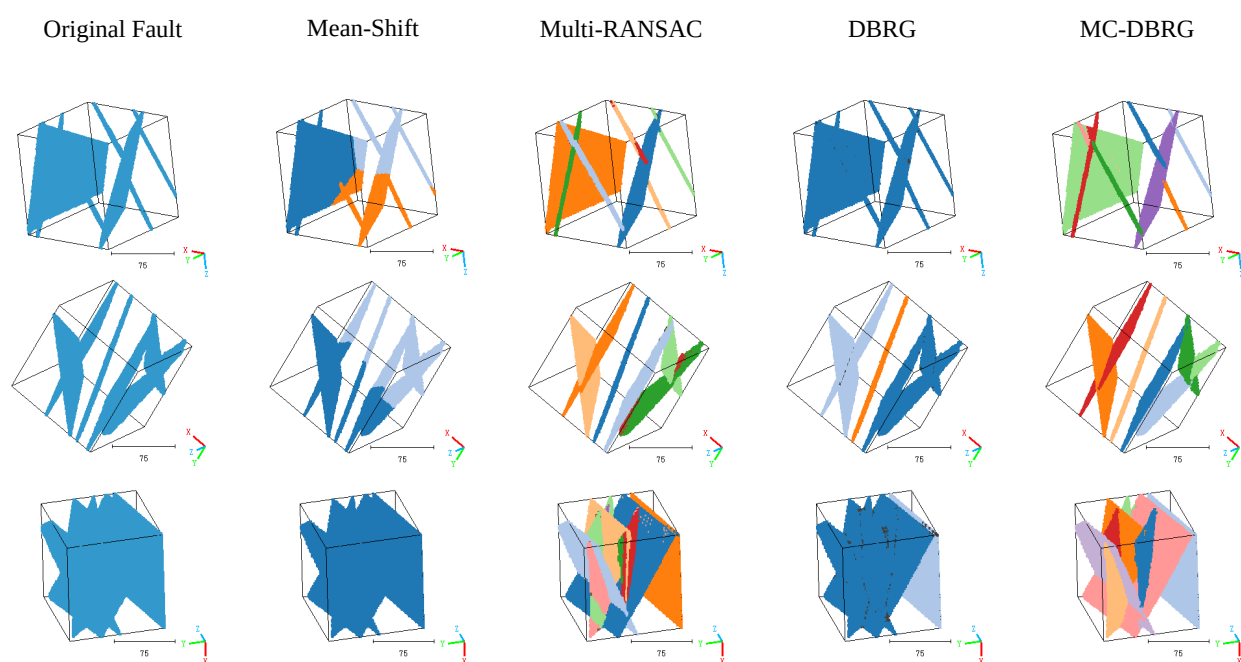


Figure 6. Segmentation results of four methods on selected synthetic seismic fault volume test samples, where different colors represent different fault instances. Distinct colors in the fault point cloud indicate that they have been segmented into different instance classes.

Regarding the standardization of macroscopic spatial morphology, MC-DBRG achieves the highest geometrical anisotropy index of 0.758, significantly outperforming Mean-Shift (0.422) and DBRG (0.231). It also surpasses Multi-RANSAC (0.686), which attains its score largely due to an underlying mathematical assumption that forces absolute plane fitting. Geological faults, however, naturally exhibit spatial bending, and forced planarization smooths out these essential non-linear geological features. Consequently, segmentation results on real datasets (Figure 7) expose the disastrous performance and limited generalization capability of Multi-RANSAC. By successfully accommodating the natural bending of fault manifolds, our method not only secures the highest quantitative score but also demonstrates clear visual superiority on real datasets. These findings indicate that the fault models generated by our method macroscopically converge to regular spatial surfaces with high aspect ratios, preserving excellent topological integrity to provide high-fidelity skeleton models for subsequent automated geological modeling.

