



Article

Improving Geological Remote Sensing Interpretation via Optimal Transport-Based Point–Surface Data Fusion

Jiahao Wu , Wei Han , Jia Chen * and Sheng Wang

School of Computer Science, China University of Geosciences, Wuhan 430078, China; jiahao.wu@cug.edu.cn (J.W.); weihan@cug.edu.cn (W.H.); ws@cug.edu.cn (S.W.)

* Correspondence: chen_jia@cug.edu.cn

Abstract: High-quality geological remote sensing interpretation (GRSI) products play a vital role in a wide range of fields, including the military, meteorology, agriculture, the environment, mapping, etc. Due to the importance of GRSI products, this research aimed to improve their accuracy. Although deep-learning (DL)-based GRSI has reduced dependence on manual interpretation, the limited accuracy of multiple geological element interpretation still poses a challenge. This issue can be attributed to small inter-class differences, the uneven distribution of geological elements, sensor limitations, and the complexity of the environment. Therefore, this paper proposes a point–surface data optimal fusion method (PSDOF) to improve the accuracy of GRSI products based on optimal transport (OT) theory. PSDOF combines geological survey data (which has spatial location and geological element information called point data) with a geological remote sensing DL interpretation product (which has limited accuracy and is called surface data) to improve the quality of the resulting output. The method performs several steps to enhance accuracy. First, it calculates the gray-scale correlation feature information for the pixels adjacent to the geological survey points. Next, it determines the distribution of the feature information for geological elements in the vicinity of the point data. Finally, it incorporates complementary information from the survey points into the geological elements' interpretation boundary, as well as calculates the optimal energy loss for point–surface fusion, thus resulting in an optimal boundary. The experiments conducted in this study demonstrated the superiority of the proposed model in addressing the problem of the limited accuracy of GRSI products.

Keywords: geological interpretation; optimal transport; multi-source fusion; remote sensing; data alignment



Citation: Wu, J.; Han, W.; Chen, J.; Wang, S. Improving Geological Remote Sensing Interpretation via Optimal Transport-Based Point–Surface Data Fusion. *Remote Sens.* **2024**, *16*, 53. <https://doi.org/10.3390/rs16010053>

Academic Editors: Biswajeet Pradhan and Sicong Liu

Received: 23 September 2023

Revised: 23 November 2023

Accepted: 18 December 2023

Published: 22 December 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

Over the past decade, exponential growth in remote sensing data generated by ground-based sensors has been observed. These data have been widely utilized in geological remote sensing interpretation (GRSI) [1]. In the context of regional lithological mapping at large scales, GRSI refers to the process of identifying various geological features based on their characteristics as remote sensing images [2–4]. It is an integral component of regional lithological mapping, and it holds significant importance. Compared to traditional manual geological surveys, GRSI offers the advantage of conducting large-scale geological investigations in a cost-effective manner. As such, it has the potential to serve various related applications, such as geological mapping, urban and regional planning, environmental protection, and disaster assessment. Hence, the procurement of high-quality GRSI products is of utmost importance for conducting effective research in these fields.

To improve the accuracy of GRSI products, researchers have proposed various geological interpretation models. Early national and international geological remote sensing practitioners developed interpretation models based on the spectral characteristics of rocks, minerals, and geological tectonic features [5,6]. In Ref. [7], a novel band ratio image was

generated by utilizing the desired band from the ASTER data and by effectively combining it with geological field observations to accurately map lithological units in the Wadi Kid region. Another approach to the use of spectral information is the spectral angle mapper (SAM) [8]. However, these geological interpretation models are heavily reliant on manual labor and have a low degree of automation. Moreover, the spectral response of the same type of rock varies considerably, thus resulting in spectral uncertainty. This phenomenon is referred to as spectral variability, which is widely observed in many scenarios where the spectral characteristics of pure material substances vary in the acquired hyperspectral remote sensing images. Spectral variability can be caused by illumination and environmental factors, as well as by the inherent properties of the materials [9,10]. For example, variations in the orientation, strike, and dip of a single rock can result in differences in how the rock absorbs and reflects solar radiation, thus leading to distinct pieces of spectral information, which are captured by remote sensing sensors. These issues constrain the development of GRSI products.

In the last few years, there has been a shift toward the use of automated machine learning (ML) methods to extract geological element features from remotely sensed imagery. These methods address the limitations of manual interpretation models. The first common type of ML method is dimensionality reduction, which includes independent component analysis (ICA) [11], principal component analysis (PCA) [12], and the minimum noise fraction (MNF) [13]. Another type of method is classification, such as the support vector machines (SVMs) method [14]. For example, Ref. [15] utilized ASTER spectral data and SVMs to classify the lithology on ASTER images. The results showed that the SVM algorithm performed well. The classification maps obtained were consistent with both the field survey and officially published geological maps. There are also a number of related platforms, like the GoldenEye project. However, as the temporal, spatial, and spectral resolution of remote sensing data continues to rise and as the information of remote sensing data becomes more abundant [16], ML-based feature-extraction methods face challenges in terms of feature representation.

Since 2012, a multitude of advanced deep learning (DL) methods have emerged in the field of image processing. DL techniques can automatically model large-scale and complex datasets [17], and they have shown exceptional performance in various computer vision tasks. Among these, convolutional neural networks (CNNs) have gained widespread attention in remote sensing due to their effectiveness in image processing [18]. For instance, Ref. [19] applied CNNs to map geological target features and to classify Landsat images. Similarly, Ref. [20] employed both CNNs and traditional ML methods, such as SVMs and multilayer perceptrons, to map lithological units in mineral-rich areas of southeastern Iran. Ref. [21] proposed a high-resolution mapping approach using unmanned aerial vehicles (UAVs) to obtain data and DL algorithms to extract target features. Ref. [22] proposed a multistage self-guided separation network (MGSNet) that enhances the discriminability of targets and backgrounds in remote sensing scenes through the utilization of a target-background separation strategy, contrastive regularization, and self-guided networks. Ref. [23] proposed a representation-enhanced status replay network (RSRNet). This approach addresses representation bias, classifier bias, and insufficient information interaction through the combined augmentation of feature representation, a status replay strategy, and cross-modal interactive fusion. Ref. [24] proposed a structural optimization transmission framework, namely a structural optimization transmission network (SOT-Net). This method effectively utilizes the reflectance-specific information from HSIs, as well as the detailed edge representations from multiple sources, to enhance feature extraction and classification. Ref. [25] proposed a spatial-logical aggregation network (SLA-NET) which leverages morphological transformations and trainable structuring elements to extract fine-grained morphological structures from hyperspectral images. The method aims to enhance the classification of tree species and has shown superior performance compared to other state-of-the-art classifiers. Ref. [26] proposed an innovative method called Large kernel Sparse ConvNet weighted by Multi-frequency Attention (LSCNet) to overcome the

limitations of traditional CNNs in remote sensing scene understanding. Overall, DL-based methods have shown remarkable potential in extracting semantic features of geological elements, such as lithology [27], minerals [28], glaciers [29], soils [30], and geological formations [31]. However, the accuracy of these interpretations is often inadequate due to various factors. These factors can be summarized as follows:

1. **Small inter-class differences:** Certain geological elements have blurred boundaries and similar imaging features due to various physical and chemical effects, such as weathering, erosion, and biological activity. Distinguishing between these elements in satellite imagery is difficult, as the color, shape, and structural features can be challenging to classify accurately.
2. **Unevenly distributed geological elements:** The interpretation model is faced with higher requirements due to the variety of geological elements and the different sizes of the areas covered by these elements. It is noteworthy that those geological elements with a wider area of coverage, such as soil and water bodies, are relatively easier to identify.
3. **Sensor limitations:** Sensor aberrations, changes in the operating conditions, and the movement of the Earth can cause image distortions. Natural phenomena, such as clouds and fog, can also obscure the sensor's view of the Earth's surface, thus making it difficult to obtain clear remote sensing images.

These factors pose significant challenges in improving the accuracy of GRSI data. The limitations in the acquisition and processing of remote sensing data result in insufficient information content in GRSI data. Therefore, relying solely on remote sensing data for interpretation often fails to meet the high accuracy requirements. However, experts obtain point data through geological surveys, which have higher precision. Therefore, it is better to combine the complementary information from both point and surface data for geological analysis. The fusion of point and surface data enables the acquisition of more-comprehensive and -accurate interpretation results. However, due to the heterogeneity of data from different sources, each with its own data domain, the fusion of point–surface data often faces the challenge of data domain misalignment.

To address the issues of the limited accuracy of GRSI products caused by small inter-class differences, the uneven distribution of geological elements, sensor limitations, and the misalignment of data from multiple sources, this paper proposes a point–surface data optimal fusion method (PSDOF), based on optimal transport (OT) theory, to improve the accuracy of GRSI products. PSDOF combines geological survey point data (which have spatial location and geological element information and are called point data) with a geological remote sensing DL interpretation product (which has limited accuracy and is called surface data) to improve the quality of the resulting output, as well as to introduce OT to facilitate heterogeneous data alignment. The method presented includes two primary stages: (1) the extraction of high-precision location information from the point data and (2) point–surface fusion, which is where the OT model incorporates point data information into the accuracy-constrained GRSI products to achieve an information gain. To demonstrate the effectiveness of the method, experiments were conducted with PSDOF using GRSI products over the Pamir Plateau in the southern part of the Tianshan Mountains in China. Our work made the following contributions:

1. A new fusion method, named PSDOF, in the field of GRSI that is designed to aid with the fusion of heterogeneous geological survey point data and GRSI data.
2. By fusing the high-precision location information extracted from geological survey point data, the PSDOF achieves information gains and effectively enhances the accuracy of GRSI products.
3. PSDOF employs the concept of OT to address the challenge of data misalignment between geological survey point data and GRSI data, thus achieving a significant improvement in accuracy.

