



Article

A Slope Structural Plane Extraction Method Based on Geo-AINet Ensemble Learning with UAV Images

Rongchun Zhang^{1,2} , Shang Shi¹, Xuefeng Yi^{3,4}, Lanfa Liu^{5,*} , Chenyang Zhang⁶, Meiru Jing¹ and Junhui Li^{1,7}

¹ School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

² School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

³ School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China

⁴ School of Network and Communication, Nanjing Vocational College of Information Technology, Nanjing 210023, China

⁵ Hubei Provincial Key Laboratory for Geographical Process Analysis and Simulation, Central China Normal University, Wuhan 430079, China

⁶ School of Civil Engineering and Architecture, Changzhou Institute of Technology, Changzhou 213032, China

⁷ Jiangsu Geologic Surveying and Mapping Institute, Nanjing 211102, China

* Correspondence: lanfa@cnu.edu.cn

Abstract: In the construction of large-scale water conservancy and hydropower transportation projects, the rock mass structural information is often used to evaluate and analyze various engineering geological problems such as high and steep slope stability, dam abutment stability, and natural rock landslide geological disasters. The complex shape and extremely irregular distribution of the structural planes make it challenging to identify and extract automatically. This study proposes a method for extracting structural planes from UAV images based on Geo-AINet ensemble learning. The UAV images of the slope are first used to generate a dense point cloud through a pipeline of SfM and PMVS; then, the multiple geological semantics, including color and texture from the image and local geological occurrence and surface roughness from the dense point cloud, are integrated with Geo-AINet for ensemble learning to obtain a set of semantic blocks; finally, the accurate extraction of structural planes is achieved through a multi-semantic hierarchical clustering strategy. Experimental results show that the structural planes extracted by the proposed method perform better integrity and edge adherence than that extracted by the AINet algorithm. In comparison with the results from the laser point cloud, the geological occurrence differences are less than three degrees, which proves the reliability of the results. This study widens the scope for surveying and mapping using remote sensing in engineering geological applications.

Keywords: UAV; digital images; slope; rock mass; structural planes; ensemble learning



Citation: Zhang, R.; Shi, S.; Yi, X.; Liu, L.; Zhang, C.; Jing, M.; Li, J. A Slope Structural Plane Extraction Method Based on Geo-AINet Ensemble Learning with UAV Images. *Remote Sens.* **2023**, *15*, 1441. <https://doi.org/10.3390/rs15051441>

Academic Editors: Basem Zoheir and Ashraf Emam

Received: 1 January 2023

Revised: 19 February 2023

Accepted: 2 March 2023

Published: 4 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In significant engineering projects in hydropower, transportation, mining, etc., the geological environment is complex, and there may be various engineering geological problems, such as high and steep slope stability and dam stability. Therefore, it is necessary to obtain the rock mass structural information in time to evaluate and analyze these problems. The accurate acquisition of structural planes is an essential basis for rock mass structural information analysis. The rock mass structural plane is a planar geological interface produced in the rock mass under tectonic stress, and its physical and mechanical properties determine the nature of the rock mass. Many regular geometric characteristics exist in the artificial environment, including regular planes, straight lines, circles, curved surfaces, and other manifold structures, which provide available features for object recognition and

extraction. By contrast, the rock mass structural planes perform extremely complex formations, and experience tectonic movements of different natures in different periods. Therefore, various natural forms and complicated distribution make it difficult to extract the structural planes automatically. The high-level and low-level features from images and point clouds can be accurately extracted by deep learning technology, which has significant advantages in structural plane recognition and extraction. In addition, hydropower, transportation, and other engineering projects are primarily located in high mountains and valleys, which are highly dangerous and inaccessible, making it difficult to achieve comprehensive on-site geological surveys. UAV image acquisition can quickly reach full coverage of complex terrain, and this non-contact photogrammetry technology has better applicability to geological surveys. Therefore, this paper provides research on the slope structural plane extraction from UAV images based on the ensemble deep learning strategy.

Photogrammetry and 3D laser scanning techniques provide non-contact measurement methods for extracting rock mass structural information. Photogrammetry can extract rock mass structural information through image interpretation or 3D model measurement and analysis [1–3]. The 3D model of the rock mass surface is generated from multi-view images based on photogrammetry or computer vision methods [4–6]. The 3D laser scanning technology can efficiently and quickly acquire massive 3D point clouds of the target surface in high precision and high spatial resolution. It has been widely applied in geological engineering fields [7]. Currently, many studies focus on the planar geometric feature extraction of rock mass [8–11].

In the 1970s, Ross-Brown et al. first used calibrated images to interpret the direction and trace length of structural planes, pioneering the application of photogrammetry technology to engineering geology [1]. Lee et al. trained a classifier based on the DeepLabV3+ network to detect joint traces from digital images and to calculate their length using point cloud data generated by the stereo photogrammetry technique [12]. Xu et al. proposed a fast-fuzzy clustering method for discontinuity sets using a point cloud model generated by close-range photogrammetry technology [13]. Kong et al. researched rock mass discontinuity identification and clustering using 3D point clouds, and proposed a method for calculating normal vectors, made by clustering a density peaks algorithm, and obtained several discontinuity parameters [14]. Li et al. also proposed a method for measuring the occurrence of structural planes based on the principle of a central projection vanishing line and vanishing point [15]. Leu et al. presented a rock mass structural feature extraction method based on image processing technology. The faults, joints and fissures of a tunnel's surface can be extracted to assist geologists in analyzing and evaluating the tunnel's excavation face [16]. Wang and Liu established the object-image relationship model of the slope through the digital photogrammetry system and combined it with the structural plane trace visualization model to obtain the trace length and occurrence information of the structural planes [17,18]. Bi et al. used aerial images to obtain the geometric morphology of formation faults in the local area of the Altyn fault area and the micro-fault geomorphological features near the faults based on the Structure from Motion technique [19]. Xiao et al. extracted the contour information of cracks in a dangerous rock mass based on the grayscale and spatial features presented on UAV images [20].

The TLS technology was first used by Kocak et al. in the outcrop exploration of seabed rock formations [21]. Feng et al. successively applied TLS technology to the measurement of an exposed rock mass surface and the roughness and trace measurement of the structural planes [22]. Slob et al. carried out triangulation reconstruction on the point cloud, and then the structural planes were automatically extracted through fuzzy K-means clustering [23]. Riquelme et al. determined the plane equation of the structural plane through the coplanar test of adjacent points and identified the exposed rock mass' structural planes [24]. DK et al. proposed a structural plane extraction method. In this method, the normal vector is first calculated by iterative weighted plane fitting, and structural planes clustering is performed by combining it with fast search and density peaks. Then, the structural plane fitting is finished using random sampling consistency [25]. Battulwar et al. compared the

automatic extraction methods of rock discontinuity features based on 3D surface models. They concluded that the region growth method is faster and more accurate for joint detection [8].

In summary, the existing non-contact rock mass structural information extraction methods mainly use point clouds or images as the data source, and different methods have pros and cons. For example, the three-dimensional measurement method based on a laser point cloud mainly determines the plane by selecting coplanar points to extract structural planes [26–28]. Due to the significant difference in the exposure range of the rock mass' structural planes, and some structural surfaces are curved, undulating, and rough, it is difficult to determine the optimal parameters for the plane fitting method, which may lead to incorrect extraction results. The method based on a single image considering a vanishing point is more suitable for outcrops with enough thickness and extension length. The photogrammetry method based on stereopairs mostly depends on a human–computer interaction mode [6,29,30].

Empirically, image processing methods have greater advantages when extracting the two-dimensional geometric features of rock masses, and the rich and intuitive color and texture play an important role for image segmentation. Superpixel segmentation provides an efficient solution for image segmentation and has been extensively studied [31,32]. Traditional superpixel segmentation methods can be roughly divided into two categories, i.e., gradient-based segmentation and graph-based segmentation [33]. Simple Linear Iterative Clustering (SLIC) is a classic superpixel segmentation algorithm, in which a k-means iterative clustering is performed to achieve superpixel segmentation. The iterative process mainly includes two steps: (i) pixel–superpixel hard association; (ii) superpixel center update. In recent years, the research and application of deep learning technology in computer vision and other fields has proliferated, and related research on combining deep neural networks and superpixels has gradually emerged. However, since the definition of standard convolution operations is performed under regular grids in most deep network architectures, the processing efficiency of irregular grid units will be greatly affected. Moreover, most of the current superpixel algorithms are not differentiable. Therefore, it is necessary to add nondifferentiable modules to combine superpixels with neural networks, such as end-to-end trainable networks.

Superpixel segmentation with fully convolutional networks (FCN) provide a solution [34]. However, there is a skip-connect operation in the full convolutional network. Moreover, the low-level pixel–pixel relationship is introduced into the superpixel segmentation algorithm. Therefore, they both have bad effects on the segmentation results. Wang et al. proposed an AINet superpixel segmentation algorithm, which integrated the Association Implantation (AI) Module into the fully convolutional network to directly predict the relationship of pixel–superpixel [35]. This algorithm effectively improves the segmentation efficiency. In addition, a new loss function considering the boundary-perceiving loss is proposed in the algorithm, and this helps to improve the edge consistency of superpixels [35]; however, AINet superpixel segmentation only considers the 2D features of the image. In this study, an improved Geo-AINet method for slope structural plane extraction from UAV images is proposed. Both 2D and 3D semantics are used to divide the rock slope into a series of small blocks with multi-dimensional feature perception capabilities. Then, structural planes can be extracted through multi-dimensional semantic hierarchical clustering. The proposed method fully integrates both the 2D and 3D multiple features of the rock slope to measure the similarity of small blocks, which can effectively improve the accuracy of structural plane extraction. Moreover, compared with a single feature, multiple features have more significant impacts on the identification of the structural plane.