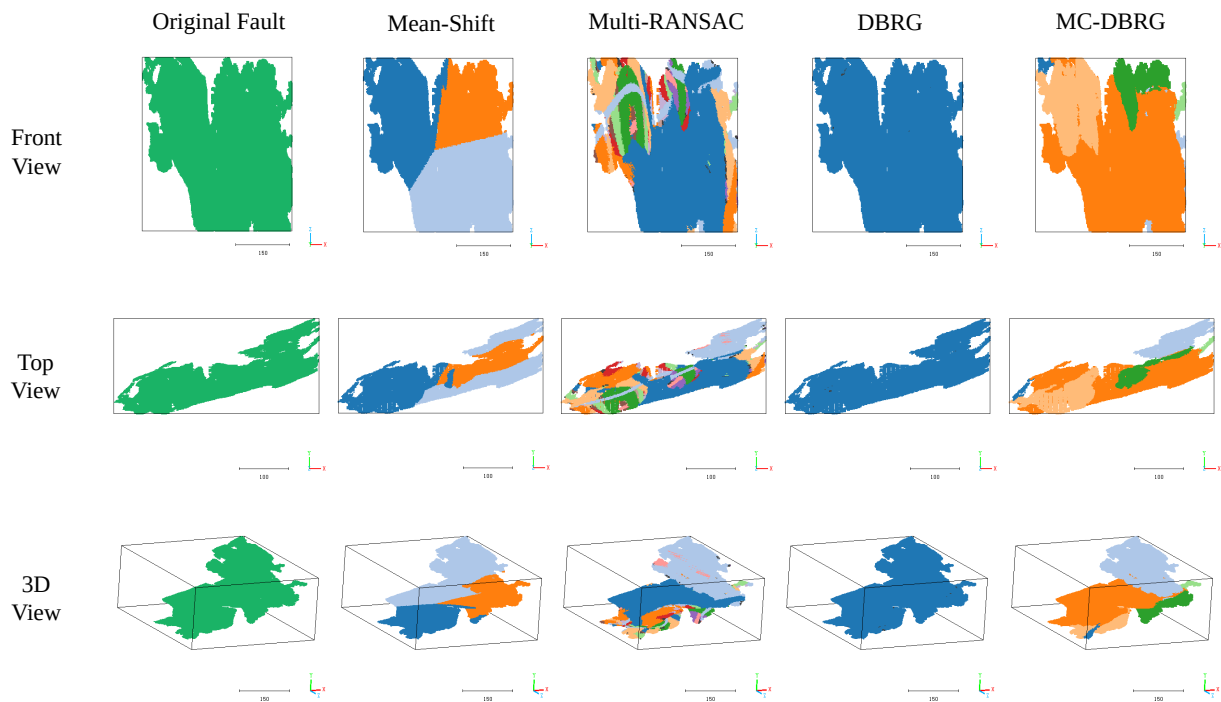


Figure 7. Segmentation results of four methods on a real fault dataset (size $415 \times 193 \times 479$), displayed in front, top, and 3D views. Distinct colors in the fault point cloud indicate that they have been segmented into different instance classes.

3.3. Parameter Sensitivity Analysis

To understand how core hyperparameters affect the multi-surface instance segmentation performance, we conduct a univariate sensitivity analysis on key geometric and topological variables, as shown in Table 2. The default hyperparameter configuration includes a feature extraction neighborhood K_{extract} of [15, 25, 35], a contour extraction confidence threshold τ of 0.75, a topological closing voting threshold α_{vote} of 0.5, a directional similarity angle θ_{thresh} of 25° , an area mutation threshold τ_a of 2.5, a density mutation threshold τ_d of 1.2, a backfilling retrieval neighborhood K_{backfill} of 30, and a trend dominant weight w_{trend} of 10.

Parameters in the contour point recognition and filling stage determine the construction accuracy and spatial closure of the physical isolation walls, while the local feature extraction neighborhood defines the window size for local covariance matrix estimation. As the neighborhood size increases, all metrics deteriorate. This confirms that an excessively large neighborhood window causes severe over-smoothing, blurring the local spatial heterogeneity at multi-surface fault intersections and impairing the internal manifold purity. However, varying the contour confidence threshold τ and the topological closing voting threshold α_{vote} has minimal impact on the synthetic dataset, likely because it contains less noise.

Growth stage parameters guide the spatial evolution of the region growing wave within the closed topological manifold. Varying the directional similarity angle θ_{thresh} , the area mutation threshold τ_a , and the density mutation threshold τ_d exerts little influence on the overall segmentation results. This is primarily because the growth stage operates on isolated fault point clouds that have already been separated by the isolation walls.

Parameters in the secondary merging and backfilling stage determine the connectivity of the macroscopic manifold reconstruction and the label fidelity of the thick-wall boundary attribution. As the backfilling retrieval neighborhood K_{backfill} increases from 10 to 30, all three evaluation metrics improve consistently. This indicates that expanding the retrieval

neighborhood provides a more robust label pool, enabling ambiguous boundary points to absorb the macroscopic skeleton information of the formed fault bodies and filtering out interference from isolated noise. Furthermore, setting the trend dominant weight w_{trend} to an extremely low value of 0.5 significantly degrades segmentation quality due to the weakened macroscopic trend constraint. This suggests that relying solely on spatial nearest-neighbor distances for backfilling in ambiguous thick-wall intersection zones leads to label mismatches. When w_{trend} increases to 10, the ICV metric improves substantially, dropping to 5.94×10^{-4} . This confirms the value of introducing a global trend vertical extrapolation penalty into the fitness function.

Table 2. Quantitative results of the univariate sensitivity analysis on key geometric and topological hyperparameters. Here, K_{extract} denotes the feature extraction neighborhood, τ is the contour extraction confidence threshold, α_{vote} represents the topological closing voting threshold, θ_{thresh} is the directional similarity angle, τ_a and τ_d denote the area and density mutation thresholds respectively, K_{backfill} represents the backfilling retrieval neighborhood, and w_{trend} is the trend dominant weight.

Parameter	Value/Range	ICV	ICC	GAI
K_{extract}	[10, 15, 20]	5.53×10^{-4}	0.988	0.756
	[15, 25, 35]	5.94×10^{-4}	0.986	0.758
	[30, 40, 50]	7.05×10^{-4}	0.983	0.737
τ	0.5	5.99×10^{-4}	0.986	0.758
	0.75	5.94×10^{-4}	0.986	0.758
	0.9	5.91×10^{-4}	0.986	0.758
α_{vote}	0.35	6.34×10^{-4}	0.983	0.742
	0.5	5.94×10^{-4}	0.986	0.758
	0.75	5.31×10^{-4}	0.989	0.761
θ_{thresh}	15°	5.76×10^{-4}	0.987	0.759
	25°	5.94×10^{-4}	0.986	0.758
	35°	6.08×10^{-4}	0.986	0.757
τ_a	1.5	5.72×10^{-4}	0.987	0.754
	2.5	5.94×10^{-4}	0.986	0.758
	3.5	6.06×10^{-4}	0.986	0.756
τ_d	1.1	5.72×10^{-4}	0.987	0.755
	1.2	5.94×10^{-4}	0.986	0.758
	1.3	6.12×10^{-4}	0.986	0.756
K_{backfill}	10	6.83×10^{-4}	0.984	0.750
	20	6.27×10^{-4}	0.985	0.755
	30	5.94×10^{-4}	0.986	0.758
w_{trend}	0.5	7.95×10^{-4}	0.983	0.752
	2	6.27×10^{-4}	0.986	0.757
	10	5.94×10^{-4}	0.986	0.758

3.4. Structural Contribution of Proposed Modules

To rigorously evaluate the structural contributions and physical rationale of each proposed module, an ablation study is conducted on the synthetic dataset. As detailed in Table 3, the framework is evaluated through progressive module additions. The baseline method utilizes the basic DBRG without geometric constraints. The second configuration introduces the Gaussian Mixture Model (GMM) boundary constraint. The third configuration adds secondary geometric merging. Finally, the complete framework includes label backfilling.

