remote sensing

MDPI

Article

RockCloud-Align: A High-Precision Benchmark for Rock-Mass Point-Cloud Registration

Yunbiao Wang †, Dongbo Yu †, Lupeng Liu * and Jun Xiao

School of Artificial Intelligence, University of Chinese Academy and Sciences, No. 19 Yuquan Road, Beijing 100049, China; wangyunbiao@ucas.ac.cn (Y.W.); yudongbo@ucas.ac.cn (D.Y.); xiaojun@ucas.ac.cn (J.X.)

* Correspondence: liulupeng@ucas.ac.cn

† These authors contributed equally to this work.

Abstract: Rock-mass point-cloud registration is a critical yet challenging task in the fields of geology and engineering. Currently, the lack of dedicated datasets for rock-mass point-cloud registration significantly limits the development and application of advanced algorithms in this area. To address this gap, we introduce RockCloud-Align, a large-scale dataset specifically designed for rock-mass point-cloud registration. Created using high-resolution LiDAR scans, this dataset covers a wide range of geological scenarios with varying densities and includes over 14,000 meticulously curated point-cloud pairs. RockCloud-Align provides a comprehensive benchmark for evaluating registration algorithms, along with a robust evaluation protocol to standardize the assessment of these methods. Building upon this dataset, we propose a novel registration method that eliminates the dependence on feature points and random sampling consensus, ensuring high efficiency and precision across diverse scenes and densities. Extensive experiments demonstrate that the proposed method significantly outperforms existing approaches in both accuracy and computational efficiency.

Keywords: rock-mass dataset; point-cloud registration; deep learning; benchmark

check for updates

Academic Editor: Giuseppe Casula

Received: 20 November 2024
Revised: 11 January 2025
Accepted: 13 January 2025
Published: 20 January 2025

Citation: Wang, Y.; Yu, D.; Liu, L.; Xiao, J. RockCloud-Align: A High-Precision Benchmark for Rock-Mass Point-Cloud Registration. *Remote Sens.* **2025**, *17*, 345. https://doi.org/10.3390/rs17020345

Copyright: © 2025 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

Rock-mass point-cloud registration is a fundamental yet challenging task in geology and engineering. It aims to align multiple independently acquired point clouds to a common coordinate system, which is crucial for applications such as rock-mass stability analysis, tunnel engineering, and mining operations. Recent advancements in deep-learning and probabilistic models have significantly enhanced the accuracy and robustness of these methods, making them more reliable for practical applications [1–4]. However, rock-mass point clouds present unique challenges due to irregular surface geometries, variable point densities, and complex geological structures.

Based on the underlying principles and methodologies, registration algorithms can be divided into traditional methods and deep-learning methods. Traditional point-cloud registration methods predominantly rely on handcrafted features and iterative algorithms. The Fast Point Feature Histogram (FPFH) [5], a widely used feature descriptor, plays a key role in identifying feature correspondences in noisy data. Typically paired with iterative algorithms such as the Iterative Closest Point (ICP) [6], FPFH facilitates the refinement of point-cloud alignment. However, this combination often suffers from computational inefficiencies and limited robustness in dynamic or complex environments. The Normal Distributions Transform (NDT) [7] is efficient for large-scale data due to its probabilistic approach, yet it requires careful tuning of parameters and may be less effective for highly

irregular surfaces. Probabilistic registration methods, such as Gaussian Mixture Models (GMM) [8] and Bayesian Coherent Point Drift (BCPD) [9], gained popularity due to their robustness to noise and outliers, though they can be computationally intensive and may require good initialization. Overall, these methods often struggle when applied to rock-mass scenarios due to their high computational complexity and reliance on empirically designed feature descriptors.

In recent years, deep-learning techniques have revolutionized point-cloud processing. Innovative architectures [10–12] have emerged that enable direct learning on unstructured point clouds, and significant progress has been made in improving registration efficiency and extracting high-quality features.

Specifically, rather than relying on iterative refinement, some methods adopt a process referred to as "single-step estimation" to compute the transformation matrix. It is important to clarify that this process consists of two distinct steps: (1) feature pair matching and (2) robust transformation estimation to compute the final registration parameters. The term "single-step" highlights the avoidance of iterative refinement seen in methods like ICP.

For instance, Ref. [13] employs AlexNet to extract 3D features from RGB-D data, while PPFNet [14] introduces a local Point Pair Feature (PPF) based on neighboring point distributions, which is subsequently fed into the network for deep feature learning. Additionally, Ref. [15] proposes a rotation-invariant handcrafted feature, which is processed by a neural network to learn robust correspondences. These approaches uniformly utilize deep learning as a feature extraction tool, leveraging advanced architectures or tailored loss functions to learn distinctive features for establishing accurate correspondences, avoiding the need for iterative optimization.

Many other methods focus on learning feature descriptors, with some also addressing keypoint detection. For example, refs. [16–18] primarily extract distinctive features from 3D points to identify accurate correspondences, which are then employed in a separate transformation estimation step. Works such as [19–22] emphasize both feature learning and keypoint detection, yet the later stages of the process, namely nearest neighbor matching and RANSAC for transformation estimation, remain mostly unchanged.

Recent approaches have begun to leverage transformer models [23–27], which excel at capturing long-range dependencies and contextual information, therefore achieving high-quality feature extraction. Additionally, several end-to-end registration frameworks have emerged, such as DeepVCP [28] and FCGF [29], which integrate correspondence refinement and feature learning into a single pipeline to improve registration accuracy.

Other methods focus on coarse-to-fine strategies to improve robustness and scalability, such as TEASER++ [30], which uses a robust solver for outlier rejection, and RPM-Net [31], which employs iterative refinement to align point clouds. PointDSC [32] further enhances outlier rejection by incorporating dynamic similarity-based correspondence clustering. These approaches demonstrate the growing diversity of techniques in the registration process, from robust transformation estimation to advanced feature learning, highlighting the need for tailored methods for specific applications, such as rock-mass registration.

Despite these advancements, the application of deep-learning methods to rock-mass point-cloud registration remains underexplored, largely due to the scarcity of large-scale, domain-specific datasets that reflect the diverse and unique characteristics of geological structures. RockCloud-Align aims to address this gap by providing a comprehensive dataset that captures a wide range of geological features. Unlike urban or indoor datasets, rock-mass point clouds offer a distinctive set of attributes:

- Irregular Geometries: Geological structures feature diverse and intricate shapes, such as fractures, folds, and faults, which serve as valuable test cases for robust algorithm development.

- Variable Densities: Point-cloud densities vary due to differences in scanning equipment, environmental conditions, and distances, offering an opportunity to evaluate adaptive processing techniques.
- Geological Complexity: Rock masses display unique and chaotic structures with overlapping textures, noise, and outliers, providing a benchmark for point-cloud registration methods.

By addressing these attributes, the RockCloud-Align dataset creates a robust platform for testing and refining registration algorithms. While existing methods demonstrate strong performance in some scenarios, this dataset extends its applicability by introducing geological-specific conditions that are often absent in traditional datasets. Instead of focusing solely on the challenges, RockCloud-Align highlights the opportunities to advance registration techniques and develop new methods tailored to geoscientific applications.

The scarcity of domain-specific datasets has significantly hindered progress in rock-mass registration. Existing point-cloud registration datasets can be broadly categorized into the following types:

(1) Simple Models and Indoor Scenes: Early datasets in this category primarily focused on providing 3D models and indoor reconstructions for various computer vision tasks, including registration. For instance, ModelNet [33] introduced a comprehensive collection of 3D CAD models across multiple object categories, while ShapeNet [34] expanded this concept with a large-scale repository of 3D models enriched with semantic annotations. For more registration-oriented tasks, datasets like 3DMatch [13] and ScanNet [35] provided RGB-D reconstructions of indoor scenes, specifically aimed at learning local geometric descriptors for registration purposes. These datasets offer a controlled environment with well-defined geometries, making them suitable for testing registration algorithms under simplified conditions. While these datasets have advanced the development of feature descriptors and registration techniques, they fail to capture the irregular geometries, unstructured surfaces, and noise characteristics inherent in outdoor geological environments. The lack of diversity in scene types limits their applicability to complex, real-world scenarios like rock-mass registration.

(2) Urban Outdoor Scenes: To address the limitations of indoor datasets, numerous efforts have been made to capture point clouds from complex urban environments. For instance, Semantic3D [36] introduced a large-scale benchmark for point-cloud classification, which has also been utilized for registration tasks. Similarly, the KITTI dataset [37], while primarily designed for autonomous driving applications, has become a standard benchmark for various 3D vision tasks, including point-cloud registration in urban and highway settings. Although these datasets incorporate real-world outdoor scenes, their focus remains on urban structures, such as roads, buildings, and vehicles, rather than geological features. As such, they lack the unstructured and irregular characteristics of rock-mass datasets, making them less relevant for tasks that involve natural terrain or geological complexities.