2. Methodology

The flow chart of the Geo-AINet rock mass structural plane extraction method proposed in this study is shown in Figure 1.

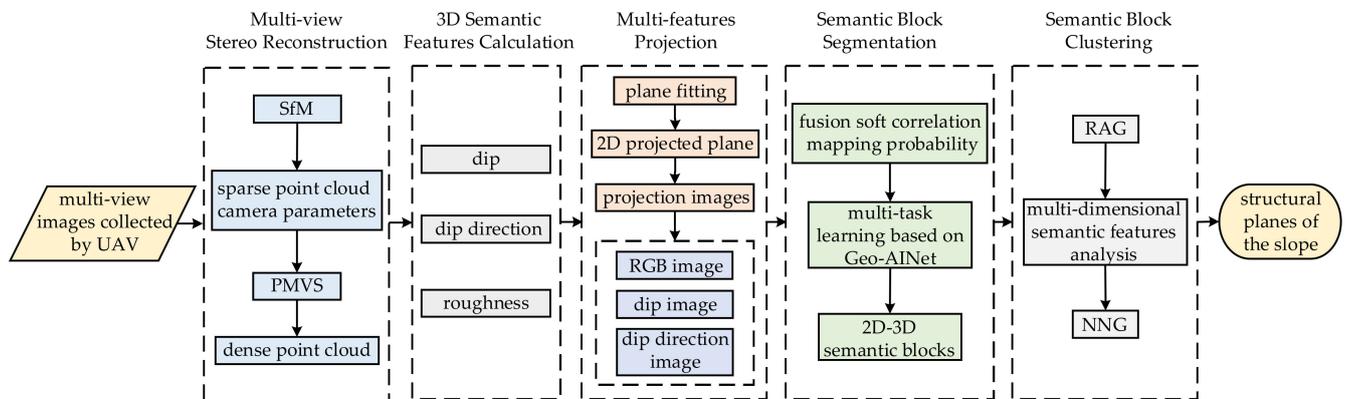


Figure 1. The flowchart of the proposed Geo-AINet rock mass structural plane extraction method.

The method is comprised of five parts: (1) Multi-view Stereo Reconstruction: the UAV multi-view images of the slope are collected and used to estimate the camera parameters and generate a sparse point cloud by the Structure from Motion technique, and the Patch-based Multi-view Stereo (PMVS) method is then used for dense reconstruction to generate a dense point cloud of the slope; (2) 3D Semantic Features Calculation: in this study, 3D geological semantics including dip, dip direction, and roughness are selected for structural plane extraction. For each point, the dip and dip direction can be calculated according to the normal vector of the local best fitting plane obtained by its nearest neighbors, and the roughness can be obtained by the open source software CloudCompare; (3) Multi-features Projection: a projection plane is defined according to the spatial orientation of the dense point cloud of the slope, and the 2D RGB and 3D geological semantics of the dense point cloud are, respectively, projected onto the 2D plane to obtain 2D projection correlation images (Details will be introduced in Section 2.2); (4) Semantic Block Segmentation: using AINet as the basis function to establish a Geo-AINet model for ensemble learning, then the slope is divided into a set of semantic blocks; (5) Semantic Block Clustering: both the Region Adjacent Graph (RAG) and the Nearest Neighbor Graph (NNG) involving multi-dimensional geological semantics, including RGB, dip, dip direction, and roughness, are established according to the 2D projection association images to complete the structural surface clustering.

2.1. Multiple Geological Semantic Features of the Structural Plane

1. The color and texture features from the 2D pixel: the color and texture can be used to identify geological and non-geological bodies (such as vegetation, buildings, etc.). Generally, the color and texture are diverse for structural planes. For example, black organic films are attached to some rock mass structural surfaces; yellow and red muddy fillings may exist in weak interlayers; calcareous and siliceous fillings are generally white; some friction marks may adhere to the surface of the structural face, which form various textures; some rocks, such as granite, have different colors and textures. Therefore, the color and texture are essential semantic elements for segmentation and classification. In this study, the color and texture features will be used for projection, segmentation and vegetation filtering from the clustering results.
2. The geological occurrence features from the 3D space: the occurrence of structural planes can reflect the spatial distribution of structural planes and is an important parameter for rock mass stability analysis. Generally, the occurrence is characterized by three parameters; i.e., dip, dip direction, and strike. The dip direction and strike are related, and there is a 90 degree difference between them. Therefore, in this study,

only dip and dip direction are selected as the two important geological semantic features. There is a spatial relationship between the occurrence and the normal vector of a structural plane. The open source software CloudCompare provides several methods to obtain the normal vectors of a point cloud that can be converted to dip and dip direction. In this study, the dip and dip direction are mainly used for projection, segmentation and clustering.

3. The roughness from the 3D morphological features: the surface of a rock mass may be smooth, rough, or slightly rough. Taking roughness as a geological semantic feature may play an important role in the identification of structural planes. In this study, the roughness can be obtained by the open source software CloudCompare. For each point in a point cloud, the roughness value refers to the distance between this point and the best fitting plane computed on its nearest neighbors. In this study, the roughness is mainly used for non-structural planes recognition from merging results.

2.2. Multi-Features Semantic Association Projection Images Generation

The characteristics of structural planes depend on 2D and 3D multiple semantic features, while traditional image segmentation methods do not fully consider the comprehensive influence of multi-dimensional information. In this study, a dense point cloud is projected onto a 2D plane to generate a group of multi-feature semantic association projection images; i.e., the RGB projection image and the geological occurrence images (the dip projection image and the dip direction projection image). Multiple semantic features between images are associated pixel by pixel, which can provide input data with multi-features for multi-task learning based on Geo-AINet. The generation of the association projection images is as follows:

1. Definition of a 2D projection plane of a rock mass slope

A 2D projection plane is established according to the spatial orientation and distribution of the slope. For the dense point cloud of a slope, the object space coordinates of a 3D point P in the $O - XYZ$ coordinate system are denoted as (X, Y, Z) . A 2D projection plane is obtained by plane fitting of the dense point cloud of the rock slope along an approximate XOY coordinate plane. The geometric relationship between the three-dimensional coordinate system of the dense point cloud and the 2D coordinate system of the projected plane is shown in Figure 2. From Figure 2 it can be seen that the 2D projection plane is parallel to the fitting plane of the point cloud of the slope, and it is apparently that the plane is not unique. It is worth mentioning that for the irregular shape of the rock slope, the occlusion problem should be avoided during projection as much as possible.

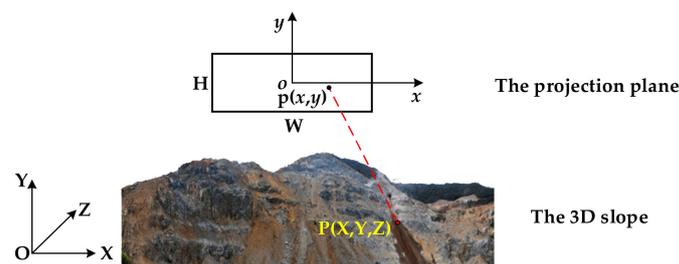


Figure 2. The schematic diagram of the geometric relationship between the 3D space coordinate system of the dense point cloud and the 2D projected plane coordinate system.

2. Determination of the size of the projection plane

The size of the 2D projection plane should be appropriate in order to obtain good projection resolution. Generally, the width and the height are determined by the spatial resolution of the point cloud as well as the size of the slope. Define a coordinate system $o - xy$ in the 2D projected plane, which is shown in Figure 2. Let the spatial resolution of the dense point cloud be δ , and the maximum and minimum values of the X-coordinate

and Y -coordinate of the dense point cloud be X_{max} , X_{min} , Y_{max} , and Y_{min} , respectively. Therefore, the projection coordinates can be calculated by Equation (1):

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (X_{max} - X)/\delta \\ (Y - Y_{min})/\delta \end{bmatrix}, W = \Delta X/\delta; H = \Delta Y/\delta \quad (1)$$

where $\Delta X = |X_{max} - X_{min}|$, $\Delta Y = |Y_{max} - Y_{min}|$. W and H refer to the width and the height of the projection plane.

3. Multi-features semantics projection

The dense point cloud generated by the multi-view stereo reconstruction pipeline carries RGB features, and the local geological occurrence semantic features can also be calculated from the normal vector of each 3D space point. The multi-dimensional semantics of the dense point cloud are, respectively, projected onto the 2D plane. Therefore, the multi-feature semantic association projection images of the slope façade are generated. The multi-dimensional semantic features, e.g., RGB, dip, and dip direction, are, respectively, assigned to the corresponding pixels of the 2D projection image to obtain multi-feature semantic association projection images of the rock slope. Therefore, the mapping relationship between the RGB and the spatial semantic features is established, which provides crucial data for semantic block segmentation by ensemble learning based on Geo-AINet.

