remote sensing

MDPI

Article

A Sub-Second Method for SAR Image Registration Based on Hierarchical Episodic Control

Rong Zhou [1,†], Gengke Wang [2,†], Huaping Xu [2], Zhisheng Zhang [1,*]

[1] Mechanical Engineering School, Southeast University, Nanjing 210096, China; zhourong@seu.edu.cn
[2] School of Electronic and Information Engineering, Beihang University, Beijing 100191, China; gkwang@buaa.edu.cn (G.W.); xuhuaping@buaa.edu.cn (H.X.)
* Correspondence: oldbc@seu.edu.cn
† These authors contributed equally to this work.

Abstract: For Synthetic Aperture Radar (SAR) image registration, successive processes following feature extraction are required by both the traditional feature-based method and the deep learning method. Among these processes, the feature matching process—whose time and space complexity are related to the number of feature points extracted from sensed and reference images, as well as the dimension of feature descriptors—proves to be particularly time consuming. Additionally, the successive processes introduce data sharing and memory occupancy issues, requiring an elaborate design to prevent memory leaks. To address these challenges, this paper introduces the OptionEM-based reinforcement learning framework to achieve end-to-end SAR image registration. This framework outputs registered images directly without requiring feature matching and the calculation of the transformation matrix, leading to significant processing time savings. The Transformer architecture is employed to learn image features, while a correlation network is introduced to learn the correlation and transformation matrix between image pairs. Reinforcement learning, as a decision process, can dynamically correct errors, making it more-efficient and -robust compared to supervised learning mechanisms such as deep learning. We present a hierarchical reinforcement learning framework combined with Episodic Memory to mitigate the inherent problem of invalid exploration in generalized reinforcement learning algorithms. This approach effectively combines coarse and fine registration, further enhancing training efficiency. Experiments conducted on three sets of SAR images, acquired by TerraSAR-X and Sentinel-1A, demonstrated that the proposed method's average runtime is sub-second, achieving subpixel registration accuracy.

Keywords: reinforcement learning; episodic control; synthetic aperture radar; image registration

check for updates

Citation: Zhou, R.; Wang, G.; Xu, H.; Zhang, Z. A Sub-Second Method for SAR Image Registration Based on Hierarchical Episodic Control. *Remote Sens.* **2023**, *15*, 4941. https://doi.org/10.3390/rs15204941

Academic Editor: Piotr Samczynski

Received: 28 August 2023
Revised: 6 October 2023
Accepted: 8 October 2023
Published: 12 October 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

Researchers have an ongoing commitment to monitor and study the Earth's complex surface and its changes. As an effective means of remote sensing, SAR images are indispensable in various fields, such as ecological development, environmental protection, resource exploration, and military reconnaissance. Research involving change detection, information extraction, and image fusion using multiple SAR images can provide additional information that a single image cannot convey. This necessitates a more-concise and -efficient high-precision image-registration process initially.

Existing SAR-image-registration methods can be categorized into traditional methods and deep-learning-based methods. Traditional methods mainly fall into two categories: gray scale-based methods and structural-feature-based methods. Grayscale registration methods utilize the intensity values of image pixels. They are computationally intensive and are susceptible to image quality issues, noise, and geometric distortion. The registration methods based on structural features typically include the processes of feature extraction, feature matching, fitting the transformation matrix, and interpolation resampling. Among

them, feature points extracted based on SIFT [1] or SAR-SIFT [2] have certain invariance to various changes such as rotation, position, scale, and gray scale and have been widely used [3–6]. Such methods can usually extract a significant number of feature points; for instance, SIFT can extract about 2000 feature points from images with a size of 500 × 500 [1].

Feature matching involves complex mathematical calculations, and its time and space complexity depend on the number of feature points extracted, the dimension of the feature descriptors, and the matching algorithm used. This process requires substantial computing resources.

When employing feature-point-based methods for SAR image registration, the accuracy can be compromised due to the impact of speckle noise on the primary orientation of traditional feature descriptors [7,8]. There are two primary approaches to address this issue. The first approach involves enhancing feature-based registration techniques such as SIFT [8] or SAR-SIFT [9]. Alternatively, the second approach utilizes neural networks [7].

In recent years, deep neural networks have found wide application in SAR image registration. These networks can flexibly extract multidimensional and deeper features, achieving promising and robust results [7,10]. The common processing flow for deep-learning-based registration involves applying traditional feature-extraction algorithms to obtain image feature points, extracting image blocks based on these points, using deep learning networks to learn the feature and matching labels of image patch pairs, employing constraint algorithms to eliminate mismatches, and calculating transformation matrices based on matching point pairs.

While deep-learning-based SAR image registration holds promise, the scarcity of open-source SAR datasets poses challenges, as creating such datasets requires specialized personnel and resources. A common workaround is to perform self-learning using existing images, involving multiple affine transformations to generate a large training dataset with known correspondences. Despite this, many deep-learning-based SAR registration studies still rely on traditional methods for matching processing. These methods have high time and space complexity, often involving iterative computations and significant computing resource requirements.

Reinforcement learning, a branch of machine learning, has found extensive application in areas such as robot control and intelligent decision-making. Reinforcement learning adjusts model behavior dynamically according to rewards, offering more-flexible error correction compared to supervised learning. Although reinforcement-learning-based computer vision applications have been proposed, they remain relatively unexplored in the realm of SAR image registration.

It is worth noting that mainstream reinforcement learning needs to strike a balance between exploration and exploitation. However, in computer vision application scenarios, extensive exploration might not be necessary. Therefore, the reinforcement learning framework based on Episodic Memory is better suited for computer vision applications. Hierarchical reinforcement learning can further enhance training efficiency, especially in scenarios with significant state differences.

This paper applied the OptionEM-based reinforcement learning framework to achieve end-to-end SAR image registration. Feature extraction was accomplished using the Transformer [11], while a correlation network was introduced to learn correlations and transformation matrices between sensed and reference image pairs. In order to compare with the existing registration methods, this paper proposes re-registration, involving registering the SAR-RL-registered image with the reference image. We adopted the indicators in Goncalves et al. [12] to quantitatively evaluate image registration.

The main contributions of this article are as follows:

First, it introduces an end-to-end architecture that directly outputs affine transformation matrices and registered images, significantly reducing the processing time compared to multi-step registration algorithms.

Second, reinforcement learning's dynamic decision-making with error correction mechanisms enhances efficiency and robustness when compared to deep learning frameworks.

Third, a hierarchical reinforcement learning framework is introduced, combining Episodic Memory to address the inherent invalid exploration issue in generalized reinforcement learning algorithms, resulting in faster training times.

Fourth, the use of hierarchical reinforcement learning further improves training efficiency and effectively combines coarse and fine registration.

In the experiments, a self-learning method was employed for dataset generation. The test dataset comprised three sets of SAR images acquired by TerraSAR-X and Sentinel-1A. The experimental results demonstrated that our method achieved not only an average running time at the sub-second level, but also a superior registration performance.

2. Related Work

2.1. Deep Learning

Image registration based on the deep learning framework centers on image feature extraction, leveraging booming neural network architectures such as the Transformer. Previous research has indicated that applying deep neural networks to the registration of complex and diverse SAR image pairs can yield more-accurate matching features compared to manually designed feature extraction algorithms, showcasing their promising performance and applicability [7,13,14]. These methods require a sufficient number of samples for training. However, several challenges remain, including the limited availability of publicly accessible datasets, the scarcity of labeled data [15,16], the substantial computational and time costs during the training phase, and the need for high-performance computer hardware. Moreover, local similarities may lead to mistaken matches. Addressing these challenges represents critical research areas when applying deep learning to SAR image registration.

Neural-network-based SAR image registration [17] falls under the umbrella of feature-based registration [7], transcending the limitations of manually designed features. It can extract multi-level features that reflect distributional, structural, and semantic characteristics. Various researchers have explored this approach, employing methods such as correlating coefficients and neural networks [18,19], utilizing Deep Convolutional Networks (CNNs) and Conditional Generative Adversarial Networks (CGANs) to extract geographic features [20], applying Pulse-Coupled Neural Networks (PCNNs) [21] for edge information [22], and combining SIFT algorithms with deep learning [23]. Fang Shang [24] constructed position vectors and change vectors that cleverly characterize image pixels and classified Polarimetric Synthetic Aperture Radar (PolSAR) images of complex terrain by a Quaternion Neural Network (QNN), which is not influenced by height information. Moreover, advanced techniques integrate self-learning with SIFT feature points for near-subpixel-level registration [7], employ deep forest models to enhance robustness [13], utilize unsupervised learning frameworks for multiscale registration [25–27], and leverage Transformer networks for efficient and accurate registration [28–33]. Deng, X. [13] employed a unique approach where each key point serves as a distinct class in the design of their multi-class model. This approach effectively circumvents the challenge of constructing matched-point pairs typically encountered in two-classification registration models. In a similar vein, S. Mao [31] introduced an adaptive self-supervised SAR-image-registration method that achieved comparable results. Meanwhile, Li, B. [29] presented a novel Siamese Dense Capsule Network designed to facilitate a more-even distribution of correctly matched keypoint pairs in SAR images featuring complex scenes. Fan, Y. [28] introduced an advanced and high-precision dense matching technique, specifically tailored for registering SAR images in situations characterized by weak texture conditions. The approaches of B. Zou [34] and Ming Zhao [35] involve the adoption of a pseudo-label-generation method, eliminating the need for additional annotations. Y. Ye [26] and D. Quan [36] separately built coarse-to-fine deep learning image registration framework based on stacking several deep models, which can significantly improve the multimodal image registration performances.

In summary, deep-learning-based registration methods for SAR images can leverage multi-level, latent, and multi-structural features to capture complex data variations. They

guide feature extraction using registration results, eliminating the need for manually set metrics. These methods have demonstrated favorable accuracy and applicability. However, they require a substantial number of training samples and high computational power during the training phase.

Regarding registration or matching, it plays a significant role in the entire image registration process. This process is used to identify misregistrations between two images or two patches, and for two images, it detects their mapping matrix, ultimately transforming one image to match the other. When it comes to two patches cropped from key points, matching classification performs well. Quan et al. [37] introduced a deep feature Correlation learning network (Cnet) along with a novel feature correlation loss function for multi-modal remote sensing image registration. The experiments demonstrated that the well-designed loss function improved the stability of network training and decreased the risk of overfitting. Li, L. [38] and D. Xiang [39] utilized networks to extract feature information and generate descriptors, which can be used to obtain more-correct matching point pairs.

2.2. Reinforcement Learning

Blundell and colleagues introduced the Model-Free Episodic Control (MFEC) algorithm [40] as one of the earliest episodic reinforcement learning algorithms. In comparison to traditional parameter-based deep reinforcement learning methods, MFEC employs non-parametric Episodic Memory for value function estimation, resulting in higher sample efficiency compared to DQN algorithms. Neural Episodic Control (NEC) [41] introduced a differentiable neural dictionary to store episodic memories, enabling the estimation of state–action value functions based on the similarity between stored neighboring states.