This paper is divided into six sections. Section 2 examines the related works. Section 3 defines the problem and objectives, and presents the data used in this study, while Section 4 explains the PSDOF method, including its principles and features. Next, Section 5 covers the experimental setup and results. Finally, Section 6 summarizes the findings and future research implications.

2. Related Work

This section has two parts and focuses on domain adaptation and optimal transport.

2.1. Domain Adaptation

Domain adaptation can be categorized into two groups: semi-supervised and unsupervised techniques. The former utilizes labels in the target domain (TD), whereas the latter does not.

In the field of semi-supervised domain adaptation, many methods have been proposed by researchers. For instance, Ref. [32] used a linear conversion to map characteristics from the TD onto the source domain (SD) to acquire representations. In terms of non-linear transport learning across domains, Ref. [33] proposed a method to transfer object models from one dataset to another. Ref. [34] proposed a technique that adjusts the object models obtained in a specific visual domain to new imaging situations by training a transformation that lessens the impact of the feature distribution changes induced by the domain. Ref. [35] presented a kernel method, called the kernel method for manifold alignment (KEMA), for aligning manifolds that can match any number of data sources without requiring corresponding pairs. The approach only needs a few labeled examples in all domains.

In fact, due to the difficulty in acquiring TD labels, unsupervised domain adaptation methods have received more attention. For example, Ref. [36] suggested a dimensionality reduction approach to achieve domain adaptation by decreasing the inter-domain distance. Ref. [37] proposed a neural-network-based model that learns transferable features by jointly learning from the unlabeled data in the TD and the labeled data in the SD. Ref. [38] mapped multiple pieces of data to a reduced dimensionality while preserving the neighborhood relationships in each dataset, and this approach represented a new stream-alignment technique. Ref. [39] described a method through which to learn the projection of the data in a low-dimensional space. The empirical distribution distance between the source and target data was kept to a minimum during the projection, and it was a domain-invariant type of projection.

2.2. Optimal Transport

OT has demonstrated effectiveness as a domain-adaptation technique in both unsupervised and semi-supervised modes. For instance, a regularized unsupervised OT model was proposed by [40] to pair the probability distributions of the two domains while keeping samples of the same category in the SD closest in distance during transmission. To detect flood hazards, Ref. [41] employed an OT model to fuse different types of remote sensing and social media data. Building on this work, Ref. [42] proposed a geo-optimal transport model (GOT) that addressed the problem of the geographical bias of social media tweets during OT. Ref. [43] proposed an optimal transmission method that incorporates regularization in multimodal 2-dimensional and 3-dimensional facial expression recognition (FER) (which was unsupervised). In Ref. [44], OT was utilized to achieve a semi-supervised adaptation in heterogeneous domains. It ensured that samples belonging to the same class were constrained to exhibit similar distributions during the transfer process from the SD to the TD.

3. Problem Definition

Consult Table 1 for the primary symbols utilized in this research. PSDOF is described as follows: input a set of Landsat 8 images of size $\{r \times c \times W\}$ (r and c denote the quantity

of pixels along the rows and columns of the image, respectively, and W equals the quantity of channels in the image) with the surface data and the point data Q .

Table 1. Explanation of the primary symbols used in this paper.

Symbols	Definition
r	The number of pixels on the rows of the image
c	The number of pixels on the columns of the image
X_n	Characteristic point set for n pixels
Y_m	Characteristic surface for m pixels
D_s	Source domain for the X_n set
D_t	Target domain for the Y_m set
C	Cost matrix for the OT plan
d	The step distance for a gray-level co-occurrence matrix
θ	The angle for a gray-level co-occurrence matrix

$$Cor_Q \in \{Cor_1, Cor_2, \dots, Cor_{r \times c}\} \quad (1)$$

where Cor_Q represents the gray-scale correlation feature value of the Landsat 8 image pixel corresponding to the point data. We aimed to use $Cor_{r \times c}$ to reflect the information about the geological elements of each pixel of the image. After calculation and expert screening ($A = 0.05$), it was determined that pixels with feature values in the range $Cor_Q \pm A$ have similar geological elements and that these pixels are called similar feature pixels.

$$Source D_s : \{X_1, \dots, X_n\} \quad (2)$$

$$Target D_t : \{Y_1, \dots, Y_m\} \quad (3)$$

where the SD D_s is the finite set of all similar feature pixels X_n of the point data in the remote sensing image. The TD D_t is the set of pixels Y_m within the geological element boundaries (i.e., the boundaries obtained by fitting similar pixel curves) in the vicinity of the observation point of the decoded product. The inter-domain transport scheme Tp can be written as follows:

$$Tp : P(X_n, Y_m) \in \{P_1, \dots, P_\infty\} \quad (4)$$

In the OT model, the SD samples and TD samples are the input, and the OT cost of the sample domain is calculated to output the fused optimal bound. The OT plan can be written as follows:

$$OT : arg \min W(X_n, Y_m) \quad (5)$$

Finally, the GRSI products are updated according to the optimal bounds needed to obtain a higher-quality interpretation product. In summary, our proposed method can effectively fuse two sample domains in the information space D at a minimal cost, thus making it suitable for multimodal remote sensing data fusion. Our approach extracts complementary information from the point data through the calculation of gray-scale correlation features, as well as leverages lithological and positional information from the geological survey point data to fuse them into boundary-ambiguous GRSI products, thus resulting in higher accuracy.

Materials

The study area selected for this experiment is situated on the Pamir Plateau in the southern region of the Tianshan Mountains in China. The area has a highland mountain climate characterized by cold weather and numerous rock glaciers:

1. Remote sensing data: These are a critical component of this study. Landsat 8 is a multispectral satellite that covers a total of 11 bands, thus making it an ideal data source for this study. To better distinguish the spectral characteristics of the geological elements, this work utilized the three RGB bands of Landsat 8 to synthesize color images, as shown in Figure 1.

2. GRSI data: These were derived from the DL interpretation products obtained from Landsat 8 satellite imagery in the study area when using a multi-species semantic segmentation model. Separately, they were FCN [45], DeepLabV3 [46], DANet [47], OCNNet [48], PSPNet [49], and AdvSemi-OCGNet [50]. The geological elements extracted are characterized by nine categories: glacier, granitic rock, lakes, carbonate rocks, slates, sandstone, volcanic debris, schist, and soil bodies. Table 2 shows the accuracy of the AdvSemi-OCGNet interpretation products, and through that information, we can conclude the following: the classification accuracy of glaciers (Acc: 89.6%), lakes (Acc: 95%), and soil bodies (Acc: 89.9%) ranked in the top three of the nine geological elements for the best classification accuracy. Meanwhile, the three geological elements with the worst classification accuracy were as follows: granitic rocks, sandstone, and volcanic debris. To assess the efficacy of our model, three of these localized areas (i.e., where geological survey data existed) were selected for experimentation.
3. Geological survey data: In this study, these mainly consisted of point data. These points were selected by geologists within the study area through manual effort, thus resulting in a sparse dataset with precise spatial location and lithological information. After careful screening and confirmation, only point data within the interpretation boundary were used. The distribution of the selected point data is illustrated in Figure 2. To improve the interpretation accuracy, these point data were fused with the GRSI products to provide complementary information.
4. Ground truth data: These are essential in assessing the precision of the GRSI model. In this study, the GRSI products were obtained from Landsat 8 satellite imagery, and they were manually labeled by experts with geographic alignment and cropping. The ground truth data represent the distribution of the geological elements in the study area, and they were used as a reference to evaluate the performance of the GRSI model. The annotated map in Figure 3 shows the nine categories of the geological elements included in the ground truth data, which are spatially consistent with the Landsat 8 satellite imagery.

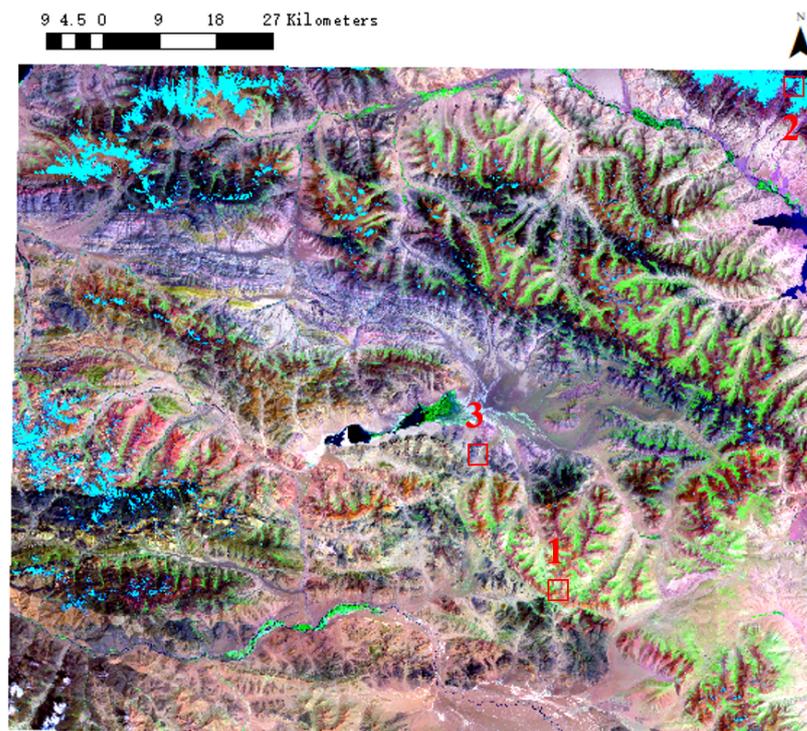


Figure 1. Remote sensing data: Landsat 8 image. The locations of the three study areas are marked in red boxes on the map.

Table 2. AdvSemi-OCGNet product accuracy.