2.3. Semantic Block Segmentation Based on Geo-AINet Ensemble Learning

Superpixel segmentation has apparent advantages over traditional pixel-based segmentation algorithms, and superpixel segmentation considers the correlation between pixel features, which can improve segmentation accuracy [36]. Generally, superpixel segmentation is achieved by dividing the image into a series of regular grid cells and estimating the relationship between each pixel and its adjacent grid cells, the accuracy of which has a great effect on the segmentation results. Similar to a superpixel, a semantic block with multi-dimensional feature perception capability is defined in this study, which is composed of a set of pixels with similar semantics features for the structural plane or the non-structural plane. The semantics of pixels are similar inside each semantic block and different between semantic blocks.

2.3.1. The Traditional Superpixel Segmentation Algorithms Based on Deep Learning

An innovative FCN superpixel segmentation algorithm was first proposed in 2020 [34]. In this method, the FCN is adopted for deep learning under regular grids. In the initialization step, the traditional superpixel strategy is applied to associate pixels with regular grid units. Superpixel segmentation is completed by finding the association score between image pixels and regular grid units. A simple and standard FCN structure is used for superpixel segmentation under regular grid cells. In the network, the traditional down-sampling and up-sampling convolution calculation are replaced by the scheme based on a superpixel, which can more effectively retain the target edge details and efficiently improve the segmentation efficiency.

In a word, the superpixel segmentation based on the FCN refers to obtaining a soft association matrix with a dimension of $H \times W \times 9$ through a U-Net. Here, H and W are the height and width of the input image, respectively. This matrix quantitatively reflects the relationship between the pixel and its surrounding nine superpixels. It can also be regarded as a probability that the current pixel belongs to each surrounding superpixel.

First, the attribute $h(s)$ of each superpixel is estimated according to the colors and the positions of the pixels inside it.

$$h(s) = \frac{\sum_{p:s \in N_p} l(p) \cdot q(p,s)}{\sum_{p:s \in N_p} q(p,s)} \quad (2)$$

where $l(p)$ refers to the attribute of a pixel, N_p denotes a superpixel set around a pixel p , and $q(p, s)$ denotes the probability that the pixel p is assigned to its surrounding superpixels s .

Then, the reconstructed attribute $l'(p)$ of this pixel is calculated as Equation (3).

$$l'(p) = \sum_{s \in N_p} h(s) \cdot q(p,s) \quad (3)$$

The training loss is expressed as the distance between the ground-truth attribute and the reconstructed attribute, and it can be described as Equation (4).

$$\mathcal{L}(Q) = \sum_p \text{dist}(l(p), l'(p)) \quad (4)$$

On the basis of the above algorithm, the AINet superpixel segmentation algorithm was proposed [35], and its basic principal framework is shown in Figure 3. The input of the network is an image, and the output is an association map Q . First, convolution operations are performed to extract the pixel embedding and the superpixel embedding, which are then fed into the AI module. Then, the corresponding neighborhood superpixel features are implanted around each pixel embedding for expansion. Finally, a convolution with a kernel size of 3×3 is performed to achieve knowledge propagation and obtain the pixel–superpixel associations.

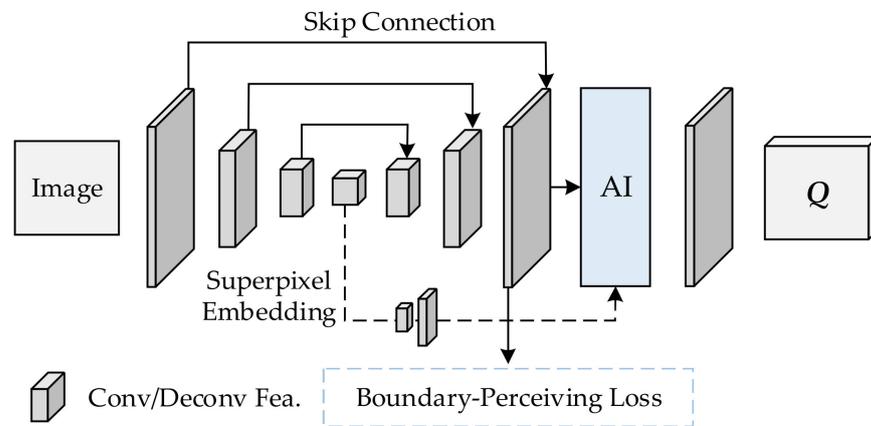


Figure 3. The basic principal framework of the AINet superpixel segmentation algorithm [35].

2.3.2. The Geo-AINet Ensemble Learning Superpixel Segmentation Algorithm

The semantic segmentation process is an important prerequisite for the subsequent accurate clustering of structural planes. The network architecture of the Geo-AINet ensemble learning is shown in Figure 4.

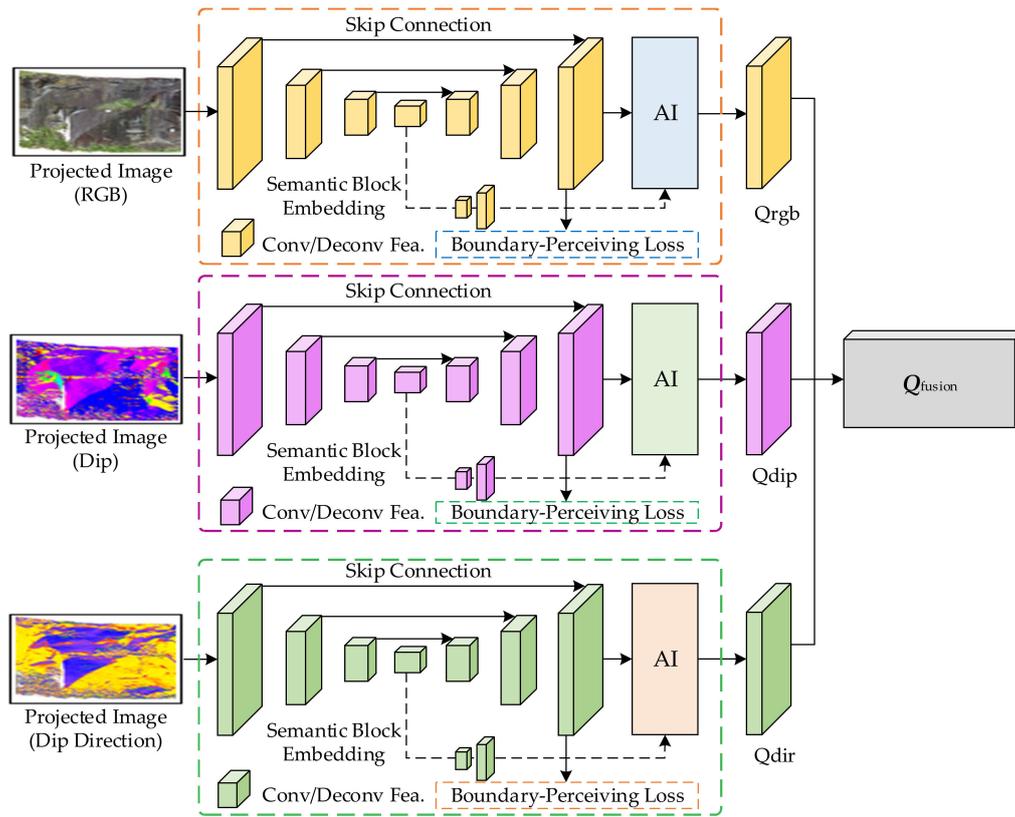


Figure 4. The network architecture of Geo-AINet ensemble learning.

In the proposed algorithm, three groups of pixel embeddings, respectively, corresponding to multiple semantic features (i.e., color and texture, dip, and dip direction) are obtained through a deep neural network, which are denoted as $E_{rgb} \in R_{rgb}^{H \times W \times D}$, $E_{dip} \in R_{dip}^{H \times W \times D}$, and $E_{dir} \in R_{dir}^{H \times W \times D}$, respectively. For a pixel p , its embeddings involving the above semantic features can be represented by $e_p^{rgb} \in R_{rgb}^D$, $e_p^{dip} \in R_{dip}^{dip}$ and $e_p^{dir} \in R_{dir}^{dir}$, which is shown in Figure 4. Let the sampling interval be S , and the input image compressed by multiple operations with convolutions and max poolings to generate three feature maps of grid cells with multi-dimensional semantics; i.e., $M_{rgb} \in R_{rgb}^{h \times w \times D}$, $M_{dip} \in R_{dip}^{h \times w \times D}$, and $M_{dir} \in R_{dir}^{h \times w \times D}$, where $h = H/S$, $w = W/S$.