Savinov et al. [42] utilized Episodic Memory to devise a curiosity-driven exploration strategy. Episodic Memory DQN (EMDQN) [43] combined parameterized neural networks with non-parametric Episodic Memory, enhancing the generalization capabilities of Episodic Memory. Generalizable Episodic Memory (GEM) [44] parameterized the memory module using neural networks, further enhancing the generalization capabilities of Episodic Memory algorithms. Additionally, GEM extended the applicability of Episodic Memory to continuous action spaces.

These algorithms represent significant advancements in the field of episodic reinforcement learning, offering improved memory and learning strategies that contribute to more-effective and -efficient training processes.

3. Deep-Reinforcement-Learning-Based SAR Image Registration

3.1. SAR-RL

The registration of reference and sensed image pairs is approached as a sequential decision-making process. In this process, the sensed images undergo a transformation based on an action, referred to as a time step, which involves adjusting the transformation parameter. The resulting transformed image, achieved through image resampling, is then registered with the reference image to yield a reward value. This reward value typically indicates whether the latest image resampling has brought the sensed image closer to the reference image. If the transformation executed at that particular time step results in the resampled image being closer to the true value compared to the previous time step, a positive reward is received. Conversely, if the proximity decreases, a negative reward is obtained.

In this research, the state is the gray scale image of the reference and sensed image pairs. The sensed image is resampled according to the affine transformation parameters output by the agent, generating a new sensed image.

The action space consists of the affine transformation parameters of the sensed image, along with an additional trigger action. The agent executes the chosen action, causing the sensed image to undergo the corresponding affine transformation and generating a new sensed image. In this experiment, the discrete action space is defined based on transformation parameters with two scales of low and high precision. This design aims

to avoid both registration failures caused by very low precision and high interaction costs resulting from very high precision. The latter scenario might require numerous iterations before achieving a successful registration. The effective action set consists of 16 elements, each corresponding to a specific affine transformation: translate left 1, 10 px, translate right 1, 10 px, translate up 1, 10 px, translate down 1, 10 px, rotate clockwise 1°, 10°, rotate counterclockwise 1°, 10°, zoom in scale 0.1, 0.01, zoom out scale 0.1, 0.01. The trigger action represents that registration is complete, and further transformations are unnecessary.

The reward function is directly proportional to the enhancement in the registration of reference and sensed image pairs achieved by the agent through its actions. Intuitively, this improvement can be quantified using the Euclidean distance between the current transformation matrix and the ground truth. However, a challenge arises from differences in parameter units, as the scaling units are smaller compared to translation and rotation. To address this, salient points in the image—such as corner points—are employed to formulate the reward function.

Salient points are identified as key points detected by the maximum value of the Difference of Gaussians (DoG). These key points compose the salient point reference set P_G, derived from the ground truth of the sensed/reference image. In each training episode, the transform set $\tilde{P}_G$ of salient points is generated by applying the inverse matrix of the transformation matrix.

Subsequently, for each action, the distorted landmarks are transformed using the given transformation matrix. The reward for the action is defined based on the Euclidean distance D between the transformed landmarks and their corresponding landmark references. The entire process is depicted in Figure 1.

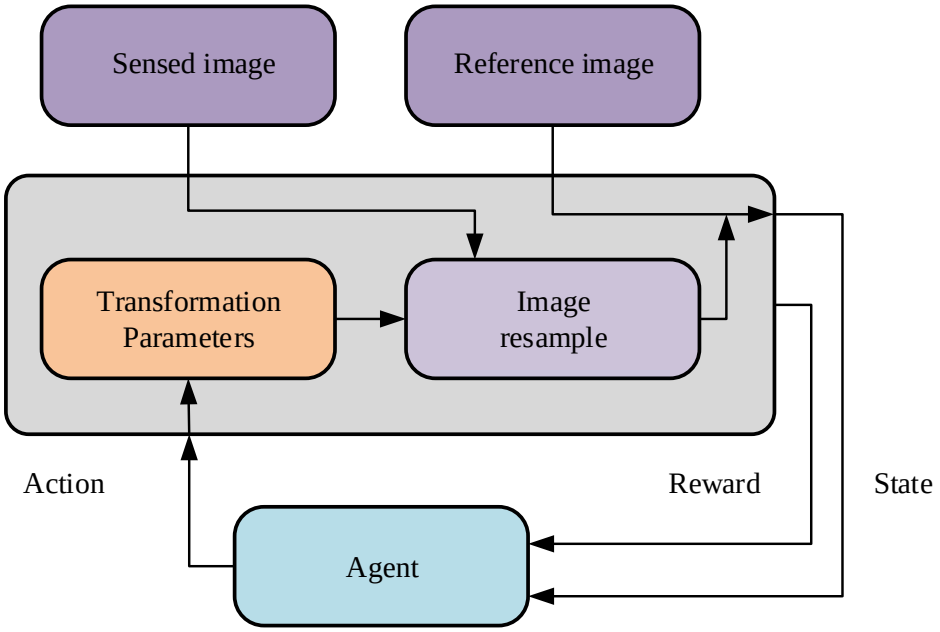


Figure 1. Image registration based on reinforcement learning.

3.2. Model

The neural network architecture is illustrated in Figure 2. The left side of the network features a Transformer layer, responsible for conducting feature learning through self-attention on the input sensed and reference image pairs. Meanwhile, the right side of the network comprises a correlation layer that captures position-related details between the sensed and reference image pairs. Ultimately, the network produces the parameters necessary for the affine transformation.

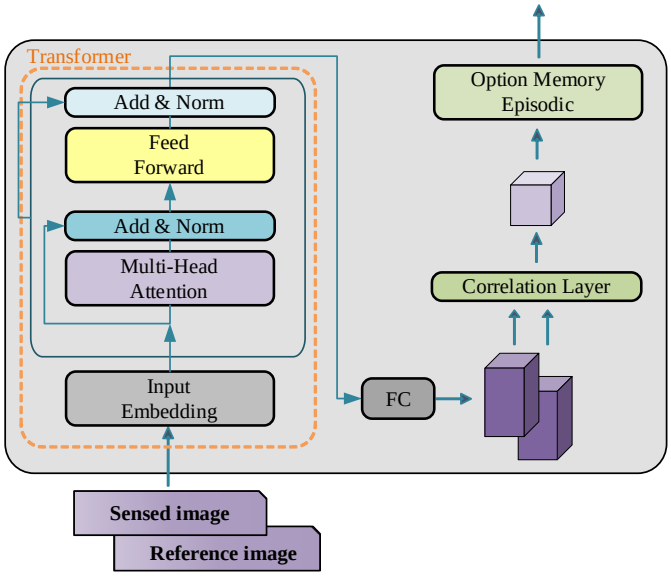


Figure 2. Agent network structure.

3.2.1. Transformer

The core module of the Transformer is the Multi-Head Attention (MHA) module. The definition of the dot-product attention is as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

Here, $Q \in \mathbb{R}^{T \times d_k}$ is the query matrix, $K \in \mathbb{R}^{T \times d_k}$ is the key matrix, and $V \in \mathbb{R}^{T \times d_k}$ is the value matrix. T and M represent the sequence lengths of queries and keys; d_k is the feature size of queries and keys; d_v is the feature size of V. In visual tasks, Q and K are typically reshaped query and key feature maps, where $T = M = h \times w$, with h and w being the height and width of the feature maps. The definition of Multi-Head Attention is as follows:

$$\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right) \tag{2}$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots, \text{head}_H)W^O \tag{3}$$

where $W_i^Q \in \mathbb{R}^{d \times d_k}$, $W_i^K \in \mathbb{R}^{d \times d_k}$, $W_i^V \in \mathbb{R}^{d \times d_v}$, and $W^O \in \mathbb{R}^{hd_v \times d}$ are parameter matrices and H is the number of heads. In the case of the Multi-Head Self-Attention (MHSA) in the encoder, $Q = K = V$.

The architecture of the Transformer encoder, which excludes positional encoding, is illustrated on the left side of Figure 2. This encoder includes Multi-Head Self-Attention (MHSA) and a feed-forward layer. The feed-forward layer initially increases the feature dimension from d to D and, subsequently, reduces it back to d. For added depth, the encoder can be stacked a total of N times, where N represents the number of encoder layers. Notably, in this study, a configuration akin to ViT was adopted, employing solely the Transformer encoder and incorporating additional positional encoding.

In the original Transformer design, a decoder with Multi-Head Cross-Attention (MHCA) is employed. Initially, the query consists of a learnable query embedding, which is then replaced by the output of the preceding decoder layer. The key and value components are derived from the outputs of the encoder layers. Similar to the encoder, the decoder can also be stacked N times.

3.2.2. Correlation Layer

The architecture of the correlation layer draws inspiration from the research conducted by Rocco et al. [45]. The objective is to effectively capture both positional and spatial correlation details within image features. The configuration of the correlation layer is depicted on the right side of Figure 2.

The feature maps f_A and f_B are in $\mathbb{R}^{h\times w\times d}$; the correlation map c_{AB} is in $\mathbb{R}^{h\times w\times d}$; the output of the correlation layer consists of scalar products of individual descriptors at each position:

$$c_{AB}(i,j,s) = \mathbf{f}_B(i,j)^T \mathbf{f}_A(i_s,j_s) \tag{4}$$

where (i,j) and (i_s,j_s) represent individual feature positions in the $h \times w$ dense feature map and $s = h(j_s-1)+i_s$ is an auxiliary index variable for (i_s,j_s).

The architecture of the correlation layer is illustrated in Figure 3. The right side of the illustration represents the correlation feature c_{AB}, which quantifies the similarity between features in f_B and f_A at a specific position (i,j). It is important to emphasize that c_{AB} is distinct from c_{BA}. The computation of similarity through the correlation feature addresses the challenge of ambiguous matches and necessitates further processing.

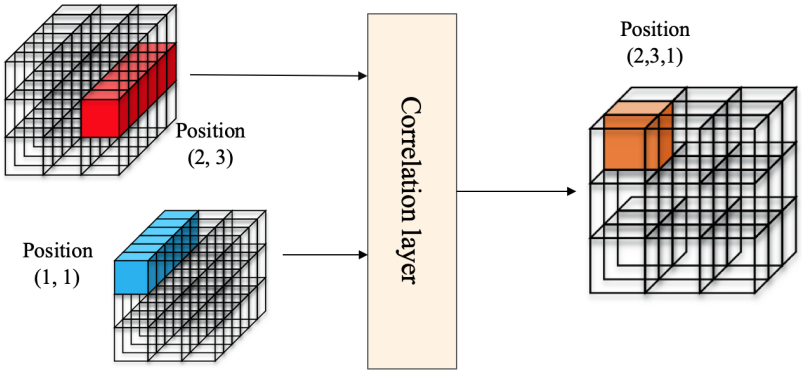


Figure 3. Architecture of the correlation layer.

