

Article

Intelligent Classification and Segmentation of Sandstone Thin Section Image Using a Semi-Supervised Framework and GL-SLIC

Yubo Han and Ye Liu *

School of Computer Science, Xi'an Shiyou University, Xi'an 710065, China; 17791974959@163.com

* Correspondence: yeliu@xsyu.edu.cn

Abstract: This study presents the development and validation of a robust semi-supervised learning framework specifically designed for the automated segmentation and classification of sandstone thin section images from the Yanchang Formation in the Ordos Basin. Traditional geological image analysis methods encounter significant challenges due to the labor-intensive and error-prone nature of manual labeling, compounded by the diversity and complexity of rock thin sections. Our approach addresses these challenges by integrating the GL-SLIC algorithm, which combines Gabor filters and Local Binary Patterns for effective superpixel segmentation, laying the groundwork for advanced component identification. The primary innovation of this research is the semi-supervised learning model that utilizes a limited set of manually labeled samples to generate high-confidence pseudo labels, thereby significantly expanding the training dataset. This methodology effectively tackles the critical challenge of insufficient labeled data in geological image analysis, enhancing the model's generalization capability from minimal initial input. Our framework improves segmentation accuracy by closely aligning superpixels with the intricate boundaries of mineral grains and pores. Additionally, it achieves substantial improvements in classification accuracy across various rock types, reaching up to 96.3% in testing scenarios. This semi-supervised approach represents a significant advancement in computational geology, providing a scalable and efficient solution for detailed petrographic analysis. It not only enhances the accuracy and efficiency of geological interpretations but also supports broader hydrocarbon exploration efforts.



Citation: Han, Y.; Liu, Y. Intelligent Classification and Segmentation of Sandstone Thin Section Image Using a Semi-Supervised Framework and GL-SLIC. *Minerals* **2024**, *14*, 799. <https://doi.org/10.3390/min14080799>

Academic Editor: Sebastian Iwaszenko

Received: 10 June 2024

Revised: 28 July 2024

Accepted: 1 August 2024

Published: 5 August 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/>).

Keywords: semi-supervised learning; GL-SLIC algorithm; petrographic analysis; sandstone thin sections; computational geology

1. Introduction

Sandstone thin section identification plays a crucial role in geological research and the exploration of oil and gas resources. Traditionally, experts manually observe and analyze the composition and structure of sandstone thin sections, marking the boundaries between mineral particles and pores to identify the mineral and pore characteristics of the target rock. This manual process is not only time-consuming but also susceptible to subjective bias, making the study of automated thin section analysis based on image analysis particularly significant [1–3]. Automated analysis can enhance efficiency, reduce researchers' workloads, and minimize subjective biases, making the identification results more quantitative and reliable.

However, sandstone thin section images present numerous challenges for image analysis due to their numerous targets, significant scale differences, complex component shapes, crowded detection targets, and similar features among targets. Figure 1 depicts thin section images of sandstone from the Yanchang Formation in the Ordos Basin under plane-polarized light (LP). The main minerals visible include quartz, kaolinite, and lithic fragments, with iron calcite cement evident in some areas. The contact between grains varies from point contacts to long contacts, which adds to the complexity of the image

analysis. The scale of the images is 40 μm , highlighting the fine granularity that must be accurately identified and segmented.

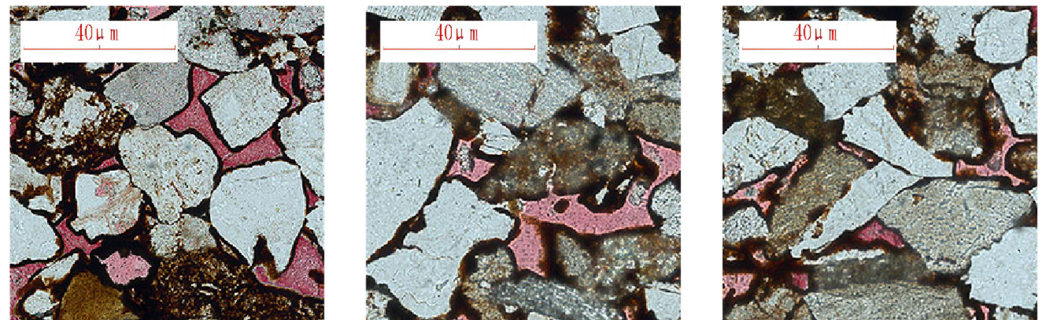


Figure 1. Thin section image of sandstone under plane-polarized light: the main components are quartz, kaolinite, matrix, pores and lithic fragments.

Thin section image analysis can be divided into two-phase segmentation-recognition processing and end-to-end recognition methods. In the two-phase segmentation-recognition processing, traditional segmentation methods such as superpixel segmentation are used to preprocess images into monolithic rock images, followed by recognition using deep learning models. This process requires high-quality segmentation to be effective. End-to-end recognition methods have proven effective in semantic segmentation and enhancing pixel classification accuracy using deep learning models like U-Net and SegNet. However, these models, which integrate segmentation and classification tasks and depend on high-quality data, sometimes overlook the detailed geological understanding essential for accurate image analysis. Notably, deep learning systems face challenges in handling more detailed and complex geological images [4].

Recent advancements in the field have addressed some of these challenges. For instance, Ren et al. introduced a multi-channel attention transformer model that significantly improves instance segmentation in sandstone images, surpassing previous methods in mean average precision (mAP) [5]. Similarly, Liu et al. applied artificial intelligence techniques for the automated identification of sandstone thin sections, enhancing both accuracy and efficiency [3]. Moreover, Shebl et al. demonstrated the application of cognitive image recognition to automate the description of thin section images, providing significant improvements in the analysis of carbonate rocks [6]. Yalamanchi et al. used machine learning algorithms and scanning electron microscope (SEM) images to predict the pore structure and permeability of carbonate reservoirs [7].

Additionally, Chen et al. developed a multiangle polarized imaging-based method tailored for practical production use in sandstone thin sections, integrating petrology with advanced image processing techniques [8]. Liu et al. explored the use of super-resolution techniques to enhance the resolution of thin section images, allowing for more detailed mineral segmentation [9]. Furthermore, Vellappally et al. presented a machine learning-based approach for grain segmentation and mineral classification in Jurassic sandstones, demonstrating the potential for high-accuracy automated analysis [10].

Further supporting this progress, Dabek et al. proposed a superpixel-based grain segmentation method specifically designed for sandstone thin sections, significantly enhancing the accuracy of grain boundary recognition [11]. Dong et al. introduced a high-accuracy image segmentation approach based on a hybrid attention mechanism, effectively addressing the complex shapes of sandstone particles [12]. Caja et al. leveraged high-resolution thin section scan images for digital rock physics in cuttings, providing a comprehensive framework for rock property analysis [13]. Additionally, Visalli et al. developed a semi-automated edge detector using ArcGIS for quantitative microstructural analysis of rock thin sections, offering a robust tool for detailed geological analysis [14].

These advancements highlight the ongoing evolution of automated image analysis in geological studies. The integration of sophisticated algorithms and AI techniques has

significantly improved the accuracy and efficiency of thin section analysis, addressing the critical challenge of insufficient labeled data and enhancing the model's ability to generalize from minimal initial input. This progress marks a significant advancement in computational geology, providing a scalable and efficient solution for detailed petrographic analysis and aiding in the broader efforts of hydrocarbon exploration.

Object and Novelty

A fundamental challenge in machine learning-based methods for geological image analysis is the labor-intensive and error-prone process of data annotation. This issue is particularly pronounced in geological image analysis, where the diversity and complexity of rock thin sections exacerbate the difficulties of manual labeling. Accurate data labels are crucial for training effective machine learning models; however, traditional fully supervised methods are hindered by substantial time and resource commitments, as well as potential human error.

This study addresses these issues through two key innovations:

Semi-Supervised Learning Framework: This novel framework integrates a limited amount of labeled data with a larger volume of unlabeled data. Utilizing a discriminator model to refine predictions based on limited ground truth significantly enhances the training process. By reducing reliance on extensive labeled datasets, this approach mitigates the challenges associated with manual annotation and improves the model's ability to generalize from minimal initial input. This method effectively tackles the challenge of insufficient labeled data, a common limitation in geological studies, thereby enabling more robust and scalable model training.

GL-SLIC Algorithm: An advanced pre-segmentation technique combining Gabor filters and Local Binary Patterns (LBP) is introduced to improve feature extraction. This enhancement allows the algorithm to capture unique textural patterns and structural complexities of sandstone more accurately. The GL-SLIC algorithm ensures better adherence to natural boundaries within rock sections, significantly reducing misalignment issues commonly observed with other algorithms and enhancing the precision of subsequent segmentation processes. Precise segmentation is crucial for accurate mineral and pore identification, directly impacting the reliability of petrographic analyses.

2. Methodology

2.1. Data Preparation

The dataset incorporated in this study is derived from the Yanchang Formation within the Ordos Basin. This dataset is instrumental in understanding the porosity and permeability challenges of the tight sandstone reservoir. The formation's reservoir quality is quantified by a porosity average of 9.66% and permeability of 0.05 millidarcies among the collected samples. The samples were taken perpendicular to the bedding. They comprise mainly quartz, kaolinite, and clay minerals, with carbonates, silica, and iron oxides as cements. Grain sizes range from fine to medium with point and long-grain contacts. The depositional environment includes fluvial channels, floodplains, and lacustrine settings. Diagenesis involves early compaction and cementation, followed by mineral dissolution and secondary porosity formation, and late-stage deep burial processes.