The three feature maps $M_{rgb} \in R_{rgb}^{h \times w \times D}$, $M_{dip} \in R_{dip}^{h \times w \times D}$, and $M_{dir} \in R_{dir}^{h \times w \times D}$ are transformed into further, new feature maps through the 3×3 convolution operation, and are represented by $\hat{M}_{rgb} \in R_{rgb}^{H \times W \times D}$, $\hat{M}_{dip} \in R_{dip}^{H \times W \times D}$, and $\hat{M}_{dir} \in R_{dir}^{H \times W \times D}$, respectively. Therefore, the embedding of the nine grid cells around pixel p are defined according to Equation (5), which directly associates the pixel with the semantic block.

$$SP = \begin{bmatrix} \hat{m}_{tl} & \hat{m}_t & \hat{m}_{tr} \\ \hat{m}_l & \hat{m}_c + e_p & \hat{m}_r \\ \hat{m}_{bl} & \hat{m}_b & \hat{m}_{br} \end{bmatrix} \quad (5)$$

where, $SP = [SP_{rgb} \quad SP_{dip} \quad SP_{dir}]$, $\hat{m}_{|\cdot|} = [\hat{m}_{|\cdot|}^{rgb} \quad \hat{m}_{|\cdot|}^{dip} \quad \hat{m}_{|\cdot|}^{dir}]$.

The association map can be predicted using a 3×3 convolution and Equation (6).

$$\begin{cases} e_p^{rgb} = \sum_{ij} SP_{rgb}^{ij} \times \omega_{ij} + b \\ e_p^{dip} = \sum_{ij} SP_{dip}^{ij} \times \omega_{ij} + b \\ e_p^{dir} = \sum_{ij} SP_{dir}^{ij} \times \omega_{ij} + b \end{cases} \quad (6)$$

The proposed method adopts the same loss function as the AINet superpixel segmentation method, which includes three items of the cross-entropy loss, the reconstruction losses of pixels (Equation (4)) and the boundary-perceiving loss. The loss function for the Geo-AINet is expressed as Equation (7).

$$\mathcal{L} = \sum_p CE(l_s(p), l'_s(p)) + \lambda \|p - p'\|_2^2 + \alpha \mathcal{L}_B \tag{7}$$

where, λ and α are weight factors for a tradeoff of the loss items; \mathcal{L}_B represents a classification loss term used to enhance feature discrimination, which can effectively improve the edge accuracy of the semantic block.

A new set of embedded pixels $E' = \{E_{rgb}, E_{dip}, E_{dir}\}$ can be calculated by Equations (5) and (6), which directly reflects the pixel-superpixel associations on color and texture, dip, and dip direction. In the proposed method, the AINet is used as the base learner, and multi-feature semantic association projection images are used as multiple inputs of Geo-AINet. Therefore, the association maps Q_{rgb} , Q_{dip} , and Q_{dir} can be predicted, which are further integrated according to Equation (8). Then, a soft association map Q_{fusion} considering multi-feature semantics is obtained. Finally, a group of semantic blocks can be extracted from Q_{fusion} .

$$Q_{Fusion} = \lambda_1 Q_{rgb} + \lambda_2 Q_{dip} + \lambda_3 Q_{dir} \tag{8}$$

where λ_1 , λ_2 , and λ_3 denote the weight factors for the three association maps Q_{rgb} , Q_{dip} , and Q_{dir} , respectively, and they satisfy $\lambda_1 + \lambda_2 + \lambda_3 = 1$.

The detailed calculation process of the soft association map Q_{Fusion} is shown in Figure 5. In Figure 5, r_1-r_9 , d_1-d_9 , and v_1-v_9 , respectively, refer to three probability distributions, which reflect the similarities in the RGB, dip, and dip direction semantics between the pixel p in the i -th row and j -th column and its surrounding nine neighboring semantic blocks. f_1-f_9 represents the probability distribution of the pixel $p(i, j)$ on multi-feature semantics. The initial semantic block centers are defined by regular pixel blocks, and the centers will be optimized according to the association between each pixel and its surrounding pixel blocks. The optimization is achieved iteratively [34]. Label mapping can be obtained by taking the maximum values of nine probabilities for each pixel, which corresponds to the result of the semantic block segmentation.

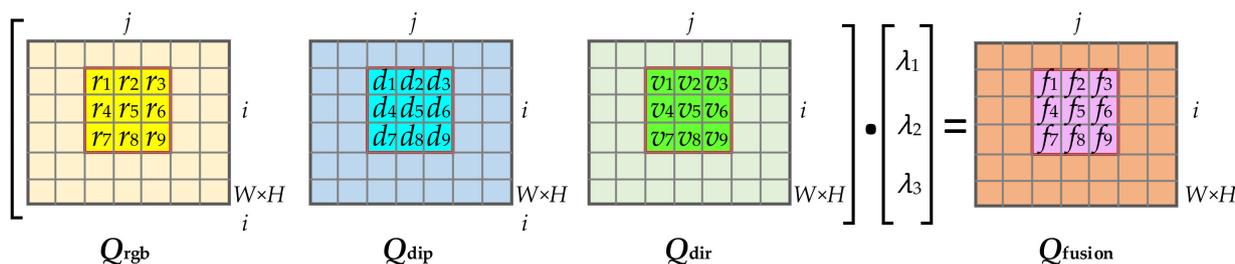


Figure 5. The detailed calculation process of the soft association map Q_{Fusion} on the pixel $p(i, j)$.

2.4. Semantic Block Clustering and Structural Plane Extraction

The semantic blocks generated from Section 2.3 are over-segmented results, which should be further clustered to obtain the extraction results of the rock slope’s structural planes. In this study, the topological adjacency of all semantic blocks is firstly expressed by a RAG proposed by [37]. Then, the multi-dimensional geological semantics are used to define a region dissimilarity, which measures the similarity of adjacent semantic blocks in the RAG. The RAG is further transformed into an NNG, where efficient and fast clustering of semantic blocks is achieved by merging connected bidirectional edges.

Figure 6a–d, respectively, shows the schematic diagram of the segmented semantic blocks, the corresponding RAG and NNG, and the merging results. Figure 6a shows a superimposed display of segmented semantic blocks (blue and red curves) and an image. The numbers ①–⑳ in Figure 6b,c represent nodes of RAG and NNG, respectively. Each node denotes a semantic block, and two adjacent semantic blocks connect with an edge, the length of which reflects the similarity of the two semantic blocks. In Figure 6d, the blue and pink regions represent two structural planes. The semantic blocks belonging to the two regions should be merged, respectively, and then the structural planes can be extracted.

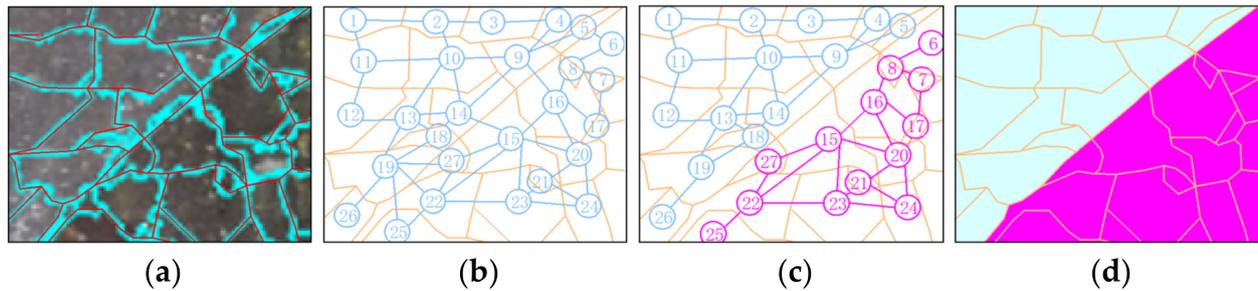


Figure 6. The schematic diagram of the semantic blocks' clustering: (a) the segmented semantic blocks; (b) RAG; (c) NNG; (d) the merging results.

Semantic blocks are comprised of valid and invalid structural planes, and the latter may include vegetation and non-structural plane rock masses. Therefore, it is necessary to perform a further filtering process on the merging results to eliminate these invalid structural planes. For example, the RGB semantic can be used to identify green vegetation areas, and the roughness semantic can be used to recognize non-structural planes. Finally, the structural planes can be extracted successfully.

3. Experiment and Analysis

In this study, the slope of a dismissed quarry in Australia was used for the experiments. The multi-view image sequence was collected by DJI UAV. The primary optical axes of the images were perpendicular to the rock slope surface. A total of 98 digital images of the rock slope were acquired, which have enough overlap (more than 80%) to guarantee multi-view stereo reconstruction. The camera model was an FC300X, and the image resolution was 2000×1500 pixels. The focal length f and ISO were set to 2.8 and 100, respectively. A total of 18 coded and 13 natural features were used as control points, and their coordinates were measured using two Leica TS11 reflectorless total stations. Therefore, the dense point cloud generated by multi-view images could be scaled. A laser scanning point cloud was acquired using a Leica ScanStation C10 TLS, and the accuracy was in the range of ± 4 mm. It was used to provide the ground truth for a comparison of the geological occurrence extracted by the proposed method. The width and height of the whole rock wall was, respectively, 80 m and 6 m. Figure 7 shows two parts of the study area of the slope, which were used for our experiments. The considered area of the wall was about 20 m long and 5 m high.