To enhance the correlation feature, channelwise normalization is applied to the correlation feature at each position, yielding a normalized feature f_{AB} in the form of a correlation map. Subsequently, Softmax and L2 normalization are employed for the correlation map f_{AB}. This normalization approach emphasizes the scores of favorable matches. For instance, when only one feature in f_B exhibits a strong correlation with f_A, this method resembles nearest-neighbor matching in classical geometry computation [45]. Additionally, in scenarios where descriptors in f_B match with multiple features in f_A, resulting from noise or repetitive patterns, the matching scores are down-weighted—a concept similar to the second nearest-neighbor test.

Both the correlation and normalization operations are differentiable in relation to the input features, enabling effective backpropagation for end-to-end learning. Notably, this study employed the self-attention mechanism of the Transformer within the correlation layer to facilitate extensive feature matching across longer ranges.

3.3. Option Episodic Memory

The Option Episodic Memory framework (OptionEM) introduced by Zhou et al. [46] is a comprehensive hierarchical episodic control framework, which, to a certain extent, tackles the challenge of extensive sample requirements, which posed difficulties for the initial generations of deep reinforcement learning models. This framework employs hierarchical Episodic Memory, updated via implicit memory planning, to estimate the optimal rollout value for each state–action pair. Algorithm 1 provides pseudo-code for this algorithm.

Algorithm 1 Option Episodic Memory (OptionEM).

Initialize the Episodic Memory network and option network
Initialize the target network parameters $\theta'_{(1)} \leftarrow \theta_{(1)}, \theta'_{(2)} \leftarrow \theta_{(2)}, \alpha'_{(1)} \leftarrow \alpha_{(1)}, \alpha'_{(2)} \leftarrow \alpha_{(2)}, \phi' \leftarrow \phi, \zeta' \leftarrow \zeta, \eta' \leftarrow \eta$
Initialize the Episodic Memory $\mathbb{M}$
for $t = 1, \ldots, T$ **do**
  Choose option ω, execute action a
  receive reward r and next state s'
  store tuple $(s, \omega, a, r, s', \omega', \beta)$ in memory $\mathbb{M}$
  for i in $\{1, 2\}$ **do**
    sample N tuples $(s, a, r, s', \beta, R_t^i)$ from memory $\mathbb{M}$
    if $\beta = 0$ **then**
      update $\theta_{(i)} \leftarrow \min_{\theta_{(i)}} \sum \left(R_t^{(i)} - M_\theta^i(s, \omega, a) \right)^2$
    else
      update $\alpha_{(i)} \leftarrow \min_{\alpha_{(i)}} \sum \left(R_t^{(i)} - M_\alpha^i(s, \omega) \right)^2$
    end if
  end for
  if $t \; mod \; u = 0$ **then**
    $\theta'_{(i)} \leftarrow \tau\theta_{(i)} + (1-\tau)\theta'_{(i)}$
    $\alpha'_{(i)} \leftarrow \tau\alpha_{(i)} + (1-\tau)\alpha'_{(i)}$
    $\eta'_{(i)} \leftarrow \tau\eta_{(i)} + (1-\tau)\eta'_{(i)}$
    $\phi'_{(i)} \leftarrow \tau\phi_{(i)} + (1-\tau)\phi'_{(i)}$
    update Episodic Memory according Algorithm 2
  end if
  if $t \; mod \; p = 0$ **then**
    update ϕ, $\nabla_\phi J(\phi) = \nabla_a M_{\theta_1}(s, \omega, a)\big|_{\omega=\pi_\omega, a=\pi_\phi(s)} \nabla_\phi \pi_{\phi(s)}$ according to policy gradient
    update ζ, $\nabla_\zeta J(\zeta) = \nabla_\omega M_{\alpha_1}(s, \omega)\big|_{\omega=\pi_\omega} \nabla_\zeta \pi_{\zeta(s)}$ according to policy gradient
  end if
  if $t \; mod \; q = 0$ **then**
    update η, $\frac{\partial M_\alpha(\omega_0, s_1)}{\partial \eta} = -\sum_{s',\omega} \mu_\Omega(s', \omega \mid s_1, \omega_0) \frac{\partial \beta_{\omega,\eta}(s')}{\partial \eta} A_\Omega(s', \omega)$ according to policy gradient
  end if
end for

Algorithm 2 Update memory.

for stored trajectories τ **do**
  for one of the trajectories τ: $t = T \rightarrow 1$ **do**
    according to the current chosen option ω, executing $\tilde{a}_{t+1} \sim \pi_{\omega,\theta}(a \mid s)$
    if $\beta = 0$ **then**
      computing $M_\theta^{(1,2)}(s_{t+1}, \omega, a_{t+1})$
    else
      computing $M_{\Omega,\alpha}^{(1,2)}(s_{t+1}, \omega)$
    end if
    for $h = 0 : T - t$ computing $V_{t,h,\omega}^{(1,2)}$ according to (5)
    computing $R_{t,h,\omega}^{(1,2)}$ according to (6)
    saved into the memory
  end for
end for

At each step, the cumulative reward along the trajectory up to that point is compared with the value derived from the memory module. The greater value between these two is chosen. The memory module affiliated with an option encompasses both an option value memory module and an option internal memory module. The decision regarding which to choose is determined by the termination equation. The notations M_θ and $M_{\Omega,\alpha}$ arise from analogous experiences and represent value estimations for counterfactual trajectories linked to options. This process recursively establishes an implicit planning scheme within the Episodic Memory, aggregating experience longitudinally and across trajectories.

The complete backpropagation process can be formulated using Equation (5).

$$R_t = \begin{cases} r_t + \gamma \max(R_{t+1}, M_\theta(s_{t+1}, \omega_{t+1}, a_{t+1})) & \text{if } t < T, \beta_{\omega,\eta}(s_{t+1}) = 0 \\ r_t + \gamma \max(R_{t+1}, M_{\Omega,\alpha}(s_{t+1}, \omega)) & \text{if } t < T, \beta_{\omega,\eta}(s_{t+1}) = 1 \\ r_t & \text{if } t = T \end{cases} \tag{5}$$

where t denotes the step along the trajectory and T represents the episode length. The backpropagation process in Equation (5) can be expanded and rewritten as Equation (6).

$$\begin{aligned} V_{t,h} &= \begin{cases} r_t + \gamma V_{t+1,h-1} & \text{if } h > 0 \\ M_{\Omega,\alpha}(s_{t+1}, \omega) & \text{if } h = 0, \beta = 1 \\ M_\theta(s_t, \omega_t, a_t) & \text{if } h = 0, \beta = 0 \end{cases} \\ R_t &= V_{t,h^*} \end{aligned} \tag{6}$$

where $h^* = \arg\max_{h>0} V_{t,h}$.

The parameterized neural network M_θ is introduced to represent the parameterized option internal memory, while the parameterized neural network M_α serves as the parameterized option memory. Both networks are trained from a tabular memory M. In order to harness the generalization capability of M_θ and M_α, an augmented reward is propagated along trajectories using value estimates from M_θ and M_α, along with the actual rewards from M. This approach aims to determine the optimal value across all potential rollouts. During training, the improved target is regressed to train the versatile memories M_θ and M_α, with the chosen value guided by the termination equation. This refined target then guides policy learning and establishes fresh learning objectives for OptionEM.

A key challenge within this learning process is the potential overestimation stemming from identifying the best value along a trajectory. During the backpropagation process, overestimated values can persist and hinder efficient learning. To mitigate this challenge, a Siamese Network, akin to the concept of Double Q-Learning, is employed to refine the backpropagation of value estimates. Traditional reinforcement learning algorithms with function approximation are prone to overestimating values, which makes addressing this tendency critical. The Siamese Network structure is leveraged to render value estimates from M_θ more conservative. Training involves updating the memory network, the termination function, and the option policy using three distinct timescales.

3.4. Self-Learning

Compared to RGB image datasets, SAR image datasets are relatively limited in availability, and acquiring labeled datasets can be expensive. To tackle this challenge, this study employed a self-learning strategy. The core concept is to apply various transformation matrices—such as translation, scaling, rotation, etc.—to images. Consequently, a substantial number of corresponding sensed and reference image pairs are generated from the original images and their transformed versions. In the context of SAR image registration, the training dataset can be produced by applying affine transformations to one of the image pairs.

In this scenario, the initial large-scale image undergoes an affine transformation. A 240 × 240 px image patch is selected from the original image as the reference image. Utilizing the coordinates of the reference image's center point within the original image,

the coordinates for this point are calculated within the transformed large-scale image using the transformation matrix. Subsequently, a 240 × 240 px image patch is cropped to serve as the sensed image, with its center point determined by the computed coordinate values. This self-learning approach eliminates image patches that might exhibit black boundaries after the image patch's affine transformation, ensuring dataset quality. The process of self-learning is visually represented in Figure 4.

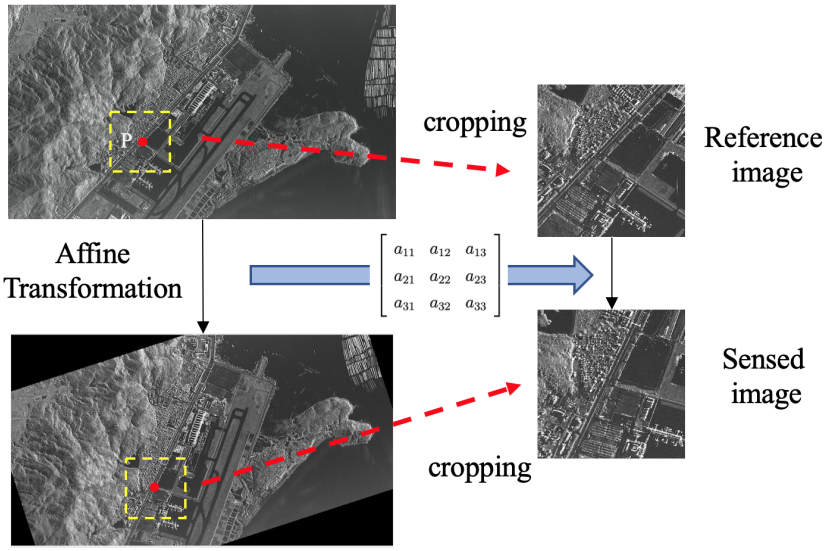


Figure 4. Self-learning.

This self-learning strategy proves effective in alleviating the need for extensive sample data. Shuang Wang et al.'s work already illustrated the method's success in SAR image registration through a deep learning framework [7]. However, their primary objective was to acquire localized feature information and details about feature point neighborhoods. To achieve this, they harnessed self-learning to generate a significant quantity of sample data, focusing on feature points and their surrounding localized information.