(3) Geological and Related Datasets: Although not explicitly designed for rock-mass registration, several datasets contain scenarios with relevant and similar scenes. For instance, the UNAVCO TLS Archive [38] provides terrestrial laser scanning (TLS) data of geological outcrops and rock surfaces, offering potential use for rock-mass studies. The WHU-TLS dataset [39] includes a "mountain" category featuring point clouds of rocky terrain, closely resembling some of the challenges encountered in rock-mass registration. Similarly, the NASA Planetary Pits and Caves Analog Dataset contains 3D point clouds of planetary-like rocky environments, making it highly relevant for tasks involving unstructured and complex rock scenes. Although these data sets provide valuable geological contexts, they exhibit several limitations that

restrict their applicability to advancing rock-mass registration techniques. These datasets are generally smaller in scale compared to standard registration benchmarks, making them less suitable for training large-scale machine learning models. In addition, their scenes are often captured using uniform equipment and under similar geological conditions, resulting in limited variability in point-cloud density, texture, and structural complexity.

While urban and indoor datasets have seen significant growth, datasets specifically tailored for geological applications, particularly rock-mass registration, remain limited. One notable effort to address this gap is the RockBench dataset [40], which established a common repository and standards for assessing rock-mass characteristics using LiDAR and photogrammetry. This dataset has served as a foundation for numerous studies in rock-mass point clouds. However, the dataset is currently no longer accessible, which is a significant loss for the research community.

These datasets highlight the diversity of point-cloud data and offer opportunities to explore registration techniques in geological contexts. However, they lack the scale and specificity of datasets such as our proposed RockCloud-Align, which focuses exclusively on rock-mass registration and includes diverse geological structures.

Despite these efforts, the availability of open-source rock-mass point-cloud registration datasets for unified evaluation remains scarce. The scarcity of large-scale, diverse rock-mass datasets highlights a significant opportunity for future work. Developing more comprehensive datasets that capture the unique challenges of rock-mass environments would greatly benefit the advancement of registration techniques in geological applications.

In this paper, we introduce RockCloud-Align (Figure 1), a large-scale, accurate, and open-source dataset for rock-mass point-cloud registration. This dataset covers diverse geological sites across different regions, encompassing a wide range of rock types, structural complexities, and environmental conditions, and further research was conducted on rock-mass registration methods based on this dataset. The main contributions of this article are as follows:

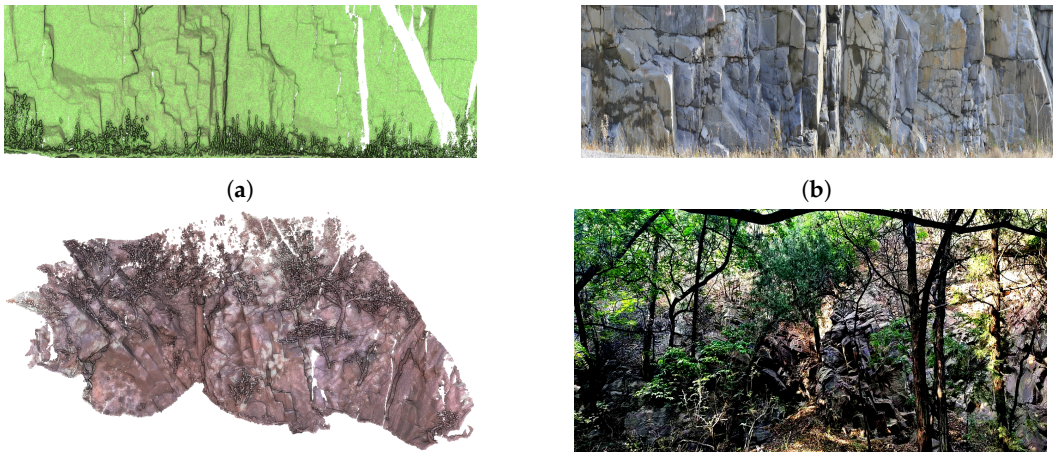


Figure 1. Examples of rock-mass point clouds from different regions used in our study. Subfigures (**a**,**b**) correspond to Ontario, Canada, while (**c**,**d**) correspond to Beijing, China.

1. We introduce RockCloud-Align, the first comprehensive benchmark that covers a wide range of geological scenarios for rock-mass point-cloud registration.
2. We propose a novel point-cloud registration method, which incorporates adaptive feature extraction, an overlap-aware transformer mechanism, and an end-to-end pipeline

to address challenges such as irregular geometries, varying densities, and sparse overlaps, ensuring high efficiency and accuracy across diverse geological scenarios.
3. We conduct a thorough evaluation of both conventional methods and several recent deep-learning approaches on our benchmark, establishing baseline results for future research.

2. Materials and Methods

2.1. Construction Principles of Rock Cloud-Align

Our goal is to provide a dedicated dataset focusing on rock-mass point-cloud registration. Several principles were considered in constructing this dataset to ensure its comprehensiveness, diversity, and utility.

2.1.1. Comprehensive Transformations

The *RockCloud-Align* dataset was collected using two high-precision laser scanning systems: the Leica P30 (Leica Geosystems, Heerbrugg, Switzerland) and the Leica HDS6000 (Leica Geosystems, Heerbrugg, Switzerland). These instruments offer complementary capabilities, enabling the collection of both large-scale and fine-detailed geological data.

The Leica P30 is renowned for its exceptional accuracy in scanning large-scale geological features. The specifications for the Leica P30 are as follows:

- Range Accuracy: 1.2 mm + 10 ppm over the full scanning range.
- Angular Accuracy: 8″ (both horizontal and vertical).
- 3D Position Accuracy: 3 mm at 50 m and 6 mm at 100 m.
- Target Acquisition Standard Deviation: 2 mm at 50 m.
- Effective Scanning Range: Up to 270 m at 34% reflectivity.

The Leica P30 is ideally suited for capturing expansive geological structures due to its ability to scan over long distances while maintaining a high level of accuracy.

The Leica HDS6000 is optimized for capturing detailed scans of mid-sized geological features, offering ultra-high-speed scanning capabilities. The key specifications for the Leica HDS6000 are:

- Scanning Speed: Up to 500,000 points per second.
- Modeled Surface Precision/Noise:
  - 2 mm at 25 m and 4 mm at 50 m (for surfaces with 90% albedo).
  - 3 mm at 25 m and 7 mm at 50 m (for surfaces with 18% albedo).
- Dual-Axis Sensor Resolution: 3.6″.
- Effective Scanning Range: Up to 79 m at 90% reflectivity.

The Leica HDS6000 is particularly effective for capturing intricate geological features such as rock fractures and smaller-scale rock formations, ensuring high precision even on surfaces with varying reflectivity.

By leveraging the capabilities of both the Leica P30 and Leica HDS6000, the *RockCloud-Align* dataset captures a diverse range of geological features, from expansive cliff faces to intricate rock fractures. The use of these two complementary systems ensures the robustness and diversity of the dataset, making it suitable for a wide range of research applications in rock-mass point-cloud registration.

While the variability in device configurations may introduce differences in point-cloud density and resolution, a uniform preprocessing pipeline was applied to normalize the data. This normalization minimizes inconsistencies and ensures that the dataset is reliable and suitable for evaluating point-cloud registration methods. We encourage users to consider these factors when interpreting results or applying the dataset to specific applications.

The RockCloud-Align dataset includes a wide range of rotations, translations, and varying degrees of overlap, ensuring thorough coverage of possible scenarios in rock-mass point-cloud registration. These diverse transformations allow for robust testing and evaluation of various registration methods, addressing the unique challenges posed by irregular geometries and sparse correspondences in geological data. Ground-truth preparation varied across datasets to ensure high accuracy:

Ground-truth transformations for the BJ dataset were derived using calibration points included during the data collection phase. For the QH and ON datasets, point clouds were initially coarsely registered and then refined using the Iterative Closest Point (ICP) algorithm. However, for datasets generated through mesh sampling, the ground-truth was indirectly derived. Specifically, a mesh model was first created from high-resolution images, and the point cloud was then sampled from this mesh. Consequently, the accuracy of the ground-truth in these datasets is primarily determined by the resolution and quality of the mesh model rather than the physical measurement precision of the scene. Additionally, for some datasets, ground-truth transformations were directly generated through the application of random rotations and translations. In this approach, point clouds were randomly rotated and translated within predefined ranges, with the applied transformations recorded as precise ground-truth values. This method ensures that the true transformations are both accurate and free from any external measurement or modeling limitations, providing a reliable reference for evaluating registration performance.