A sparse reconstruction was performed using Agisoft Photoscan software. First, feature extraction and matching were performed on multi-view images, and the essential matrix was calculated via bundle adjustment. Therefore, both interior and exterior orientation elements of the camera were estimated, and a sparse point cloud was generated with a total of 28,933 points. The sparse point cloud and camera positions are shown in Figure 8.



Figure 7. The experimental areas of the quarry slope in Australia.

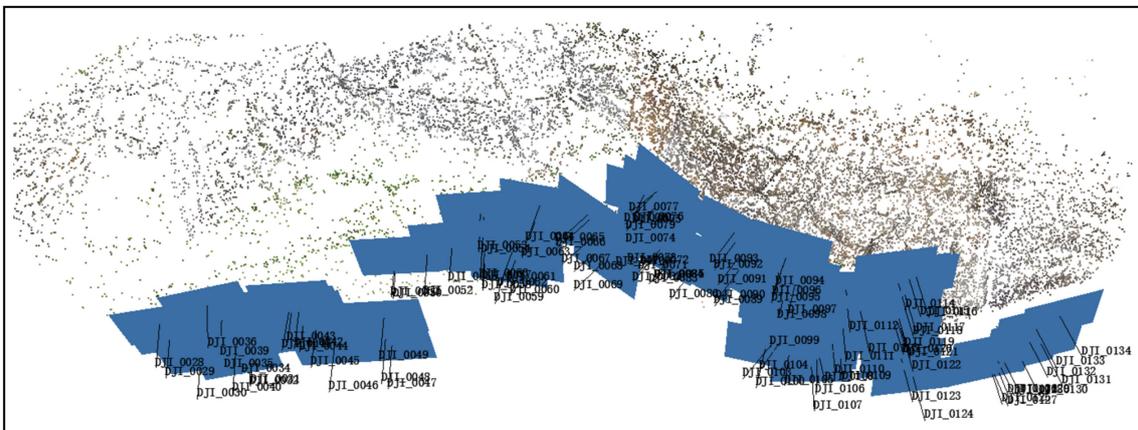


Figure 8. The visualization of the sparse point cloud and cameras.

Based on the sparse reconstruction results, a PMVS dense reconstruction method was performed. A group of depth maps corresponding to the original multi-view images were generated and optimized by propagation; an example is shown in Figure 9. The grey scale values of the depth image was in the range from 0 to 255. The values 0 and 255 represent black and white, respectively.



Figure 9. An example of the original image and the corresponding depth map: (a) the RGB image; (b) the depth image.

After the patch growing and expanding process, a textured dense point cloud of the rock slope was obtained. The number of points in the dense point cloud was 4,754,284; the visualization of the dense point clouds is shown in Figure 10.



Figure 10. The visualization of the dense point cloud.

For any point in the dense point cloud, its local geological semantics involving dip, dip direction, and roughness were, respectively, calculated according to the spatial relationship between the 3D point and its surrounding local neighbor spatial points. The front view of the dense point cloud of the slope was selected for plane fitting, and then a 2D projection plane was determined. According to the projection method proposed in Section 2.2, multiple geological semantic features were projected onto a two-dimensional plane to obtain multi-feature semantic association projection images; as shown in Figure 11. Figure 11a–c show projection images of RGB, dip, and dip direction semantics, respectively.

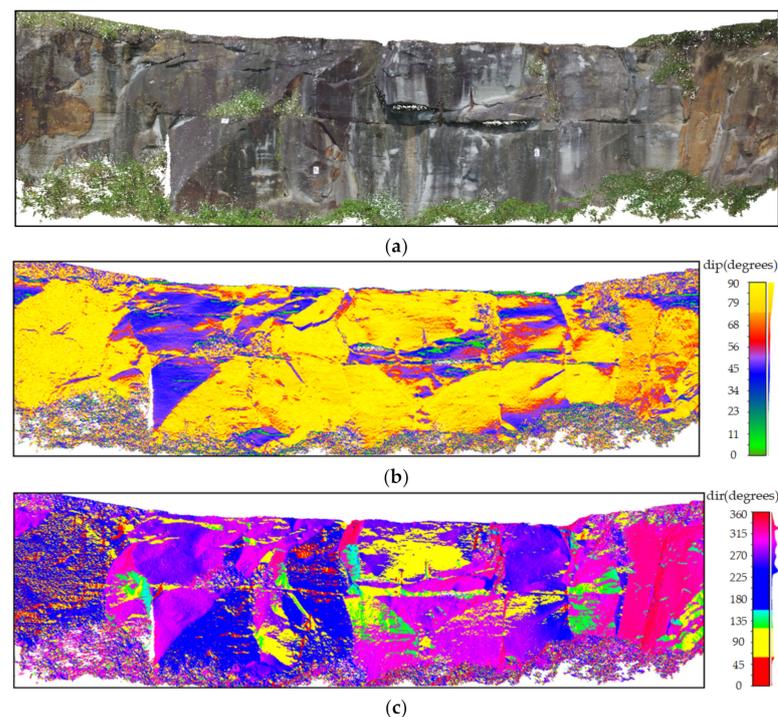


Figure 11. Multi-feature semantic association projection images: (a) the RGB projection image; (b) the dip projection image; (c) the dip direction projection image.

The multi-feature semantic association projection images were taken as inputs, and a semantic block set with similar geological features was generated using the Geo-AINet semantic block segmentation method proposed in Section 2.3. Figure 12a–c show the visualizations of the semantic block segmentation results from the proposed method on the multi-feature semantic projection images, respectively. The color range in Figure 12b,c reflect the dip (0–90 degrees) and dip direction (0–360 degrees), respectively. The color scales correspond to Figure 11b,c.

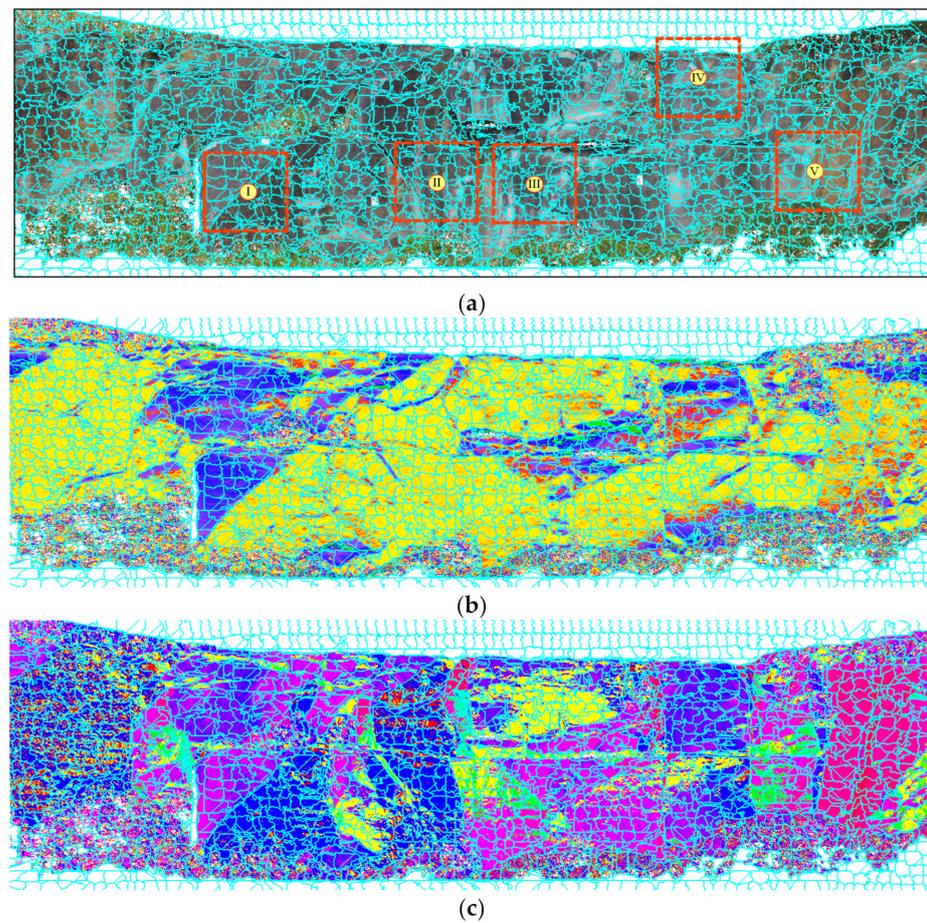


Figure 12. Visualizations of the semantic block segmentation results with the Geo-AINet method proposed in this study: (a) the segmented semantic block overlays on the RGB projection image, and five regions marked by red dashed boxes (numbered I–V) perform relative evident differences of geological features; (b) the segmented semantic block overlays on the dip projection image; (c) the segmented semantic block overlays on the dip direction projection image.

Figure 13 shows the segmentation results generated using the AINet method proposed in [35]. The segmentation results from five regions numbered I–V (red boxes marked in Figure 13) are compared with the corresponding results obtained by the proposed Geo-AINet method (red boxes marked in Figure 12a). Figure 14 shows the details of the comparison.