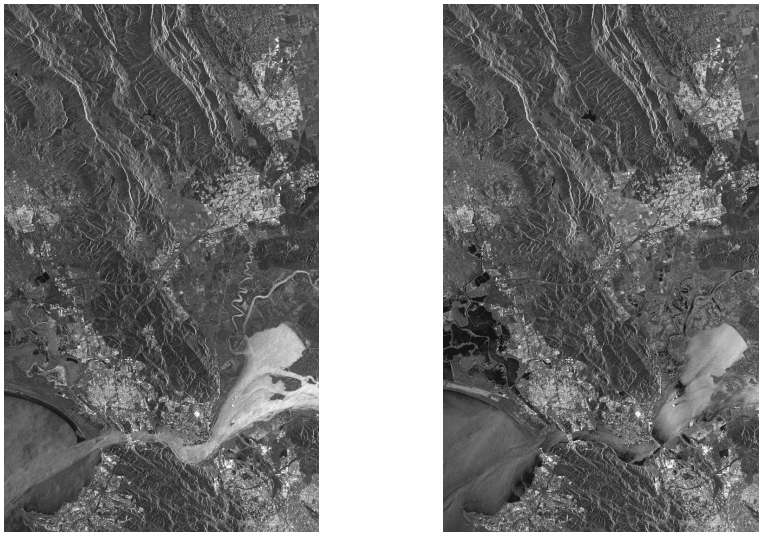
In contrast, this current study employed self-learning to create an extensive amount of global sample data. The self-learning samples were designed to construct new sensed–reference image pairs on a larger scale, rather than primarily concentrating on localized feature points.

4. Experiments

4.1. Dataset

The training image dataset for this experiment was generated using the right-looking descending image from the TerraSAR-X satellite over Napa, USA (the original images were downloaded from: https://download.geoservice.dlr.de/supersites/files/, accessed on 25 October 2021). These images were captured on 8 September and 30 September 2014, respectively, as depicted in Figure 5. By observing the images, evident changes in the gray scale and coverage can be noticed, particularly over the water bodies. It is essential to recognize that the presented images are quick-look previews, which are scaled-down versions of the original full-scene images. The resolution of these previews is approximately ten-times smaller than that of the actual data.

In comparison to the SAR image taken on 30 September, the SAR image captured on 8 September exhibits an offset of approximately 60 px in the range direction. The SAR reference image utilized during the network training phase was derived from a pre-processed version of the original image taken on 30 September 2014.



(**a**) (**b**)

Figure 5. Quick-look of the descending orbit right-looking SAR images acquired by the TerraSAR-X satellite for different periods in the Napa region, USA. (**a**) 8 September 2014. (**b**) 30 September 2014.

Due to the substantial size of SAR images and their inherent characteristics, including speckle noise and geometric distortions, a series of preprocessing steps are applied to the original images. These steps encompass block division, multilook, filtering, downsampling, and differencing.

Acquiring a significant volume of annotated remote sensing images, as opposed to optical image datasets, poses challenges due to the specialized domain knowledge required and the associated costs. To address this challenge, a self-learning strategy was employed. Given that SAR images can be captured under diverse conditions—such as varying incidence angles, distinct orbital directions, and diverse azimuth angles—various fundamental transformation types are employed for the affine transformation of the reference image. These fundamental transformations include translation, scaling, rotation, and flipping. This approach aims to examine the resilience of image features to these core transformation types. The parameters for these transformations are uniformly sampled within predefined ranges, and each affine transformation is created by combining these fundamental transformations.

Once the training dataset comprising sensed and reference image pairs is generated, SAR images undergo standardization and normalization. Image standardization, a widely used preprocessing technique, involves centering the data by subtracting the mean. This registration with convex optimization theory and knowledge about the data probability distribution enhances the suitability of the centered data distribution for generalization after training. Standardization results in a zero-mean centered data distribution, enhancing the effectiveness of gradient descent algorithms. Furthermore, it eliminates common image characteristics while accentuating individual differences. As a result, neural networks can learn more distinctive features, thereby fostering rapid learning, iteration, optimization, and improved learning efficiency. In this experiment, the standardization parameters the mean and standard deviation were set to $mean = 0.485$ and $std = 0.229$.

For generating the training data, 60 small images of size 240×240 px were cropped from the larger source image. Each of these small images underwent 700 affine transformations. The scaling factor ranged from 0.71 to 1.5 (uniformly sampled), and the rotation angle spanned from -180 to 180° with a precision of 1° (uniformly sampled). No flipping transformations were applied [7].

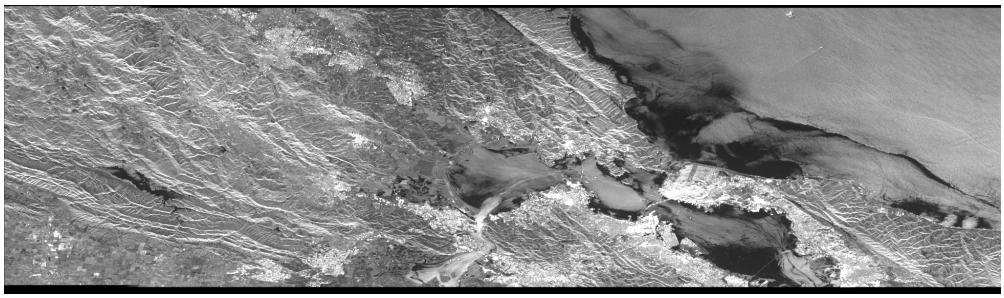
In this paper, we utilized three distinct datasets for testing:

1. TerraSAR-X Napa 2014 dataset: This was already introduced in the training dataset section, consisting of 110 pairs of 240×240 px image blocks.

2. Sentinel-1A Napa 2014 dataset: The source images were obtained from the Sentinel-1A SAR satellite images of the Napa region in the western United States on 30 October and 6 October 2014. A total of 18 pairs of 240 × 240 px image blocks were cropped. (as depicted in Figure 6, the original images were downloaded from: https://scihub.copernicus.eu/dhus//home, accessed on 10 November 2021).
3. TerraSAR-X Zhuhai 2016 dataset: The source images were obtained from the TerraSAR-X satellite images of the Zhuhai region in southern China, captured on 3 November 2016. A total of 62 pairs of 240 × 240 px image blocks were cropped. (as depicted in Figure 7).



(**a**)



(**b**)

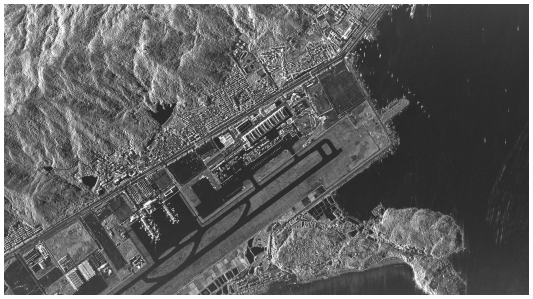
Figure 6. Descending pass right-looking SAR images from different periods of the Sentinel-1A satellite over the Napa region, USA. (**a**) 6 October 2014. (**b**) 30 October 2014.

The initial test dataset employed original images that aligned with the training dataset. Given the relative consistency in the imaging geometry of SAR-sensed and reference images—featuring minor variations in geometric structural characteristics—the evaluation of the proposed registration method's resilience to translation, scale, rotation, and other geometric transformations necessitated the application of random affine transformations to the SAR-sensed images. The types of transformations and the parameter ranges used for preprocessing remained consistent with those applied to the training dataset. Notably, the preprocessing methods employed across all three sets of test data were identical.

It is evident that the Sentinel-1A Napa 2014 image dataset contains more noise and pronounced dissimilarities in image features—such as water bodies and ridges—when contrasted with the training dataset derived from TerraSAR-X Napa 2014. Furthermore, the brightness distribution of images demonstrates considerable variations. Conversely, the TerraSAR-X Zhuhai 2016 dataset displays a greater presence of linear features, and its overall brightness distribution closely resembles that of the TerraSAR-X Napa 2014 dataset.



(**a**)



(**b**)

Figure 7. Right-looking SAR images from the TerraSAR-X satellite over the Zhuhai region, China. (**a**) Descending pass (3 November 2016). (**b**) Ascending pass (3 November 2016).

4.2. Implementation Details

The Backbone Network adopted the VGG16 [47] architecture to extract image features. The "Pool4 layer" was chosen as the output layer, while the structure of the other network layers remained unchanged. The input dimensions were set at 240×240, and the output dimensions were 512. Training started from scratch, and the status of the Batch Normalization (BN) [48] layer was fixed. Within the Option Episodic Memory, the internal policy network consisted of 3 convolutional layers followed by a fully connected layer, with output dimensions of $[2,\ 4]$. The convolutional kernel sizes were 7, 7, and 5, with the channel numbers being 225, 128, and 64, respectively. Padding was set to 0 for these convolutional layers. The Option Policy utilizes the same network structure as the internal policy network within the Option Episodic Memory. The output dimension was set to 2, representing the number of options. Some parameters in the identical network structure were shared in the implementation. Other parameter settings are in Table 1.

Table 1. Hyperparameter settings.

Hyperparameter	Value
Memory learning rate	1×10^{-5}
Actor learning rate	1×10^{-5}
Optimizer	Adam
Target network update frequency	0.6
η	0.3
Epochs	1000
Maximum steps	20
Memory size	1×10^{5}
Number of options	2
Batch size	32
γ	0.99

4.3. Option Analysis

Table 2 illustrates the correlation between the normalized Root-Mean-Squared Error (RMSE) and options. An option is displayed when the occurrence probability surpasses 75% within the RMSE range of statistically distinct score intervals across the three test datasets. Conversely, it is denoted as "–" when the option's proportion falls within the range of 25–75%. For instance, consider the TerraSAR-X Zhuhai 2016 dataset, which consists of 62 pairs of test sets. In the testing process, among 238 instances of image pairs with a normalized RMSE in the range of $[0, 0.2)$, the agent chose Option 0 a total of 223 times, resulting in a proportion of $193/238 = 93\%$. Therefore, in the table, it appears as Option 0. On the other hand, for 502 image pairs with a normalized RMSE in the range of $[0.4, 0.6)$, the agent selected Option 0 in 337 cases, accounting for a proportion of $337/502 = 67\%$. Consequently, this is indicated as "–" in the table.

Table 2. Relationship between the normalized root-mean-squared error and options.

	[0, 0.2)	**[0.2, 0.4)**	**[0.4, 0.6)**	**[0.6, 0.8)**	**[0.8, 1)**	**[1,)**
TerraSAR-X Napa 2014	Option 0	Option 0	Option 0	-	Option 1	Option 1
Sentinel-1A Napa 2014	Option 0	-	-	-	Option 0	Option 1
TerraSAR-X Zhuhai 2016	Option 0	Option 0	-	Option 1	Option 1	Option 1

The data imply that, across the three test datasets, the agent is more inclined to opt for fine registration when the option is 0, while it tends to select coarse registration when the option is 1. Notably, the complexity of the data in the Sentinel-1A Napa 2014 dataset resulted in unstable performance. Subsequent experimental results demonstrated that this dataset yielded poorer registration outcomes relative to the other two test datasets.