A total of 1000 original thin section images of sandstone were used. These images were captured at a magnification of 200× under single-polarized light to accurately capture the textural details. The segmentation of these original images resulted in 7920 minerals and pores fragments, which were categorized into six distinct classes. A deliberate effort was made to label 100 samples from each class to provide a basis for the semi-supervised learning model. This strategic labeling allows the model to recognize and categorize the diverse pore structures present within the formation.

2.2. Methodology Sequence

The methodological sequence, as depicted in the Figure 2, begins with the preprocessing of original rock thin section images through our novel GL-SLIC algorithm. This initial phase involves an over-segmentation step, where the GL-SLIC algorithm decomposes the image into superpixel clusters. Subsequently, a region fusion process is employed. This process combines nearby superpixel fragments that have similar features to achieve a more coherent segmentation of mineral and pore fragments.

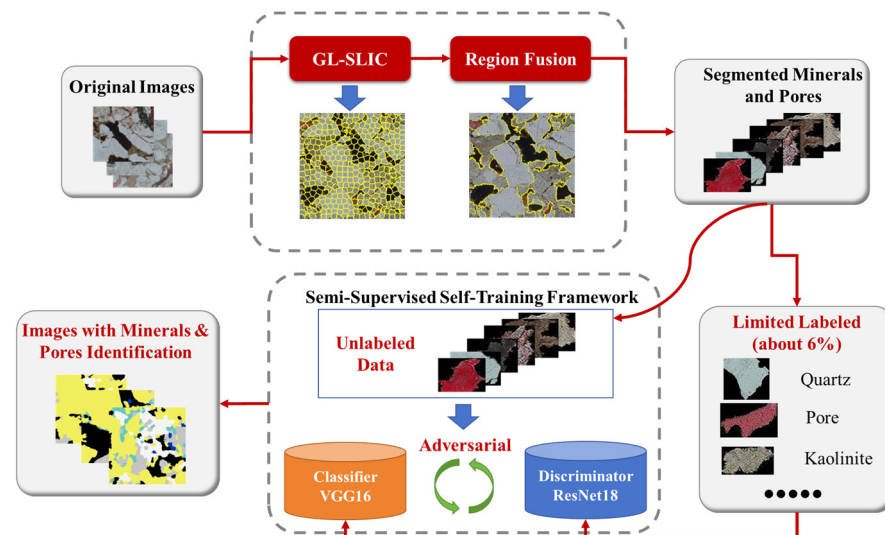


Figure 2. Workflow for recognizing minerals using GL-SLIC segmentation and semi-supervised training.

The next pivotal step involves manual annotation. Approximately 6% of the samples are selected randomly and labeled to represent categories such as quartz, pores, kaolinite, etc. These annotated samples are instrumental in training both the classifier, utilizing a VGG16 architecture, and the discriminator, based on a ResNet18 model.

Under the semi-supervised learning framework we propose, these models process the unlabeled data fragments. Through iterative adversarial training, the models learn to accurately classify the remaining unlabeled data. This semi-supervised approach allows for comprehensive labeling of all fragments. As a result, the segmented output accurately identifies mineral constituents in the original thin section images.

2.3. GL-SLIC Algorithm

This section introduces the GL-SLIC algorithm, a novel approach designed to enhance the segmentation and classification of sandstone thin section images. This is achieved through the methodical integration of textural and color features into superpixel segmentation. The GL-SLIC algorithm is structured into three primary stages: GLBP Feature Construction, GL-SLIC Superpixel Segmentation, and Region Merging.

GLBP Feature Construction: The foundation of this algorithm lies in GLBP (Gabor-Local Binary Pattern) features, which synergistically combine texture and color information.

GL-SLIC Superpixel Segmentation: Building upon the GLBP features, this step employs the SLIC (simple linear iterative clustering) algorithm for superpixel segmentation. This method effectively partitions the image into numerous coherent superpixels, each closely corresponding to natural divisions within the image based on the enhanced GLBP features.

Region Merging: Following segmentation, this process merges adjacent superpixels exhibiting similar characteristics. This merging step addresses the issue of over-segmentation, where individual superpixels may be too small or fragmented to be useful in isolation.

2.3.1. GLBP Feature Extraction

This section introduces a novel texture feature extraction algorithm that integrates the local binary pattern (LBP) operator with Gabor filters, leveraging both spatial and frequency domain characteristics, as shown in Figure 3. This hybrid approach robustly extracts features from sandstone thin section images, effectively minimizing the impact of external factors such as noise, illumination, rotation, and shadows.

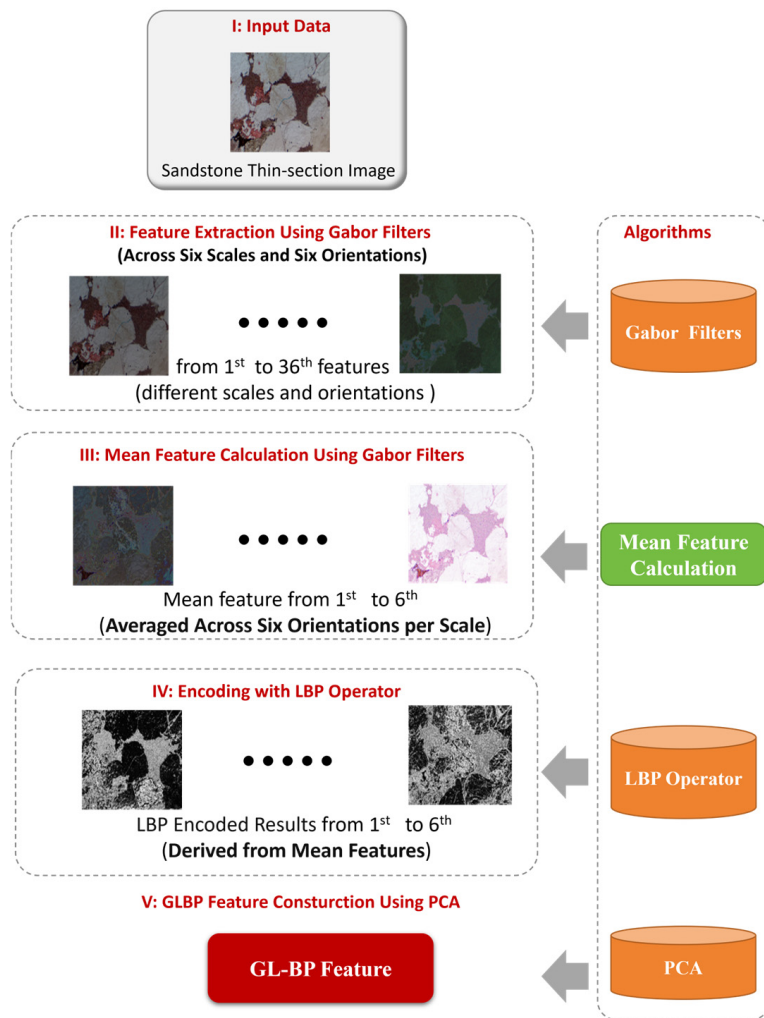


Figure 3. GL-BP feature extraction workflow integrating LBP operator and Gabor filters for sandstone thin section images.

Operational Steps:

1. **Input Data:** The initial step involves inputting the sandstone thin section image data designated for segmentation.
2. **Feature Extraction Using Gabor Filters:** The algorithm applies Gabor filters at various scales and orientations to the original sandstone thin section images. Specifically, the filters are set at six different scales and six orientations, creating a total of 36 distinct Gabor filters. These filters systematically extract features across the images, resulting in 36 Gabor feature vectors. Figure 4 illustrates the features extracted by the Gabor filters at each scale and orientation. The scales are shown vertically from top to bottom (1 through 6), and the orientations are displayed horizontally.

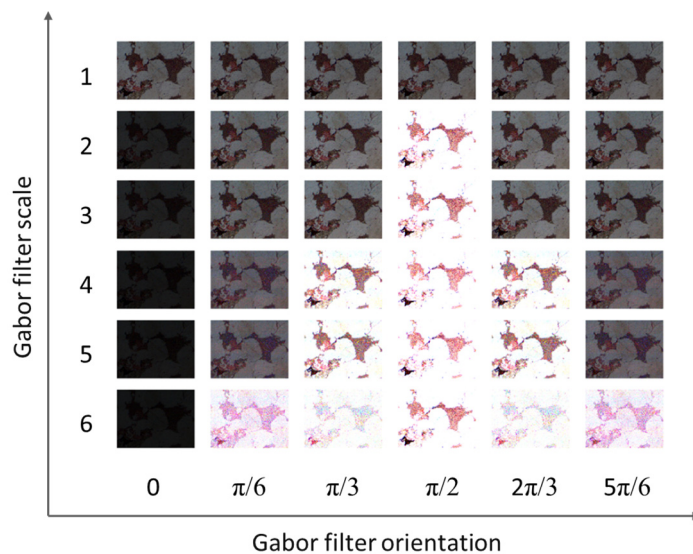


Figure 4. Feature extraction visualization using Gabor filters at various scales and orientations for sandstone thin section images.