Class	No.	IOU (%)	Acc (%)
Glacier	0	78.461	89.628
Granitic rock	1	53.162	67.310
Lakes	2	90.610	95.076
Carbonate rocks	3	59.681	71.612
Slate	4	70.996	85.678
Sandstone	5	43.650	57.188
Volcanic debris	6	55.917	69.794
Schist	7	58.392	71.108
Soil bodies	8	78.319	89.905

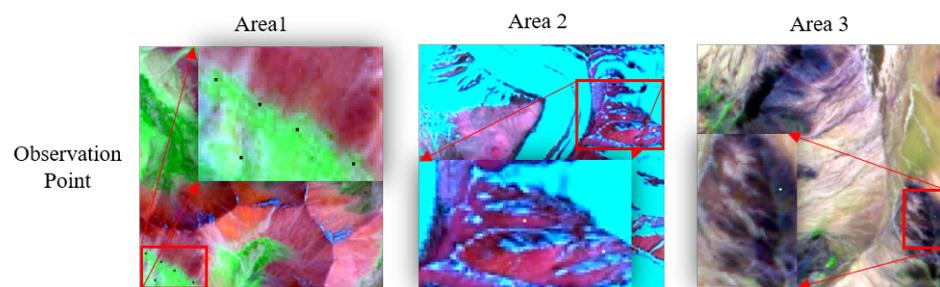


Figure 2. The spatial distribution of the observation points in the study area, where the red boxes indicate the area covered by the observation points. These points are associated with three distinct local areas that correspond to certain geological elements, such as sandstone, volcanic debris, and granitic rock.

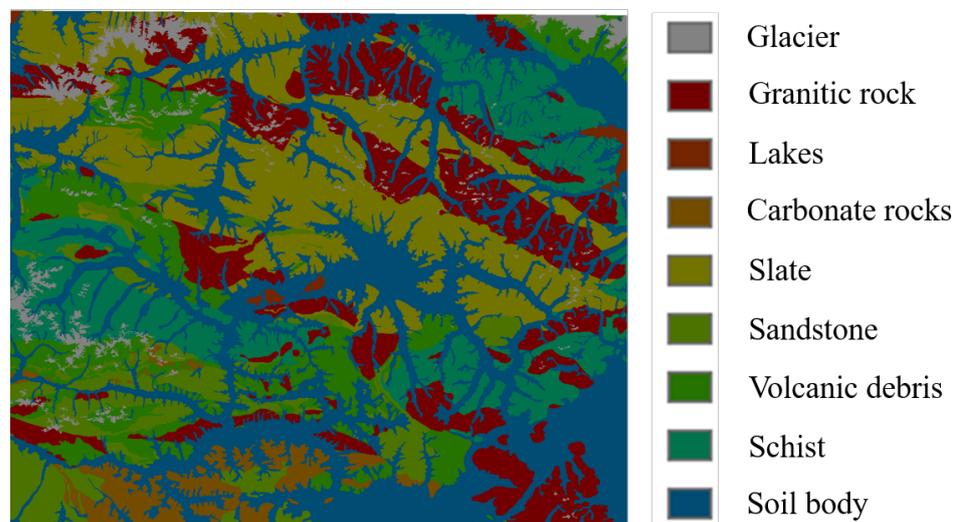


Figure 3. A map of the study area with labeled features.

4. Methodology

This section introduces the PSDOF method for the fusion of the point data and surface data. Figure 4 shows a diagrammatical flowchart of PSDOF for the fusion of surface data and point data.

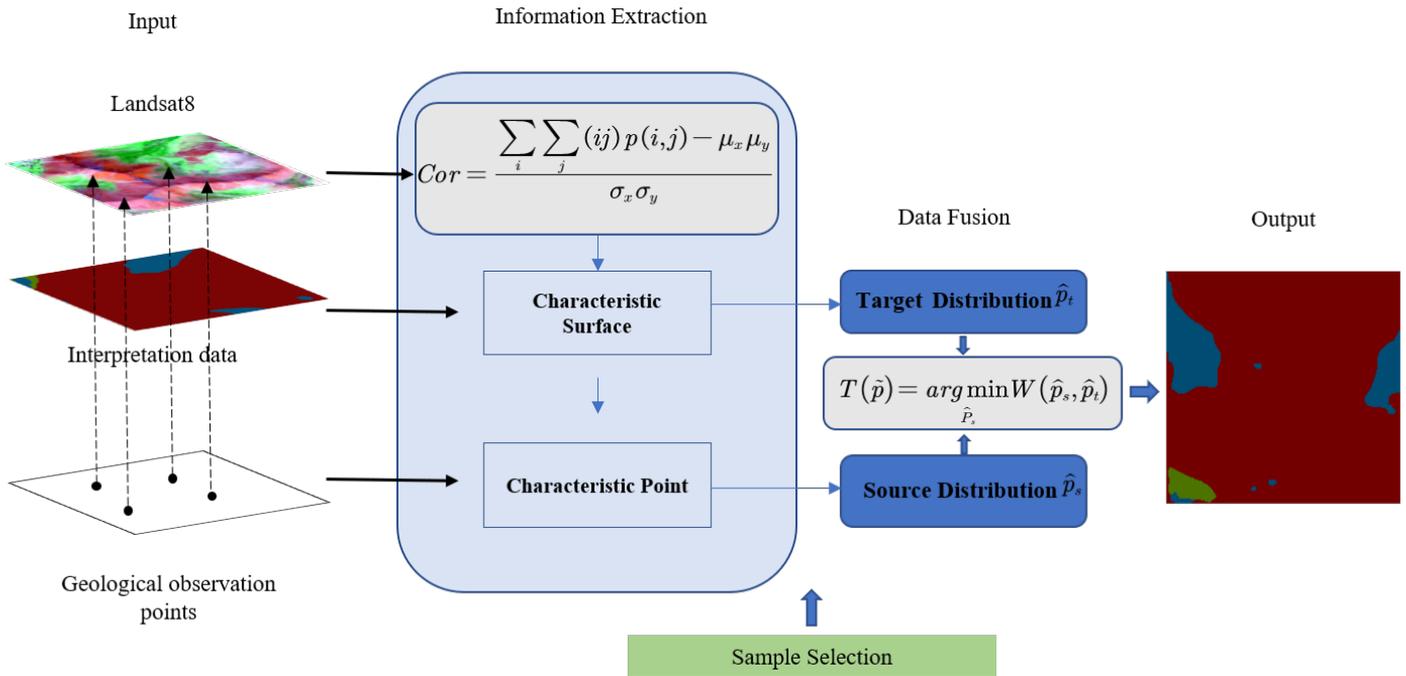


Figure 4. Diagrammatical flowchart of PSDOF for the fusion of surface data and point data.

4.1. Information Extraction

Geological survey point data are sparse, which raises an additional problem: a lack of samples and labels. To address this problem, PSDOF uses Landsat 8 images and a gray-level co-occurrence matrix (GLCM) to generate point samples and labels.

The GLCM is defined as follows [51]. Suppose our image (e.g., Landsat 8 images and GRSI products) is square, with c representing the columns and r the rows. The gray values of each pixel are quantized into different levels N_g , $G \in \{0, 1, \dots, N_g - 1\}$, thus creating a set of weighted gray values. The image can be depicted as a function that allocates a gray value G to couple $L_r \times L_c$ coordinates. $p(a, b)$ is the co-occurrence frequency matrix of the two gray values a and b , which are detached by the step distance d in the image. Moreover, it is a function of the distance and angle relationship between neighboring pixels, which, thus, reveals information about the texture of the image [52].

For d and angle θ , the normalized frequency is defined by the following equation:

$$\begin{aligned}
 p(a, b \mid d, \theta) &= \left\{ \begin{array}{l} (r, c) \mid f(r, c) = a, f(\Delta r + r, \Delta c + c) = b \\ r, c = 0, 1, \dots, N - 1 \end{array} \right\} \\
 d &= \sqrt{d_r^2 + d_c^2}; \\
 \Delta r &= r + d_r; \\
 \Delta c &= c + d_c; \\
 \theta &\in (0^\circ, 45^\circ, 90^\circ, 135^\circ).
 \end{aligned}
 \tag{6}$$

where N , represents the size of the image. Define $p(a, b)$ as the value of the (a, b) -th entry in the normalized GLCM. The averages of the rows and columns—as denoted by μ_r and μ_c , respectively—as well as their corresponding standard deviations can be described as follows:

$$\begin{aligned}
 \mu_r &= \sum_a \sum_b a \cdot p(a, b), \mu_c = \sum_a \sum_b b \cdot p(a, b) \\
 \sigma_r &= \sum_a \sum_b (a - \mu_r)^2 \cdot p(a, b) \\
 \sigma_c &= \sum_a \sum_b (b - \mu_c)^2 \cdot p(a, b)
 \end{aligned}
 \tag{7}$$

For different texture features, which represent different meanings, this study used only the gray-scale correlation features. These features are calculated as follows:

$$Cor(r, c) = \frac{\sum_a \sum_b (ab)p(a, b) - \mu_r \mu_c}{\sigma_r \sigma_c} \quad (8)$$

where $a, b \in \{0, \dots, 255\}$ corresponds to the number of gray values in the image pixels and $r, c \in \{0, \dots, 223\}$ corresponds to the number of ranks of the pixels. $Cor(r, c)$ represents the linear correlation of the gray-scale value of each pixel of the image; the higher the value, the stronger the correlation is [53].

The setting of the parameters d and θ is particularly important according to Equation 6. Fine textures require relatively small distance values, and ensuring this will produce more texture information than would be the case otherwise, as it is difficult to represent them if the distance values are too large. θ is relatively less important in co-occurrence matrices. Many authors have used the average [54]. After extensive experimental analysis, we chose the best parameters of d and θ for the surface data, where the step distance was $d = 2$. Each texture feature was rotated by $0^\circ, 45^\circ, 90^\circ$, and 135° , and the average feature value of the four angles was taken.