It is important to highlight that there exist numerous metrics for measuring image pair similarity, with the Root-Mean-Squared Error (RMSE) being a straightforward indicator. However, the RMSE's emphasis on pixel-level discrepancies can lead to challenges when assessing SAR image pairs, which may exhibit significant differences in scale, noise, gray scale, structural features, and more, thereby substantially influencing the accuracy of RMSE.

Indeed, there are instances where image pairs might share similar gray scale values, but fail to match in other aspects [7]. Moreover, the root-mean-squared error can also be influenced by factors such as the quantity, distribution, and precision of the feature point pairs. The examples provided in this section serve the purpose of illustrating the general connection between image disparities and options. However, it is crucial to note that these examples do not constitute quantitative analyses.

Figure 8 visually portrays the registration process of three image pairs, offering a comprehensive view of the reinforcement-learning-based registration method's entire path and more intuitively showcasing the option strategy employed throughout the registration process.

An observation can be made that, even in cases where a large difference exists between the sensed image and the reference image—such as the $T = 0$ time step in the second image pair—the occurrence of Option 0 did not adversely impact the subsequent registration process. Across the first and second image pairs, the agent tended to favor the selection of the coarse registration option, involving significant rotation (e.g., the $T = 0$ time step in the first image pair and the $T = 1$ time step in the second image pair). Subsequently, the agent executed scaling and rotation to achieve fine registration (e.g., $T = 1, 2, 3, 4, 5$ time steps for the first image pair and $T = 2, 3, 4, 5$ time steps for the second image pair).

In the case of the third image pair, the agent's decision-making process appeared to exhibit some instability (e.g., the $T = 5$ time step). However, it is worth noting that the outputs for the $T = 1, 2, 3$ time steps aligned with the expectations, indicating a certain level of consistency.

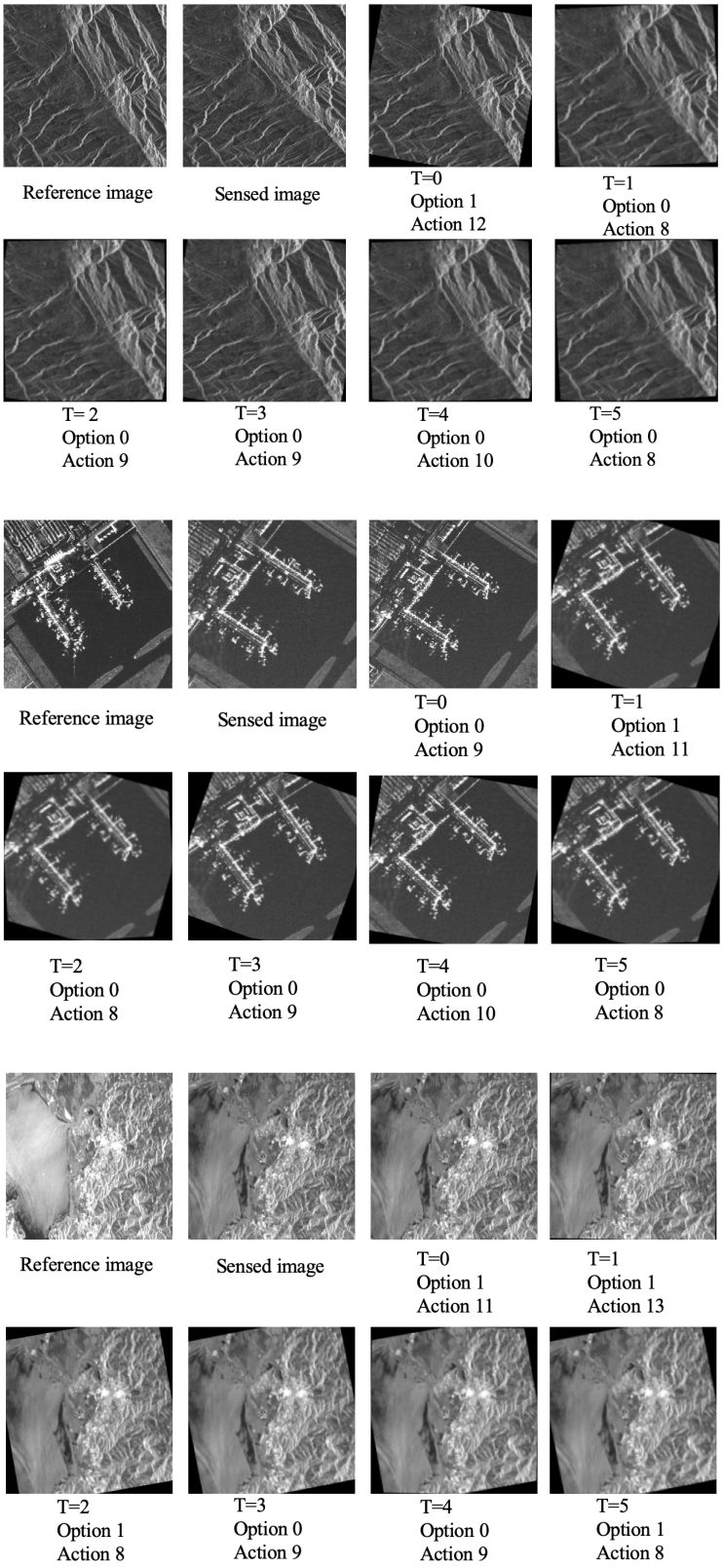


Figure 8. Options and actions in testing process.

4.4. SAR Image Registration Results' Analysis

4.4.1. Single-Registration and Re-Registration

In this study, we introduced a comprehensive image-registration approach that delivers the affine transformation matrix and registered images directly from input sensed

and reference images. Unlike methods with distinct stages such as feature point extraction, direction assignment, and descriptor construction, our algorithm does not offer a direct calculation of quantitative measures relying on feature points.

To enable a meaningful comparison with conventional registration techniques, we devised an indirect strategy. This approach involves two steps: first, the initial registration of sensed and reference images; second, a subsequent re-registration wherein the sensed image is registered with the SAR-Reinforcement Learning (RL) registered image. Quantitative metrics are then computed for both scenarios. The underlying assumption is that if the re-registration outcome exhibits a smaller Root-Mean-Squared Error (RMSE) value and elements of the affine transformation matrix that are closer to zero, this implies an improved SAR-RL registration result, in line with the expectations.

4.4.2. Qualitative Evaluation Results

The qualitative assessment primarily involved visually inspecting the spatial geometric registration results of the registered images [7]. This can be achieved by overlaying the registered reference image onto the sensed image using distinct colors to gauge the extent of overlap and differences in content. Another technique involves dividing the images into smaller blocks and overlapping them in a checkerboard pattern to create a mosaic map. By examining the continuity of structural features (such as stitching boundaries, rivers, bridges, roads, ridges, etc.) and assessing the overlap of common areas, the effectiveness of registration can be assessed. While this approach is straightforward and effective, it might not be suitable for large-scale automated applications.

Figure 9 provides examples of checkerboard-mosaicked images for the three test datasets. Due to space limitations, one example was selected from each dataset: a land–water interface area containing man-made structures and mountains, a mountainous region, and an airport area. Column *a* is the Reference image; The Column *b* is Sensed image; The images in Column *c* were Sensed images processed by SAR-RL. Column *d* is the Mosaic image of Column *c* and Column *b*. As evident from Column *d* in the figure, the sensed images and reference images from the three scene sets demonstrated successful registration. Most regions of the images overlapped well, and the stitching boundaries were consistent.

Furthermore, it is noticeable from Columns *d* and *c* that the sensed images produced by the SAR-RL technique exhibited a high visual similarity to the reference images in terms of geometric structural features. This aspect lays the groundwork for achieving effective registration.

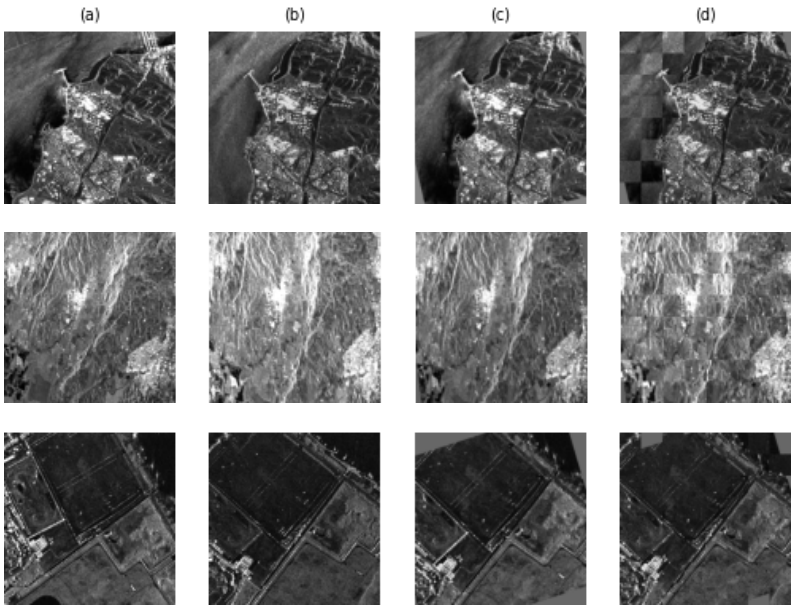


Figure 9. The checkerboard-mosaicked image examples of registered images from the test datasets.

4.4.3. Quantitative Evaluation Results

Quantitative evaluation involves expressing registration outcomes in numerical terms. Commonly employed metrics for quantitative assessment encompass the count of control point pairs (N_{red}), the Root-Mean-Squared Error (RMSE) of all control point residuals ($RMSE_{all}$), and the RMSE calculated using the leave-one-out method ($RMSE_{Loo}$). These metrics are frequently normalized to the pixel size to facilitate comparisons [49].

The transformation matrix (T) for control point pairs is determined through the least-squares method, and based on this matrix, the residuals of control point pairs are computed, giving rise to $RMSE_{all}$ and $RMSE_{Loo}$. The presence of more-accurately matched point pairs translates into a better determination of the parameters in the geometric transformation model, thereby yielding enhanced performance [28]. However, it is important to note that an increase in the number of correctly matched point pairs acquired through a registration method does not necessarily guarantee a decrease in the root-mean-squared error. This is because errors can still persist in the point pairs matched by the algorithm.

Gonçalves et al. took into account the distribution of control points and residuals and put forward various evaluation metrics, which included the quadrant residual distribution (P_{quard}), the proportion of poorly matched points with residuals (norm) exceeding 1.0 (Bad Point Proportion (BPP)), denoted as BPP(1.0), the detection of a preferred axis in the residual scatter plot (S_{kew}), and the statistical attributes of the control point distribution in the image (S_{cat}), all of which were considered in conjunction with N_{red}, $RMSE_{all}$, and $RMSE_{Loo}$. Among these metrics, except for N_{red}, smaller values in the other evaluation metrics signify improved registration performance [7].

These seven metrics are not completely independent of one another. For instance, a higher value of N_{red} along with a lower value of $RMSE_{all}$ indicates greater accuracy in point matching. Conversely, if both metrics are either high or low, this suggests lower accuracy.