The baseline method, relying solely on the basic DBRG algorithm, operates efficiently with an average computational time of 1.67 s and a low Unclassified Ratio (UR) of 0.46%. However, the structural quality of this baseline is poor, yielding the highest ICV of 1.53×10^{-3} and the lowest GAI of 0.231. These results indicate severe under-segmentation and topological leakage at complex intersections, where the algorithm greedily connects distinct fault planes.

By incorporating the GMM boundary constraint, a strict physical isolation wall is enforced. This addition substantially improves the purity of the local manifold, increasing the ICC to 0.996 and decreasing the ICV to 5.50×10^{-5} . However, this conservative hard blocking strategy results in significant over-segmentation. Consequently, the UR increases to 28.08%, because transition zone points are deliberately isolated to prevent erroneous topological propagation.

The introduction of the secondary merging module addresses this fragmentation issue. This step successfully reconnects the pure sub-clusters based on macroscopic geometrical trends, which improves the GAI from 0.568 to 0.681 while maintaining the high ICC and low ICV. Nevertheless, the UR remains constant at 28.08%, as the merging process only fuses existing classified fragments without recovering the discarded contour points within the damage zones.

Finally, the wavefront label backfilling mechanism completes the framework. This step reassigns the unclassified contour points back to the primary fault planes by leveraging macroscopic projection distances. Consequently, the UR drops precipitously from 28.08% to 0%, resolving the missing data issue at complex thick-wall intersections. Furthermore, the GAI reaches a value of 0.758, achieving the highest macroscopic geometric integrity. Although incorporating these chaotic boundary points slightly increases the ICV to 5.94×10^{-4} and lowers the ICC to 0.986, this trade-off is necessary to ensure the physical completeness of the geological fault models. The overall computational times for the extended configurations remain stable between 5.23 and 5.42 s, demonstrating that robust accuracy improvements are achieved at a reasonable computational cost.

Table 3. Quantitative results of the progressive ablation study demonstrating the structural contribution of each proposed module. The Unclassified Ratio (UR) represents the percentage of spatial points left unassigned, and Avg Time (s) denotes the average computational time required per dataset.

Method	ICC	ICV	GAI	UR (%)	Avg Time (s)
DBRG	0.940	1.53×10^{-3}	0.231	0.46	1.67
+ GMM boundary constraint	0.996	5.50×10^{-5}	0.568	28.08	5.23
+ Secondary merging	0.996	5.60×10^{-5}	0.681	28.08	5.31
+ Label backfilling	0.986	5.94×10^{-4}	0.758	0	5.42

3.5. Performance on Real Datasets

To verify the robustness and generalization capability of the MC-DBRG framework in practical applications, we conduct experiments on two representative real 3D seismic survey datasets. The first is the classic Netherlands North Sea F3D survey dataset, which has a grid size of $951 \times 651 \times 362$ voxels. The second derives from a real 3D seismic probability volume estimated using the fault detection technique proposed by Wu et al. [38], originally sized at $201 \times 831 \times 851$ voxels. To specifically evaluate our algorithm's ability to process densely intersecting faults, we extract a $200 \times 200 \times 200$ core sub-region from the second dataset, which features complex topological structures and dense multi-surface intersections.

The F3D survey dataset poses significant challenges, as it is affected by a large-scale irregular gas chimney structure in its central region, introducing severe spatial noise into the probability volume. Furthermore, the area features multiple sets of dense conjugate

intersecting faults distributed longitudinally, where individual fault planes exhibit typical non-linear listric bending. As illustrated in Figures 8 and 9, the proposed MC-DBRG method demonstrates superior segmentation performance on this dataset. Compared to the standard Region Growing (RG) algorithm and the combined RG+RANSAC approach, our method successfully maintains the global manifold integrity of the faults despite the strong spatial noise and trace discontinuities. It fuses geologically consistent fragments into unified surfaces with high topological completeness, avoiding the fragmentation and over-segmentation commonly observed in the baseline methods.