Pixels with similar $Cor(r, c)$ values around the geological survey points were selected as a set of point samples X_n for $n \in \{1, \dots, r \times c\}$, and the experiment showed that the distribution of the point sets matched well with the distribution of the geological elements in the ground truth data.

The GLCM method utilizes the correlation of neighboring pixels' gray-scale values in an image to represent its texture features. It is useful in addressing the issue of inadequate point samples in the application of the PSDOF method.

4.2. Sample Selection

How to map the complementary information of heterogeneous data into fusion space is a challenge. Specifically, we need to find an effective set of characteristic data and a mapping method that represents these data. This section introduces our data-processing method:

Characteristic point: This study selected pixel points in the neighboring area of the geological elements, which are represented by the geological survey data that have a consistent gray-scale correlation with the geological survey point data, as point samples. The point samples' set X_n was determined by the texture features of the geological elements. Since the geological survey point data had high-precision location information, all of the point samples could be generated with labels (with both location and geological element information). Regardless of the heterogeneity and origin of the point data, after information extraction, the point samples can all be converted into labels.

Characteristic surface: In the experiment, the characteristic surface was obtained by combining the textural features of the geological elements. We denote the $Cor(r, c)$ pixel distribution surface (using curve fitting) of the geological elements as P and the classification boundary of the surface data as K . The characteristic surface was $Y_m = P \cup K$ for $m \in \{1, \dots, r \times c\}$.

In principle, PSDOF requires a mapping of n labels among L modalities. Suppose there exist L different modalities of heterogeneous data (e.g., geological survey point and GRSI data), which are denoted as the mapping f_L with their own domain D_l and with the ranges V_l , respectively:

$$f_L : D_l \rightarrow V_l. \quad (9)$$

where D_l is a measurable probability space of the l types of multimodal data (including point data and surface data), and V_l can have values taken in the field of real numbers \mathbb{R} , complex values \mathbb{C} , integers \mathbb{Z} , etc. In our experiments, different geological survey point data represent different mappings f_L , thus revealing the types and spatial location information of the geological elements.

The goal of PSDOF is to homogenize the domains of the data of different modalities into the common space $D = \{x_1, \dots, x_n\}$ with $x_i \in \mathbb{R}^2$ for $i \in \{1, \dots, n\}$:

$$\phi_l : D_l \rightarrow D, \quad (10)$$

where the mapping ϕ_l is a non-linear many-to-one relationship and is metrizable. In the study, the remote sensing data were associated with $15 \times 15 \text{ m}^2$ grid cells, wherein one can consider those cells as elements in D . Thus, ϕ_l serves as a mapping from the positions of point samples to their nearest unit centers. This mapping relationship, when taken, allows for heterogeneous domains to be aligned and fused into a common domain:

$$\phi = \arg \min \phi_l \quad (11)$$

The fusion of multi-source heterogeneous information requires solving the alignment problem of the data domain. The domain adaptation technique provides a feasible solution: the source samples (characteristic point set X_n) and the target samples (characteristic surface Y_m) are correlated in a unified information space D . Furthermore, the data deviations between the domains can be addressed by solving the correlation mapping ϕ .

4.3. Data Fusion

To obtain the optimal mapping ϕ of the data domain, the PSDOF method introduces OT theory.

In the domain adaptation research area, OT provides a scheme for transferring SD distributions to the TD. Specifically, OT aims to estimate transport plans that minimize transport costs [55]. This fusion model considers only a finite domain $D = \{x_1, \dots, x_n\}$ with $x_i \in \mathbb{R}^2$. The source and target distributions can be expressed as $D_s = \{x_1^s, \dots, x_{n_s}^s\}$ (e.g., geological survey point data) and $D_t = \{y_1^t, \dots, y_{m_t}^t\}$ (e.g., GRSI products), where n_s and m_t are the count of units in D_s and D_t , respectively. Then, the expected distribution of the two domains P_s and P_t can be written as follows:

$$P_s = \sum_{x_i \in D} \omega_i^s \delta_{(x_i)}, \quad P_t = \sum_{x_i \in D} \omega_i^t \delta_{(y_i)} \quad (12)$$

where $\delta_{(x_i)}$ and $\delta_{(y_i)}$ are the Dirac function at x_i and y_i , respectively, and ω_i^s and ω_i^t are the factors of the unit simplex, with $\sum_{i=1}^{n_s} \omega_i^s = 1$ and $\sum_{i=1}^{m_t} \omega_i^t = 1$.

In the case of discrete OT, the empirical distributions \hat{P}_s and \hat{P}_t are estimates of P_s and P_t for the discrete data points in D_s and D_t , respectively. As such, we have

$$\hat{P}_s = \sum_{i=1}^{n_s} \hat{\omega}_i^s \delta_{(x_i^s)}, \quad \hat{P}_t = \sum_{i=1}^{m_t} \hat{\omega}_i^t \delta_{(y_i^t)}. \quad (13)$$

where $\hat{\omega}_i^s$ and $\hat{\omega}_i^t$ are the probability masses of the unit simplex $\delta_{(x_i^s)}$ and $\delta_{(y_i^t)}$, respectively, with $\sum_{i=1}^{n_s} \hat{\omega}_i^s = 1$ and $\sum_{i=1}^{m_t} \hat{\omega}_i^t = 1$. The best-matching π^* between \hat{P}_s and \hat{P}_t is calculated by Kantorovich's formula [56] as follows:

$$\pi^* = \arg \min_{\pi \in \beta} \langle \pi, C \rangle_F \quad (14)$$

where C is the matrix of the cost, $C(i, j) = c(x_i^s, y_j^t)$ indicates the cost of transporting the probability mass from x_i^s to y_j^t , and $\langle \cdot, \cdot \rangle_F$ is the Frobenius inner product. β is the set of associative discrete couplings between \hat{P}_s and \hat{P}_t , such that

$$\beta = \left\{ \pi \in (\mathbb{R}^+)^{n_s \times m_t} \mid \pi 1_{m_t} = \hat{P}_s, \pi^T 1_{n_s} = \hat{P}_t \right\} \quad (15)$$

where 1_n represents a vector of ones with length n .

The probability density of each position in the target distribution should be determined based on the geological element distribution, which is unknown and more difficult to obtain. This work assumed that the probability density of each position of the target distribution follows a uniform distribution—in other words, the probability density of x_i^s and y_j^t is $1/n_s$ and $1/n_t$, respectively. The function T_{OT} can then be stated as follows:

$$T_{OT}(x_i^s) = \sum_{j=1}^{n_i} n_s \pi^*(i, j) y_j^t \quad (16)$$

The corrected empirical distribution can be stated as

$$\tilde{P}_c = \sum_{i=1}^{n_s} \frac{1}{n_s} \delta_T(x_i^s) \quad (17)$$

Brenier [57] proved that the Kantorovich formula can be solved by a system of linear partial differential equations when the cost function is quadratic and the domain of definition D is a Euclidean space. This further optimizes $\langle \pi, C \rangle_F$, and \tilde{P}_c , which can be approximated as

$$\tilde{P}_c = \arg \min W(\hat{P}_s, \hat{P}_t) \quad (18)$$

where $W(\cdot)$ is the Wasserstein distance. The approach can be viewed as a local second-order approximation, which can be used to minimize the Wasserstein distance $\langle \pi, C \rangle_F$ at $\{x_i^s\}_{i=1}^{n_s}$.

In a given geometric space, X_n and Y_m can be regarded as the SD and TD, respectively. The optimal coupling can be calculated from Equation (14), where the cost matrix $C(X_n, Y_m)$ is the Euclidean distance of the uniformly discrete feature pixels in the sample space X_n and Y_m :

$$C(X_n, Y_m) = \|x_n^s - y_m^t\|_2^2 \quad (19)$$

The transport costs of the OT model can be seen in solving for the two-Wasserstein distance:

$$\tilde{P} = \arg \min W_2(\hat{P}_{x_n^s}, \hat{P}_{y_m^t}) \quad (20)$$

To address the issue of data domain misalignment during the fusion of heterogeneous data from multiple sources, the data fusion module introduces the OT method. The experiment demonstrated the superiority of PSDOF in the fusion process of geological survey point data and GRSI products.

5. Experimentation

This section describes the experimental methodology and results of the study. This includes the experimental setting and parameters, the evaluation metrics employed, an analysis of the experimental results, and a comparison of the outcomes obtained with different GRSI models.

5.1. Experimental Setting and Parameters

In this experiment, we aligned Landsat 8 image and ground truth data, as well as cropped them into 1200 images of a 224×224 px in size. The training and testing data were divided into an 80%:20% ratio. We used six semantic segmentation models for geological remote sensing interpretation: FCN, DeeplabV3, OCNNet, DANet, PSPNet, and AdvSemi-OCGNet. To ensure fairness in the experiment, all of these models were trained using ResNet50 as the backbone network, as well as by utilizing the official pre-trained models provided by PyTorch. In the comparative experiments, the geological remote sensing interpretation models were initialized with a learning rate of 2.5×10^{-4} and were trained for a total of 10,000 iterations with a batch size of 24. The optimizer was set to SGD. All of the experiments were conducted on a workstation equipped with an Intel i7 11700k CPU, an NVIDIA RTX 3090 GPU, and code methods that were utilized in Python.