Comparing and analyzing the quantitative results in these tables, several observations can be made: The N_{red} value obtained by re-registration is comparable to that of SAR-SIFT, indicating that a similar number of control point pairs are correctly matched. However, our re-registration achieved more-accurate registration results. The N_{red} value obtained by re-registration was lower than that of DL-WangS, but the accuracy achieved was comparable. The S_{cat} value obtained by re-registration was comparable to that of both SAR-SIFT and DL-WangS, suggesting that the distribution of matched feature point pairs obtained by the three methods was similar. It can also be observed that the running time of SAR-SIFT was in the order of tens of seconds, while DL-WangS took seconds to run. In contrast, our re-registration method had a significantly faster registration speed, with running times in the sub-second range.

Tables 3 and 4 present the median and mean values of the quantitative analysis results for SAR-SIFT, DL-WangS [7], and our SAR-RL model after re-registration was applied to the three datasets. The analysis was conducted using the seven metrics introduced by Gonçalves et al. (2009) to quantitatively evaluate the registration results for all image pairs in the test datasets.

These tables provide a comprehensive overview of the quantitative evaluation results and highlight the strengths of our re-registration method in terms of accuracy and efficiency compared to the other methods.

It should be emphasized that the re-registration approach does not directly measure the performance of the SAR-RL method. The calculation of the metrics still relied on the SIFT and SAR-SIFT algorithms. However, the metrics indicated that SAR-RL registration was significantly more accurate than these two traditional methods.

Table 3. Quantitative analysis of median results of the metrics for re-registration, SAR-SIFT and DL-WangS.

TerraSAR-X Napa 2014 dataset								
	N_{red}	$RMSE_{all}$	$RMSE_{Loo}$	P_{quard}	$BPP(1.0)$	S_{kew}	S_{cat}	running time
SAR-SIFT	12	0.7170	0.9021	0.0000	0.5000	0.4843	1.0000	28.4611 s
DL-WangS	22	0.3290	0.3977	0.0000	0.0498	0.3012	1.0000	9.5026 s
Re-registration	7	0.3257	0.4331	0.0000	0.0317	0.2943	0.9999	-
Sentinel-1A Napa 2014 dataset								
	N_{red}	$RMSE_{all}$	$RMSE_{Loo}$	P_{quard}	$BPP(1.0)$	S_{kew}	S_{cat}	running time
SAR-SIFT	6	0.8263	1.0324	0.6340	0.5000	0.0143	1.0000	17.3443 s
DL-WangS	20	0.3310	0.3449	0.1182	0.0023	0.1293	1.0000	9.2473 s
Re-registration	4	0.3714	0.4064	0.0931	0.0317	0.1169	0.9999	-
TerraSAR-X Zhuhai 2016 dataset								
	N_{red}	$RMSE_{all}$	$RMSE_{Loo}$	P_{quard}	$BPP(1.0)$	S_{kew}	S_{cat}	running time
SAR-SIFT	23	0.7669	0.8945	0.4045	0.7011	0.2349	0.9000	32.8332 s
DL-WangS	30	0.3591	0.4156	0.2812	0.0101	0.1058	0.9999	10.2314 s
Re-registration	26	0.3402	0.4911	0.1298	0.0061	0.011	1.0000	-

Table 4. Quantitative analysis of mean results of the metrics for re-registration, SAR-SIFT and DL-WangS.

TerraSAR-X Napa 2014 dataset								
	N_{red}	$RMSE_{all}$	$RMSE_{Loo}$	P_{quard}	$BPP(1.0)$	S_{kew}	S_{cat}	running time
SAR-SIFT	13	0.8007	0.9451	0.0001	0.4201	0.3801	0.9997	32.1185 s
DL-WangS	20	0.2912	0.3071	0.0000	0.0032	0.2189	0.9999	9.8215 s
Re-registration	10	0.3092	0.3471	0.0000	0.0712	0.1324	0.9995	-
Sentinel-1A Napa 2014 dataset								
	N_{red}	$RMSE_{all}$	$RMSE_{Loo}$	P_{quard}	$BPP(1.0)$	S_{kew}	S_{cat}	running time
SAR-SIFT	6	0.8019	0.9733	0.5215	0.5991	0.1021	0.9999	22.4352 s
DL-WangS	18	0.2431	0.3130	0.1572	0.0169	0.0035	0.9999	9.0089 s
Re-registration	6	0.3119	0.4204	0.1874	0.0245	0.0097	0.9999	-
TerraSAR-X Zhuhai 2016 dataset								
	N_{red}	$RMSE_{all}$	$RMSE_{Loo}$	P_{quard}	$BPP(1.0)$	S_{kew}	S_{cat}	running time
SAR-SIFT	20	0.7395	0.862	0.4121	0.2422	0.2911	0.9333	30.9462 s
DL-WangS	30	0.3292	0.2213	0.3119	0.0102	0.051	1.0000	10.1931 s
Re-registration	25	0.3371	0.24324	0.2454	0.0014	0.068	1.0000	-

In order to compare our method with traditional and deep-learning-based approaches, we selected SAR-SIFT and Wang Shuang's research (referred to as DL-WangS) [7] as the benchmarks. As supplements to the tables, Figures 10–12 illustrate the interval statistical results of the Root-Mean-Squared Error (RMSE) values for SAR-SIFT, DL-WangS [7], and our re-registration results for the three datasets.

The intervals depicted in these figures are $[0, 0.01)$, $[0.01, 0.1)$, $[0.1, 0.2)$, $[0.2, 0.3)$, $[0.3, 0.4)$, $[0.4, 0.5)$, $[0.5, 0.6)$, $[0.6, 0.7)$, $[0.7, 0.8)$, $[0.8, 0.9)$, $[0.9, 1)$, and $[1, \infty)$.

These can be seen in the figures; the re-registration results after the initial registration of the sensed and reference image pairs used for testing were significantly better. The results from the TerraSAR-X Napa 2014 dataset and TerraSAR-X Zhuhai 2016 dataset outperformed those from the Sentinel-1A Napa 2014 dataset. About 24% of re-registrations in the TerraSAR-X Napa 2014 dataset achieved RMSE values below 0.01, while only 5% in the Sentinel-1A Napa 2014 dataset achieved this. Similarly, 24% of re-registrations in the TerraSAR-X Zhuhai 2016 dataset achieved RMSE values below 0.01. The RMSEs obtained by SAR-SIFT were concentrated around 1 px, with a small portion larger than 1 px. The RMSEs obtained by DL-WangS were mostly smaller than 0.6 px, but for the Napa region images, some were close to 1 px. The RMSEs obtained by our proposed re-registration

method were all less than 1 px and were concentrated near 0. This suggested that our re-registration achieved sub-pixel level accuracy and outperformed SAR-SIFT, while being comparable or even slightly better than DL-WangS.

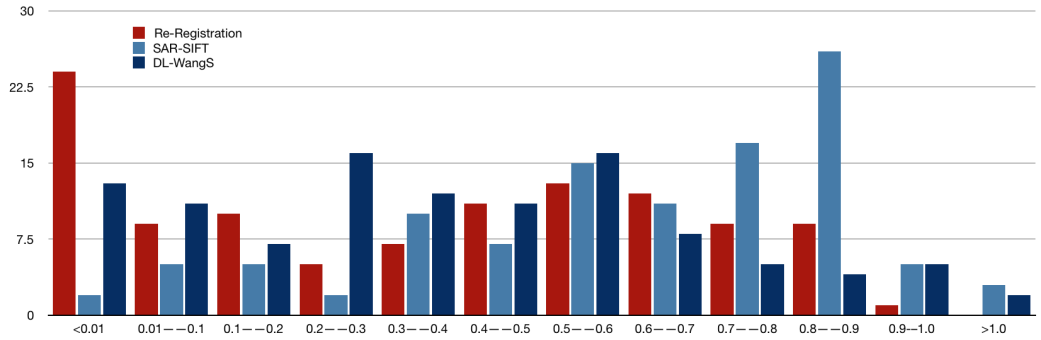


Figure 10. Statistical analysis of RMSE values for 110 pairs of TerraSAR-X Napa 2014 images using SAR-SIFT.

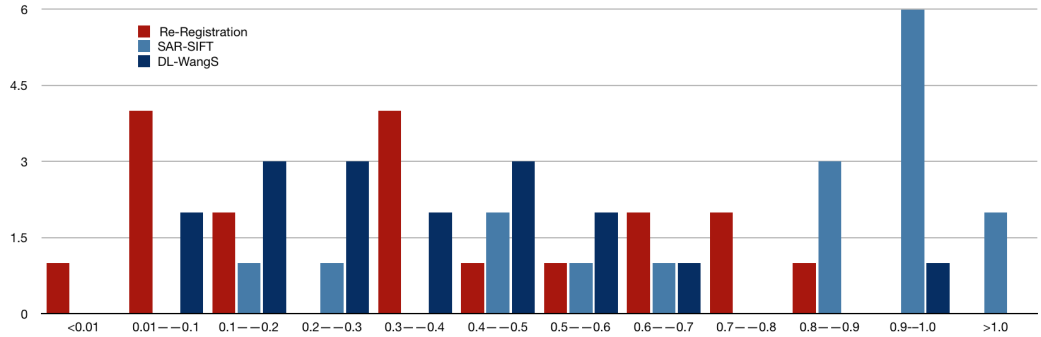


Figure 11. Statistical analysis of RMSE values for 18 pairs of Sentinel-1A Napa 2014 images using SAR-SIFT.

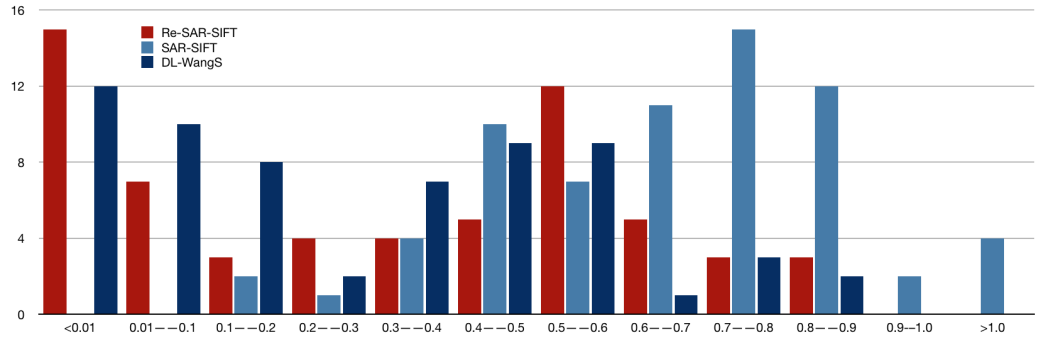


Figure 12. Statistical analysis of RMSE values for 62 pairs of TerraSAR-X Zhuhai 2016 images using SAR-SIFT.