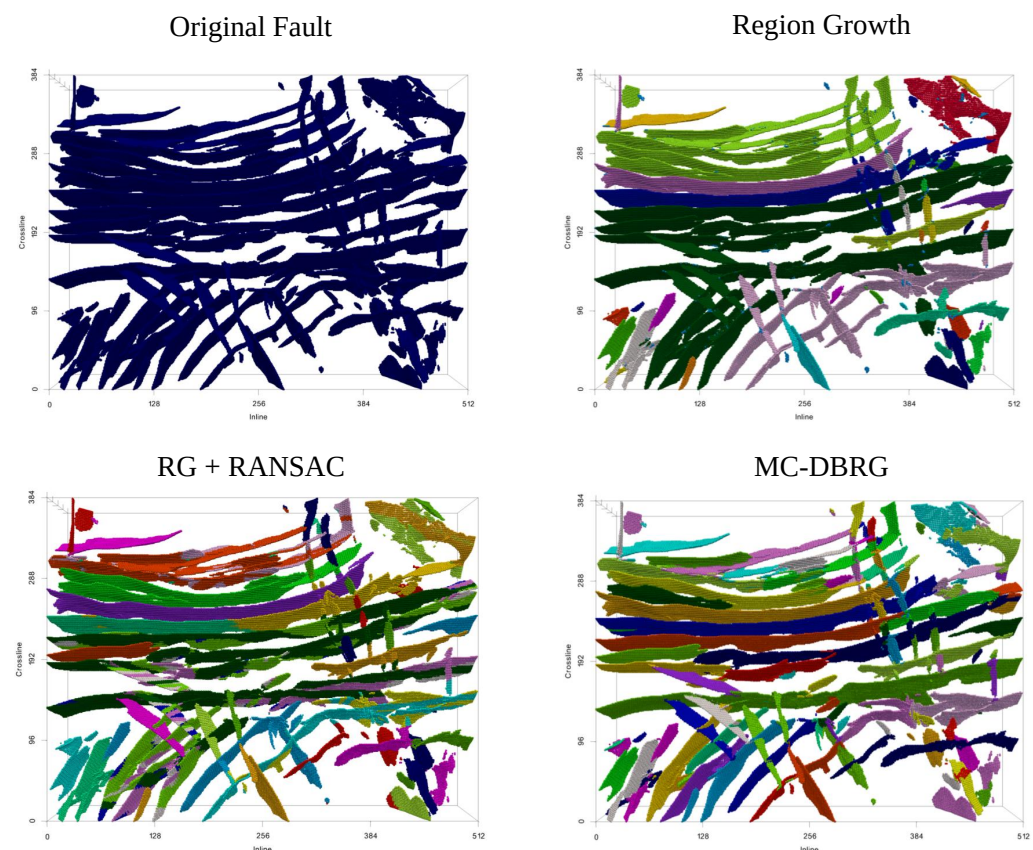


Figure 8. Segmentation results of three methods on the classic Netherlands North Sea F3D survey dataset (top view), where “RG+RANSAC” denotes the segmentation method proposed by Zou et al. [15]. Distinct colors in the fault point cloud indicate that they have been segmented into different instance classes.

For the cropped core sub-region dataset, the evaluation focuses on decoupling complex intersection networks. Although the front-end detection algorithm extracts the spatial locations of the faults, it fails to distinguish individual fault instances. As shown in Figures 10 and 11, our method exhibits clear advantages in preserving manifold independence. When processing the dense Y-shaped and X-shaped intersection ridges, traditional RG and RG+RANSAC frequently suffer from erroneous spatial merging. MC-DBRG, however, robustly separates these intersecting faults, ensuring clear boundaries and strictly independent instances.

To provide a quantitative assessment, we compute the unsupervised evaluation metrics—Intracluster Coplanarity Consistency (ICC), Intracluster Curvature Variance (ICV), Global Anisotropy Index (GAI), and Unclassified Ratio (UR)—for all three methods, as summarized in Table 4. Interestingly, the quantitative results reveal a notable discrepancy when compared with the visual evaluation. Although MC-DBRG achieves the most geologically meaningful segmentation in visual terms, it does not yield the best scores for ICC

and ICV. For instance, on the F3D dataset, the basic Region Growth method achieves the lowest ICV and a highly competitive ICC, whereas MC-DBRG yields a relatively higher ICV and a slightly lower ICC.

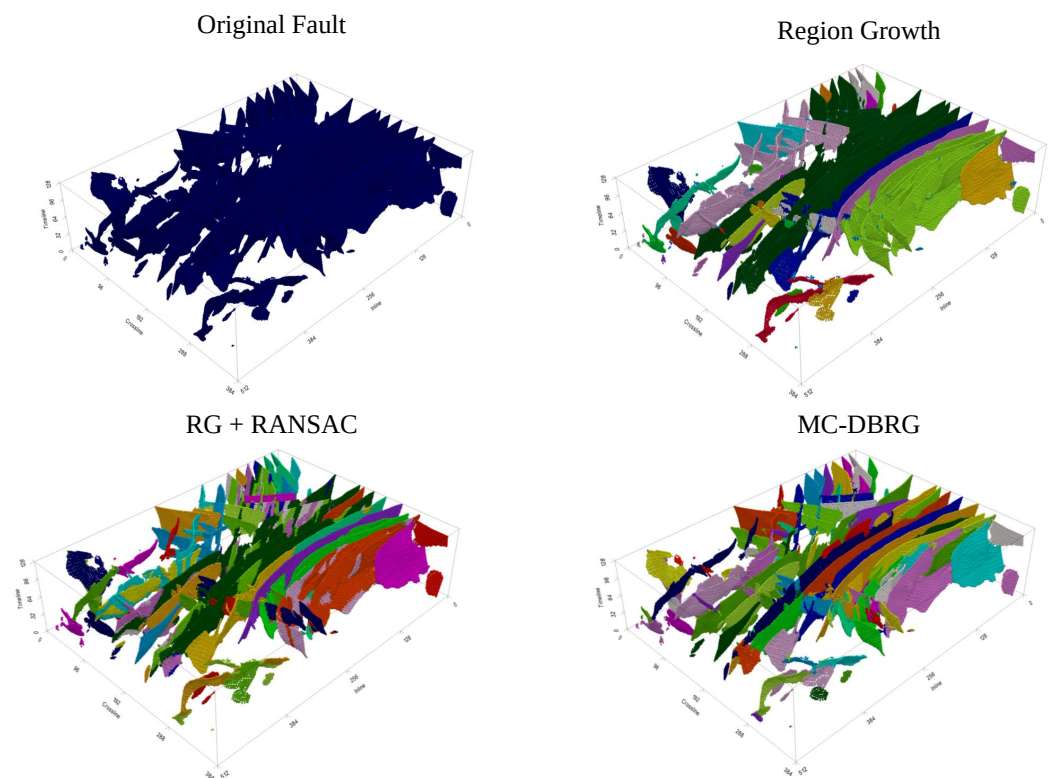


Figure 9. Segmentation results of three methods on the classic Netherlands North Sea F3D survey dataset (3D view), where “RG+RANSAC” denotes the segmentation method proposed by Zou et al. [15]. Distinct colors in the fault point cloud indicate that they have been segmented into different instance classes.