3. Mean Feature Calculation Using Gabor Filters.

Each Gabor feature extracted from different orientations and scales is averaged to calculate the mean feature for each respective setting. The Gabor filters are applied to six different scales and orientations, resulting in 36 distinct Gabor filters. The mean of the features extracted by each Gabor filter at these scales and orientations is computed as follows:

$$\left\{ \begin{array}{l} GM_{1_1} = \frac{1}{6} \sum_{i=1}^6 G_{1_1, \theta_i} \\ GM_{1_2} = \frac{1}{6} \sum_{i=1}^6 G_{1_2, \theta_i} \\ GM_{1_3} = \frac{1}{6} \sum_{i=1}^6 G_{1_3, \theta_i} \\ GM_{1_4} = \frac{1}{6} \sum_{i=1}^6 G_{1_4, \theta_i} \\ GM_{1_5} = \frac{1}{6} \sum_{i=1}^6 G_{1_5, \theta_i} \\ GM_{1_6} = \frac{1}{6} \sum_{i=1}^6 G_{1_6, \theta_i} \end{array} \right.$$

Figure 5 visually demonstrates the effectiveness of Gabor filter transformations in extracting distinctive textural features from sandstone thin sections at various scales and orientations. Each panel within the figure highlights the nuanced differences between quartz grains and other geological components, which are not easily discernible in the original images. In panels such as Figure 5a,b, red boxes identify two quartz grains that appear similar in the raw image but are distinctly separated in the processed images due to the enhanced textural contrasts provided by the Gabor features. Similarly, Figure 5d,f use red and blue boxes to differentiate between quartz and pores with enhanced clarity. This shows that even subtle textural differences become pronounced when viewed through Gabor-filtered images. The distinct representation of quartz in Figure 5f, where its color appears markedly darker than surrounding materials, exemplifies how effectively these features can isolate specific components.

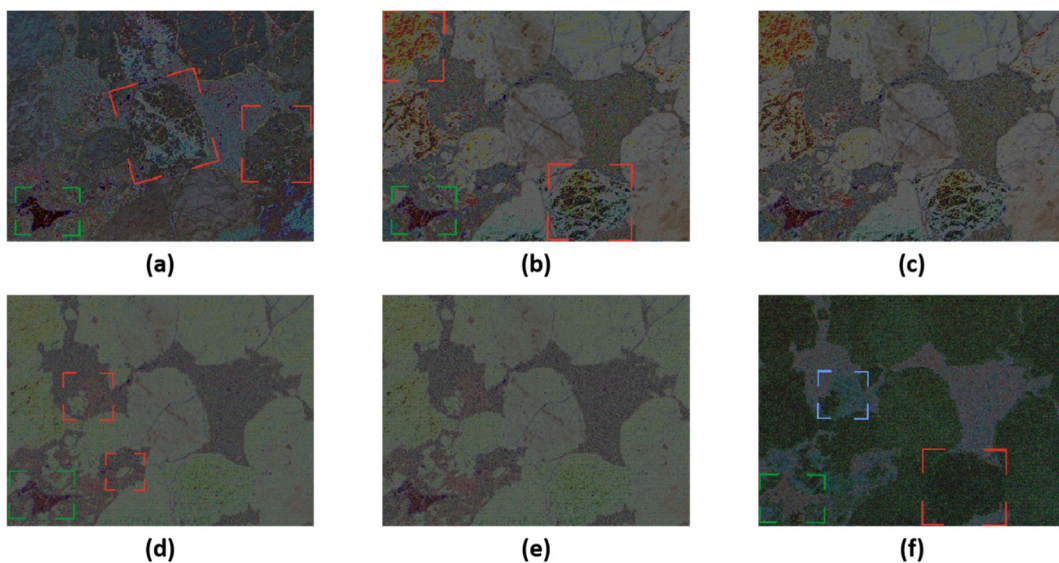


Figure 5. Mean feature comparison chart: (a) mean feature chart at scale 1; (b) mean feature chart at scale 2; (c) mean feature chart at scale 3; (d) mean feature chart at scale 4; (e) mean feature chart at scale 5; (f) mean feature chart at scale 6.

4. Encoding with LBP Operator

The mean feature images obtained from the Gabor filters are then encoded using the LBP operator. This operator emphasizes local textural patterns by comparing each pixel with its neighborhood. Figure 6 displays the LBP-encoded images corresponding to the mean features.

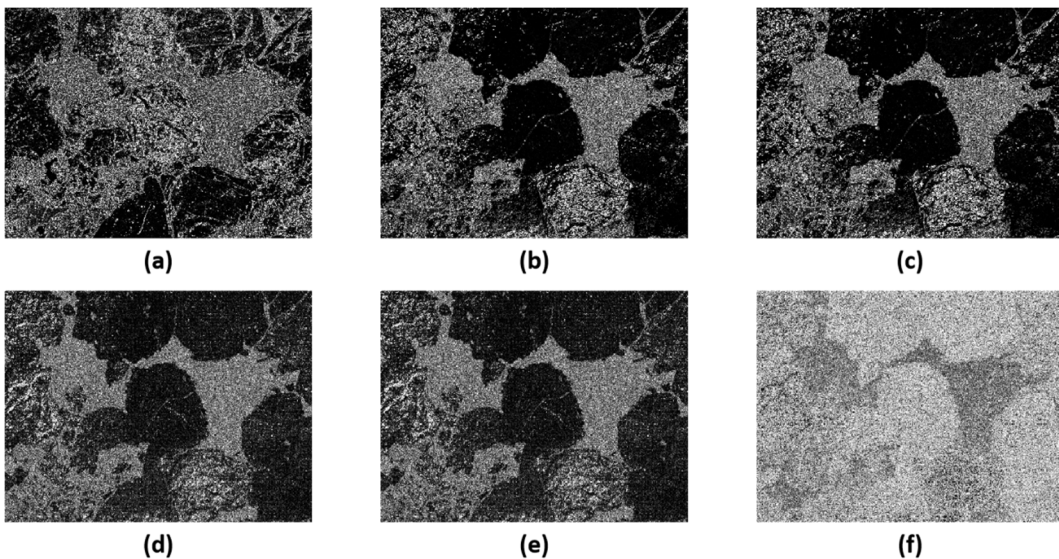


Figure 6. LBP Feature Extraction: (a) mean feature chart at scale 1; (b) mean feature chart at scale 2; (c) mean feature chart at scale 3; (d) mean feature chart at scale 4; (e) mean feature chart at scale 5; (f) mean feature chart at scale 6.

5. GLBP Feature Construction Using PCA

The six image matrices generated from the LBP encoding are then subjected to principal component analysis (PCA). This step reduces dimensionality and extracts the most significant features, forming a feature vector that succinctly represents the textural characteristics of the sandstone images.

2.3.2. Enhanced SLIC Superpixel Segmentation Algorithm: GL-SLIC

In this section, we introduce the GL-SLIC algorithm, an advanced version of the Simple Linear Iterative Clustering (SLIC) superpixel segmentation method. This algorithm incorporates texture features derived from GLBP (Gabor–local binary pattern) attributes. This enhanced algorithm aims to produce superpixels that more accurately conform to the edges of mineral particles within sandstone images, capturing both local and global features effectively.

Process and Methodology:

1. Feature Extraction: Initially, the GLBP texture features of the image are extracted using Gabor filters and LBP operators. The image is then transformed from RGB color space to CIE-Lab color space to prepare it for segmentation. This transformation enhances the distinction based on lightness and color components, which align more closely with human visual perception.

2. Initialization of Cluster Centers: The image is divided into K superpixel grids, with each grid's step size proportional to the total number of pixels in the image. The initial cluster centers are chosen by placing them at certain intervals, ensuring that each superpixel represents a region of approximately equal size within the image. Each initial cluster center corresponding to a superpixel center is denoted as C_k . C_k is mapped to a six-dimensional vector $C_k = [l_k, a_k, b_k, x_k, y_k, GL_k]$, where x_k and y_k are spatial positions, l_k , a_k , and b_k are the color features in the CIE-Lab color space, and GL_k represents the GLBP feature value for $k = 1, 2, \dots, K$.

3. Optimization of Initial Cluster Centers: To enhance the accuracy of clustering and segmentation, an optimization strategy for initial cluster centers is employed. This process involves adjusting the cluster centers to areas with the lowest gradient within their neighborhoods. This helps avoid placing them on edges or noisy regions where abrupt changes in image properties occur. As a result, the algorithm ensures that superpixels are more likely to encapsulate homogeneous regions, thereby improving the overall segmentation quality.

The calculation of the gradient at a pixel position (x, y) is crucial for this optimization process and is defined by the following formula:

$$G(x, y) = \|l(x+1, y) - l(x-1, y)\|^2 + \|l(x, y+1) - l(x, y-1)\|^2$$

where $l(x, y)$ represents the first three components of the feature vector C_k at location (x, y) , specifically l_k, a_k , and b_k . This gradient computation integrates the local differences in luminance and color channels, effectively capturing edge intensity which guides the relocation of cluster centers to less variant regions.

4. Initialization of Distances and Labels: Each pixel is initially assigned a label $labels(i) = -1$ and a distance $d(i) = \infty$. This setup prepares for the subsequent clustering process by ensuring all pixels can be evaluated for their proximity to potential new cluster centers.

5. Distance Calculation Between Pixels and Cluster Centers: For each pixel within the neighborhood of a cluster center, the distance to the cluster center is computed. This calculation incorporates a combination of Euclidean distances based on spatial, color, and texture features. The overall distance metric used for superpixel segmentation is composed of three parts: color feature distance, spatial feature distance, and texture feature distance.