The accuracy of the ground-truth was validated by visual inspections and by monitoring the convergence criteria of the ICP algorithm. Across all datasets, the final alignment error was consistently within acceptable thresholds for registration tasks.

2.1.2. Diverse Geological Structures

Rock masses exhibit complex and varied structures, making their geological features less uniform compared to generic point clouds. Our dataset captures this diversity by including features such as fractures, folds, bedding planes, and joints. These elements are critical in representing the structural characteristics of rock masses. By incorporating more fractured and folded rock masses, we simulate conditions typically associated with complex geological environments. This diversity presents significant challenges for registration algorithms, therefore testing their effectiveness in real-world scenarios.

2.1.3. Multi-Scale Spatial Coverage

Rock-mass point clouds typically do not adhere to a fixed spatial scale, leading to a richer and more complex set of geometric information. Some geological structures span wide areas within the point clouds, while others cover only smaller regions. To ensure the preservation of geological features and to test the performance of registration methods under varying conditions, we chose not to downsample the point clouds to a uniform size, such as 100,000 points. This allows the dataset to reflect the true scale and complexity of geological structures, providing a more realistic and challenging environment for registration methods. Figure 2 illustrates this multi-scale nature of our dataset, showcasing samples with varying spatial dimensions and point densities.

By adhering to these principles, the RockCloud-Align dataset offers a comprehensive, diverse, and realistic platform for the development and evaluation of rock-mass point-cloud registration methods. The dataset's detailed annotations and rigorous construction process ensure it meets the high standards required for advanced research in this field.

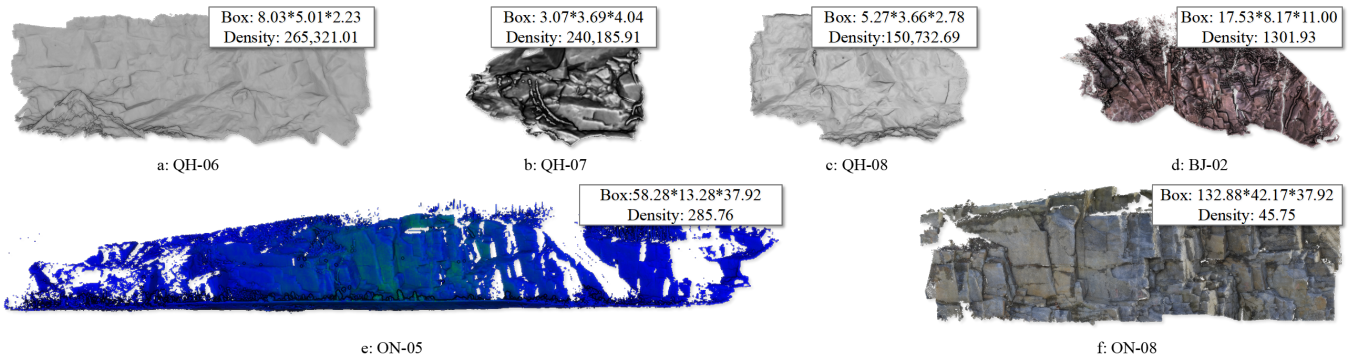


Figure 2. Representative samples from the RockCloud-Align dataset illustrating multi-scale spatial coverage and diverse geological structures. (**a**–**f**) show point clouds from different locations (QH: Qinghai, China; BJ: Beijing, China; ON: Ontario, Canada) with varying bounding box dimensions (in meters) and point densities (in points/m^3). The samples demonstrate significant variations in spatial scale, point density, and geological features, ranging from densely packed rock formations (**a**–**c**) to expansive cliff faces (**d**–**f**). This diversity in scale and structure presents unique challenges for point-cloud registration algorithms in geoscience applications.

2.2. Data Statistics

Our dataset is primarily acquired through terrestrial laser scanning, encompassing multiple scenes from Beijing, China; Qinghai, China; and Ontario, Canada. The terrestrial laser scanning process involved using high-precision equipment to capture detailed 3D representations of rock masses. The scanning parameters were carefully calibrated to ensure data accuracy and completeness, and preprocessing steps, such as noise reduction and outlier removal, were performed to enhance data quality.

Each scene includes between 1 million and over 10 million points, covering areas ranging from tens to hundreds of meters. After identifying target geological feature categories, such as fractures, folds, bedding planes, and joints, we selected representative regions, extracted consecutive registration pairs, and cropped sub-regions containing the target features from the original point clouds. Ultimately, we constructed a rock-mass point-cloud registration dataset consisting of 14,139 registration pairs.

Following the construction principles outlined in Section 2.1, we performed precise annotations for all registration pairs. These 14,139 registration pairs are divided into 11,383 training sets and 2756 testing sets, each with varying rotation angles and translation distances. Additionally, we accurately annotated points located in the overlapping regions to provide precise information for subsequent studies on overlap areas. The RockCloud-Align dataset consists of point clouds sampled to the same quantity, with an average of around 20,000 points per cloud. The smallest point cloud contains 10,000 points, while the largest contains 30,000 points. Figure 2 presents representative samples from our dataset, demonstrating the diversity in geological structures, spatial scales, and point densities across different locations.

To provide a comprehensive overview of our dataset, we present detailed statistics in Table 1. This table includes information on the average number of points, total point-cloud count, overlap range, mean overlap, and overlap standard deviation for each dataset. Additionally, the distribution of the dataset into training and testing sets is provided for each scene location. The dataset IDs are categorized by their respective scenes, specifying whether they belong to the training or testing set along with the corresponding set size.

Table 1. Detailed Statistics for the RockCloud-Align dataset.

Location	ID	Total Count	Mean Overlap (%)	Overlap SD (%)	Set Size	Type	Generation
Beijing, China	BJ-01	166,639	43.79	16.09	7	train	Leica P30 System
	BJ-02	2,051,090	43.19	16.62	107	train	
Ontario, Canada	ON-01	8,005,601	42.57	16.00	390	train	Leica HDS6000
	ON-02	7,131,878	42.62	16.13	350	train	
	ON-03	8,245,452	43.89	16.12	390	train	
	ON-04	8,151,500	43.71	16.77	407	train	
	ON-05	8,386,620	42.01	16.01	401	train	
	ON-06	6,517,786	42.41	16.90	320	train	
	ON-07	8,765,102	42.30	15.80	420	test	
	ON-08	9,721,227	42.18	15.72	468	test	
Qinghai, China	QH-01	16,724,982	44.27	15.55	807	train	MeshSampling
	QH-02	11,558,045	43.47	15.93	565	train	
	QH-03	17,780,166	44.11	15.97	863	train	
	QH-04	14,083,280	43.57	15.94	672	train	
	QH-05	16,748,272	42.88	15.72	818	train	
	QH-06	23,802,895	42.73	15.89	1166	train	
	QH-07	10,992,428	42.46	15.93	544	train	
	QH-08	8,082,467	42.62	16.15	389	train	
	QH-09	16,801,435	42.69	16.14	817	train	
	QH-10	11,527,740	42.74	16.15	554	train	
	QH-11	8,424,664	41.46	15.89	416	train	
	QH-12	15,238,810	43.37	15.59	731	train	
	QH-13	13,553,131	41.76	15.50	669	train	
	QH-14	10,900,741	41.95	16.18	538	test	
	QH-15	15,850,606	43.22	15.82	783	test	
	QH-16	11,143,133	43.26	15.62	547	test	

2.3. Comparison with Other Existing Datasets

RockCloud-Align exhibits several distinctive characteristics that differentiate it from other point-cloud datasets:

1. Existing datasets are typically collected using standardized equipment in similar scenes, resulting in data with consistent density distributions. In contrast, RockCloud-Align data are often captured under varying conditions and with non-uniform equipment, leading to heterogeneous density distributions and more complex data structures.
2. Unlike indoor or urban environments, which feature predictable geometric structures such as walls and buildings, RockCloud-Align's rock-mass data lacks such regularity. This irregularity introduces additional complexity in the processing and analysis of rock-mass point clouds.
3. The inherent variability of rock-mass features, such as fractures and voids, poses significant challenges for achieving precise and reliable registration, further distinguishing RockCloud-Align from other datasets.

To achieve a more robust rock-mass point-cloud registration method, RockCloud-Align does not rely on a single acquisition device or similar scenes, ensuring a diverse and comprehensive representation of rock-mass characteristics. This approach contrasts with the limitations observed in traditional datasets mentioned above.

2.4. Methodology

In this section, we present our proposed transformer-based framework for rock-mass point-cloud registration. The overall network architecture is shown in Figure 3. Our method

comprises three main steps. First, we employ an adaptive-density feature extraction module to enhance feature generalization. Second, we utilize a transformer-based feature encoding process to effectively capture and represent the extracted features. Third, we implement an overlap-aware positional decoding mechanism to accurately determine the positions in the overlapping regions.