Then, we analyzed the experimental parameters of the PSDOF method based on the interpretation results of the AdvSemi-OCGNet method:

1. Table 3 shows the grading of the characteristic points X_n^i and characteristic surface Y_m^i for the three study areas. After our analysis of the data from the different study areas, the levels were divided into 11 classes, where $i \in \{1, 2, \dots, 11\}$. The number of feature pixel points corresponding to each level of the feature surface showed a linear distribution, thereby reflecting the distribution of the 11 geological elements.
2. X_n^i and Y_m^i are understood as discrete samples, where $i \in \{1, 2, \dots, 11\}$. The minimum cost of transport of the OT model after performing domain fusion is shown in Table 4. For the transport costs, Table 4 yields the following conclusions: When the costs corresponding to level 2 in Area 1 and level 7 in Areas 2 and 3 are the respective optimal transport costs, the minimum transport costs are 0.4409×10^{-6} , 0.2180×10^{-6} , and 0.4092×10^{-6} , respectively. In addition, we could update the GRSI products according to the optimal transport rates for each of the three areas.

Table 3. Characteristic surface division.

Level (A)	Points		
	Area 1	Area 2	Area 3
1. Cor ± 0.001	13	18	5
2. Cor ± 0.050	52	84	14
3. Cor ± 0.010	121	179	25
4. Cor ± 0.015	185	267	32
5. Cor ± 0.020	254	358	42
6. Cor ± 0.025	349	444	52
7. Cor ± 0.030	410	538	64
8. Cor ± 0.035	487	634	75
9. Cor ± 0.040	552	711	-
10. Cor ± 0.045	614	796	-
11. Cor ± 0.050	704	873	-

Table 4. Transport costs.

Level	Cost		
	Area 1	Area 2	Area 3
1	1.2230×10^{-6}	2.2230×10^{-6}	1.9357×10^{-6}
2	0.4409×10^{-6}	0.8954×10^{-6}	0.0096×10^{-6}
3	0.6152×10^{-6}	0.4241×10^{-6}	0.7318×10^{-6}
4	1.0146×10^{-6}	70.2747×10^{-6}	0.7429×10^{-6}
5	0.9241×10^{-6}	0.2277×10^{-6}	0.7186×10^{-6}
6	1.5249×10^{-6}	0.2209×10^{-6}	0.5095×10^{-6}
7	1.5358×10^{-6}	0.2180×10^{-6}	0.4092×10^{-6}
8	2.2975×10^{-6}	0.3129×10^{-6}	0.4552×10^{-6}
9	2.4674×10^{-6}	0.2982×10^{-6}	-
10	2.6349×10^{-6}	0.3071×10^{-6}	-
11	1.7774×10^{-6}	0.2985×10^{-6}	-

5.2. Evaluation Indicators

In this section, the evaluation metrics used to quantitatively assess the accuracy of our model are presented, followed by a discussion of the important parameter settings for the experimental area used in this research. Finally, a presentation and analysis of the experimental results is provided.

To evaluate the accuracy improvement of the GRSI product in our model, the evaluation metric used was the intersection over union [58], which can be derived by calculating the confusion matrix.

Intersection over Union (IOU): The IOU of nine geological elements is computed by the model as the ratio of the intersection and union of the anticipated and actual pixels within a group:

$$IOU = \sum_{i=0}^k \frac{p_{ij}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ij} - p_{ii}} \quad (21)$$

$$MIOU = \frac{1}{k+1} IOU \quad (22)$$

where k indicates the category of the geological elements, p_{ii} denotes the total number of true pixels of class i that are correctly predicted as i , and p_{ij} denotes the total number of true pixels of class i that are incorrectly predicted as j . The accuracy of the model is determined by the IOU. The larger the IOU, the greater the overlap between the predicted pixels and the ground truth labels is, thus indicating a more-accurate model.

5.3. Effectiveness Assessment

We used the AdvSemi-OCGNet PSDOF model for the geological interpretation of the remote sensing data to evaluate the effect of PSDOF on the accuracy improvement of the geological interpretation product. As shown in Figure 5, the geological element of the observation point was sandstone, and the classification result of the interpretation model in Area 1 was soil. The IOU accuracy of the sandstone was only 8.1%, and the interpretation accuracy was poor. Our PSDOF combines the geological element information around the observation point (red rectangular boxed area) into the interpretation of the product. It utilizes this information to update the optimal interpretation boundary based on the minimum transport cost, thus resulting in a more-accurate GRSI product. Based on the accuracy assessment data in Table 5, the IOU of the updated interpreted product metamorphic sandstone in Area 1 improved by 53%. Moreover, as can be seen in Table 6, there was an overall improvement of 19% in the MIOU of its nine geological elements.

In Area 2, the geological element of the observation site (the red rectangular boxed area) was volcanic debris, and the GRSI product identified only a small amount of volcanic debris within this area (with the remainder being classified as extensive glaciation). According to Table 5, the IOU accuracy for volcanic debris in the area was 68.4% (which improved to 73.4% after our PSDOF model), and the classification accuracy of the volcanic debris improved by 5.6%. According to Table 6, the overall MIOU of the nine geological elements improved by 2%.

The information gathered from the observation points (within the red rectangular boxed area) in Area 3 suggests a significant presence of granite. Although the geological interpretation products yielded a relatively clear granite–soil boundary, we estimated that it can still be optimized. Table 5 displays the outcomes, where the granite IOU within the observation site improved from 71.1% to 73.2% (with an improvement in the classification accuracy by 21.1%). According to Table 6, the MIOU of the nine geological elements in Area 3 improved by 1%. The results demonstrated the correctness of our optimization scheme in Area 3.

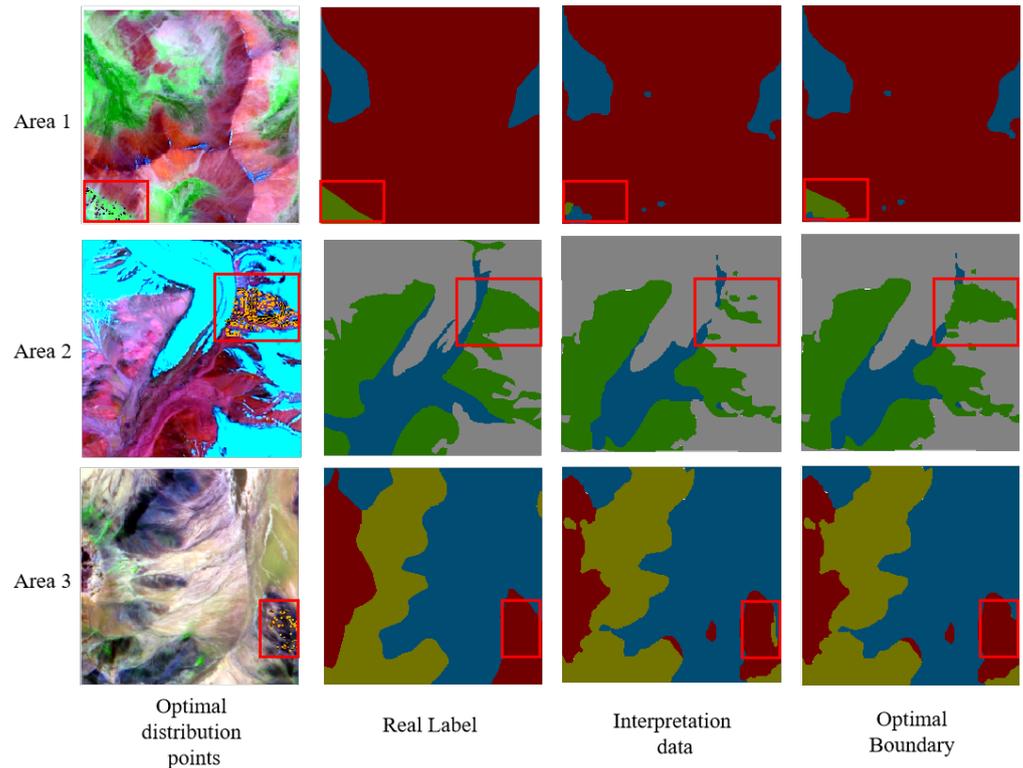


Figure 5. Results obtained by AdvSemi-OCGNet and PSDOF.

Table 5. IOU quantitative assessment.

Area	IOU (%)	
	AdvSemi-OCGNet	PSDOF
1	8.11	61.18
2	68.42	73.39
3	71.11	73.22

Table 6. MIOU quantitative assessment.

Area	IOU (%)	
	AdvSemi-OCGNet	PSDOF
1	59.93	78.78
2	66.45	68.66
3	81.02	82.03

5.4. Comparative Experiments

To verify the accuracy improvement effect of PSDOF for different DL-based GRSI products, GRSI was carried out under the following classical and representative semantic segmentation models: FCN, DeepLabV3, DANet, OCGNet, and PSPNet. Our PSDOF model was added for comparison experiments in three study areas, and the visual effects are shown in Figures 6–8:

1. FCN and PSDOF: As shown in Figure 6, the processed Landsat 8 satellite images were fed into an FCN interpretation model to classify the nine geological elements at the pixel level in order to obtain the interpretation products. As can be seen from Figure 6, the interpretation accuracy for all three geological elements was unsatisfactory. The sandstone and volcanic debris in the red boxes in the figure were located, but with inaccurate boundaries (the specific IOU accuracy can be seen in Table 7). The

distribution of what was originally thought to be granite was determined as slate instead, which was completely inconsistent with the observations (and the annotated diagram also fully supports this view). The PSDOF method resulted in the correct repositioning of the three lithologies in the vicinity of the observation point, as evident from the visual effect. It fully combined the high-precision spatial information of the observation point data and achieved a higher positional accuracy.