Color Feature Distance d_c :

$$d_c = \sqrt{(l_i - l_k)^2 + (a_i - a_k)^2 + (b_i - b_k)^2}$$

Spatial Feature Distance d_s :

$$d_s = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$$

Texture Feature Distance d_{gl} :

$$d_{gl} = \sqrt{(g^l_i - g^l_k)^2}$$

Overall Distance Metric D' :

$$D' = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{s}\right)^2 + g(d_{gl})^2}$$

where m and S are normalization factors for spatial and color distances, respectively, and g is a weighting function for the texture gradient.

6. Pixel Classification and Iteration: Pixels are classified into the nearest cluster based on the calculated distances. Cluster centers are then recalculated as the mean of the features within each cluster, and this reassignment continues iteratively until the cluster centers stabilize, indicating that a robust segmentation has been achieved.

7. Recalculation of Cluster Centers: The process of recalculating the cluster centers involves updating their positions based on the mean of the feature vectors of all pixels currently assigned to each cluster. This is achieved by performing the following steps:

Compute Mean Feature Vectors: For each cluster, calculate the mean of the feature vectors of all assigned pixels. This mean vector becomes the new position of the cluster center.

Reassign Pixels: With the newly calculated cluster centers, reassign each pixel to the nearest cluster based on the comprehensive distance metric previously defined. This involves recalculating the distances for all pixels relative to each new cluster center and updating their cluster assignments accordingly.

Iteration: This process of recalculating cluster centers and reassigning pixels is iterated. After each iteration, check for significant changes in the positions of the cluster centers.

Convergence Criterion: The iteration stops when the changes in the cluster centers' positions between successive iterations fall below a predefined threshold, indicating convergence. This threshold ensures that the algorithm terminates when cluster centers stabilize, reflect.

8. Finalization of Clusters: The process concludes with a final adjustment of clusters to ensure that each superpixel optimally represents a distinct region of the image, with minimal intra-cluster variance and maximized inter-cluster distinction.

The GL-SLIC algorithm leverages the textural and color features encoded by the GLBP method, resulting in superpixels that are more aligned with natural divisions in the image, such as mineral boundaries. This improved segmentation forms a critical foundation for accurate subsequent analyses, including region merging and classification tasks.

2.3.3. Superpixel Merging Strategy

To enhance the quality of data for training classification models, a coarse merging of generated superpixels is necessary. Initially, an adjacency matrix is constructed based on the average color features of each superpixel block. The dissimilarity between two adjacent regions is measured by the Euclidean distance between the average color values of the superpixels. Regions with smaller dissimilarities are more similar and thus are candidates for merging. Given a threshold, superpixel merging is conducted using a breadth-first traversal algorithm tailored for graph-based segmentation. The range for the threshold ψ is determined by the variability in the color features across different rock images, which can be significant.

To set the threshold ψ , consider the following inequalities which balance between merging too many dissimilar superpixels and maintaining meaningful segmentation.

Normalized Difference Threshold:

$$\frac{w_{ij} - \min(w_{ij})}{\max(w_{ij}) - \min(w_{ij})} < \psi$$

where w_{ij} is the weight of the edge between superpixels i and j , representing the dissimilarity based on color features. $\min(w_{ij})$ and $\max(w_{ij})$ are the minimum and maximum weights in the adjacency graph, respectively.

Weighted Threshold Formulation:

$$\psi \times \max(w_{ij}) + (1-\psi) \times \min(w_{ij}) > w_{ij}$$

Alternative Formulation:

$$\psi_1 \times \max(w_{ij}) + \psi_2 \times \min(w_{ij}) > w_{ij}$$

where ψ_1 and ψ_2 are parameters that you can adjust to fine-tune the merging criteria, ideally within a range of [0, 1].

Given the diversity in rock images, a more relaxed merging strategy might lead to excessive merges; hence, the threshold settings are critical to ensure that only truly similar superpixels are combined. Moreover, while the GL-SLIC algorithm primarily utilizes the CIE-Lab color space for distance metrics, in this context, the YCBCR color space is also employed to cater to different chromatic and luminance variations more effectively.

2.4. Semi-Supervised Self-Training Framework

2.4.1. Method Process

This section introduces a semi-supervised self-training model for classifying petrographic thin section images, as shown in Figure 7. This model emphasizes pseudo-labeling to augment the training dataset effectively. The process is iterative and refines classification accuracy through repeated training cycles. Here are the detailed steps:

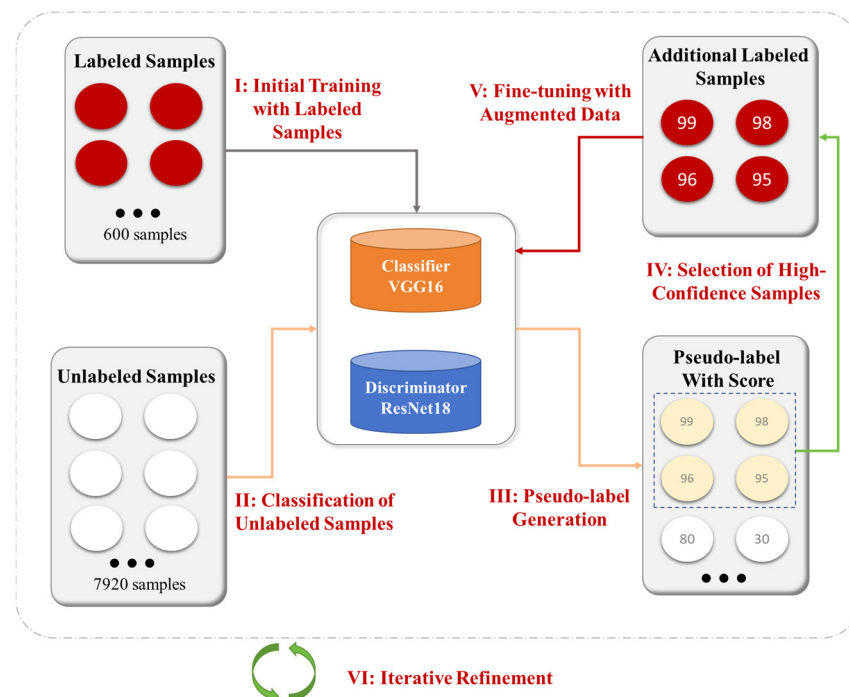


Figure 7. Semi-supervised self-training process.

1. Initial Training with Labeled Samples

The model initialization begins with the primary classifier (VGG16) being trained using 600 manually labeled samples. This foundational training establishes the initial learning parameters that guide the classification of unlabeled samples.

2. Classification of Unlabeled Samples

The primary model processes the remaining 7920 unlabeled samples, classifying them based on the features learned during initial training. Each sample receives a preliminary label and a confidence score.

3. Pseudo-Label Generation

The primary model assigns pseudo-labels to the unlabeled samples. Only samples with a confidence score greater than 94% are considered. These high-confidence pseudo-labels indicate that the model is highly certain of their classifications.

4. Selection of High-Confidence Samples

Samples that meet the high-confidence threshold are selected for further processing. These samples are considered accurate enough to be used for additional training. The discriminator model (ResNet18) reviews these pseudo-labeled samples to confirm their validity.

5. Fine-Tuning with Augmented Data

The training cycle enters the fine-tuning phase. The model is refined using both the original labeled samples and the new high-confidence samples identified in the previous step. This step improves the model's accuracy by integrating more verified data into the training set.

6. Iterative Refinement

Steps 2 through 5 are repeatedly executed to gradually expand the dataset with high-confidence pseudo-labeled samples. This iterative refinement boosts the capabilities of both the classifier and discriminator, enabling the system to reliably annotate the entire dataset automatically. Each cycle aims to not only increase the quantity of data being used for training but to improve the overall precision and reliability of the model.

Through these steps, the model leverages both labeled and high-quality pseudo-labeled data to enhance its learning efficacy and accuracy. This method reduces dependency on extensive labeled datasets while ensuring the classifier remains robust and capable of handling diverse data scenarios. Each cycle aims to increase the pool of high-confidence data, progressively enhancing the model's performance until no significant gains can be derived from additional unlabeled data.

2.4.2. Classifier Architecture

The architecture of the Classifier is detailed in Figure 8. It utilizes a modified VGG16 framework, incorporating the first 14 layers with batch normalization. The model is structured into two segments, each comprising two convolutional blocks. Each block includes a convolution operation, followed by batch normalization and a ReLU activation function, and concludes with a max-pooling layer. The input data, with dimensions (3, 32, 32), traverse these layers, resulting in a tensor of (128, 8, 8). This tensor is subsequently flattened into a one-dimensional array of 8192 elements before passing through two fully connected layers. The output is processed via a softmax function to produce probability distributions across N classes. The highest probability determines the class label.

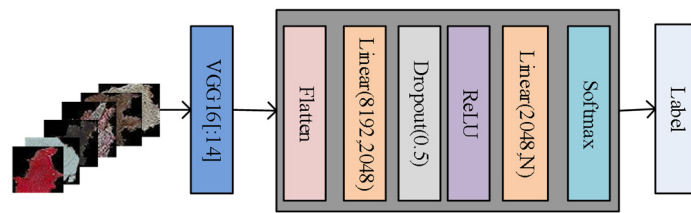


Figure 8. Modified VGG16 Classifier Architecture.