This counter-intuitive phenomenon highlights a critical limitation of these unsupervised geometric metrics when applied to real datasets. Synthetic datasets typically contain perfectly flat, isolated, and rigidly planar fault models. In contrast, real faults exhibit significant non-linear characteristics, such as listric bending, rippling, and topological folding. Unsupervised metrics inherently reward strict coplanarity. By contrast, MC-DBRG successfully recovers the complete, large-scale curved manifold of the fault. By grouping these naturally curved and undulating points into a single instance, the internal curvature variance inevitably increases, and the strict coplanarity decreases. Thus, in the context of real geological structures, traditional planar-based metrics fail to effectively reflect the true topological integrity and the validity of the segmentation methods.

Finally, regarding computational efficiency, the basic Region Growth method is extremely fast, requiring only around 4 to 5 s per dataset due to a simple local angle-checking mechanism. The RG+RANSAC approach requires significantly more time (33.57 s and 40.32 s) due to the iterative nature of multi-plane fitting and point-to-plane distance calculations. The proposed MC-DBRG method requires the highest execution time (65.39 s for the F3D dataset and 53.58 s for the A2 dataset). This increased computational cost is primarily attributed to the complex multi-scale feature extraction, GMM contour modeling, global Delaunay triangulation, and wave-front label backfilling processes. Nevertheless, considering the substantial qualitative improvements in solving dense intersections and preserving manifold integrity, this trade-off in execution time remains highly acceptable for off-line seismic interpretation tasks.

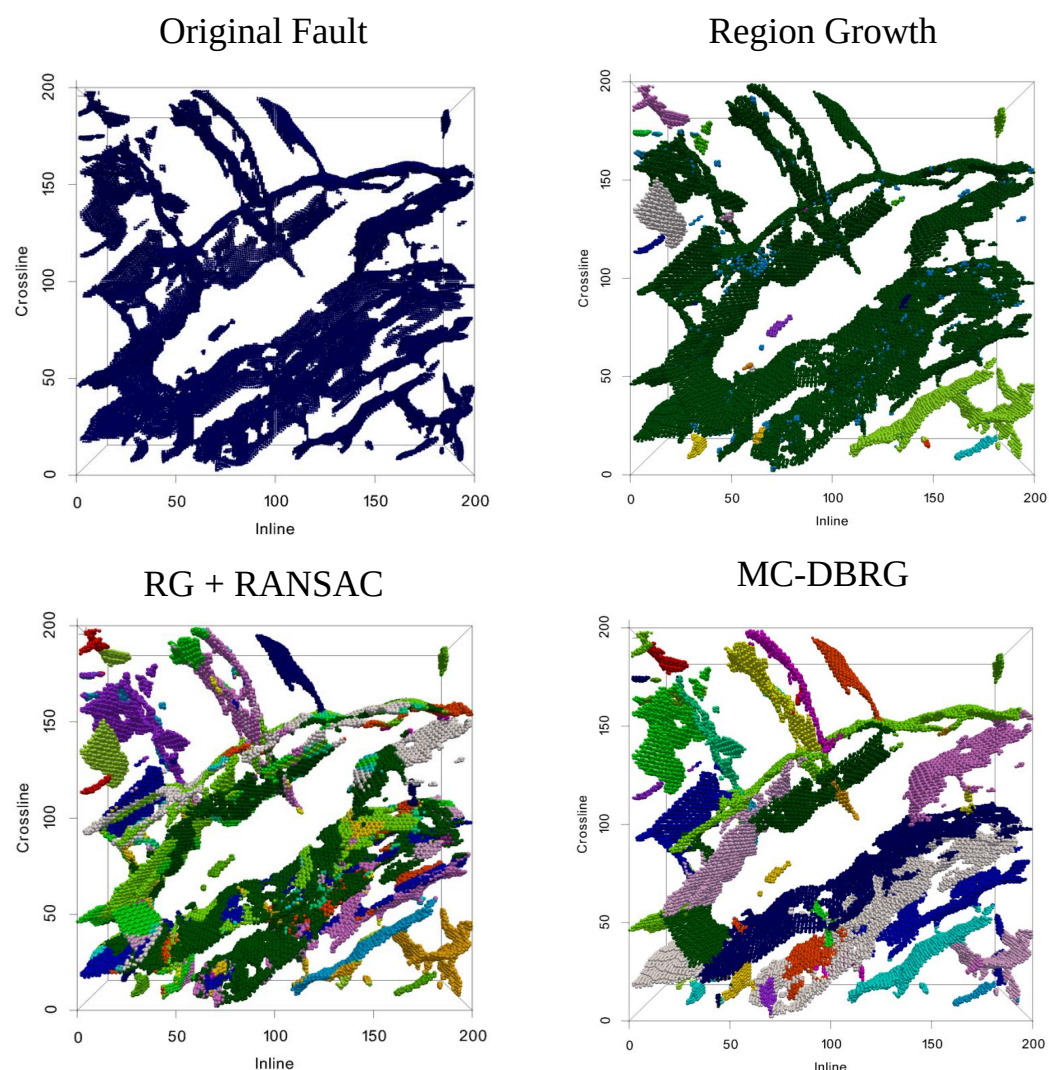


Figure 10. Segmentation results of three methods on the second real fault dataset (top view), where “RG+RANSAC” denotes the segmentation method proposed by Zou et al. [15]. Distinct colors in the fault point cloud indicate that they have been segmented into different instance classes.

Table 4. Quantitative comparison of different segmentation methods on the F3D and A2 datasets. The Unclassified Ratio (UR) represents the percentage of unassigned spatial points, and Time (s) denotes the computational time required per dataset.