2. DeepLabv3 and PSDOF: Figure 6 shows the interpretation products obtained from the DeepLabv3 semantic segmentation model for the remote sensing images of the three study areas. Unlike the classical FCN model, the Deeplabv3 model uses an improved atrous spatial pyramid pooling (ASPP) approach. It demonstrated better performance in pixel-level classifications; thus, it outperformed the FCN model overall. However, as can be seen from Figure 7, the results were no better than the FCN model for the boundary localization of sandstone, granite, and volcanic debris, and the accuracy needs to be improved. Our experimental results demonstrated that PSDOF effectively improves the accuracy of the GRSI products generated by the Deeplabv3 model. Within the observation point range, correct boundary relocations were obtained for the three lithologies.
3. DANet and PSDOF: The dual-attention network (DANet) has a location attention module to capture spatial information. DANet introduces a channel attention module that is used to integrate the relevant features between all channel mappings, and it overall outperformed the FCN network. The performance of DANet was poor in the extraction of the following three lithological features: sandstone, granite, and volcanic debris. As shown in Figure 7, where the metamorphic sandstone and volcanic debris in the red boxes of the interpretation results were located, granite was incorrectly located, and the segmentation accuracy of all of the three geological elements was not high. After applying PSDOF, the accuracy of the three geological elements in the product improved.
4. OCNet and PSDOF: The neural network model OCNet is a target semantic network for scene parsing in the form of semantic aggregation, where instead of pixel-by-pixel prediction, similar pixel points are aggregated and, then, semantically segmented. The experimental results in the three study areas showed that the classification of sandstone, granite, and volcanic debris, as shown in Figure 7, was poor. The three geological elements of the scene showed a fragmented and scattered distribution after classification, as well as blurred geological element boundaries. In contrast, after incorporating the complementary observation point information, PSDOF was able to incorporate it into the interpretation product and relocate the boundaries of the geological elements at a local scale. The experimental results also showed that PSDOF can perform well for geological elements with poor accuracy.
5. PSPNet and PSDOF: PSPNet is a scenario-analysis network built using the pyramid pooling module; furthermore, it outperforms traditional FCNs, and its overall performance was relatively good in this experiment. The PSPNet model in the study area had a good feature extraction capability for sandstone, as shown in Figure 8, and it ranked high in overall performance for the classification of granite and metamorphic sandstone. However, there was a large error in the boundary positioning of the three geological elements at the observation point range, but our PSDOF successfully handled and corrected this error. Moreover, the resulting interpretation products after processing also validated the good performance of PSDOF.

To more-accurately represent the enhancement effect of the different geological interpretation models with the addition of PSDOF, the IOU accuracy of the individual geological elements were ascertained in each of the three study areas, as is presented in Tables 7–11. It can be observed that the five DL interpretation frameworks had different performances for different geological elements. However, it was found that it was more difficult to learn the features of granite, sandstone, and volcanic debris. The main reason for this is that granite and volcanic clasts are both magmatic rocks with similar morphological structures

in the image. It is difficult to distinguish between sandstone, metamorphosed granite, and fine-grained granite, as they are similar. In a comparison of the three study area ranges, the PSPNet model was the most accurate in classifying sandstone with an IOU of 72%, and the OCNet model performed the worst with an accuracy of 0%. In terms of extracting volcanic debris features, the DANet model had the best IOU accuracy of 65%, while the OCNet model was less effective in classification. In terms of extracting features from granite, all five models had an IOU accuracy of around 30%, and none of them showed satisfactory performance.

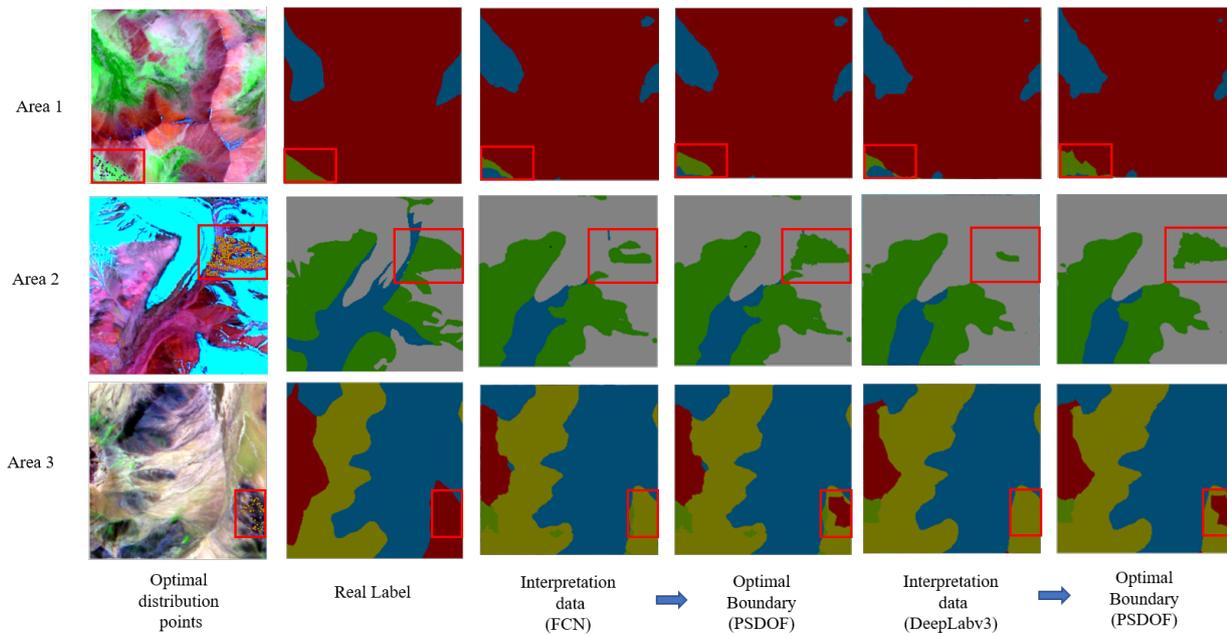


Figure 6. Results obtained by FCN, DeepLabv3, and PSDOF.

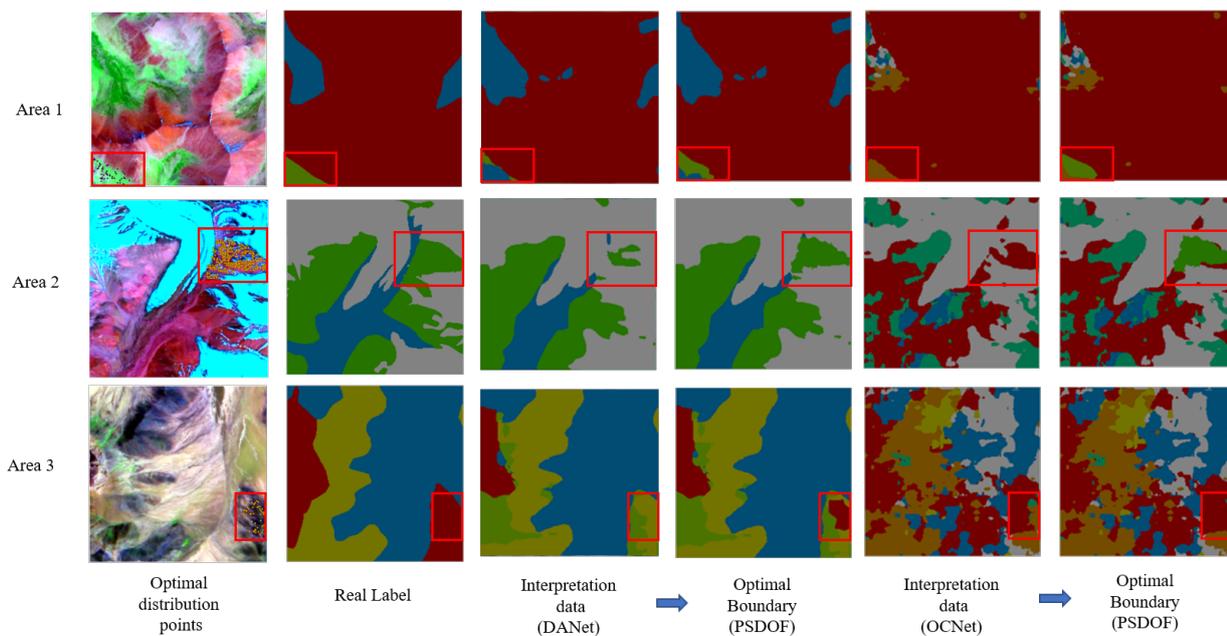


Figure 7. Results obtained by DANet, OCNet, and PSDOF.

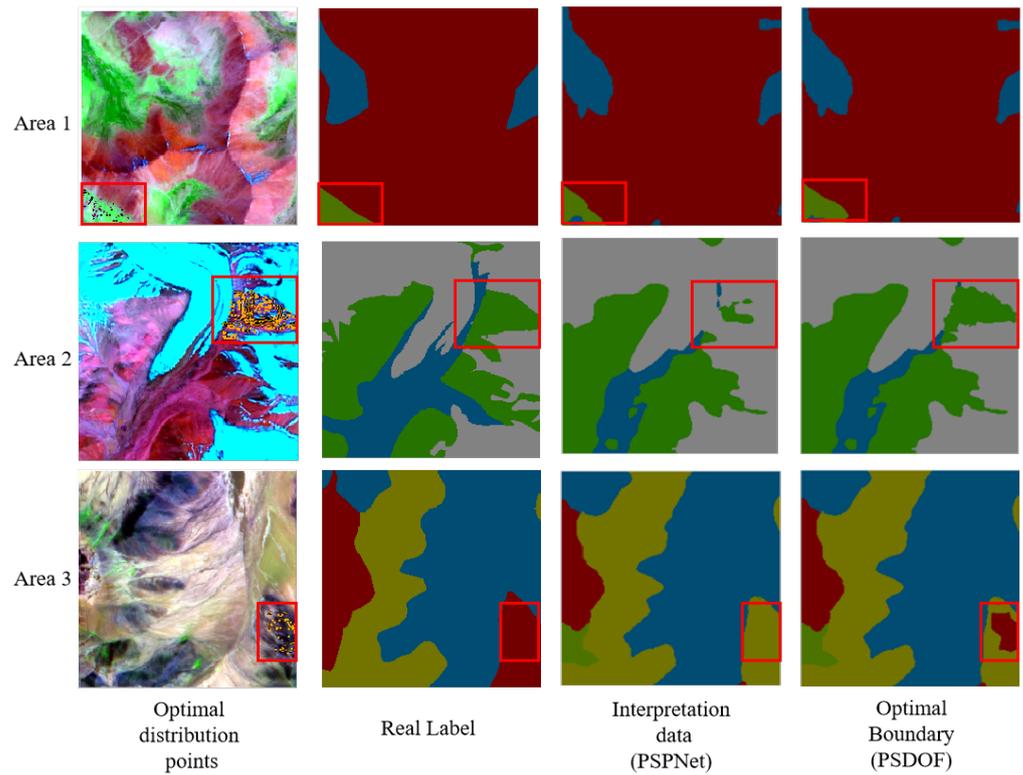


Figure 8. Results obtained by PSPNet and PSDOF.