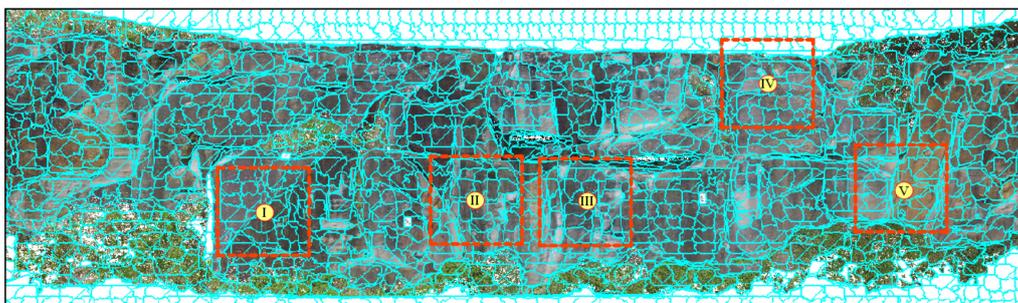


Figure 13. A visualization of the AINet-based segmentation results.

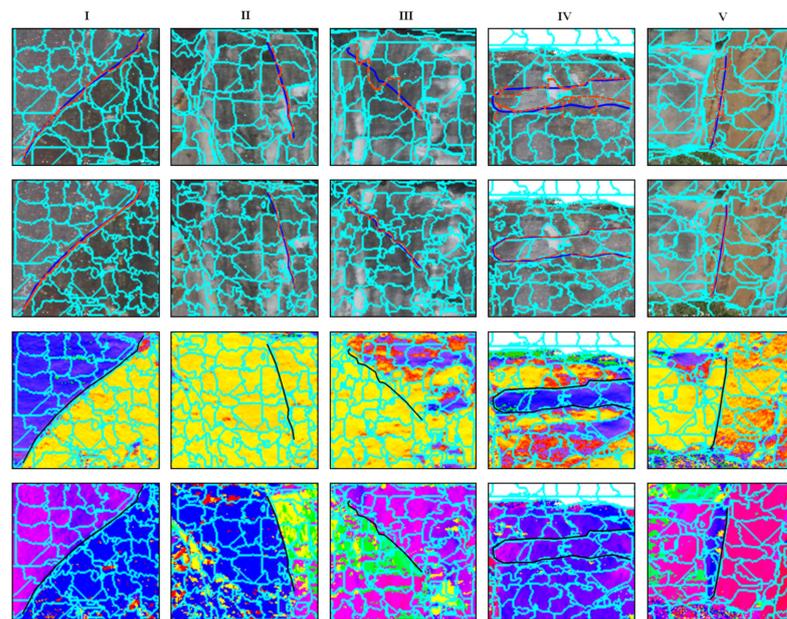


Figure 14. The local detail comparisons of the segmentation results from the two methods: the first row shows the segmentation results from the AINet-based method overlaying the RGB projection image. The last three rows show the segmentation results from the Geo-AINet-based method overlaying on RGB, dip, and dip direction projection images, respectively. The black curves represent the ground truth of the structural plane boundaries. The red solid curves marked in the first row, and the red dotted line marked in the second row, represent the segmented labels obtained by the two methods, respectively.

In Figure 14, five examples of the local detail comparisons of the segmentation results are listed in columns. The two structural planes in region I have noticeable brightness differences caused by illumination, and both methods achieved ideal results. For region II, III, and V, the boundaries of the structural planes in the dip direction projection images are more distinctive. Similarly, for region IV, the boundaries of the structural planes in the dip projection image are easier to distinguish. In comparison with the 2D color and textures features, the 3D geological semantic features contribute to improve the segmentation accuracy of Geo-AINet. In regions II, III, and V, the structural planes have similar RGB and dip characteristics, but significantly different dip directions. It is difficult for the traditional AINet method to distinguish different structural planes accurately. Geo-AINet obtains better segmentation results because it considers the three semantics. In region IV, different structural planes can only be distinguished by the dip semantic instead of similar texture and dip direction. Therefore, the semantic blocks segmented by the Geo-AINet-based method perform better edge adherence. These experimental comparisons demonstrate, fully, that the proposed method integrates multiple geological semantics for structural plane extraction, which can effectively improve accuracy and reliability.

The semantic blocks were first merged using the clustering method described in Section 2.4. The RAG and NNG were generated by semantic blocks with a number of 4327. Each edge connected by two adjacent semantic blocks represents a distance measurement, which was calculated by the values of the dip and the dip direction of the semantic blocks [37]. For a semantic block, its occurrence was obtained according to the normal vector of the plane fitted by the 3D points corresponding to all the pixels in it. Therefore, the length of the edge reflects the similarity of the geological occurrence of the two semantic blocks, i.e., the smaller the distance was, the more similar the two semantic blocks. The clustering criteria is that each semantic block will be merged with the one with the closest distance to itself. Finally, those semantic blocks with the most similarity should be merged. The accuracy of the clustering results has an effect on the integrities of the extracted structural planes. Figure 15a,b show the merging results of semantic blocks under different

perspectives. In Figure 15, the multi-feature semantic blocks have been merged successfully, and the merging results include geological structural planes and invalid surfaces, which are represented by different colors.

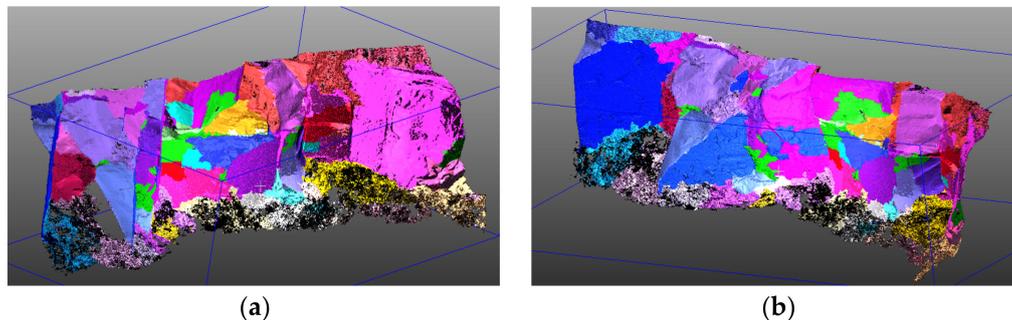


Figure 15. The results from semantic block merging under different perspectives: (a) the left view; (b) the right view.

For the merging results, the color and roughness were used to filter out the vegetation and invalid structural planes regions, respectively. Finally, twenty-seven valid structural planes were successfully recognized, which are shown in Figure 16a–c, providing a comparison of some of the structural planes displayed in the images.

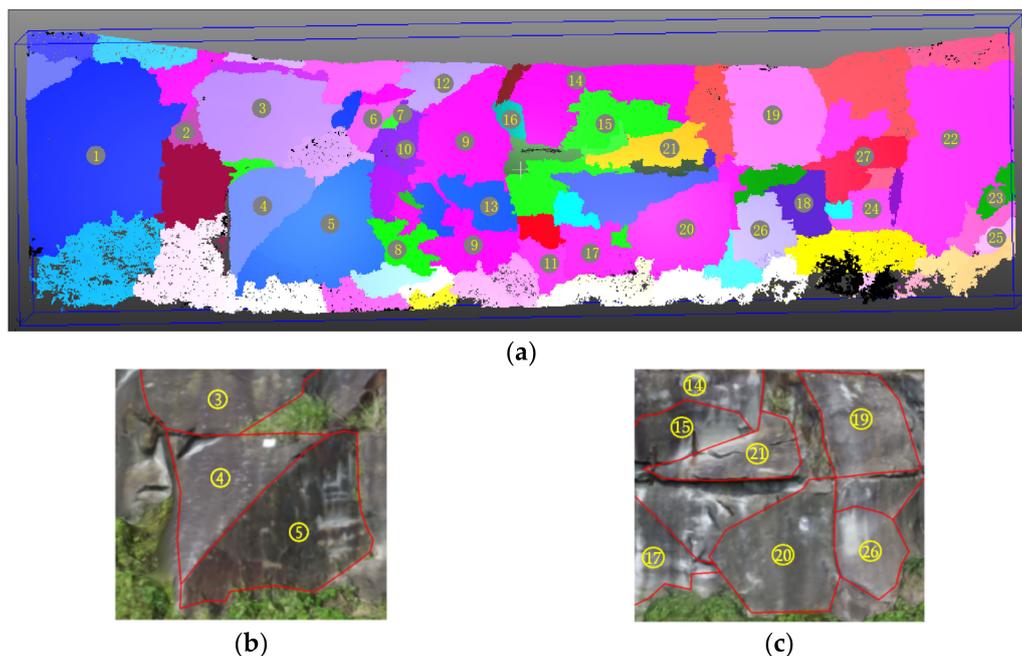


Figure 16. Visualization of the structural plane extraction results: (a) the whole results; (b) and (c): distributions of some structural planes in images.

4. Evaluation and Discussion

To verify the accuracy of the segmentation, four classic accuracy metrics, including Under-segmentation Error (UE), Boundary Recall (BR), Achievable Segmentation Accuracy (ASA), and Mean Distance to Edge (MDE), were used to evaluate the performance of the semantic blocks [38–42]. UE represents the ratio between pixels inside the semantic block but outside the ground truth and all pixels in the ground truth. BR reflects the consistency of the boundaries of superpixels with the ground truth boundaries. A higher BR score indicates that the superpixel has better edge adherence. ASA quantifies the segmentation performance of areas that are semantic block-based instead of pixel-based, and a higher ASA score corresponds to the more accurate segmentation of semantic blocks.