Datasets	Methods	ICC	ICV	GAI	UR (%)	Time (s)
F3D	Region Growth	0.935	2.73×10^{-4}	0.140	4.62	4.8
	RG+RANSAC	0.939	6.44×10^{-4}	0.212	0	33.57
	MC-DBRG	0.906	7.17×10^{-4}	0.386	0	65.39
A2	Region Growth	0.818	1.64×10^{-3}	0.306	6.29	3.96
	RG+RANSAC	0.789	1.90×10^{-3}	0.413	0	40.32
	MC-DBRG	0.822	2.04×10^{-3}	0.116	0	53.58

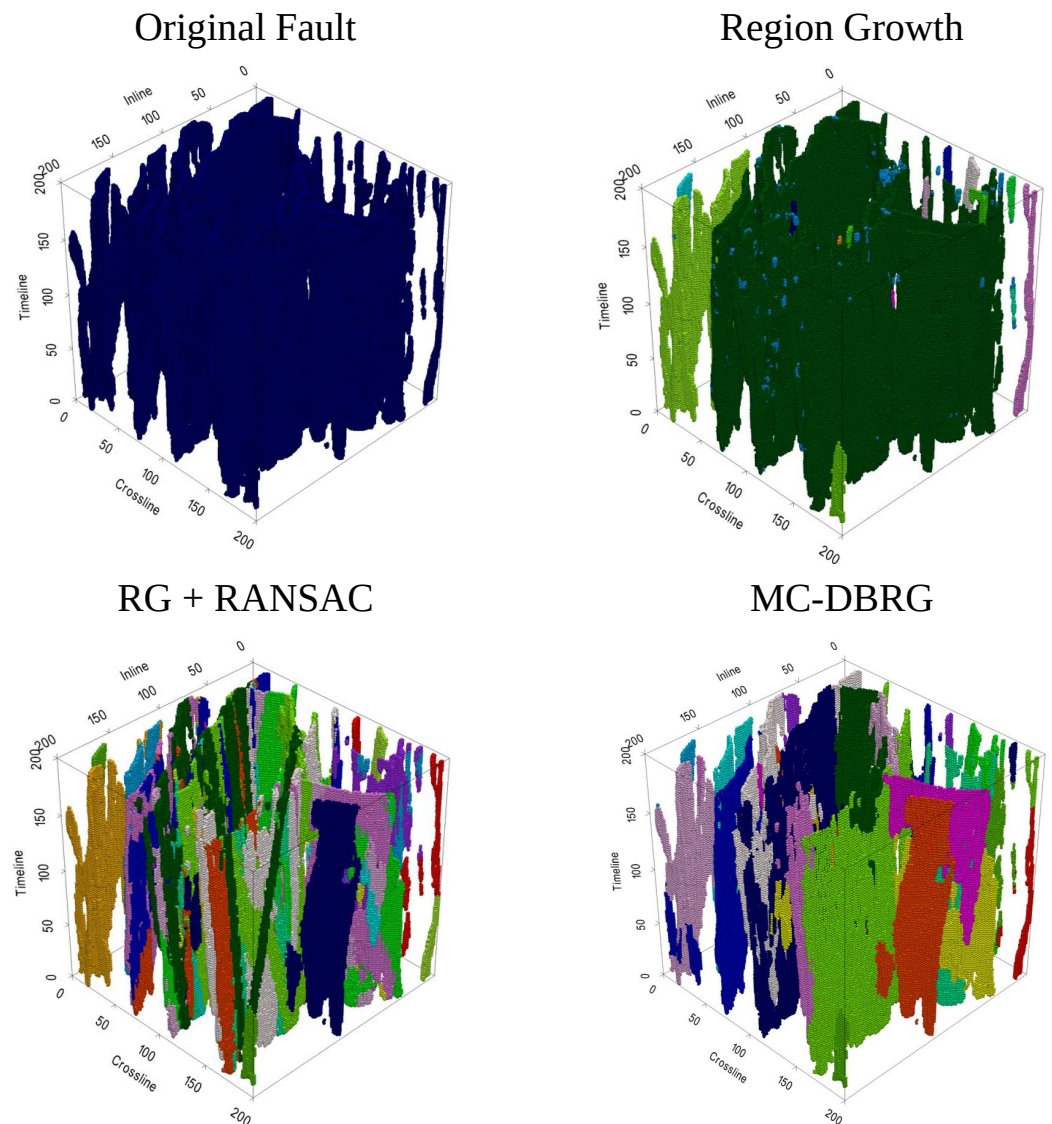


Figure 11. Segmentation results of three methods on the second real fault dataset (3D view), where “RG+RANSAC” denotes the segmentation method proposed by Zou et al. [15]. Distinct colors in the fault point cloud indicate that they have been segmented into different instance classes.

4. Conclusions

To resolve the boundary ambiguity and erroneous spatial merging that frequently occur in the edge transition zones of complex intersecting faults during intelligent geological modeling, this paper proposes a multi-surface instance segmentation method for 3D fault point clouds based on multi-scale contour constraints and Delaunay triangulation-based region growing. Unlike traditional global homogeneous growth algorithms, the proposed method decouples the multi-instance segmentation task into four progressive stages: feature-driven boundary recognition and wall reinforcement, multi-constraint spatial controlled growth, global manifold geometric merging, and trend-guided stripping and backfilling. This approach explicitly separates ambiguous boundary divergence zones from stable manifold core regions. Consequently, the segmented individual faults exhibit high coplanar consistency, accurate morphology, and low intra-instance curvature volatility. Furthermore, introducing a Delaunay triangulation-based region growing algorithm to 3D seismic fault interpretation enables $\mathcal{O}(1)$ topological network adjacency queries, significantly improving the computational throughput for massive fault point clouds. Finally, we construct an unsupervised geometric feature field evaluation system based on the

intra-instance local structure tensor. This system overcomes the lack of quantitative metrics caused by the absence of absolute geometric ground truth for underground geological bodies, providing a reliable data foundation for subsequent automated geological modeling and numerical simulations.