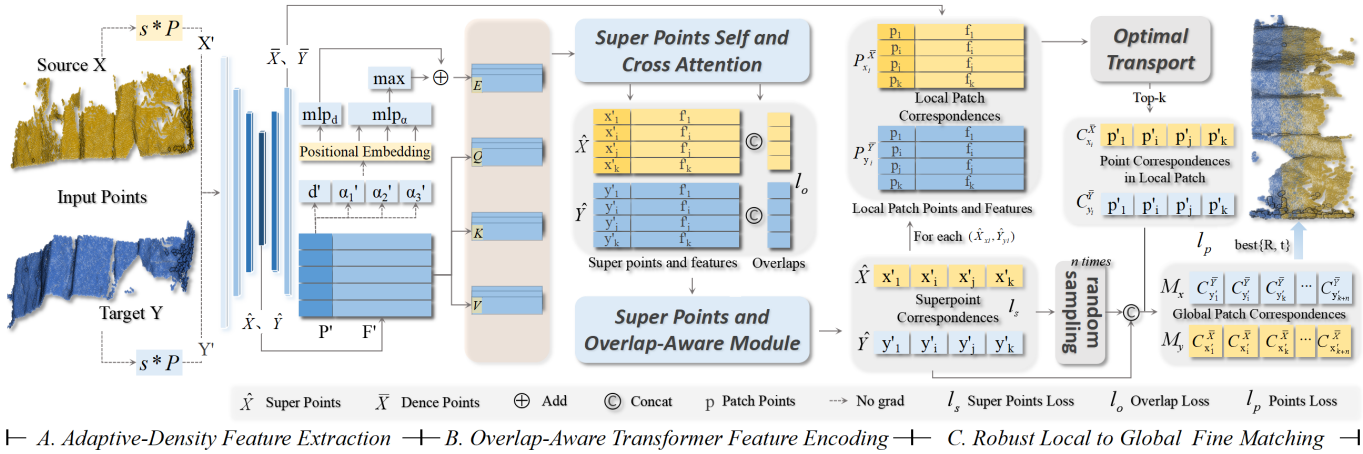


Figure 3. Illustrates our proposed point-cloud registration framework. The framework consists of three key stages: (**A**). Adaptive-Density Feature Extraction, which preprocesses and samples the input point clouds via adaptive sampling and multi-level feature encoding; (**B**). Iteratively encode intra-cloud geometric structures and inter-cloud consistency, generating high-quality superpoint correspondences; and (**C**). Propagates superpoint correspondences to dense points, combining local patch features with global overlap information. The framework is optimized using multiple loss functions l_s, l_o, l_p to address superpoint matching, overlap handling, and dense point alignment.

2.4.1. Adaptive-Density Feature Extraction

Adaptive-density feature extraction (Figure 3A) is a crucial step in our framework to address the varying density distributions present in rock-mass point-cloud data. The need for adaptive density arises because traditional point-cloud data, when collected using similar equipment in similar scenes, often exhibits uniform density. However, our RockCloud-Align Dataset, collected from different scenes, shows significant variations in density and scale. If not properly managed, these variations may prevent the network from consistently learning point-cloud features.

To mitigate this issue, we preprocess the raw and target point clouds by scaling them in all directions to ensure that the point density within a unit cubic meter remains consistent. Inspired by KPConv [12], which introduces convolution operations over irregular point clouds, we propose a density-aware preprocessing strategy that dynamically adjusts the scale of point clouds based on their local density. This preprocessing step helps maintain a uniform scale during the sampling and learning processes, therefore facilitating the extraction of consistent features across different densities. The scaling factor s is computed as follows:

$$s = \left(\frac{N}{V \times \rho_t} \right)^{\frac{1}{3}}, \tag{1}$$

where N is the number of points in the point cloud, V is the volume of the axis-aligned bounding box containing the point cloud, and ρ_t is the target density per unit cubic meter.

The transformation of the coordinates (x, y, z) after this scaling process is given by:

$$(x', y', z') = (s \times x, s \times y, s \times z). \tag{2}$$

After applying the adaptive-density transformation, we perform random sampling on the point clouds to obtain a set of super points and their corresponding features. In this process, we adopt the Kernel Point Convolution (KPConv) [12] backbone for feature extraction. The KPConv backbone employs a series of ResNet-like blocks and strided convolutions to transform each input point cloud into a reduced set of superpoints $\hat{X} \in \mathbb{R}^{M_0 \times 3}$, $\hat{Y} \in \mathbb{R}^{N_0 \times 3}$ and their associated features $F_{\hat{X}} \in \mathbb{R}^{M_0 \times D}$, $F_{\hat{Y}} \in \mathbb{R}^{N_0 \times D}$ and a dense-points set $\bar{X} \in \mathbb{R}^{M \times 3}$, $\bar{Y} \in \mathbb{R}^{N \times 3}$ and their associated features $F_{\bar{X}} \in \mathbb{R}^{M \times D}$, $F_{\bar{Y}} \in \mathbb{R}^{N \times D}$.

KPConv introduces convolutional operations over irregular point clouds by defining kernel points. Our method extends this concept by incorporating a density-aware sampling strategy to handle the highly variable densities of geological point clouds. Specifically, we introduce a scaling factor that dynamically adjusts the kernel size based on local point densities, ensuring consistent feature extraction across regions with varying densities. This modification significantly enhances performance in datasets with non-uniform density distributions.

This adaptive-density approach ensures that the network can learn robust features despite variations in point-cloud density, ultimately leading to improved performance in rock-mass point-cloud registration.

2.4.2. Overlap-Aware Transformer Feature Encoding

While capturing global context is essential for point-cloud registration, it is equally important to focus on the specific regions where the source and target point clouds overlap. These overlapping regions contain the true correspondences, making them critical for accurate registration. Inspired by transformer-based models [25], we incorporate an overlap-aware attention mechanism to prioritize these regions. Existing methods, however, often treat all parts of the point clouds with equal importance, neglecting the unique significance of overlapping regions. This oversight exacerbates matching ambiguities, as non-overlapping areas introduce noise and irrelevant information, further complicating the registration process. By explicitly emphasizing overlapping regions, our method reduces noise interference and enhances alignment accuracy, particularly in challenging scenarios with sparse overlaps.

To address these challenges, we propose an *Overlap-Aware Transformer Feature Encoding* (Figure 4) mechanism that emphasizes the overlapping regions between source and target point clouds during feature learning. By assigning higher weights to features within these critical regions, our method enhances the discriminative power and robustness of the learned representations, facilitating more precise and reliable point-cloud registration.

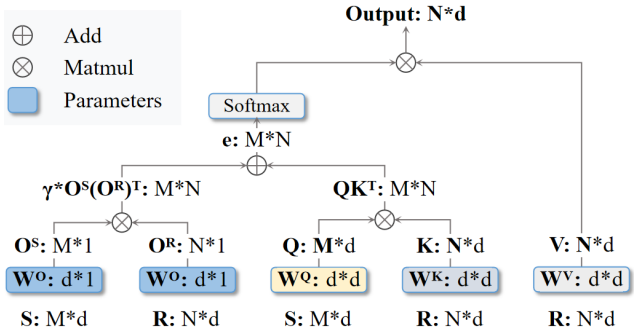


Figure 4. The structure of overlap-aware attention module. The use of the same color indicates that parameters are shared.

Given a point cloud $\mathcal{P} = \{p_i\}_{i=1}^{N}$ with associated feature descriptors $F = \{\mathbf{f}_i\}_{i=1}^{N}$, we first process the features using GeoTransformer. The GeoTransformer extracts global contextual and geometric features, which are then divided into two parts: one part is passed

through a multilayer perceptron (MLP) to estimate the overlap score o_i for each point, indicating the probability of point p_i belonging to an overlapping region. The overlap score is computed as follows:

$$o_i = \sigma(\mathrm{MLP}(\mathbf{f}_i)), \tag{3}$$

where $\sigma(\cdot)$ denotes the sigmoid activation function.

The other part of the features extracted by GeoTransformer is concatenated with the overlap scores to form the final superpoint feature representation:

$$\mathbf{f}_s = \mathrm{concat}(\mathbf{f}_i, o_i). \tag{4}$$

Finally, the cross-attention mechanism is adapted to integrate these overlap scores, ensuring that the attention computation is enhanced by focusing on points within close proximity and with high overlap probability:

$$\mathbf{z}_i = \sum_{j=1}^{N} \alpha_{ij} \mathbf{W}_V \mathbf{f}_j, \tag{5}$$

with attention weights α_{ij} defined as:

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{N} \exp(e_{ik})}, \tag{6}$$

and attention scores e_{ij} computed by:

$$e_{ij} = \frac{(\mathbf{W}_Q \mathbf{f}_i)^\top (\mathbf{W}_K \mathbf{f}_j)}{\sqrt{d}} + \gamma \cdot (o_i \cdot o_j), \tag{7}$$

The weighting factor γ balances feature and overlap similarities.