The network architecture of the primary model is depicted in Figure 8. It uses the first 14 layers of a batch-normalized VGG16 network. This network is partitioned into two segments, each comprising two convolution blocks. Each block includes a convolution operation, batch normalization, and a ReLU activation, followed by a max-pooling layer. The input tensor size is (3, 32, 32). After processing through these layers, it outputs a tensor of size (128, 8, 8). Before entering the fully connected layers, the tensor is flattened to (1, 8192). It then passes through two dense layers and concludes with a softmax function that outputs the probabilities for N classes. The class with the highest probability is selected as the label.

2.4.3. Discriminator Architecture

Figure 9 illustrates the Discriminator model, which employs a subset of the ResNet18 architecture. This model begins with an initial convolution, followed by batch normalization and max pooling. The subsequent layers include Layer 1 to Layer 3, each consisting of two basic block types. Basic Block1 is composed of two repetitive sets of convolution, batch normalization, and ReLU activation, without any downsampling or increase in dimensionality. Basic Block2 follows a similar structure for the first two sets but includes a downsampling and dimensionality increase in the third set. The output tensor, resized to (256, 3, 3), is flattened to (1, 2304). It then proceeds through a final fully connected layer. A softmax function calculates the probability distributions for N classes, aiming to fine-tune the classification results derived from the primary model.

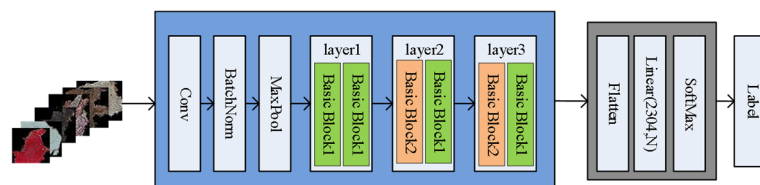


Figure 9. Discriminator model architecture.

2.4.4. Evaluation Metrics

To quantitatively assess the proposed segmentation algorithm, this study utilizes three established metrics: under segmentation error (UE), boundary recall (BR), and precision. Each metric provides insights into different aspects of the segmentation quality.

Under Segmentation Error (UE)

Under segmentation error evaluates the extent to which the segmentation fails to cover the entire region of interest. It is defined as the average proportion of the union minus the intersection over the segmentation regions compared to the true segments. A lower UE indicates higher accuracy in covering the true segments fully. The formula for UE is expressed as:

$$UE = \frac{1}{N} \sum_{i=1}^N (G_i \cap S_i) \cup S_i - G_i$$

where G_i represents the ground truth region, and S_i denotes the segmented superpixel region.

Boundary Recall (BR)

Boundary recall measures how well the segmentation boundaries align with the actual boundaries of the region. Higher values indicate better alignment and segmentation accuracy. The formula for BR is:

$$BR = \frac{TP}{TP + FN}$$

where TP (true positives) are the correctly identified boundary pixels, and FN (false negatives) are the missed boundary pixels.

Precision

Precision quantifies the accuracy of the boundary delineation, where a higher score indicates more precise boundary segmentation. It is calculated as:

$$Precision = \frac{TP}{TP + FP}$$

where FP (false positives) represents the incorrectly identified boundary pixels outside the true boundary.

F1 Score

F1 score is the harmonic mean of precision and recall, useful when seeking a balance between precision and recall performance, particularly when they might be contradictory. It is especially beneficial in situations with an uneven class distribution. F1 score is calculated as:

$$F1_score = \frac{2Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FN + FP}$$

Mean F1 Score

To evaluate the model's performance across multiple classes, the mean F1 score across all classes is computed, providing a holistic measure of the model's overall accuracy:

$$\overline{F1_score} = \frac{1}{n} \sum_{i=1}^n F1_score_i$$

These metrics collectively enable a comprehensive evaluation of the segmentation algorithm, detailing its effectiveness in accurately delineating and recognizing regions within images.

3. Experiments and Results

3.1. GL-SLIC Segmentation Results

3.1.1. Comparison of Superpixel Segmentation Algorithms

To establish the efficacy of the GL-SLIC algorithm, a comparative analysis was conducted against mainstream superpixel segmentation algorithms: LSC, SLIC, FH, QS, SEEDS, and Watershed. The segmentation outcomes on sandstone thin section images were observed, and the results generated by each algorithm are presented in Figure 10.

Sandstone thin sections, characterized by an abundance of mineral grains with closely associated features and indistinct boundaries, present a higher degree of segmentation complexity compared to other natural images. A visual assessment reveals that while some algorithms perform admirably, others fall short. The FH, QS, SEEDS, and Watershed algorithms produce irregularly shaped superpixels of uneven sizes. Specifically, the FH algorithm shows redundant boundaries and multiple excessively small superpixels. Watershed aligns more closely with mineral boundaries but generates small, isolated areas within mineral particles that fail to merge. QS displays instances of over-segmentation, and SEEDS—although closely aligning with the mineral edges—lacks uniformity in shape and size.

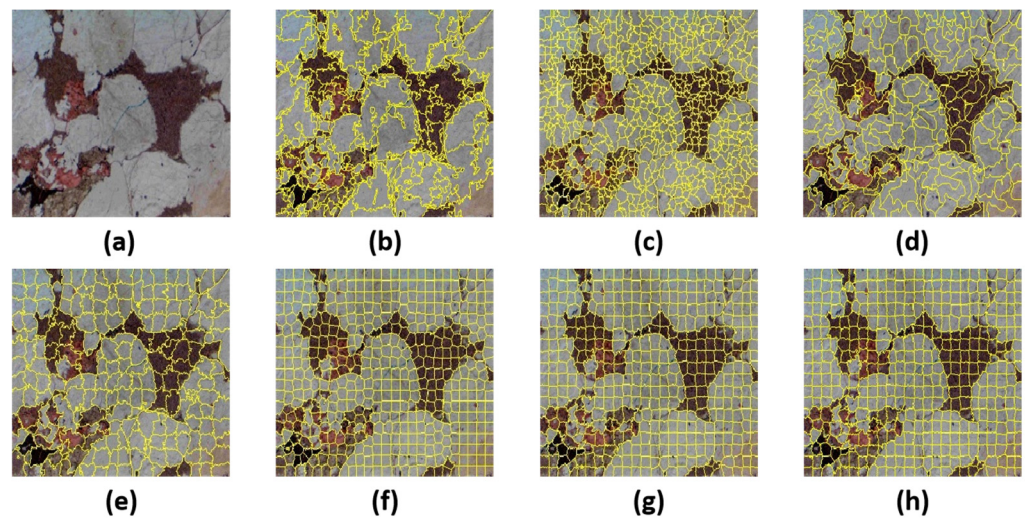


Figure 10. Comparison of superpixel segmentation algorithms on sandstone images: (a) original sandstone image; (b) FH; (c) QS; (d) SEEDS; (e) Watershed; (f) LSC; (g) SLIC; (h) GL-SLIC.

The SLIC and LSC algorithms, along with the proposed GL-SLIC algorithm, yield more regularly shaped, evenly sized, and compact superpixels. Notably, the LSC algorithm produces the most regular shapes but has lower adherence to the true mineral edges. Both SLIC and GL-SLIC create well-defined shapes with SLIC producing consistently regular superpixels.

In Figure 11, the red-boxed areas marked (a) and (b) are magnified to display the detailed segmentation comparison in Figure 12. In the comparison, areas (a1), (a2), and (a3)—produced by the SLIC algorithm—are juxtaposed with areas (b1), (b2), and (b3)—generated by the GL-SLIC algorithm—respectively. (a1) fails to discern the mineral grain clearly, while (b1) accurately encapsulates a smaller mineral particle. For (a2) and (b2), there is a noticeable deviation from the actual mineral boundary in (a2), whereas (b2) closely matches the real boundary. Similarly, (a3) does not recognize the protrusion in the mineral particle, which is precisely identified in (b3).

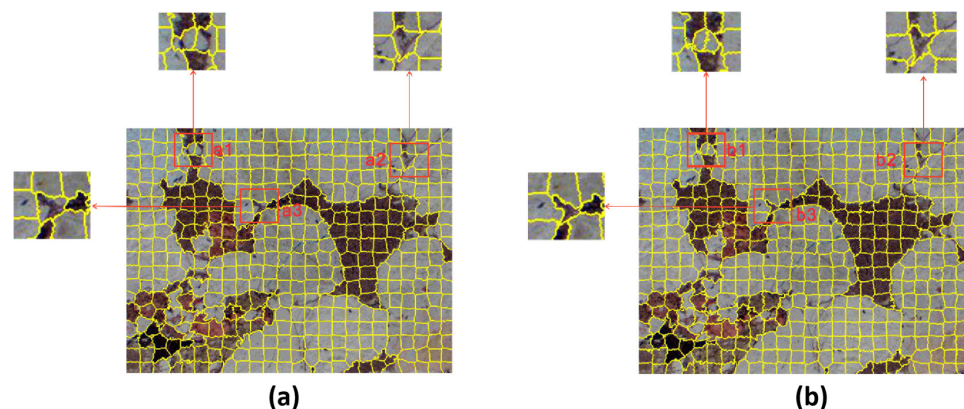


Figure 11. Comparison of segmentation results between SLIC and GL-SLIC algorithms: (a) pre-segmentation result by the SLIC algorithm; (b) pre-segmentation result using the GL-SLIC algorithm.