Table 7. IOU quantitative assessment: FCN and PSDOF.

Area	IOU (%)	
	FCN	PSDOF
1	33.89	63.99
2	64.66	67.94
3	34.24	42.78

Table 8. IOU quantitative assessment: DeepLabv3 and PSDOF.

Area	IOU (%)	
	DeepLabv3	PSDOF
1	33.86	76.44
2	56.10	62.51
3	36.69	45.18

Table 9. IOU quantitative assessment: DANet and PSDOF.

Area	IOU (%)	
	DANet	PSDOF
1	16.58	54.37
2	65.06	69.77
3	32.42	40.96

Table 10. IOU quantitative assessment: OCNet and PSDOF.

Area	IOU (%)	
	OCNet	PSDOF
1	0.00	57.98
2	0.00	10.26
3	34.37	36.69

Table 11. IOU quantitative assessment: PSPNet and PSDOF.

Area	IOU (%)	
	PSPNet	PSDOF
1	72.33	78.09
2	59.31	64.67
3	35.62	44.15

Our experimental analysis of the different geological interpretation models showed that different models had different feature-extraction performances for the same geological elements, as well as that the same models had different feature-learning abilities for different geological elements. After adding our PSDOF to the interpretation products, the interpretation accuracy of all of the models for different geological elements improved. The accuracy of the metamorphic sandstone improved the most with a 60% improvement in the IOU accuracy in the local range. The local IOU accuracy of the other two lithologies was also improved. The experimental results showed that PSDOF had an excellent effect on the improvement of the accuracy of the selected geological interpretation products.

6. Conclusions and Outlook

In this paper, we proposed the use of additional point data for the data enhancement of GRSI products so as to address the problem of their limited accuracy. A multimodal data-fusion framework base for an optimal transport model that is capable of fusing heterogeneous data with minimal cost was presented. The GRSI product for a local area of the Pamir Plateau is an example of a solution to the problem of fusing geological footprint data with interpretation data. The model uses gray-scale correlation features in the fusion task. The experiments showed that our fusion framework successfully fused point data geospatial location information and lithological information with GRSI products, thereby resulting in a higher quality.

The PSDOF method still showed certain limitations in the experiment. Due to the sparse distribution of geological survey data and the limited amount of data, the experimental study area was confined to a local region. Moreover, the fusion effect was constrained by the sample size of the point data. While the PSDOF method had a low computational cost in local regions, increasing the number of parameters will lead to a higher computational cost.

In future research, we will consider using a greater amount of multimodal remote sensing data, such as big social media and statistical data, to take full advantage of the information from different data sources in order to obtain a higher-quality GRSI product. In addition, our model was only examined experimentally on its capacity for geological remote sensing mapping; as such, in the future, we will consider more application scenarios, such as in agriculture, water bodies, wetlands, and cities, in order to combine the characteristics of different geographical elements and improve the generalization capability of the model.

Author Contributions: J.C. and J.W. conceived of the idea; J.W. verified the idea and designed the study; J.W. analyzed the experimental results; J.W. wrote the paper; S.W. provided the source of the data; J.C. and W.H. gave comments and suggestions to the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China under Grants U21A2013 and 42201415; the Open Research Program of the International Research Center of Big Data for Sustainable Development Goals under Grant CBAS2023ORP03; the Hubei Natural Science Foundation of China under Grant 2022CFB607.

Data Availability Statement: No new data were created nor analyzed in this study. Data are contained within the article.

Acknowledgments: The authors thank the National Natural Science Foundation of China under Grants U21A2013 and 42201415; the Open Research Program of the International Research Center of Big Data for Sustainable Development Goals under Grant CBAS2023ORP03; the Hubei Natural Science Foundation of China under Grant 2022CFB607 for providing support.

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

References

1. Han, W.; Zhang, X.; Wang, Y.; Wang, L.; Huang, X.; Li, J.; Wang, S.; Chen, W.; Li, X.; Feng, R.; et al. A Survey of Machine Learning and Deep Learning in Remote Sensing of Geological Environment: Challenges, Advances, and Opportunities. *ISPRS J. Photogramm. Remote Sens.* **2023**, *202*, 87–113. [[CrossRef](#)]
2. Ma, Z.; Min, X.; Lin, H.; Qian, M.; Zhang, Y. FENet: Feature enhancement network for land cover classification. *Int. J. Remote Sens.* **2023**, *44*, 1702–1725. [[CrossRef](#)]
3. Zhou, G.; Xu, J.; Chen, W.; Li, X.; Li, J.; Wang, L. Deep Feature Enhancement Method for Land Cover with Irregular and Sparse Spatial Distribution Features: A Case Study on Open-Pit Mining. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–20. [[CrossRef](#)]
4. Zhang, X.; Wang, L.; Li, J.; Han, W.; Fan, R.; Wang, S. Satellite-derived sediment distribution mapping using ICESat-2 and SuperDove. *ISPRS J. Photogramm. Remote Sens.* **2023**, *202*, 545–564. [[CrossRef](#)]
5. Inzana, J.; Kusky, T.; Higgs, G.; Tucker, R. Supervised classifications of Landsat TM band ratio images and Landsat TM band ratio image with radar for geological interpretations of central Madagascar. *J. Afr. Earth Sci.* **2003**, *37*, 59–72. [[CrossRef](#)]
6. Dalati, M. Remote sensing techniques in active faults surveying. Case study: Detecting active faulting zones NW of Damascus, Syria. In Proceedings of the 2nd International Conference on Recent Advances in Space Technologies, RAST 2005, Istanbul, Turkey, 9–11 June 2005; pp. 479–482. [[CrossRef](#)]
7. Gad, S.; Kusky, T. ASTER Spectral Ratioing for Lithological Mapping in the Arabian–Nubian Shield, the Neoproterozoic Wadi Kid Area, Sinai, Egypt. *Gondwana Res.* **2007**, *11*, 326–335. [[CrossRef](#)]
8. Kruse, F.; Lefkoff, A.; Boardman, J.; Heidebrecht, K.; Shapiro, A.; Barloon, P.; Goetz, A. The spectral image processing system (SIPS): Software for integrated analysis of AVIRIS data. In *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*; Jet Propulsion Lab.: Pasadena, CA, USA, 1992.
9. Deville, Y.; Brezini, S.E.; Benhalouche, F.Z.; Karoui, M.S.; Guillaume, M.; Lenot, X.; Lafrance, B.; Chami, M.; Jay, S.; Minghelli, A.; et al. Modeling and Unsupervised Unmixing Based on Spectral Variability for Hyperspectral Oceanic Remote Sensing Data with Adjacency Effects. *Remote Sens.* **2023**, *15*, 4583. [[CrossRef](#)]
10. Borsoi, R.A.; Imbiriba, T.; Bermudez, J.C.M.; Richard, C.; Chanussot, J.; Drumetz, L.; Tourneret, J.Y.; Zare, A.; Jutten, C. Spectral Variability in Hyperspectral Data Unmixing: A Comprehensive Review. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 223–270. [[CrossRef](#)]
11. Al-Nahmi, F.; Saddiqi, O.; Hilali, A.; Rhinane, H.; Baidder, L.; El arabi, H.; Khanbari, K. Application of Remote Sensing in Geological Mapping, Case Study Al Maghrabah Area—Hajjah Region, Yemen. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *IV-4/W4*, 63–71. [[CrossRef](#)]
12. Abdolmaleki, M.; Rasmussen, T.M.; Pal, M.K. Exploration of IOCG Mineralizations Using Integration of Space-Borne Remote Sensing Data with Airborne Geophysical Data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *XLIII-B3-2020*, 9–16. [[CrossRef](#)]
13. Traore, M.; Takodjou Wambo, J.D.; Ndepete, C.P.; Tekin, S.; Pour, A.B.; Muslim, A.M. Lithological and Alteration Mineral Mapping for Alluvial Gold Exploration in the South East of Birao Area, Central African Republic Using Landsat-8 Operational Land Imager (OLI) Data. *J. Afr. Earth Sci.* **2020**, *170*, 103933. [[CrossRef](#)]
14. Elhamdouni, D.; Karaoui, I.; Arioua, A. Automatic Geological Mapping Using Remote Sensing Data: Case of the Zgounder Deposit (Anti-Atlas, Morocco). *Appl. Geomat.* **2023**, 1–11. [[CrossRef](#)]
15. El Fels, A.E.A.; El Ghorfi, M. Using Remote Sensing Data for Geological Mapping in Semi-Arid Environment: A Machine Learning Approach. *Earth Sci. Inform.* **2022**, *15*, 485–496. [[CrossRef](#)]