Although the proposed framework improves fault topological fidelity, certain limitations remain. The algorithm's performance relies heavily on the quality of the fault probability volume generated by the front-end network. When the input data contains large areas of laterally continuous stratigraphic artifacts, the feature space classifier suffers from interference. Moreover, the back-end geometric merging and label backfilling modules depend primarily on a tangent plane linear extrapolation assumption. When processing complex fault systems with significant non-linear spatial twisting, this linear extrapolation becomes inadequate. To address these issues, future work will use large-scale raw 3D seismic volume data to construct a self-supervised pre-training task based on 3D masked voxel reconstruction. By coupling this with the implicit surface representation of 3D neural networks, we aim to dynamically infer the geometric continuity of highly non-linear boundaries, thereby achieving accurate segmentation of multiple fault instances across various spatial scales.

Author Contributions: Conceptualization, L.H. (Lu Huang) and C.Z.; methodology, L.H. (Lu Huang); software, L.H. (Lu Huang); validation, L.H. (Lu Huang), C.Z. and K.S.; formal analysis, L.H. (Lu Huang); investigation, L.H. (Lu Huang); resources, L.H. (Lu Huang); data curation, L.H. (Lu Huang); writing—original draft preparation, L.H. (Lu Huang); writing—review and editing, L.H. (Lu Huang), C.Z., J.Z., K.S. and L.H. (Liming Han); visualization, L.H. (Lu Huang); supervision, C.Z. and J.Z.; project administration, L.H. (Lu Huang); funding acquisition, C.Z. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Supported by the Open Fund of the State Key Laboratory of Intelligent Coal Mining and Strata Control (No. SKLIS202502).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: During the preparation of this manuscript, the authors used Gemini v3.6 for the purpose of language editing and polishing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

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

References

1. Saber, E.; Qu, Q.; Aminossadati, S.M.; Li, J.; Chen, Z. A new coupled approach to simulating fault reactivation during CO₂ sequestration. In Proceedings of the SPE Asia Pacific Unconventional Resources Conference and Exhibition, Brisbane, Australia, 14–15 November 2023; p. D011S004R001.
2. Aswad, M.K.; Azad, S.; Mohammed, H. Strategic Drilling Decisions Based on Seismic Fault Inclination Analysis for Enhancing Hydrocarbon Recovery in the Bijel and Harir Oil Fields, Kurdistan Region, North Iraq. *IRAQI Bull. Geol. Min.* **2025**, *21*, 381–398. [CrossRef]
3. Gersztenkorn, A.; Marfurt, K.J. Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping. *Geophysics* **1999**, *64*, 1468–1479. [CrossRef]
4. Randen, T.; Pedersen, S.I.; Sønneland, L. Automatic extraction of fault surfaces from three-dimensional seismic data. In Proceedings of the SEG International Exposition and Annual Meeting, San Antonio, TX, USA, 9–14 September 2001.
5. Van Bommel, P.P.; Pepper, R.E. Seismic Signal Processing Method and Apparatus for Generating a Cube of Variance Values. US Patent 6,151,555, 21 November 2000.