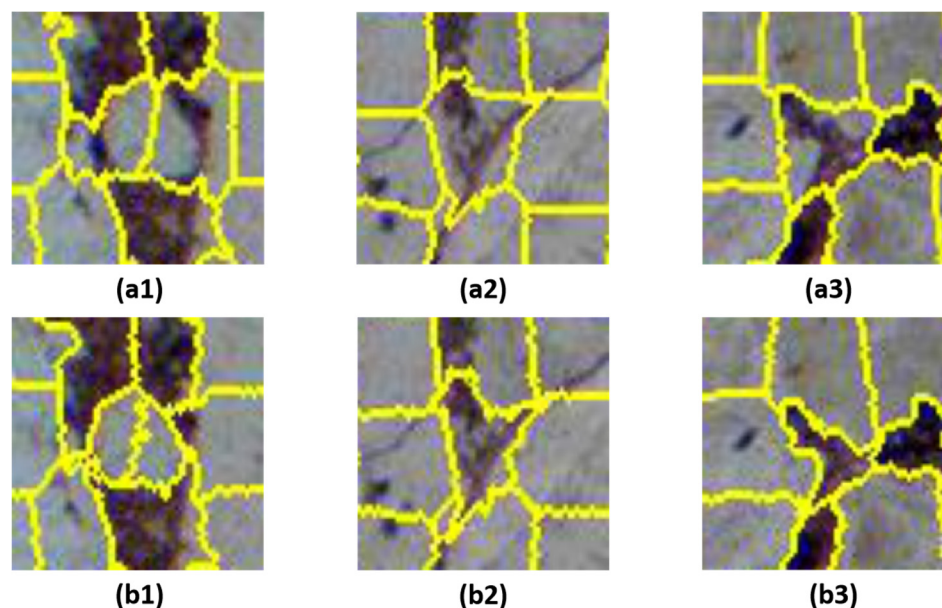


Figure 12. Detailed comparison between SLIC and GL-SLIC algorithms: (a1) detail area a1 from SLIC; (a2) detail area a2 from SLIC; (a3) detail area a3 from SLIC; (b1) detail area b1 from GL-SLIC; (b2) detail area b2 from GL-SLIC; (b3) detail area b3 from GL-SLIC.

Empirical observations demonstrate that the superpixels generated by the SLIC algorithm do not entirely conform to the actual mineral particle boundaries. In contrast, the GL-SLIC algorithm excels in aligning with the true edges. This improved adherence in GL-SLIC can be attributed to the incorporation of the GLBP texture features in the distance metric. Additionally, these features are considered during cluster center updates, making GL-SLIC more suitable for preprocessing sandstone thin section images.

Table 1 presents the averaged UE (undersegmentation error), BR (boundary recall), and precision metrics for the SLIC, LSC, QS, FH, SEEDS, and Watershed algorithms and the proposed GL-SLIC algorithm across 50 sandstone thin section images from the dataset. The GL-SLIC algorithm outperforms others on all three metrics, demonstrating its superior segmentation precision and fidelity to real mineral particle edges. This ensures more accurate mineral particle identification within the sandstone images. Overall, the GL-SLIC algorithm not only exhibits higher segmentation accuracy but also improves the fit to the true edges of mineral grains, reducing the undersegmentation error rate.

Table 1. Comparison of experimental results for superpixel segmentation algorithms.

Algorithm	UE	BR	Precision
FH	0.1345	0.5847	0.2563
QS	0.1232	0.5325	0.3648
SEEDS	0.1078	0.7453	0.5761
Watershed	0.0934	0.6147	0.5256
LSC	0.0773	0.6343	0.5272
SLIC	0.0844	0.6565	0.5846
GL-SLIC	0.0648	0.6987	0.6053

3.1.2. Validation of Region Merging Algorithm

Building upon the foundation laid by the GL-SLIC segmentation algorithm, this section validates the proposed region merging algorithm. Figure 13 illustrates the application of this algorithm on a medium-coarse-grained quartz sandstone image, showcasing the pre-segmentation and post-merging results.

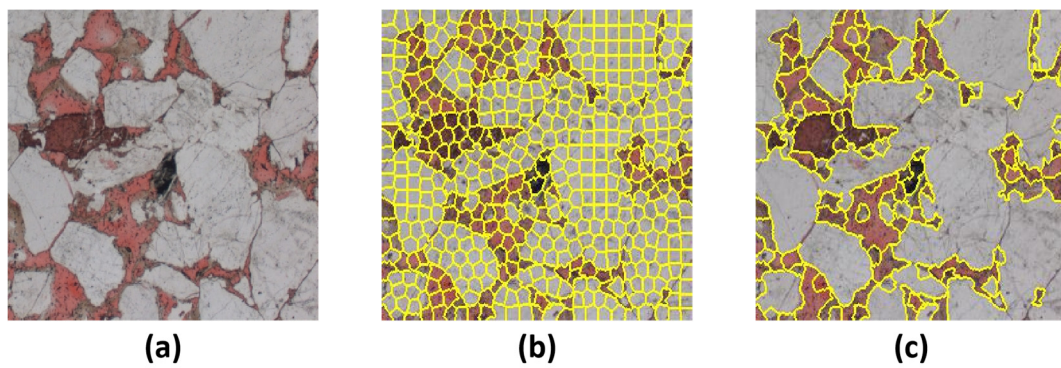


Figure 13. Comparison of superpixel merging in medium-coarse-grained quartz sandstone: (a) medium-coarse-grained quartz sandstone image; (b) pre-segmentation result; (c) result after superpixel merging.

Visually, the merged superpixels predominantly represent entire mineral grains. This indicates the algorithm’s efficacy in consolidating fragmented mineral particles and elongated pore structures into cohesive units. Figure 13c demonstrates that the superpixels after merging nearly mirror complete mineral particles, with minimal misclassification observed during the merging process.

However, the algorithm’s performance varies with the size and shape of the features. It effectively merges smaller quartz fragments and elongated pores, but larger pore spaces pose a challenge. These larger pores often contain diverse mineral inclusions, leading to significant internal feature variation, which can result in over-segmentation post-merging.

Table 2 presents the average results from 50 sandstone thin section images processed using the GL-SLIC algorithm both before and after applying the region-merging algorithm. The metrics used—undersegmentation error (UE), boundary recall (BR), and precision—show marked improvement post-merging. Notably, the precision and boundary recall values increase. This demonstrates that the merged superpixels align more closely with the true boundaries of mineral grains, resulting in more accurate representations of the actual geological structures.

Table 2. Experimental results of superpixel merging using the GL-SLIC algorithm.

Evaluation Metric	before Merging	after Merging
UE	0.0648	0.0513
BR	0.6987	0.7056
Precision	0.6053	0.7233

3.2. Experiments of Semi-Supervised Learning Framework

Figure 14 showcases the initial use of manually labeled data to prime both the primary and discriminator models. After initialization, both models exhibit testing accuracies above 90% but quickly encounter issues with overfitting. To address this, unlabeled monolithic rock data are introduced into both models to generate high-confidence data, enhancing the training dataset. By the second iteration, when the dataset size reaches 750 images, both models maintain accuracies above 90%, with overfitting issues significantly mitigated. At this stage, the classification models already show promising results. However, the monolithic rock dataset includes data beyond the manually labeled five categories, labeled as “other.” To appropriately classify this additional category, model iterations continue until the termination criteria are met. Subsequently, high-quality data from the remaining unlabeled monolithic rock data are selected as the “other” category.

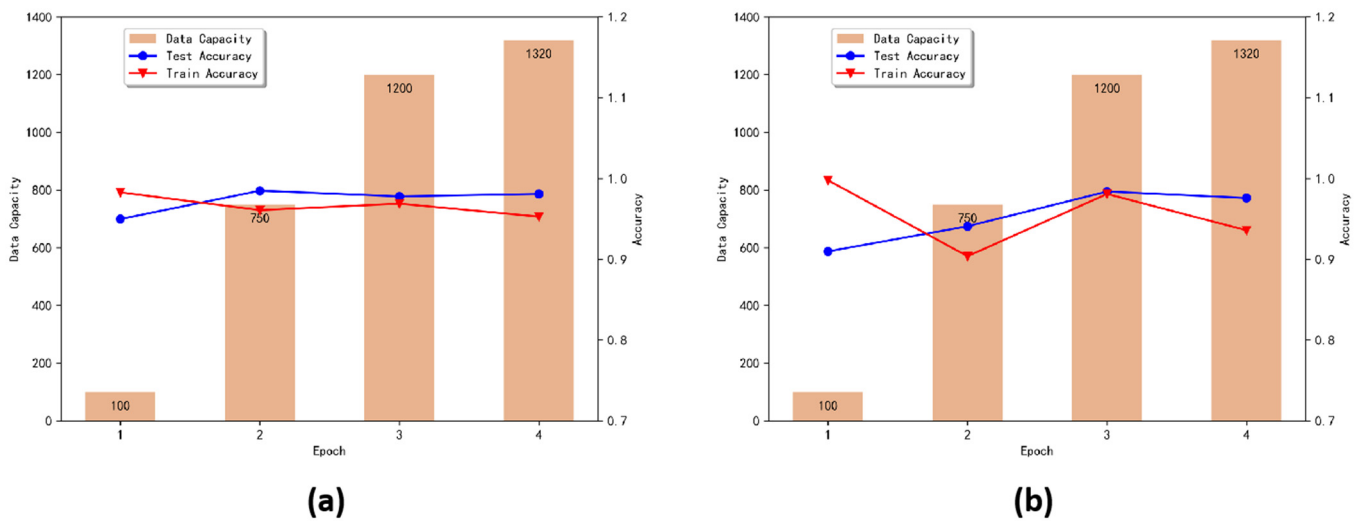


Figure 14. Iterative model training and data augmentation process using labeled and unlabeled rock data to mitigate overfitting and enhance classification accuracy: (a) primary model; (b) discriminator model.