GeoTransformer excels at capturing long-range dependencies in urban or synthetic datasets. However, its general-purpose attention mechanism does not prioritize overlapping regions critical for rock-mass registration. To address this, our overlap-aware attention mechanism introduces an overlap probability score, which is integrated into the attention computation. This innovation enhances feature alignment in regions with significant overlap while reducing the influence of non-overlapping areas.

This approach ensures that the attention mechanism prioritizes geometrically relevant and contextually significant regions, particularly the overlapping areas, leading to more accurate and reliable point-cloud registration. By focusing on these critical regions, our method enhances the discriminative power of the learned representations, resulting in better performance even in challenging scenarios with low overlap between point clouds.

2.4.3. Robust Local to Global Fine Matching

After obtaining the features of all superpoints, the correspondence between superpoints can be obtained based on the similarity of the features, represented as $S = \bigcup_{i=1}^{N_s} \{(\hat{X}_{x_i}, \hat{Y}_{y_i})\}$, where N_s is the number of valid matching pairs.

We extend the above point correspondence to the local area by assigning neighborhood points to each pair of superpoints $(\hat{X}_{x_i}, \hat{Y}_{y_i})$, which involves two steps: (1) For each superpoint S_i in the source point cloud (target point cloud), calculate its K-nearest neighbors in $\bar{X}$ ($\bar{Y}$); (2) For each point in the K-nearest neighbor, check if S_i is the nearest superpoint to it. The above steps will obtain a valid set of dense points for each pair of superpoints, denoted as $P_i = (P_{x_i}^{\bar{X}}, P_{y_i}^{\bar{Y}})$.

Similar to GeoTransformer, we construct a cost matrix C_i for each pair of patches based on point features. C_i contains the matching cost of any two points between two patches. Subsequently, we utilize the Sinkhorn algorithm [41] to calculate the final confidence matrix $\bar{C}_i$, which can be expressed as:

$$\bar{C}_i = sinkhorn(C_i). \tag{8}$$

On the basis of this confidence matrix, the dense point set with effective correspondence between local patches can be obtained using mutual top, represented as $M(x_i, y_i) = (C^{\bar{X}}_{x_i}, C^{\bar{Y}}_{y_i})$. Then, all dense corresponding point pairs contained in the point cloud $(\bar{X}, \bar{Y})$ can be represented as: $M_{N_s} = \bigcup_{i=1}^{N_s} M_{(x_i,y_i)}$.

Some existing works utilize the precise correspondence between each pair of local patches $(C^{\bar{X}}_{x_i}, C^{\bar{Y}}_{y_i})$ to calculate their respective transformation $T_i = R_i, t_i$, which can be calculated by:

$$R_i, t_i = \min_{R,t} \sum_{\left(p_{x_j}, p_{y_j}\right) \in M_{(x_i,y_i)}} \omega^i_j \left\| R \cdot p_{xj} + t - p_{yj} \right\|^2_2 \tag{9}$$

However, the above process only uses local information to achieve registration, so the performance of registration may be restricted by factors such as the density and completeness of local data. In order to improve the robustness of the method, in this paper, we merge multiple sets of dense point sets with corresponding relationships, therefore expanding the matching information from local to global so that the method can search for the optimal registration result based on both local and global information.

Specifically, we use random sampling to perform n times random samplings from the set S and merge the corresponding dense point sets in M_{N_s} according to the superpoint sampling results. Thus, n groups of global dense point sets with corresponding relationships are generated, denoted as M_n.

Finally, we expand M_{N_s} to $(M_{N_s} + M_n)$ and use Equation (9) to calculate the transformation for each pair of point sets, the obtained $N + n$ groups of transformation parameters can be used to perform rotation and translation transformation. The transformation parameters are evaluated according to the number of matched points between $R_i, t_i(\bigcup_{i=1}^{N_s} M_{x_i})$ and $\bigcup_{i=1}^{N_s} M_{y_i}$, so as to obtain the optimal matching parameters. The final result is obtained by iteratively optimizing the above best parameter using Equation (9).

2.4.4. Loss Functions

We train our network end-to-end with the ground-truth poses $\{R^*, t^*\}$ for supervision, utilizing two types of loss functions: an overlap classification loss and a point-matching loss, both adapted from existing works.

Overlap Loss

The overlap loss is framed as a binary classification task, where the predicted overlap scores are compared against ground-truth using binary cross-entropy. Specifically, for a given point cloud X, the loss is computed as:

$$\mathcal{L}^X_o = -\frac{1}{M_0} \sum_{i}^{M_0} \Big[o^*_{\tilde{x}_i} \log \hat{o}_{\tilde{x}_i} + (1 - o^*_{\tilde{x}_i}) \log(1 - \hat{o}_{\tilde{x}_i}) \Big], \tag{10}$$

where $o^*_{\tilde{x}_i}$ represents the ground-truth overlap score and $\hat{o}_{\tilde{x}_i}$ is the predicted score for each point $\tilde{x}_i$ in X. The same formulation applies to the target point cloud Y. This binary cross-entropy loss ensures that the network learns to predict accurate overlap scores between source and target point clouds.

Overlap-Aware Circle Loss

We adopt an overlap-aware circle loss, which emphasizes correspondences with higher overlap between patches. This approach adapts the circle loss from prior works [19,22,25], weighting patch pairs based on their overlap ratio to improve the matching of patches with both high and low overlap. This loss encourages the model to prioritize learning correspondences between superpoints with greater structural similarity.

Point-Matching Loss

Point matching is inherently sparse due to downsampling, and the task requires finding correspondences between a subset of points in two downsampled point clouds. We employ a point-matching loss inspired by previous work [25,42], which applies a negative log-likelihood loss on the assignment matrix $\bar{Z}_i$ for each superpoint correspondence.

For each ground-truth superpoint correspondence $\hat{C}_i^*$, the ground-truth point correspondences M_i within a matching radius τ are used to compute the loss as follows:

$$L_{p,i} = -\sum_{(x,y)\in M_i} \log \bar{z}_{ixy} - \sum_{x\in I_i} \log \bar{z}_{ix(m_i+1)} - \sum_{y\in J_i} \log \bar{z}_{i(n_i+1)y}, \tag{11}$$

where I_i and J_i are the sets of unmatched points in the source and target patches, respectively. This loss encourages the network to maximize the likelihood of correct point correspondences and penalizes unmatched points. The final point-matching loss is averaged over all sampled superpoint correspondences.

Overall Loss

The final loss function for training the network is a weighted combination of the overlap classification loss and the point-matching loss, ensuring that both the global overlap and local point correspondences are accurately predicted. This combination leads to more robust and reliable superpoint and point-level matching.

3. Results

To rigorously evaluate the efficacy of the proposed point-cloud registration method, we conducted comprehensive experiments using the RockCloud-Align dataset. Our experiments are designed to assess both the quantitative and qualitative performance of our approach in comparison to state-of-the-art techniques. This section details the experimental setup, presents the results, and discusses the findings.

3.1. Experimental Setup and Evaluation Metrics

The experiments were conducted on a range of challenging scenarios within the RockCloud-Align dataset, which encompasses a diverse set of geological features and varying point-cloud densities, providing a robust benchmark for registration performance. Each experiment was carefully designed to test the method's robustness across different conditions.

To quantitatively assess the registration performance, we employed three primary evaluation metrics, defined as follows:

- Registration Recall (RR): The proportion of correctly registered point pairs within a predefined threshold L, calculated as:

$$\mathrm{RR} = \frac{1}{M}\sum_{i=1}^{M}[\mathrm{RMSE}(T(X_i), Y_i) < L], \tag{12}$$

where RMSE is the Root Mean Square Error between the transformed points $T(X_i)$ and their corresponding points Y_i.

- Relative Rotation Error (RRE): The angular error between the estimated rotation matrix R_{est} and the ground-truth rotation matrix R_{gt}, given by:

$$\text{RRE} = \arccos\left(\frac{\text{trace}(R_{\text{gt}}^T R_{\text{est}}) - 1}{2}\right). \tag{13}$$

- Relative Translation Error (RTE): The Euclidean distance between the estimated translation vector t_{est} and the ground-truth translation vector t_{gt}, defined as:

$$\text{RTE} = \|t_{\text{est}} - t_{\text{gt}}\|_2. \tag{14}$$

3.2. Results and Analysis

Table 2 presents a detailed comparison of our proposed method against several state-of-the-art algorithms on the RockCloud-Align dataset. Our method consistently outperforms the others across nearly all metrics, underscoring its robustness and accuracy in point-cloud registration tasks.

Table 2. Comparison of different methods on the RockCloud-Align dataset.