MDE refers to the average distance between pixels in a semantic block and the nearest boundary pixels in ground truth segmentation. The smaller the MDE score, the better the segmentation results.

The semantic block segmentation results from regions I–V, marked in Figures 12a and 13, were used to evaluate the improvement in semantic block segmentation using the proposed Geo-AINet method versus the AINet method. The comparison results are listed in Table 1. The under-segmentation error rate in all regions was reduced by more than 15%. The average boundary recall rate of the five regions was improved by approximately 16%, and the highest BR rate reached 85.66%. The mean distance to edge in the five regions improved vastly, roughly ranging from 19% to 37%. Compared to the other three metrics, although the achievable segmentation accuracy demonstrates a relatively moderate increase, it also reflects the superiority of the proposed Geo-AINet-based method.

Table 1. Comparison of four accuracy metrics for evaluating semantic block segmentation with AINet and Geo-AINet.

Region	Method	UE	BR	ASA	MDE
I	AINet	0.0478	0.6791	0.9523	1.2128
	Geo-AINet	0.0383	0.7930	0.9617	0.9728
II	AINet	0.0940	0.7126	0.9060	1.2361
	Geo-AINet	0.0736	0.8312	0.9263	0.8290
III	AINet	0.0644	0.6475	0.9356	1.4652
	Geo-AINet	0.0481	0.7860	0.9519	0.9126
IV	AINet	0.1291	0.6692	0.8709	1.3503
	Geo-AINet	0.1096	0.7465	0.8904	1.0875
V	AINet	0.0533	0.7433	0.9467	1.1308
	Geo-AINet	0.0363	0.8566	0.9637	0.7211

Some conclusions can be obtained through the above evaluation. Since the AINet superpixel segmentation method only uses RGB for segmentation, it is difficult to accurately distinguish different structural planes with similar colors but obviously different geological occurrences. The integration of 2D and 3D geological semantic features makes the Geo-AINet method perform with apparent advantages in segmentation accuracy. For those complex geological environments, there may be various complex structural planes with obvious 2D differences in color texture and 3D geological occurrences, or a mixture of the two (some examples include the regions I–V in Figure 14). The proposed method comprehensively considers a variety of different features, which is more conducive to the accurate extraction of these structural planes. The Geo-AINet algorithm introduces the geological occurrence semantics and integrates them with the color and roughness semantics. From Table 1 it can be concluded that semantic block segmentation is achieved through multi-feature semantic projection and ensemble learning with Geo-AINet, which can more accurately adhere to the boundaries of complex rock mass structural planes, and performs with greater robustness. Compared with a single feature, the multiple features used for similar semantic block segmentation may improve the accuracy of structural plane extraction.

A quantitative comparison was provided to evaluate the accuracy of structural plane extraction. The dip and dip direction of the ten flat structural planes extracted by the proposed method were, respectively, calculated and compared with those correspondingly measured on the 3D laser point cloud. Firstly, an interpretation of structural planes was performed by the professional geologists according to the image and laser point cloud; then, enough laser points on the structural planes were manually selected for plane fitting; the occurrence of the structural plane was calculated according to the normal vector of the fitted plane and taken as the ground truth. It is to be mentioned that these points should be evenly distributed inside the structural surface, to try and express the spatial geometry of the structural surface. The results are listed in Table 2. It can be seen that both the average dip difference and the average dip direction difference of the structural planes are less than three degrees, and their maximum differences are no more than four

degrees. The experimental results show that the slope rock mass structural planes can be wholly and accurately identified and extracted using the proposed method. The accuracy of the results meets the relevant requirements for geology and design. Moreover, the 2D segmented results retain a mapping relationship with the 3D dense point cloud, by which the 2D results can be transmitted to the 3D structure. Deep learning can better explore the relationship between features, but it is difficult to directly perform multi-feature deep learning in 3D space due to the tremendous amount and variable dimensions of the data. Therefore, this study conducts deep learning by projecting various features into 2D space, which can effectively improve the efficiency and feasibility of the algorithm.

Table 2. The geological occurrence comparison calculated by the proposed method and measured on the 3D laser point cloud.

Label	Measured on the 3D Laser Point Cloud		Calculated by the Proposed Method		$\Delta\theta$ ($ \theta_1 - \theta_2 $)	$\Delta\alpha$ ($ \alpha_1 - \alpha_2 $)
	Dip (θ_1)	Dir (α_1)	Dip (θ_2)	Dir (α_2)		
L1	86.9	243.2	87.8	239.9	0.9	3.3
L2	46.7	272.6	48.6	275.6	1.9	3
L3	40.5	288.2	42.7	287.1	2.2	1.1
L4	81.6	235.5	82.3	237.5	0.7	2
L5	85.2	238.9	84.4	240.2	0.8	1.3
L6	87.7	273.1	86.5	270.0	1.2	3.1
L7	79.8	306.1	79.0	308.6	0.8	2.5
L8	61.0	264.4	61.5	267.7	0.5	3.3
L9	80.4	289.7	81.0	293.6	0.6	3.9
L10	72.6	328.0	73.2	330.2	0.6	2.2
$\overline{\Delta(\theta)}, \overline{\Delta(\alpha)}$					1.0	2.6

5. Conclusions

This study proposes a slope structural plane extraction method based on Geo-AINet ensemble learning, which provides primary data for the stability analysis and evaluation of rock slopes. This method uses the UAV images of the slope as the data source, and establishes multi-feature semantic association projection images through the geometric projection relationship between the 3D point cloud and the 2D plane; on this basis, Geo-AINet ensemble learning is applied for semantic block segmentation; then, the geological occurrence and roughness are adopted to perform semantic block clustering; finally the structural planes are successfully extracted. The following conclusions can be drawn from the experimental results:

1. It is difficult to guarantee sufficient precision for the sole extraction of structural planes when depending on a single semantic. Considering, fully, the joint influence of 2D and 3D features on the extraction of rock mass structural planes, RGB, geological occurrence, and roughness are integrated as the basic semantics in this study to generate multi-dimensional semantic association projection images. The 2D projection plane is defined according to the spatial distribution of the slope. The proposed method more effectively realizes the pixel-level feature association of multi-dimensional geological semantics and provides input data for multi-feature semantic association for Geo-AINet ensemble learning.
2. The Geo-AINet model is proposed to obtain a soft association map involving multi-dimensional semantics; then, the original image is divided into semantic blocks with multiple similar features. The quantitative evaluation results in Table 1 show that the extracted semantic blocks are completely and accurately attached to the true boundaries of the structural planes. The comparison of four metrics (UE, BR, ASA, and MDE) for the five regions concludes that the proposed Geo-AINet method outperforms the traditional AINet method concerning the accuracy of block semantics segmentation.

3. In this study, an RAG is established based on the topological adjacency relationship of semantic blocks. The RAG is simplified to an NNG according to the dissimilarities of the multi-dimensional features between adjacent semantic blocks. The semantic blocks are merged, and the structural planes are successfully extracted further following a filtering process aimed at identifying invalid structural surfaces.

In summary, the accuracy and completeness of the structural planes extracted from the quarry slope in the experimental area by the proposed method are satisfactory. In future work, more geological semantics will be considered to further improve the accuracy of slope structural plane extraction.

Author Contributions: Conceptualization, R.Z.; X.Y. and L.L.; methodology, R.Z., S.S. and J.L.; validation, S.S. and M.J.; formal analysis, S.S. and C.Z.; writing—original draft preparation, S.S. and L.L.; writing—review and editing, R.Z., X.Y. and L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (41901401, 42101070), the China Postdoctoral Science Foundation (Grant No. 2021M691653), Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People’s Republic of China (Grant No. KLSMNR-G202213), and the Natural Science Foundation of Jiangsu Province (Grant No. BK20190743).

Data Availability Statement: Not applicable.

Acknowledgments: We thank the University of Newcastle for providing the experimental datasets.