Towards the end, data for each of the six categories—quartz, pores, rock fragments, matrix, kaolinite, and “other”—reach approximately 1300 images each. The primary model is then trained with this varied category data. After multiple experiments, the model converges around the 120th iteration. Figure 15 illustrates the variation in training and testing accuracies, with the last iteration showing a training accuracy of 88.7% and a testing accuracy of 90.2%.

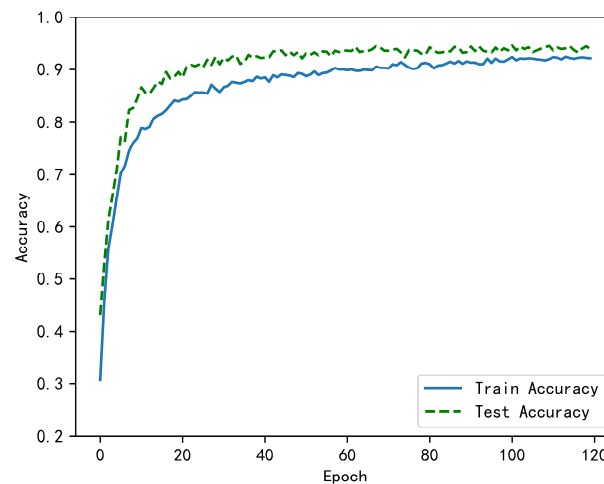


Figure 15. Curves of training and testing accuracy variation with epochs for the primary model.

To provide a detailed evaluation of the model’s performance and pinpoint the specific rock components that are problematic, confusion matrices for both the training and testing datasets are utilized, thereby illustrating the classification accuracies across all six rock types. This approach allows for a nuanced analysis of the model’s strengths and weaknesses in distinguishing between the different categories of rock components. Figure 16 shows the confusion matrices for the training and validation sets. It can be seen that, except for the “other” category, the accuracy for each class is relatively high, with an average accuracy of around 95%. However, when data from the “other” category are incorporated, the accuracy notably decreases to around 90%. This drop is primarily attributed to the inherent characteristics of the “other” category, which lacks distinctive features compared to the well-defined categories of quartz, pores, rock fragments, matrix, and kaolinite. Due to the ambiguous nature of this category, the model struggles to discern clear patterns within the

100 manually labeled samples of the “other” category, leading to a less precise learning outcome.

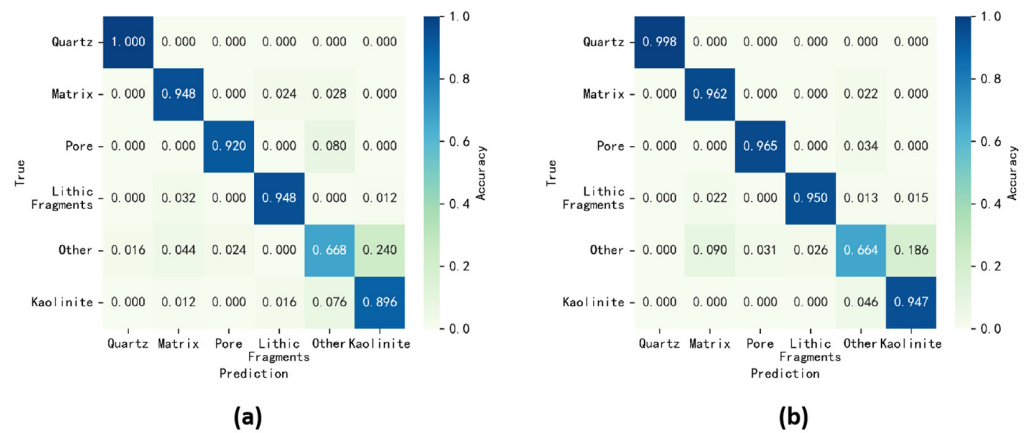


Figure 16. Classification accuracy analysis: (a) training set confusion matrix, (b) test set confusion matrix.

Analysis of Figure 16 reveals severe misclassification within the “other” category, containing data closely resembling other categories. Initially, only a limited selection of the five main rock components was manually labeled, leaving a substantial amount of data in the “other” category unsorted; the semi-supervised dual model approach attempts to address this by generating pseudolabels to expand the known categories. However, due to the complexity of rock images and the similarity among different components, the final unlabeled data assumed as “other” may still contain data from the other five categories.

To address the imbalance caused by the necessary removal of data wrongly included in the “other” category—which would reduce its volume by nearly 44%—image rotation methods are employed to enhance and balance the dataset. After cleansing and enhancing the dataset, the model converges by the 80th iteration with significantly improved accuracies: 93.7% for training and 96.3% for testing, the results are shown in Figure 17.

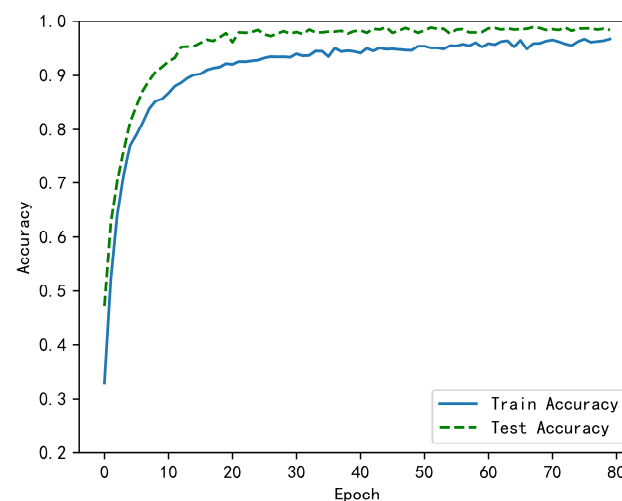


Figure 17. Improved model accuracy post dataset cleansing and enhancement.

The current model accuracy indicates significant improvement following the cleansing and enhancement of the dataset. The model’s capability to accurately classify different rock types is reassessed by generating and analyzing confusion matrices for the entire dataset and calculating the classification accuracy for each rock type. The confusion matrix presented in Figure 18 shows that while the classification accuracy for the “other” category is 89%, the accuracies for all other categories exceed 96%. This outcome confirms the

effectiveness of the strategies employed to manage the data within the “other” category. Additionally, the model achieves an F1 score of 96.4% on the test dataset, indicating that the classification performance for this monolithic rock dataset is highly satisfactory.

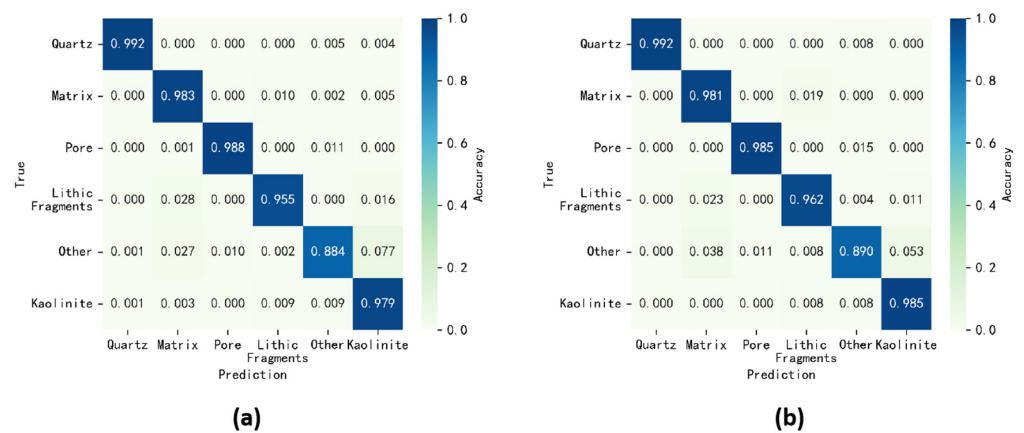


Figure 18. Final confusion matrices for model evaluation: (a) training data confusion matrix, (b) testing data confusion matrix.

In summary, this chapter integrates a semi-supervised dual-model training approach, utilizing both the primary and discriminator models to evaluate and generate high-confidence pseudolabel data through iterative refinements, ultimately achieving an expanded and highly accurate classification model for the entire monolithic rock dataset.

3.3. Component Identification Result

After achieving high classification accuracy, as evidenced by the confusion matrix and F1 score, we proceeded to integrate the classification results with segmentation to identify different mineral components within the original petrographic thin section images as in Figure 19. Comparing the method used in this study to the UNet-based semantic segmentation method, it is evident that our method yields clearer boundaries and more accurate recognition results. When calculating accuracy against manually annotated results, our method achieves an accuracy of 89.33%, whereas the semantic segmentation method only reaches 72.12%.