Method	RR@0.20	RR@0.15	RR@0.10	RR@0.05	RR@0.03	RR@0.02	RRE (∘)	RTE (m)
3DSN [15]	86.46	82.17	77.33	69.89	64.75	28.99	2.199	0.141
FCGF [29]	91.62	88.35	83.54	75.23	69.43	23.88	1.949	0.180
D3Feat [19]	89.18	85.77	80.46	72.34	67.54	31.27	3.161	0.132
Predator [22]	99.46	92.73	89.25	80.86	74.97	38.46	0.982	0.096
OMNet [43]	75.94	71.48	67.83	59.23	54.38	29.74	5.713	1.105
DGR [20]	99.76	98.29	94.95	86.15	80.85	75.43	3.702	0.067
PCAM [27]	98.18	97.44	96.38	93.02	89.76	71.63	1.327	0.134
GeoTransformer [25]	99.89	**99.86**	98.67	**98.80**	96.70	91.45	0.852	0.032
Regtr [23]	99.82	99.64	99.41	97.90	94.37	86.02	0.742	0.042
Ours	**99.96**	**99.86**	**99.82**	**98.80**	**96.84**	**91.77**	**0.618**	**0.020**

Note: Values in bold indicate the best performance. This convention is applied consistently across all subsequent tables.

The Registration Recall (RR) at various thresholds demonstrates the effectiveness of our approach. At a generous threshold of 0.20 m, our method achieves an outstanding RR of 99.96%, slightly outperforming GeoTransformer at 99.89% and Regtr at 99.82%.

The Relative Rotation Error (RRE) is another critical metric where our method excels, achieving the lowest RRE of 0.630 degrees, indicating a highly precise estimation of the rotation matrix. This is significantly better than the second-best Regtr, which has an RRE of 0.742 degrees. In contrast, methods such as OMNet and DGR exhibit much higher RRE values of 5.713 and 3.702 degrees, respectively, reflecting their limitations in handling complex rotational transformations in geological data.

Similarly, in terms of Relative Translation Error (RTE), our method achieves the lowest error at 0.020 m, highlighting its accuracy in estimating the translation vector. This is a substantial improvement over methods like FCGF and OMNet, which show higher RTEs of 0.180 m and 1.105 m, respectively. Even the high-performing GeoTransformer and Regtr methods exhibit slightly higher RTEs of 0.032 m and 0.042 m, respectively, indicating that our method's ability to estimate translations is particularly strong. Figure 5 visually demonstrates the superiority of our approach compared to other methods.

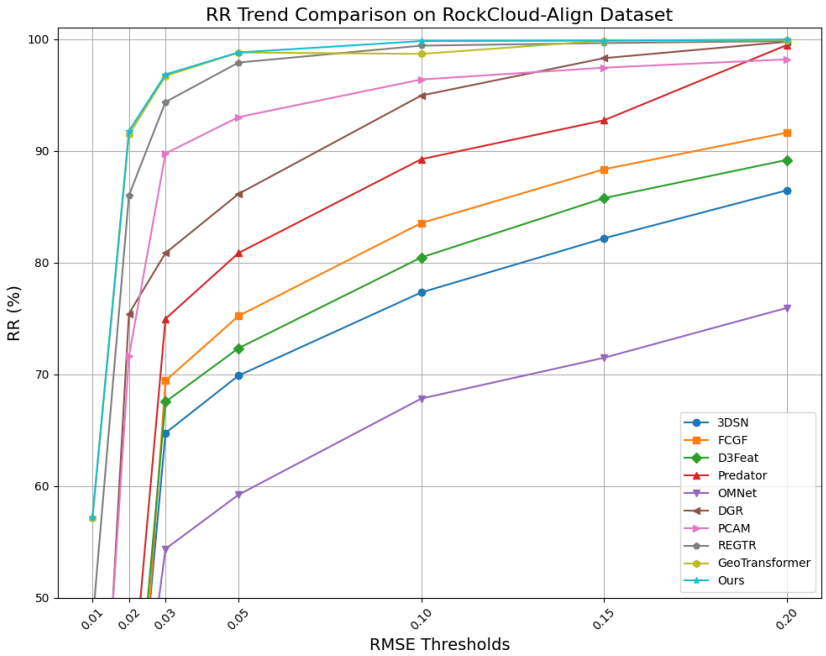


Figure 5. Comparison of Registration Recall (RR) across different thresholds on the RockCloud-Align dataset.

Overall, the results clearly demonstrate that our proposed method not only leads in most metrics but also maintains high performance across increasingly stringent thresholds, showcasing its robustness and reliability in complex registration tasks within the RockCloud-Align dataset. The combination of low RRE and RTE values further supports the method's superior accuracy, making it a powerful tool for point-cloud registration in diverse and challenging geological settings.

3.3. Qualitative Results

Figure 6 provides a visual comparison of the registration results across different methods, highlighting the effectiveness of our approach relative to two other leading algorithms: GeoTransformer and Regtr. Each row in the figure represents a distinct type of input data chosen to illustrate different challenges in point-cloud registration.

The first row showcases a relatively well-structured point cloud, where our method achieves superior alignment, closely matching the ground-truth. The results from GeoTransformer and Regtr, while still accurate, display minor misalignments, particularly in the finer details, which are more precisely captured by our method.

The second row features a point cloud with highly dispersed data points, presenting a more challenging scenario for registration algorithms. Despite the sparse distribution, our method maintains a high degree of alignment accuracy, significantly outperforming both GeoTransformer and Regtr, which struggle with dispersion, resulting in visible misalignments.

In the third row, the input data includes a clear fault line, representing a discontinuity that complicates the registration process. This scenario is particularly challenging because such geological features can disrupt the spatial continuity required for accurate registration. Our method effectively handles this discontinuity, preserving the structural integrity of the point-cloud alignment, whereas both GeoTransformer and Regtr exhibit noticeable errors in aligning the faulted regions.

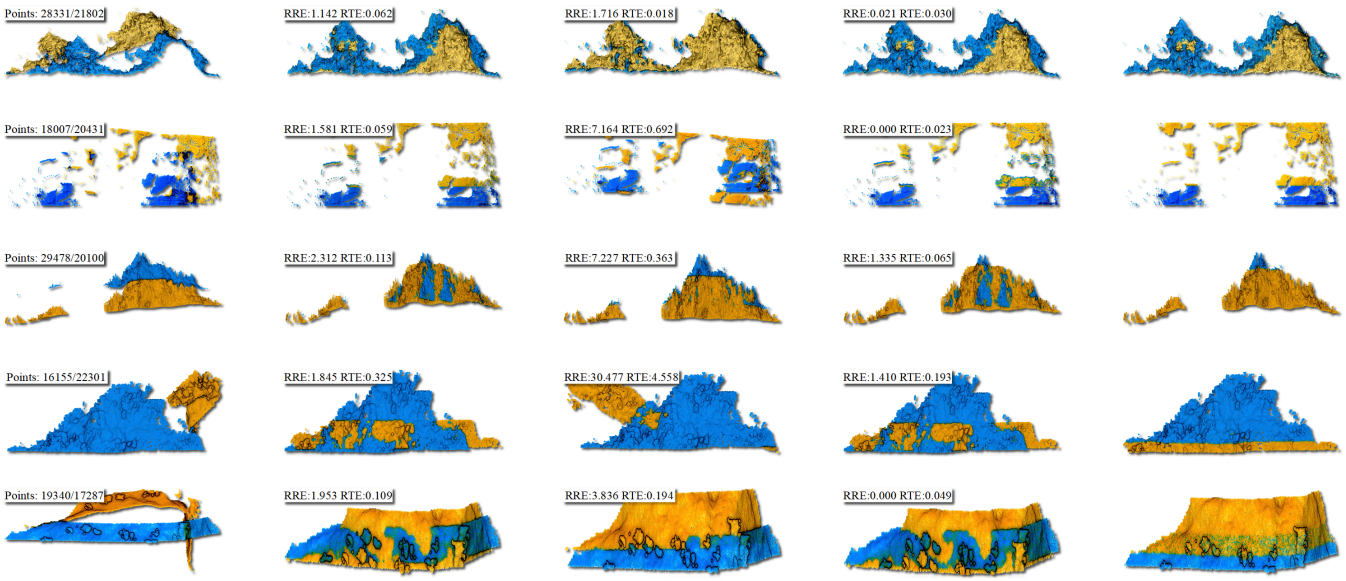


Figure 6. Visual comparison of point-cloud registration results across different methods. From left to right: (**a**) Input point clouds, (**b**) GeoTransformer, (**c**) Regtr, (**d**) Our proposed method, and (**e**) Ground-truth. Each row represents a distinct, challenging scenario, showcasing the performance of various algorithms in handling diverse point-cloud registration tasks. The blue color represents the target point cloud, while the yellow color represents the source point cloud.