16. Han, W.; Chen, J.; Wang, L.; Feng, R.; Li, F.; Wu, L.; Tian, T.; Yan, J. Methods for Small, Weak Object Detection in Optical High-Resolution Remote Sensing Images: A Survey of Advances and Challenges. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 8–34. [[CrossRef](#)]
17. Shrestha, A.; Mahmood, A. Review of Deep Learning Algorithms and Architectures. *IEEE Access* **2019**, *7*, 53040–53065. [[CrossRef](#)]
18. LeCun, Y.; Bengio, Y. Convolutional Networks for Images, Speech, and Time-Series. In *The Handbook of Brain Theory and Neural Networks*; Citeseer: Pittsburgh, PA, USA, 1995.
19. Latifovic, R.; Pouliot, D.; Campbell, J. Assessment of Convolution Neural Networks for Surficial Geology Mapping in the South Rae Geological Region, Northwest Territories, Canada. *Remote Sens.* **2018**, *10*, 307. [[CrossRef](#)]
20. Shirmard, H.; Farahbakhsh, E.; Heidari, E.; Beiranvand Pour, A.; Pradhan, B.; Müller, D.; Chandra, R. A Comparative Study of Convolutional Neural Networks and Conventional Machine Learning Models for Lithological Mapping Using Remote Sensing Data. *Remote Sens.* **2022**, *14*, 819. [[CrossRef](#)]
21. Sang, X.; Xue, L.; Ran, X.; Li, X.; Liu, J.; Liu, Z. Intelligent High-Resolution Geological Mapping Based on SLIC-CNN. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 99. [[CrossRef](#)]
22. Wang, J.; Li, W.; Zhang, M.; Tao, R.; Chanussot, J. Remote-Sensing Scene Classification via Multistage Self-Guided Separation Network. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5615312. [[CrossRef](#)]
23. Wang, J.; Li, W.; Wang, Y.; Tao, R.; Du, Q. Representation-Enhanced Status Replay Network for Multisource Remote-Sensing Image Classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–13. [[CrossRef](#)]
24. Zhang, M.; Li, W.; Zhang, Y.; Tao, R.; Du, Q. Hyperspectral and LiDAR Data Classification Based on Structural Optimization Transmission. *IEEE Trans. Cybern.* **2023**, *53*, 3153–3164. [[CrossRef](#)] [[PubMed](#)]
25. Zhang, M.; Li, W.; Zhao, X.; Liu, H.; Tao, R.; Du, Q. Morphological Transformation and Spatial-Logical Aggregation for Tree Species Classification Using Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5501212. [[CrossRef](#)]
26. Wang, J.; Li, W.; Zhang, M.; Chanussot, J. Large Kernel Sparse ConvNet weighted by Multi-frequency Attention for Remote Sensing Scene Understanding. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5626112. [[CrossRef](#)]
27. Zhao, H.; Deng, K.; Li, N.; Wang, Z.; Wei, W. Hierarchical Spatial-Spectral Feature Extraction with Long Short Term Memory (LSTM) for Mineral Identification Using Hyperspectral Imagery. *Sensors* **2020**, *20*, 6854. [[CrossRef](#)] [[PubMed](#)]
28. Shirmard, H.; Farahbakhsh, E.; Müller, R.D.; Chandra, R. A Review of Machine Learning in Processing Remote Sensing Data for Mineral Exploration. *Remote Sens. Environ.* **2022**, *268*, 112750. [[CrossRef](#)]
29. Zhang, G.; Chen, W.; Li, G.; Yang, W.; Yi, S.; Luo, W. Lake Water and Glacier Mass Gains in the Northwestern Tibetan Plateau Observed from Multi-Sensor Remote Sensing Data: Implication of an Enhanced Hydrological Cycle. *Remote Sens. Environ.* **2020**, *237*, 111554. [[CrossRef](#)]
30. Jalilvand, E.; Tajrishy, M.; Ghazi Zadeh Hashemi, S.A.; Brocca, L. Quantification of Irrigation Water Using Remote Sensing of Soil Moisture in a Semi-Arid Region. *Remote Sens. Environ.* **2019**, *231*, 111226. [[CrossRef](#)]
31. Chi, H.; Sun, J.; Zhang, C.; Miao, C. Remote Sensing Data Processing and Analysis for the Identification of Geological Entities. *Acta Geophys.* **2022**, *71*, 1565–1577. [[CrossRef](#)]
32. Hoffman, J.; Rodner, E.; Donahue, J.; Darrell, T.; Saenko, K. Efficient Learning of Domain-invariant Image Representations. *arXiv* **2013**, arXiv:1301.3224.
33. Kulis, B.; Saenko, K.; Darrell, T. What You Saw Is Not What You Get: Domain Adaptation Using Asymmetric Kernel Transforms. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; pp. 1785–1792. [[CrossRef](#)]
34. Saenko, K.; Kulis, B.; Fritz, M.; Darrell, T. Adapting Visual Category Models to New Domains. In *Computer Vision—ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Greece, 5–11 September 2010*; Daniilidis, K., Maragos, P., Paragios, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 213–226.
35. Tuia, D.; Camps-Valls, G. Kernel Manifold Alignment for Domain Adaptation. *PLoS ONE* **2016**, *11*, e0148655. [[CrossRef](#)]
36. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **2010**, *22*, 199–210. [[CrossRef](#)] [[PubMed](#)]
37. Long, M.; Zhu, H.; Wang, J.; Jordan, M.I. Unsupervised Domain Adaptation with Residual Transfer Networks. In Proceedings of the Advances in Neural Information Processing Systems 29 (NIPS 2016), Barcelona, Spain, 5–10 December 2016.
38. Wang, C.; Mahadevan, S. Manifold Alignment without Correspondence. In Proceedings of the International Joint Conference on Artificial Intelligence, Hainan Island, China, 25–26 April 2009.
39. Baktashmotlagh, M.; Harandi, M.T.; Lovell, B.C.; Salzmann, M. Unsupervised Domain Adaptation by Domain Invariant Projection. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; pp. 769–776. [[CrossRef](#)]
40. Courty, N.; Flamary, R.; Tuia, D.; Rakotomamonjy, A. Optimal Transport for Domain Adaptation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1853–1865. [[CrossRef](#)] [[PubMed](#)]
41. Wang, H.; Skau, E.; Krim, H.; Cervone, G. Fusing Heterogeneous Data: A Case for Remote Sensing and Social Media. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6956–6968. [[CrossRef](#)]
42. Liu, Z.; Qiu, Q.; Li, J.; Wang, L.; Plaza, A. Geographic Optimal Transport for Heterogeneous Data: Fusing Remote Sensing and Social Media. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 6935–6945. [[CrossRef](#)]

43. Wei, X.; Li, H.; Sun, J.; Chen, L. Unsupervised Domain Adaptation with Regularized Optimal Transport for Multimodal 2D + 3D Facial Expression Recognition. In Proceedings of the 2018 13th IEEE International Conference on Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 31–37. [[CrossRef](#)]
44. Yan, Y.; Li, W.; Wu, H.; Min, H.; Tan, M.; Wu, Q. Semi-Supervised Optimal Transport for Heterogeneous Domain Adaptation. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 13–19 July 2018; pp. 2969–2975. [[CrossRef](#)]
45. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
46. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587.
47. Fu, J.; Liu, J.; Tian, H.; Li, Y.; Bao, Y.; Fang, Z.; Lu, H. Dual Attention Network for Scene Segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019. [[CrossRef](#)]
48. Yuan, Y.; Huang, L.; Guo, J.; Zhang, C.; Chen, X.; Wang, J. OCNNet: Object Context Network for Scene Parsing. *arXiv* **2021**, arXiv:1809.00916. [[CrossRef](#)]
49. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2881–2890.
50. Wang, S.; Huang, X.; Han, W.; Li, J.; Zhang, X.; Wang, L. Lithological mapping of geological remote sensing via adversarial semi-supervised segmentation network. *Int. J. Appl. Earth Obs. Andgeoinf.* **2023**, *125*, 103536. [[CrossRef](#)]
51. Iqbal, N.; Mumtaz, R.; Shafi, U.; Zaidi, S.M.H. Gray Level Co-Occurrence Matrix (GLCM) Texture Based Crop Classification Using Low Altitude Remote Sensing Platforms. *PeerJ Comput. Sci.* **2021**, *7*, e536. [[CrossRef](#)]
52. Varish, N.; Hasan, M.K.; Khan, A.; Zamani, A.T.; Ayyasamy, V.; Islam, S.; Alam, R. Content-Based Remote Sensing Image Retrieval Method Using Adaptive Tetrolet Transform Based GLCM Features. *J. Intell. Fuzzy Syst.* **2023**, *44*, 9627–9650. [[CrossRef](#)]
53. Li, Z.; Ding, R. Vegetation Extraction in Taishan Region Based on High-Resolution Satellite Remote Sensing Images. In Proceedings of the Second International Conference on Optics and Image Processing (ICOIP 2022), Taian, China, 9 September 2022; p. 31. [[CrossRef](#)]
54. Indra, D.; Fadlillah, H.M.; Kasman; Ilmawan, L.B. Rice Texture Analysis Using GLCM Features. In Proceedings of the 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa, 9–10 December 2021; pp. 1–5. [[CrossRef](#)]
55. Ortiz-Jimenez, G.; Gheche, M.E.; Simou, E.; Maretic, H.P.; Frossard, P. Forward-Backward Splitting for Optimal Transport Based Problems. In Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2019. Available online: <http://xxx.lanl.gov/abs/1909.11448> (accessed on 13 February 2023).
56. Kantorovitch, L. On the Translocation of Masses. *J. Math. Sci.* **1958**, *133*, 1381–1382. [[CrossRef](#)]
57. Brenier, Y. Polar Factorization and Monotone Rearrangement of Vector-Valued Functions. *Commun. Pure Appl. Math.* **1991**, *44*, 375–417. [[CrossRef](#)]
58. Zhang, T.; Luo, B.; Sharda, A.; Wang, G. Dynamic Label Assignment for Object Detection by Combining Predicted IoUs and Anchor IoUs. *J. Imaging* **2022**, *8*, 193. [[CrossRef](#)] [[PubMed](#)]

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