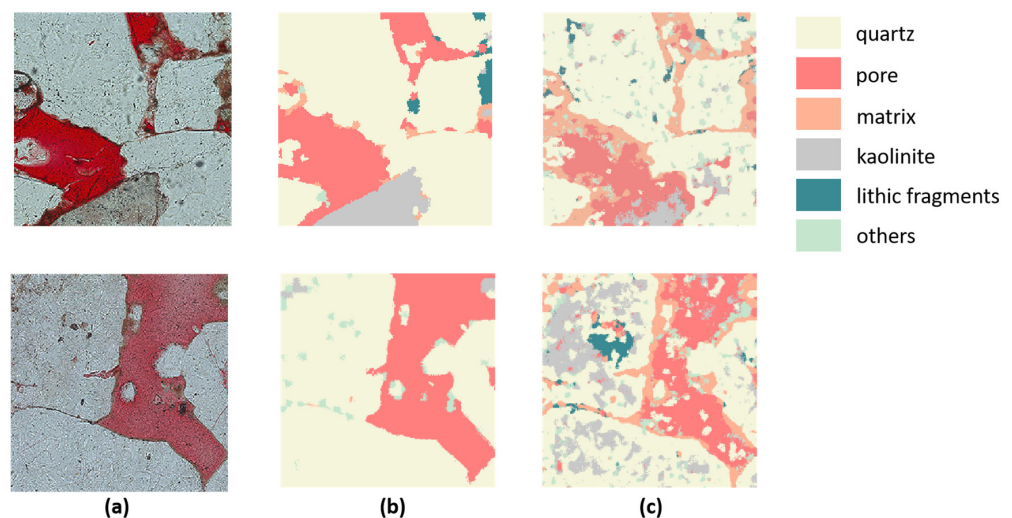


Figure 19. Component identification results: (a) original petrographic thin section images; (b) proposed method results; (c) UNet-based semantic segmentation results.

This qualitative assessment shows that the integration of classification with segmentation allows for a detailed and accurate identification of geological features within the rock samples. The clear delineation of quartz, pores, matrix, kaolinite, lithic fragments, and other components aligns well with geological expectations, providing a comprehensive understanding of the sample composition.

3.4. Discussion and Limitations

The results from our experiments highlight several key strengths and some limitations of the methods used for component identification in petrographic thin section images.

3.4.1. Strengths

High Classification Accuracy: The integrated approach combining classification and segmentation yielded high accuracy, with most rock components correctly identified. This is evident from the confusion matrix and F1 score, showing robust performance even with only 6% of the data labeled.

Clear Boundary Delineation: The boundaries between different rock components are well-defined, with minimal misclassification or merging errors. This clarity is crucial for accurate geological interpretation and analysis.

Comprehensive Component Identification: The approach effectively identifies and segments multiple components, including quartz, pores, matrix, kaolinite, and lithic fragments. This comprehensive identification supports detailed geological and petrophysical analyses.

3.4.2. Limitations

Incomplete Merging: One significant limitation of our study is the occurrence of incomplete merging, where certain rock components are not entirely segmented. This issue can affect the overall accuracy and completeness of the analysis, particularly in regions where mineral grains and pores have intricate boundaries. Future work should focus on refining the merging algorithms to better handle these complex structures.

Low Image Resolution: The relatively low resolution of the rock images used in this study may limit the level of detail captured, potentially impacting the accuracy of component identification, especially for finer geological features. Higher-resolution imaging techniques should be explored to enhance the precision of segmentation and classification tasks.

Lithological Limitations: The study's findings are based on a specific type of rock from a particular geological formation. The model's effectiveness across different lithologies remains uncertain. Therefore, further validation across various rock types and formations is necessary to ensure the generalizability and robustness of the proposed method.

Limited Data Annotation: With only 6% of the data annotated, the results are promising but may not fully represent the variability within the dataset. Expanding the annotated dataset could improve the model's performance and reliability. Additionally, the reliance on high-confidence pseudo labels might introduce biases if the initial labeled data is not sufficiently diverse.

Challenges in Feature Extraction: The effectiveness of the GL-SLIC algorithm relies heavily on the quality of feature extraction. Variations in lighting, sample preparation, and imaging conditions can affect the extracted features' quality. Enhancing the robustness of feature extraction methods to account for these variations is crucial for consistent performance.

4. Conclusions

This study presents an advancement in the automated segmentation and classification of sandstone thin section images by integrating the GL-SLIC algorithm with deep learning techniques. The semi-supervised learning framework addresses the challenge of limited labeled data by using a small set of manually labeled samples to generate high-confidence pseudo labels, thereby substantially expanding the training dataset.

The results indicate that this method enhances segmentation accuracy, particularly in delineating the intricate boundaries of mineral grains and pores. This precision is essential for detailed petrographic analysis, where accurate boundary definition is necessary for reliable component identification. The model achieves a high classification accuracy of up to 96.3% across various rock types, confirming its robustness and effectiveness in distinguishing different geological components.

Additionally, the semi-supervised approach is both efficient and scalable. It generalizes well from minimal initial input, which is particularly valuable in geological studies where labeled data is often limited. By improving the clarity and accuracy of geological interpretations, this method supports more informed decision-making in geological and petrophysical analyses.

In summary, the integration of traditional segmentation techniques with semi-supervised deep learning addresses the limitations of existing methods and achieves high accuracy in petrographic analysis. This study lays a foundation for future advancements in geological image analysis, highlighting the potential for adaptive, data-efficient methods to significantly enhance petrographic and broader geological studies.

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

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable to this article.

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

References

1. Yousef, I.; Morozov, V.P.; Kadi, M. Influence and control of post-sedimentation changes on sandstone reservoirs quality, example, upper Triassic (Mulussa F reservoir), and lower Cretaceous (Rutbah reservoir), Euphrates graben, Syria. *Russ. J. Earth Sci.* **2020**, *20*, 1–24. [[CrossRef](#)]
2. de Lima, R.P.; Duarte, D.; Nicholson, C.D.; Slatt, R.; Marfurt, K.J. Petrographic microfacies classification with deep convolutional neural networks. *Comput. Geosci.* **2020**, *142*, 104481. [[CrossRef](#)]
3. Liu, H.; Ren, Y.L.; Li, X.; Hu, Y.X.; Wu, J.P.; Li, B.; Luo, L.; Tao, Z.; Liu, X.; Liang, J.; et al. Rock thin-section analysis and identification based on artificial intelligent technique. *Pet. Sci.* **2022**, *19*, 1605–1621. [[CrossRef](#)]
4. Saxena, N.; Day-Stirrat, R.J.; Hows, A.; Hofmann, R. Application of deep learning for semantic segmentation of sandstone thin sections. *Comput. Geosci.* **2021**, *146*, 104895. [[CrossRef](#)]
5. Ren, Y.; Li, X.; Bi, J.; Zhang, Y.; Su, Q.; Wang, W. Multi-channel attention transformer for rock thin-section image segmentation. *J. Eng. Geol.* **2024**. [[CrossRef](#)]
6. Shebl, H.T.; Al Tamimi, M.A.; Boyd, D.A. Automation of Carbonate Rock Thin Section Description Using Cognitive Image Recognition. In Proceedings of the Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, United Arab Emirates, 15–18 November 2021. [[CrossRef](#)]
7. Yalamanchi, P.; Datta Gupta, S. Estimation of pore structure and permeability in tight carbonate reservoir based on machine learning (ML) algorithm using SEM images of Jaisalmer sub-basin, India. *Sci. Rep.* **2024**, *14*, 930. [[CrossRef](#)] [[PubMed](#)]
8. Chen, Y.; Yi, Y.; Dai, Y. A multiangle polarised imaging-based method for thin section segmentation. *J. Microsc.* **2024**, *284*, 20–34. [[CrossRef](#)] [[PubMed](#)]
9. Liu, Y.; Zhang, Q.; Zhang, N.; Lv, J.; Gong, M. Enhancement of thin-section image using super-resolution method with application to the mineral segmentation and classification in tight sandstone reservoir. *J. Pet. Sci. Eng.* **2022**, *212*, 110876. [[CrossRef](#)]
10. Vellappally, A.; Hou, S.; Emmings, J. Automated Grain Segmentation and Mineral Classification in Rock Thin Sections. In *85th EAGE Annual Conference & Exhibition; European Association of Geoscientists & Engineers*: Utrecht, The Netherlands, 2024. [[CrossRef](#)]
11. Dabek, P.; Chudy, K.; Nowak, I.; Zimroz, R. Superpixel-Based Grain Segmentation in Sandstone Thin-Section. *Minerals* **2023**, *13*, 219. [[CrossRef](#)]
12. Dong, L.; Gui, H.; Yu, X.; Zhang, X.; Xu, M. High-Accuracy Image Segmentation Based on Hybrid Attention Mechanism for Sandstone Analysis. *Minerals* **2024**, *14*, 544. [[CrossRef](#)]

13. Caja, M.Á.; Castillo, J.N.; Santos, C.A.; Pérez-Jiménez, J.L.; García, C. Digital Rock Physics in Cuttings Using High-Resolution Thin Section Scan Images. *Minerals* **2023**, *13*, 1140. [[CrossRef](#)]
14. Visalli, R.; Ortolano, G.; Godard, G.; Ziberna, L. Micro-Fabric Analyzer (MFA): A new semiautomated ArcGIS-based edge detector for quantitative microstructural analysis of rock thin-sections. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 51. [[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.