4.4.4. Matching Results' Visualization

The visual comparisons we provided through Figures 13–15 effectively illustrated the differences between matched feature point pairs obtained by the re-registration and single-registration methods using SIFT-extracted feature points for the three sets of test data samples. These comparisons highlighted the strengths of the re-registration approach.

In the examples shown, it is evident that the matched point pairs obtained by re-registration were more accurate and less cluttered compared to the single-registration. The re-registration approach managed to find the correct matching point pairs even in image regions with significant differences, such as the varying gray scale, structural features, and challenging image content. This was a strong indication of the effectiveness of the SAR-RL method in obtaining accurate and robust registrations, even under challenging conditions.

However, it is also important to note that even the re-registration approach might have some incorrectly matched point pairs, as seen in Figure 15. This can be attributed to factors such as differences in the satellite orbital directions and potential blurring in the SAR images. Despite these challenges, the re-registration approach consistently demonstrated superior performance compared to the single-registration method, as evidenced by the overall accuracy and quality of the matched feature point pairs.

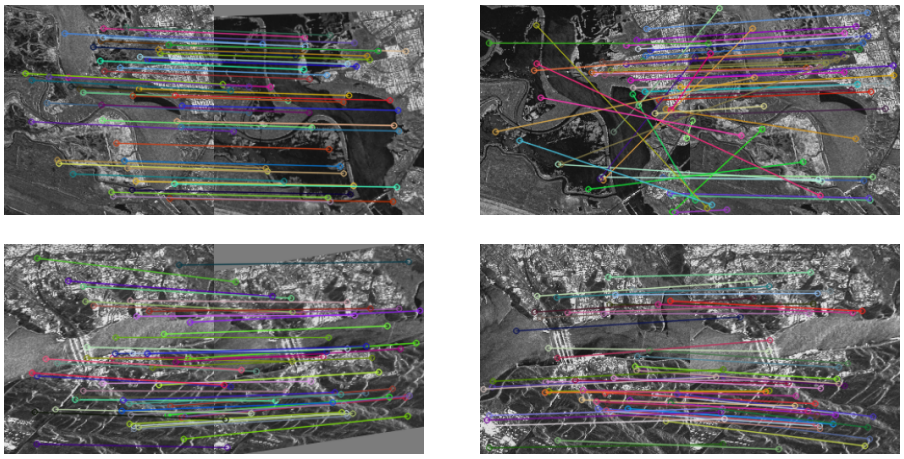


Figure 13. Examples of matching results based on SIFT by re-registration and single-registration on the TerraSAR-X Napa 2014 dataset.

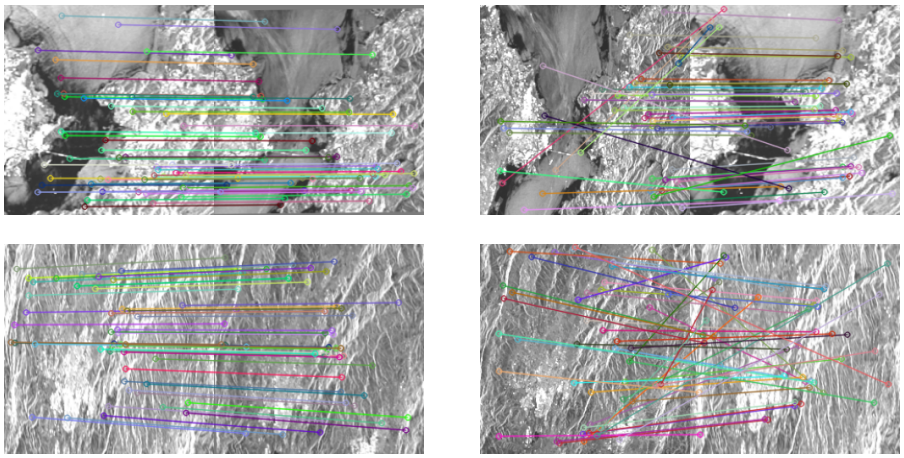


Figure 14. Examples of matching results based on SIFT by re-registration and singe-registration on the Sentinel-1A Napa 2014 dataset.

Figures 16–18 demonstrate samples of matched feature point pairs obtained through re-registration and one-time registration of feature points extracted using SAR-SIFT for the three sets of test data samples. In comparison with the results from the previous set of experiments, it was evident that the feature point pairs acquired through SAR-SIFT were fewer than those obtained using SIFT.

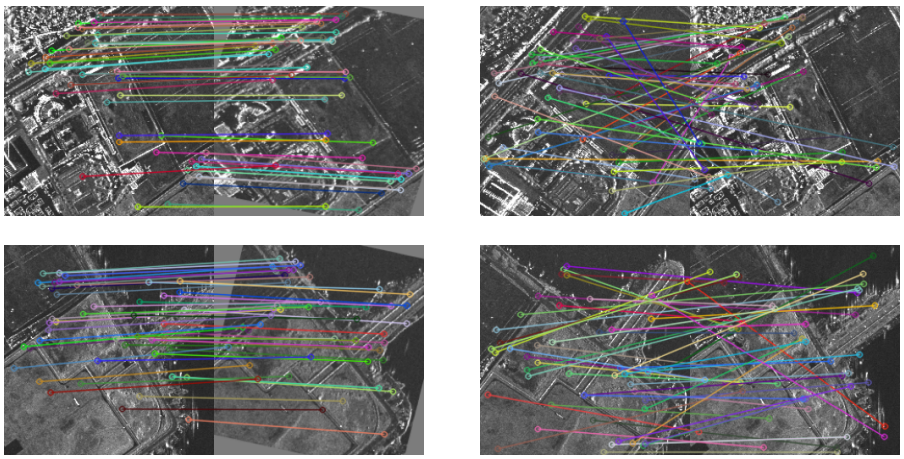


Figure 15. Examples of matching results based on SIFT by re-registration and single-registration on TerraSAR-X Zhuhai 2016.

In Figure 16, the connecting lines of the matching point pairs obtained from the single-registration of Sample 1 appear more cluttered, and all of them were incorrectly matched, resulting in a failed image registration. The outcome of the fusion processing displayed only the sensed image. The fusion process merely presented the information from one of the images. Conversely, the corresponding matching point pairs acquired through re-registration aligned with the correct positions. The fusion image exhibited continuous splicing boundaries, a substantial overlap of most image regions, and the visibility of splicing boundaries between the two images. For Sample 2, the positions of matched point pairs obtained from single-registration were accurate, resulting in successful image registration. However, the overlapping area of the fused image was blurred, which hindered the acquisition of image information. Re-registration, on the other hand, yielded more-correct matching point pairs, evenly distributed, with the fusion image displaying continuous splicing boundaries, extensive region overlap, and visible splicing boundaries and geometric structural features.

In Figure 17, feature point pairs derived from the single-registration of Sample 1 are concentrated and inaccurately matched, and the image registration failed. The fusion image primarily showed the sensed image. In contrast, feature point pairs acquired through re-registration for Sample 1 corresponded to the accurate location, with a relatively scattered distribution. The fusion image exhibits continuous splicing boundaries and substantial overlap in most image regions. For Sample 2, both re-registration and single-registration yielded accurately corresponding feature point pairs. The fusion images demonstrate extensive region overlap, with no noticeable misregistration in the splicing boundary.

The two samples in Figure 18 display evident linear structural features and gray scale similarity. In both cases, re-registration and single-registration yielded correctly corresponding feature point pairs. The fusion images revealed significant overlap in most image regions, with continuous splicing boundaries, indicating a favorable registration outcome.

It is crucial to note that, in the context of SAR images, SIFT and SAR-SIFT can sometimes extract incorrect feature points or produce mismatched feature point pairs. As a result, the comparison between single-registration and re-registration, as presented above, served as a reference. A comprehensive evaluation should encompass various indicators discussed in the preceding sections, along with other methodologies such as mosaicked images.

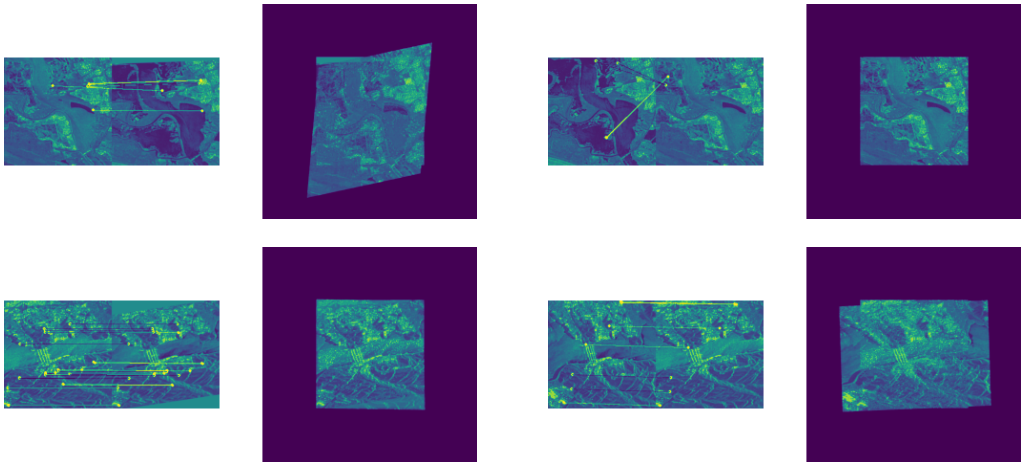


Figure 16. Examples of matching results based on SAR-SIFT by re-registration and single-registration on the TerraSAR-X Napa 2014 dataset.

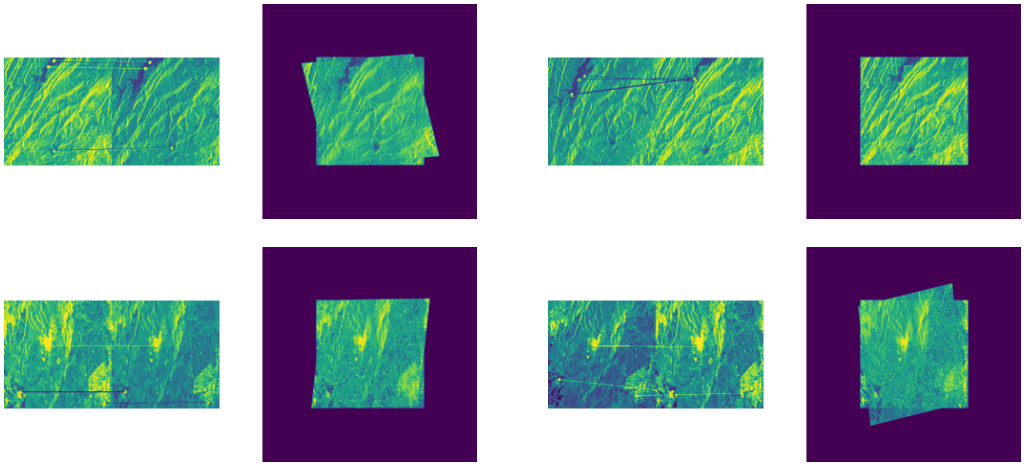


Figure 17. Examples of matching results based on SAR-SIFT by re-registration and single-registration on the Sentinel-1A Napa 2014 dataset.