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

References

1. Ross-Brown, D.M.; Atkinson, K.B. Terrestrial photogrammetry in open-pits: 1-description and use of the phototheodolite in mine surveying. *Trans. Inst. Min. Metall.* **1972**, *81*, A205–A213.
2. Li, H.; Zhang, R.C.; Yang, B.; Wu, M.F. Principle and geometric precision of photographic geological logging of tunnels. *J. Appl. Remote Sens.* **2014**, *8*, 083617. [[CrossRef](#)]
3. Deb, D.; Hariharan, S.; Rao, U.M.; Ryu, C.H. Automatic detection and analysis of discontinuity geometry of rock mass from digital images. *Comput. Geosci.* **2008**, *34*, 115–126. [[CrossRef](#)]
4. Zhou, C.L.; Zhu, H.H.; Zhao, W. Non-contact Measurement of rock mass discontinuity occurrence with binocular system. *Chin. J. Rock Mech. Eng.* **2010**, *29*, 111–117. [[CrossRef](#)]
5. Bretar, F.; Arab-Sedze, M.; Champion, J.; Pierrot-Deseilligny, M.; Heggy, E.; Jacquemoud, S. An advanced photogrammetric method to measure surface roughness: Application to volcanic terrains in the piton de la Fournaise, Reunion Island. *Remote Sens. Environ.* **2013**, *135*, 1–11. [[CrossRef](#)]
6. Vasuki, Y.; Holden, E.J.; Kovesi, P.; Micklethwaite, S. Semi-automatic mapping of geological Structures using UAV-based photogrammetric data: An image analysis approach. *Comput. Geosci.* **2014**, *69*, 22–32. [[CrossRef](#)]
7. Gigli, G.; Casagli, N. Semi-automatic extraction of rock mass structural data from high resolution LIDAR point clouds. *Int. J. Rock Mech. Min. Sci.* **2011**, *48*, 187–198. [[CrossRef](#)]
8. Battulwar, R.; Zare-Naghadehi, M.; Emami, E.; Sattarvand, J. A state-of-the-art review of automated extraction of rock mass discontinuity characteristics using three-dimensional surface models. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 920–936. [[CrossRef](#)]
9. Wang, X.; Zou, L.; Shen, X.; Ren, Y.; Qin, Y. A region-growing approach for automatic outcrop fracture extraction from a three-dimensional point cloud. *Comput. Geosci.* **2017**, *99*, 100–106. [[CrossRef](#)]
10. Li, X.; Chen, Z.; Chen, J.; Zhu, H. Automatic characterization of rock mass discontinuities using 3D point clouds. *Eng. Geol.* **2019**, *259*, 105131. [[CrossRef](#)]
11. Wu, X.; Wang, F.; Wang, M.; Zhang, X.; Wang, Q.; Zhang, S. A New Method for Automatic Extraction and Analysis of Discontinuities Based on TIN on Rock Mass Surfaces. *Remote Sens.* **2021**, *13*, 2894. [[CrossRef](#)]
12. Lee, Y.K.; Kim, J.; Choi, C.S.; Song, J.J. Semi-automatic calculation of joint trace length from digital images based on deep learning and data structuring techniques. *Int. J. Rock Mech. Min. Sci.* **2022**, *149*, 104981. [[CrossRef](#)]
13. Xu, W.; Zhang, Y.; Wang, X.; Ma, F.; Zhao, J.; Zhang, Y. Extraction and statistics of discontinuity orientation and trace length from typical fractured rock masses: A case study of the Xinchang underground research laboratory site, China. *Eng. Geol.* **2020**, *269*, 105553. [[CrossRef](#)]

14. Kong, D.H.; Wu, F.Q.; Saroglou, C. Automatic identification and characterization of discontinuities in rock masses from 3D point clouds. *Eng. Geol.* **2020**, *265*, 105442. [[CrossRef](#)]
15. Li, D.T. Vanishing line method and vanishing point deduction method for photographic survey of strike-dip of structural surfaces. *Adv. Sci. Technol. Water Resour.* **2005**, *1*, 21–24.
16. Leu, S.S.; Chang, S.L. Digital image processing based approach for tunnel excavation faces. *Autom. Constr.* **2005**, *14*, 750–765. [[CrossRef](#)]
17. Wang, F.Y.; Chen, J.P.; Fu, X.H.; Shi, B.F. Study on geometrical information of obtaining rock mass discontinuities based on virtuoso. *Chin. J. Rock Mech. Eng.* **2008**, *27*, 169–175.
18. Liu, Z.X. Research and application of rapid acquiring discontinuities information in rock mass based on digital close range photogrammetry. Ph.D. Thesis, Jilin University, Jilin, China, 2009.
19. Bi, H.Y.; Zheng, W.J.; Yu, J.X.; Ren, Z.K. Application of sfm photogrammetry method to the quantitative study of active tectonics. *Seismol. Geol.* **2017**, *39*, 656–674.
20. Xiao, G.; Wang, Q.; Liu, G.D.; Pan, Y. Method and application of extraction fracture information from high and steep dangerous rock based on UAV image. *Site Investig. Sci. Technol.* **2019**, *1*, 4–9.
21. Kocak, D.M.; Caimi, F.M.; Das, P.S.; Karson, J.A. A 3-d laser line scanner for outcrop scale studies of seafloor features. In Proceedings of the Oceans'99. MTS/IEEE. Riding Crest Into 21st Century, Seattle, WA, USA, 13–16 September 1999; Volume 3, pp. 1105–1114.
22. Feng, Q.; Fardin, N.; Jing, L.; Stephansson, O. A new method for in-situ non-contact roughness measurement of large rock fracture surfaces. *Rock Mech. Rock Eng.* **2003**, *36*, 3–25. [[CrossRef](#)]
23. Slob, S.; van Knapen, B.; Hack, R.; Turner, K.; Kemeny, J. Method for automated discontinuity analysis of rock slopes with three-dimensional laser scanning. *Transp. Res. Rec. J. Transp. Res. Board.* **2005**, *1913*, 187–194. [[CrossRef](#)]
24. Riquelme, A.J.; Abellán, A.; Tomás, R.; Jaboyedoff, M. A new approach for semi-automatic rock mass joints recognition from 3d point clouds. *Comput. Geosci.* **2014**, *68*, 38–52. [[CrossRef](#)]
25. Dk, A.; Fwa, B.; Cs, C. Automatic identification and characterization of discontinuities in rock masses from 3D point clouds. *Eng. Geol.* **2021**, *265*, 105442.
26. Slob, S.; Hack, H.R.G.K.; Feng, Q.; Roshoff, K.; Turner, A.K. Fracture mapping using 3D laser scanning techniques. In Proceedings of the 11th Congress of the International Society for Rock Mechanics, Lisbon, Portugal, 9–13 July 2007; Volume 1, pp. 299–302.
27. Deliormanli, A.H.; Maerz, N.H.; Otoo, J. Using terrestrial 3D laser scanning and optical methods to determine orientations of discontinuities at a granite quarry. *Int. J. Rock Mech. Min. Sci.* **2014**, *66*, 41–48. [[CrossRef](#)]
28. Riquelme, A.; Cano, M.; Tomás, R.; Abellán, A. Identification of rock slope discontinuity sets from laser scanner and photogrammetric point clouds: A comparative analysis. *Procedia Eng.* **2017**, *191*, 838–845. [[CrossRef](#)]
29. Nagendran, S.K.; Ismail, M.A.M.; Tung, W.Y. Photogrammetry approach on geological plane extraction using CloudCompare FACET plugin and scanline survey. *Bull. Geol. Soc. Malays.* **2019**, *68*, 151–158. [[CrossRef](#)]
30. Herrero, M.J.; Pérez-Fortes, A.P.; Escavy, J.I.; Insua-Arévalo, J.M.; De la Horra, R.; López-Acevedo, F.; Trigos, L. 3D model generated from UAV photogrammetry and semi-automated rock mass characterization. *Comput. Geosci.* **2022**, *163*, 105121. [[CrossRef](#)]
31. Ren, X.; Malik, J. Learning a classification model for segmentation. In Proceedings of the International Conference on Computer Vision and Computer Society, 13–16 October 2003; Volume 2, p. 10.
32. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2282. [[CrossRef](#)] [[PubMed](#)]
33. Stutz, D.; Hermans, A.; Leibe, B. Superpixel Segmentation Using Depth Information. Bachelor's Thesis, RWTH Aachen University, Aachen, Germany, 2014. Volume 4.
34. Yang, F.; Sun, Q.; Jin, H.; Zhou, Z. Superpixel segmentation with fully convolutional networks. In Proceedings of the International Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 13964–13973.
35. Wang, Y.; Wei, Y.; Qian, X.; Zhu, L.; Yang, Y. AINet: Association implantation for superpixel segmentation. In Proceedings of the International Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 7078–7087.
36. Stutz, D.; Hermans, A.; Leibe, B. Superpixels: An Evaluation of the State-of-the-Art. *Comput. Vis. Image Underst.* **2018**, *166*, 1–27. [[CrossRef](#)]
37. Haris, K.; Efstratiadis, S.N.; Maglaveras, N.; Katsaggelos, A.K. Hybrid image segmentation using watersheds and fast region merging. *IEEE Trans. Image Process.* **1998**, *7*, 1684–1699. [[CrossRef](#)]
38. Levinstein, A.; Stere, A.; Kutulakos, K.N.; Fleet, D.J.; Dickinson, S.J.; Siddiqi, K. TurboPixels: Fast superpixels using geometric flows. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 2290–2297. [[CrossRef](#)] [[PubMed](#)]
39. Neubert, P.; Protzel, P. Superpixel benchmark and comparison. In *Proc. Forum Bildverarbeitung*; KIT Scientific Publishing: Karlsruhe, Germany, 2012; Volume 6, pp. 1–12.

40. Martin, D.; Fowlkes, C.; Malik, J. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 530–549. [[CrossRef](#)] [[PubMed](#)]
41. Lui, M.Y.; Tuzel, O.; Ramalingam, S.; Chellappa, R. Entropy rate superpixel segmentation. In Proceedings of the International Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 2097–2104.
42. Benesova, W.; Kottman, M. Fast superpixel segmentation using morphological processing. In Proceedings of the International Conference of Machine Vision and Machine Learning, Prague, Czech Republic, 14–15 August 2014; pp. 1–9.

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.