Additionally, in each image, the top-right corner displays key metrics, such as the predicted loss for the method on that particular dataset. These metrics further quantify the visual observations, with our method consistently showing lower loss values, indicative of its superior performance. The alignment accuracy, combined with lower prediction losses, underscores the robustness of our approach across diverse geological scenarios.

Overall, Figure 6 not only illustrates the visual superiority of our method but also provides empirical evidence of its effectiveness in handling various types of point-cloud data, ranging from well-structured to highly dispersed and faulted formations. These qualitative results, coupled with the quantitative analysis presented earlier, firmly establish the reliability and precision of our proposed method in complex registration tasks.

3.4. Time Performance Analysis

The time efficiency of the proposed method is a critical factor in evaluating its practical applicability. One of the key design choices in our method is that the cross-attention mechanism in the feature extraction process operates on sampled points rather than the entire point cloud. This significantly reduces computational complexity, enabling more efficient processing without compromising performance.

To evaluate the time performance of our method, we conducted experiments on the 2756 test samples from the RockCloud-Align dataset. The average total processing time for all test samples was approximately 12 min, which translates to a mean processing time of 0.26 s per sample. These evaluations were performed on a single NVIDIA RTX 4090 GPU. It is worth noting that the processing time may vary depending on the computational resources available, such as the type of GPU or CPU used.

Overall, the results demonstrate that our method achieves high computational efficiency while maintaining robust registration performance. This balance between accuracy and efficiency makes our approach particularly suitable for large-scale geological datasets and real-world applications requiring timely processing.

3.5. Ablation Studies

To thoroughly understand the contributions of each component within our proposed method, we conducted a series of ablation studies. The results, as shown in Table 3, reveal how the performance of our method is affected when certain components are removed or modified.

Table 3. Ablation study results on the RockCloud-Align dataset.

Method	RR@0.10	RR@0.02	RRE (∘)	RTE (m)
No Adaptive Density	-	-	5.296	1.272
Without Overlap Loss	98.67	91.45	0.712	0.034
Our Baseline	98.79	91.62	0.640	0.030
Ours	**99.82**	**91.77**	**0.618**	**0.020**

3.5.1. Adaptive-Density Feature Extraction

The importance of the adaptive-density feature extraction module is evident when we observe its removal. Without this module, we were unable to compute a reliable Registration Recall (RR), and there was a dramatic increase in the Relative Rotation Error (RRE) to 5.296 degrees and the Relative Translation Error (RTE) to 1.272 m. This highlights the module's critical role in managing the variable densities in rock-mass point clouds, which is essential for maintaining accuracy in both rotational and translational estimations.

3.5.2. Overlap Loss

The exclusion of the overlap loss, which is designed to promote better alignment between corresponding regions in point clouds, results in a notable drop in performance. As shown in Table 3, removing this loss decreases RR@0.10 and RR@0.02 to 98.67% and 91.45%, respectively. Additionally, the RRE increases to 0.712 degrees and the RTE to 0.034 m. This demonstrates that the overlap loss is pivotal for enhancing the model's ability to align overlapping areas, contributing to more accurate pose estimation.

3.5.3. Our Baseline

The baseline model incorporates the overlap prediction, feature association, and adaptive-density modules. However, in overlap probability prediction, the model employs a direct regression approach rather than using the Overlap-Aware Transformer. Specifically, during the overlap probability estimation, the baseline model directly regresses the values without utilizing a sampling-based approach. As shown in Table 3, the baseline achieves an RR@0.10 of 98.79%, an RR@0.02 of 91.62%, an RRE of 0.640 degrees, and an RTE of 0.030 m. Although the performance is competitive, the higher RRE and RTE compared to the complete model highlight the advantage of incorporating the Overlap-Aware Transformer and sampling techniques for more precise rotation and translation estimations.

3.5.4. Complete Model (Ours)

Our complete model, incorporating all components, including Section 2.4.1 and the Section 2.4.2, achieves the highest performance across all metrics. Specifically, it attains an RR@0.10 of 99.82%, an RR@0.02 of 91.77%, an RRE of 0.618 degrees, and an RTE of 0.020 m. These results confirm that the integration of adaptive-density feature extraction and the overlap loss yields the best overall accuracy, achieving robust performance in both rotation and translation estimations.

4. Discussion

The results obtained from the RockCloud-Align dataset demonstrate a significant advancement in the registration of rock-mass point clouds, particularly in handling the complexities associated with geological structures. Our method consistently outperformed traditional approaches, as demonstrated by the lower prediction losses and enhanced alignment accuracy. This aligns with previous studies that emphasize the necessity for robust registration techniques capable of accommodating the intricate geometries and variable densities inherent in geological formations. The empirical evidence presented not only corroborates our working hypothesis regarding the efficacy of adaptive-density feature extraction but also highlights the critical role of multi-scale approaches in improving registration performance.

A key strength of the proposed method lies in its ability to address two specific challenges inherent to the RockCloud-Align dataset: variability in device configurations and precise overlap annotations. Unlike datasets such as WHU-TLS, which assume consistent scanner conditions, RockCloud-Align introduces variability due to the use of multiple laser scanners. Our method mitigates these differences through adaptive-density-aware feature extraction, ensuring robustness across varying point-cloud densities. Furthermore, the overlap-aware transformer framework relies on accurate overlap annotations, a feature unique to RockCloud-Align. Since such annotations are not available in existing datasets, the proposed method cannot be directly applied without modification, which limits cross-dataset evaluation. This highlights the dataset's unique contribution to enabling more precise and reliable registration in challenging geological contexts.

Moreover, the findings suggest that the integration of learning-based methods could further enhance the registration process by leveraging the unique characteristics of geological structures. This is consistent with the growing body of literature advocating for the application of machine learning techniques in geospatial data analysis. The implications of our research extend beyond mere methodological advancements; they pave the way for future studies to explore long-term registration challenges, such as tracking geological changes over time, which is crucial for applications in landslide monitoring and structural geology analysis.

In conclusion, while our study provides a solid foundation for future research, it also opens avenues for further exploration into multi-temporal registration components and the development of more sophisticated algorithms. By focusing on the dynamic nature of geological environments, future work can extend the contributions presented here to address long-term challenges in rock-mass monitoring and geospatial analysis.

5. Conclusions

We present RockCloud-Align, the first rock-mass point-cloud registration benchmark with high-quality annotated transformations. We also benchmarked 15 representative registration algorithms on RockCloud-Align and conducted an extensive evaluation and performance analysis of the algorithms according to the point-cloud pair's attributes and geological features. The evaluation and visualization results show that there is still much room for improvement in rock-mass point-cloud registration.

The main contributions of this work can be summarized as follows:

1. We introduce RockCloud-Align, a diverse and challenging dataset for rock-mass point-cloud registration, consisting of 14,139 point-cloud pairs with various geological features and challenging attributes.
2. We provide a comprehensive evaluation of 9 state-of-the-art registration algorithms on RockCloud-Align, offering insights into their strengths and weaknesses in the context of rock-mass point-cloud registration.

3. We conduct detailed analyses based on different attributes and geological features, identifying the most challenging aspects of rock-mass point-cloud registration.
4. We present visualizations of registration results, highlighting specific challenges and potential areas for improvement in registration algorithms.

In our future work, we aim to include more point-cloud pairs and registration algorithms to further extend the benchmark dataset. We also plan to develop more effective registration methods targeted at the difficulties and challenges identified by RockCloud-Align. Some potential research directions include:

1. Developing robust registration methods that can handle fine-scale features and complex geometric structures characteristic of rock masses.
2. Exploring multi-scale and hierarchical registration approaches to better handle the wide range of feature scales present in rock-mass point clouds.
3. Investigating learning-based methods that can effectively leverage the unique characteristics of geological structures for improved registration performance.
4. Addressing the challenges of long-term registration in rock masses, including handling significant deformations and occlusions over time.

Moreover, we will add a multi-temporal registration component to the RockCloud-Align dataset for further study of long-term changes in rock masses. This extension will enable research into tracking geological changes over time, which is crucial for applications such as landslide monitoring and structural geology analysis.

By providing this benchmark and analysis, we hope to stimulate further research in the field of rock-mass point-cloud registration and contribute to the development of more accurate and robust registration methods for geological applications.

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

Funding: This work is supported by the National Natural Science Foundation of China (U21A20515, 62476262, 62102393, 62206263, 62271467, 62306297, 62306296), Beijing Natural Science Foundation (4242053, L242096), China Postdoctoral Science Foundation (2022T150639) and the Fundamental Research Funds for the Central Universities.

Data Availability Statement: The data presented in this paper are openly available at https://github.com/RockAilab/RockAlign, accessed on 14 September 2025.