6. Qi, J.; Zhang, B.; Lyu, B.; Marfurt, K. Seismic attribute selection for machine-learning-based facies analysis. *Geophysics* **2020**, *85*, O17–O35. [\[CrossRef\]](#)
7. Xiong, W.; Ji, X.; Ma, Y.; Wang, Y.; AlBinHassan, N.M.; Ali, M.N.; Luo, Y. Seismic fault detection with convolutional neural network. *Geophysics* **2018**, *83*, O97–O103. [\[CrossRef\]](#)
8. Di, H.; Wang, Z.; AlRegib, G. Seismic fault detection from post-stack amplitude by convolutional neural networks. In *Proceedings of the 80th EAGE Conference and Exhibition 2018*; European Association of Geoscientists & Engineers: Houten, The Netherlands, 2018; Volume 2018, pp. 1–5.
9. Wei, X.L.; Zhang, C.X.; Kim, S.W.; Jing, K.L.; Wang, Y.J.; Xu, S.; Xie, Z.Z. Seismic fault detection using convolutional neural networks with focal loss. *Comput. Geosci.* **2022**, *158*, 104968. [\[CrossRef\]](#)
10. Huang, Z.; Zhong, Q.; Gu, Q.; Lin, Y.; Zhao, K.; Zhang, X.; Wu, Y. Numerical investigation of particle migration in fault zones during water and mud inrush using the CFD–DEM approach. *Comput. Part. Mech.* **2025**, *12*, 1821–1837. [\[CrossRef\]](#)
11. Lin, L.; Li, C.; Wei, H.; Zhong, Z.; Wang, X.; Li, Q.; Gorman, A. An improved parametric 3D geologic modeling framework for seismic structure identification using deep learning in complex geologic settings. *Geophysics* **2025**, *90*, IM81–IM102. [\[CrossRef\]](#)
12. Ryan, L.; Magee, C.; Jackson, C.A.L. The kinematics of normal faults in the Ceduna Subbasin, offshore southern Australia: Implications for hydrocarbon trapping in a frontier basin. *AAPG Bull.* **2017**, *101*, 321–341. [\[CrossRef\]](#)
13. Jiang, F.; Norlund, P.; Dietz, D. Assisted fault identification and surface extraction by machine learning—A case study from Oman. In *Proceedings of the SEG International Exposition and Annual Meeting*, Denver, CO, USA, 26 September–1 October 2021; p. D011S043R003.
14. Zhao, L.; Cai, M.; Ding, R.; Zhang, Y.; Zhao, S.; Zhang, J.; Yang, J. A 3D organized point cloud clustering algorithm for seismic fault data based on region growth. *Comput. Geosci.* **2023**, *27*, 1165–1181. [\[CrossRef\]](#)
15. Zou, Q.; Zhang, J.; Zhang, C.; Sun, K.; Tao, C.; Guo, R. Automatic fault interpretation based on point cloud fitting and segmentation. *Geophys. Prospect.* **2024**, *72*, 2599–2614. [\[CrossRef\]](#)
16. Dong, L.; Song, W.; Hu, J.; Zeng, C.; Tang, W. Fault surface construction method based on point cloud surface reconstruction. *Front. Earth Sci.* **2025**, *13*, 1561288. [\[CrossRef\]](#)
17. Liu, X.; Zhu, X.; He, X.; Wang, Y. Automatic seismic fault detection and surface construction. *Appl. Comput. Geosci.* **2025**, *17*, 100287. [\[CrossRef\]](#)
18. Algarni, M.; Sundaramoorthi, G. Surfcut: Free-boundary surface extraction. In *Proceedings of the European Conference on Computer Vision*; Springer: New York, NY, USA; pp. 171–186.
19. Kuo, C.C.; Yau, H.T. A Delaunay-based region-growing approach to surface reconstruction from unorganized points. *Comput.-Aided Des.* **2005**, *37*, 825–835. [\[CrossRef\]](#)
20. Khanna, K.; Rajpal, N. Survey of curve and surface reconstruction algorithms from a set of unorganized points. In *Proceedings of the Third International Conference on Soft Computing for Problem Solving: SocProS 2013*; Springer: New Delhi, India, 2013; Volume 1, pp. 451–458.
21. Alexander, C. Delineating tree crowns from airborne laser scanning point cloud data using Delaunay triangulation. *Int. J. Remote Sens.* **2009**, *30*, 3843–3848. [\[CrossRef\]](#)
22. Otto, G.P.; Chau, T.K. ‘Region-growing’ algorithm for matching of terrain images. *Image Vis. Comput.* **1989**, *7*, 83–94. [\[CrossRef\]](#)
23. Yang, Y.; Li, B.; Wang, X.; Qi, M.; Zhai, D. The research on intelligent evaluation methods for non-sealing faults in reservoirs based on well test analysis. *Front. Earth Sci.* **2025**, *13*, 1659387. [\[CrossRef\]](#)
24. Xiong, S.; Zhang, J.; Zheng, J.; Cai, J.; Liu, L. Robust surface reconstruction via dictionary learning. *ACM Trans. Graph. (TOG)* **2014**, *33*, 1–12. [\[CrossRef\]](#)
25. Wang, S.; Li, H.; Miao, T.; Gao, Z.; Song, A. Feature extraction of horizontal plane and optimization of 3d lidar slam in indoor environments. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 8502218. [\[CrossRef\]](#)
26. Huang, J.; You, S. Point cloud matching based on 3D self-similarity. In *Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*; IEEE: New York, NY, USA, 2012; pp. 41–48.
27. Liu, Y.; Yang, Z.; Tong, J.; Yang, J.; Peng, J.; Zhang, L.; Cheng, W. ET-PointPillars: Improved PointPillars for 3D object detection based on optimized voxel downsampling. *Mach. Vis. Appl.* **2024**, *35*, 1. [\[CrossRef\]](#)
28. Lyu, W.; Ke, W.; Sheng, H.; Ma, X.; Zhang, H. Dynamic downsampling algorithm for 3D point cloud map based on voxel filtering. *Appl. Sci.* **2024**, *14*, 3160. [\[CrossRef\]](#)
29. Li, D.; Shi, G.; Wu, Y.; Yang, Y.; Zhao, M. Multi-scale neighborhood feature extraction and aggregation for point cloud segmentation. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *31*, 2175–2191. [\[CrossRef\]](#)
30. Huang, R.; Hong, D.; Xu, Y.; Yao, W.; Stilla, U. Multi-scale local context embedding for LiDAR point cloud classification. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 721–725.
31. Reynolds, D.A. Gaussian mixture models. *Encycl. Biom.* **2009**, *741*, 3.
32. Ben-Shabat, Y.; Lindenbaum, M.; Fischer, A. 3dmfv: Three-dimensional point cloud classification in real-time using convolutional neural networks. *IEEE Robot. Autom. Lett.* **2018**, *3*, 3145–3152. [\[CrossRef\]](#)

33. Abdi, H.; Williams, L.J. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 433–459. [[CrossRef](#)]
34. Berger, M.; Tagliasacchi, A.; Seversky, L.M.; Alliez, P.; Guennebaud, G.; Levine, J.A.; Sharf, A.; Silva, C.T. A survey of surface reconstruction from point clouds. In *Proceedings of the Computer Graphics Forum*; Wiley Online Library: Hoboken, NJ, USA, 2017; Volume 36, pp. 301–329.
35. Ji, G.; Wang, Q.; Zhou, X.; Cai, Z.; Zhu, J.; Lu, Y. An automated method to build 3D multi-scale geological models for engineering sedimentary layers with stratum lenses. *Eng. Geol.* **2023**, *317*, 107077. [[CrossRef](#)]
36. Corlay, Q.; Vitel, S.; Mallison, B.; Sekar, A.; Wright, L.; Wang, E.; Amaru, M.; Hern, B.; Souche, L.; Vera, K.W.; et al. Integrating 3D neural networks fault detection and implicit fault modeling in a global seismic interpretation workflow. In *Proceedings of the SEG International Exposition and Annual Meeting*, Houston, TX, USA, 25–28 August 2025.
37. Wang, H.; Wang, H.; Liu, K.; Meng, J.; Xiao, Y.; Wang, Y. Seismic fault identification of deep fault-karst carbonate reservoir using transfer learning. *Nat. Gas Ind. B* **2025**, *12*, 174–185. [[CrossRef](#)]
38. Wu, X.; Liang, L.; Shi, Y.; Fomel, S. FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics* **2019**, *84*, IM35–IM45. [[CrossRef](#)]

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.