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

References

1. Huang, X.; Mei, G.; Zhang, J. Cross-source point cloud registration: Challenges, progress and prospects. *Neurocomputing* **2023**, *548*, 126383. [CrossRef]
2. Cui, H.; Mao, Q.; Li, J.; Hu, Q.; Tao, Y.; Ma, J.; Li, Z. Shield Tunnel Dislocation Detection Method Based on Semantic Segmentation and Bolt Hole Positioning of MLS Point Cloud. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5702815. [CrossRef]
3. Xu, N.; Qin, R.; Song, S. Point cloud registration for LiDAR and photogrammetric data: A critical synthesis and performance analysis on classic and deep learning algorithms. *ISPRS Open J. Photogramm. Remote Sens.* **2023**, *8*, 100032. [CrossRef]
4. Wu, Y.; Yuan, Y.; Fan, X.; Huang, X.; Gong, M.; Miao, Q. PCRDiffusion: Diffusion Probabilistic Models for Point Cloud Registration. *arXiv* **2023**, arXiv:2312.06063.
5. Rusu, R.B.; Blodow, N.; Beetz, M. Fast Point Feature Histograms (FPFH) for 3D registration. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 3212–3217.

6. Besl, P.J.; McKay, N.D. Method for registration of 3-D shapes. In Proceedings of the Sensor Fusion IV: Control Paradigms and Data Structures, Boston, MA, USA, 12–15 November 1991; Spie: Bellingham, WA, USA, 1992; Volume 1611, pp. 586–606.
7. Biber, P.; Straßer, W. The Normal Distributions Transform: A New Approach to Laser Scan Matching. In Proceedings of the Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, NV, USA, 27–31 October 2003; IEEE: Piscataway, NJ, USA, 2003; Volume 3, pp. 2743–2748.
8. Fitzgibbon, A.W. Robust Registration of 2D and 3D Point Sets. In Proceedings of the 2003 IEEE International Conference on Computer Vision, Nice, France, 14–17 October 2003; IEEE: Piscataway, NJ, USA, 2003; pp. 662–669.
9. Myronenko, A.; Song, X. Point Set Registration: Coherent Point Drift. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 2262–2275. [CrossRef] [PubMed]
10. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 652–660.
11. Qi, C.R.; Yi, L.; Su, H.; Guibas, L.J. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 4–9 December 2017; pp. 5099–5108.
12. Thomas, H.; Qi, C.R.; Deschaud, J.E.; Marcotegui, B.; Goulette, F.; Guibas, L.J. KPConv: Flexible and Deformable Convolution for Point Clouds. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6411–6420.
13. Zeng, A.; Song, S.; Nießner, M.; Fisher, M.; Xiao, J.; Funkhouser, T. 3DMatch: Learning Local Geometric Descriptors from RGB-D Reconstructions. In Proceedings of the CVPR, Honolulu, HI, USA, 21–26 July 2017.
14. Deng, H.; Birdal, T.; Ilic, S. Ppfnet: Global context aware local features for robust 3d point matching. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Salt Lake City, UT, USA, 18–22 June 2018; pp. 195–205.
15. Gojcic, Z.; Zhou, C.; Wegner, J.D.; Wieser, A. The Perfect Match: 3D Point Cloud Matching with Smoothed Densities. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 5545–5554.
16. Yang, J.; Zhao, C.; Xian, K.; Zhu, A.; Cao, Z. Learning to fuse local geometric features for 3D rigid data matching. *Inf. Fusion* **2020**, *61*, 24–35. [CrossRef]
17. Deng, H.; Birdal, T.; Ilic, S. 3D local features for direct pairwise registration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3244–3253.
18. Valsesia, D.; Fracastoro, G.; Magli, E. Learning localized representations of point clouds with graph-convolutional generative adversarial networks. *IEEE Trans. Multimed.* **2020**, *23*, 402–414. [CrossRef]
19. Bai, X.; Luo, Z.; Zhou, L.; Fu, H.; Quan, L.; Tai, C.L. D3feat: Joint learning of dense detection and description of 3d local features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 6359–6367.
20. Choy, C.; Dong, W.; Koltun, V. Deep Global Registration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 2514–2523.
21. Yew, Z.Y.; Lee, G.H. 3DFeat-Net: Weakly Supervised Local 3D Features for Point Cloud Registration. In Proceedings of the IEEE European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 630–646.
22. Huang, S.; Gojcic, Z.; Usvyatsov, M.; Wieser, A.; Schindler, K. Predator: Registration of 3d point clouds with low overlap. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 4267–4276.
23. Yew, Z.J.; Lee, G.H. Regtr: End-to-end point cloud correspondences with transformers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 6677–6686.
24. Qin, Z.; Yu, H.; Wang, C.; Guo, Y.; Peng, Y.; Xu, K. Geometric transformer for fast and robust point cloud registration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 11143–11152.
25. Qin, Z.; Yu, H.; Wang, C.; Guo, Y.; Peng, Y.; Ilic, S.; Hu, D.; Xu, K. Geotransformer: Fast and robust point cloud registration with geometric transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 9806–9821. [CrossRef] [PubMed]
26. Yu, H.; Qin, Z.; Hou, J.; Saleh, M.; Li, D.; Busam, B.; Ilic, S. Rotation-invariant transformer for point cloud matching. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 5384–5393.
27. Cao, A.; Puy, G.; Boulch, A.; Marlet, R. PCAM: Product of Cross-Attention Matrices for Rigid Registration of Point Clouds. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 13209–13218. [CrossRef]

28. Lu, W.; Wan, G.; Zhou, Y.; Fu, X.; Yuan, P.; Song, S. Deepvcp: An end-to-end deep neural network for point cloud registration. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 12–21.
29. Choy, C.; Park, J.; Koltun, V. Fully convolutional geometric features. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8958–8966.
30. Yang, H.; Shi, J.; Carlone, L. Teaser: Fast and certifiable point cloud registration. *IEEE Trans. Robot.* **2020**, *37*, 314–333. [CrossRef]
31. Yew, Z.J.; Lee, G.H. Rpm-net: Robust point matching using learned features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 11824–11833.
32. Bai, X.; Luo, Z.; Zhou, L.; Chen, H.; Li, L.; Hu, Z.; Fu, H.; Tai, C.L. Pointdsc: Robust point cloud registration using deep spatial consistency. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Online, 19–25 June 2021; pp. 15859–15869.
33. Wu, Z.; Song, S.; Khosla, A.; Yu, F.; Zhang, L.; Tang, X.; Xiao, J. 3D ShapeNets: A deep representation for volumetric shapes. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1912–1920.
34. Chang, A.X.; Funkhouser, T.; Guibas, L.; Hanrahan, P.; Huang, Q.; Li, Z.; Savarese, S.; Savva, M.; Song, S.; Su, H.; et al. Shapenet: An information-rich 3d model repository. *arXiv* **2015**, arXiv:1512.03012.
35. Dai, A.; Chang, A.X.; Savva, M.; Halber, M.; Funkhouser, T.; Nießner, M. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5828–5839.
36. Hackel, T.; Savinov, N.; Ladicky, L.; Wegner, J.D.; Schindler, K.; Pollefeys, M. Semantic3d. net: A new large-scale point cloud classification benchmark. *arXiv* **2017**, arXiv:1704.03847.
37. Geiger, A.; Lenz, P.; Urtasun, R. Are we ready for autonomous driving? The kitti vision benchmark suite. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 3354–3361.
38. Beckley, M.; Baker, S.; Crosby, C.J.; Okal, M.H. A Terrestrial Laser Scanning Data Archive for Multi-Temporal Mapping Across Research Disciplines. In Proceedings of the AGU Fall Meeting Abstracts, Washington, DC, USA, 10–14 December 2018; Volume 2018, p. G51C-0491.
39. Dong, Z.; Liang, F.; Yang, B.; Xu, Y.; Zang, Y.; Li, J.; Wang, Y.; Dai, W.; Fan, H.; Hyyppä, J.; et al. Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 327–342. [CrossRef]
40. Lato, M.; Kemeny, J.; Harrap, R.; Bevan, G. Rock bench: Establishing a common repository and standards for assessing rockmass characteristics using LiDAR and photogrammetry. *Comput. Geosci.* **2013**, *50*, 106–114. [CrossRef]
41. Cuturi, M. Sinkhorn distances: Lightspeed computation of optimal transport. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 5–8 December 2013; Volume 26.
42. Sarlin, P.; DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperGlue: Learning Feature Matching with Graph Neural Networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seattle, WA, USA, 13–19 June 2020; pp. 4937–4946.
43. Xu, H.; Liu, S.; Wang, G.; Liu, G.; Zeng, B. Omnet: Learning overlapping mask for partial-to-partial point cloud registration. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 3132–3141.

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.