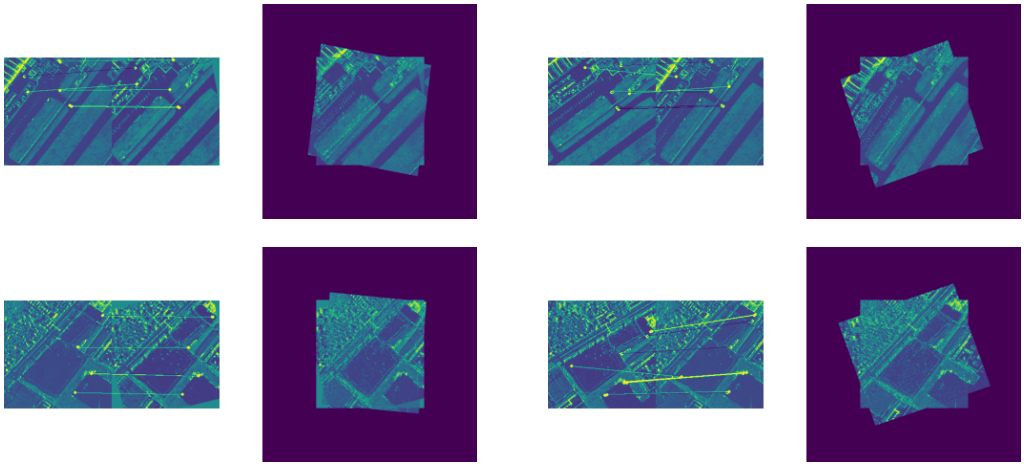


Figure 18. Examples of matching results based on SAR-SIFT by re-registration and single-registration on the TerraSAR-X Zhuhai 2016.

4.4.5. Data Visualization

The visualization results of the test dataset samples are presented in Figure 19. These figures showcase the outcomes of the attention maps, wherein each spatial location within the connectivity map encompassed all similarity scores between a feature in the reference image and all features within the sensed image. To illustrate, if a filter's central patch comprised predominantly zeros except for a peak located at the top-left corner, the filter responded positively to features within the reference image that correspond to the top-left corner of the sensed image. Likewise, when numerous spatial locations of a filter yielded similar visual responses, the filter demonstrated heightened sensitivity to spatially co-located features within the reference image that aligned with the corresponding positions in the sensed image.

The figures exhibit several attention maps that validate the presumption that the layer acquired the ability to replicate local neighborhood consensus. This was evident as certain filters exhibited strong responsiveness to spatially co-located features within the reference image that matched the consistent spatial positions in the sensed image.

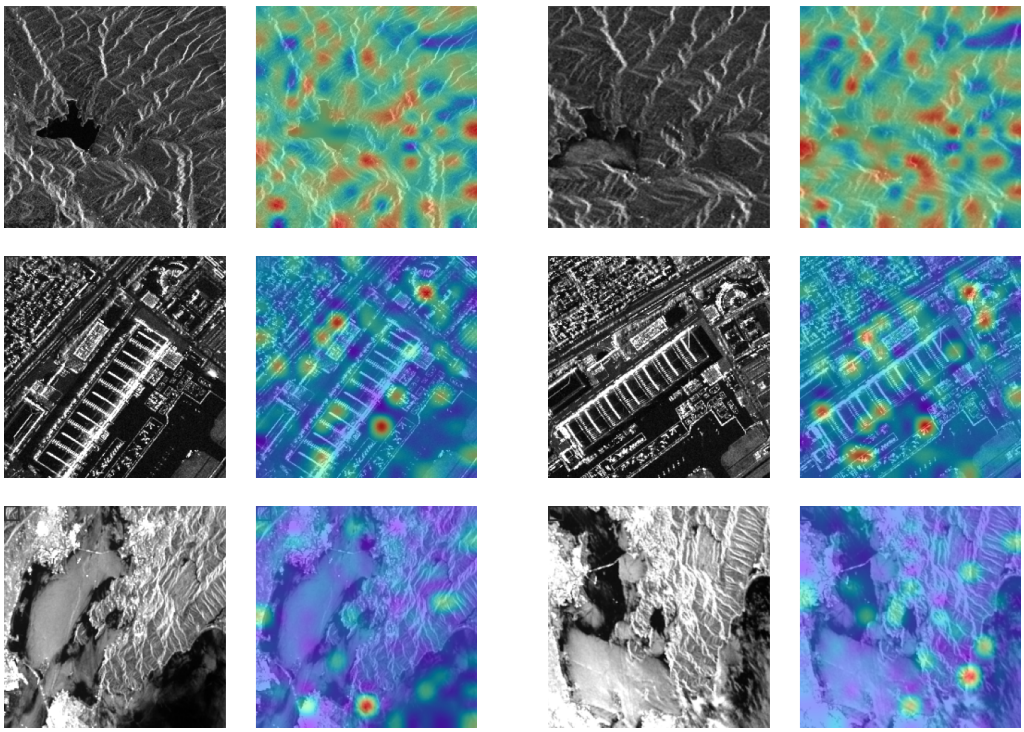


Figure 19. Visualization results of dataset test samples.

4.4.6. Analysis of Failed Cases

Figure 20 illustrates instances of failed registration.Column *a* is the Reference image; Column *b* is Sensed image. In Column *c*, the sensed image generated by the reinforcement learning framework of this paper is based on the reference image, i.e., the sensed image to be registered. Column *d* displays the outcomes of the registration using the method proposed in this paper. Variations in terrain features and imaging viewpoints between the sensed images obtained at different sampling times and the reference image pairs can lead to significant gray scale differences, as observed in the failed Sample 4, and different structural features, such as the failed Sample 3 and the failed Sample 6. In some cases, SAR images obtained from the same ground scene region may not appear intuitively identical, as demonstrated in Rows 2 and 4 of the figure. While the reinforcement learning in this paper generated sensed images that partially reduced the structural differences between them and the original reference images, it did not perform gray scale transformations on the sensed images.

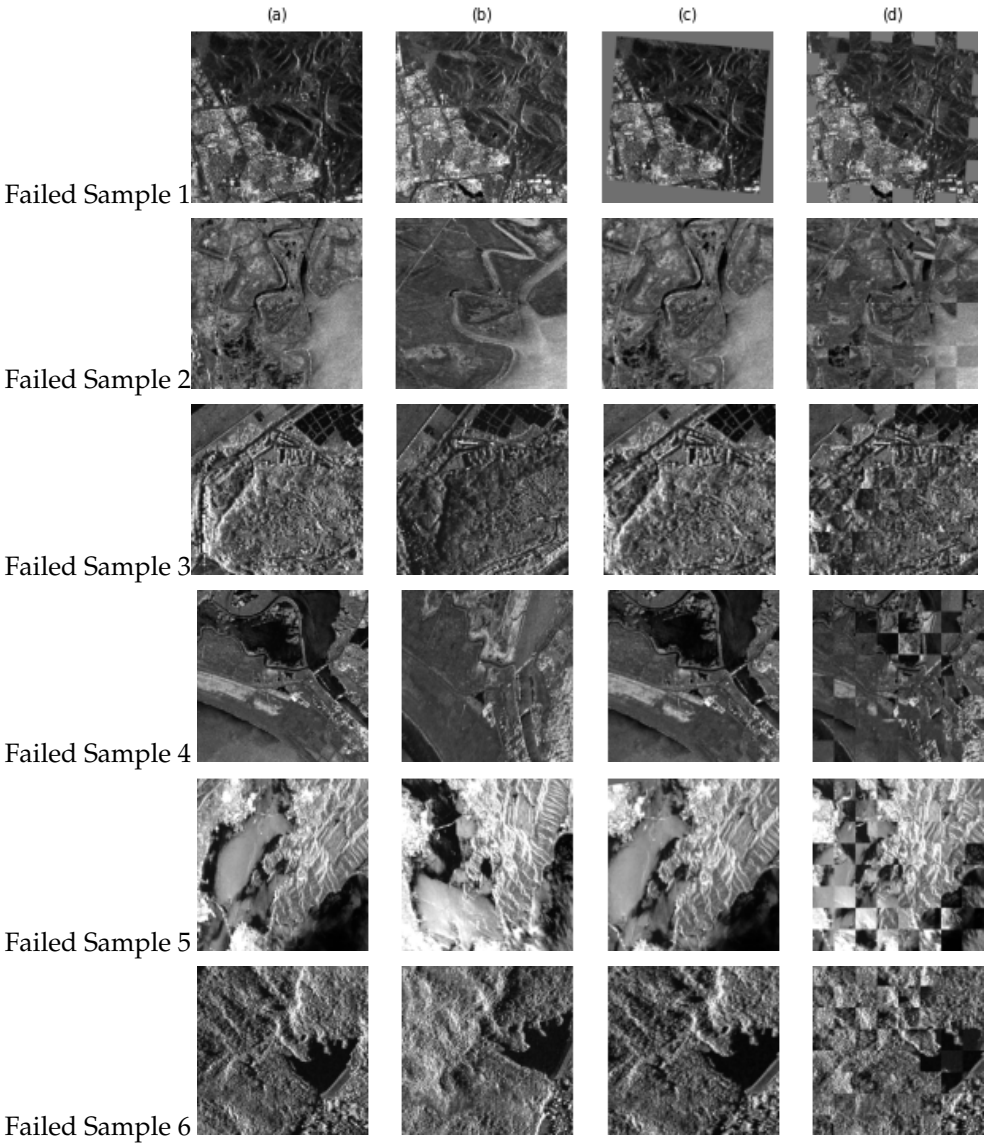


Figure 20. Failed registration samples.

The mosaic image of the sensed image and the reference image after the registration can exhibit pronounced gray scale misregistration at the splicing location, as seen in the failed Sample 5 or show insufficient smoothness in the splicing of line features and region features, as seen in the failed Sample 4. The registration scheme based on the entire image requires a certain level of gray scale and structural similarity between the sensed image and the reference image pairs to be registered. However, the reinforcement learning approach in this paper did not perform gray scale transformations on the SAR image during the transformation of the sensed image. Consequently, its applicability and scalability for registering SAR image pairs with substantial gray scale differences need to be enhanced.

Since this paper employed a self-learning method to generate training samples by applying various affine transformations to the training data, the feature differences between the sensed image and the reference image pairs that cannot be reflected by the affine transformations included significant geometric distortions, e.g., the failed Sample 3 and the failed Sample 6, and gray scale changes, e.g., the failed Sample 4. Such differences impact the correlation between image pairs, potentially leading to suboptimal performance of the registration method based on the correlation matching network presented in this paper. To address this challenge, a potential solution involves manually labeled matching labels and

a machine learning network that performs gray scale transformations, image matching, and evaluation.

Furthermore, the failed Sample 1 highlighted the significant challenge of dealing with complex and variable ground topographic features. In this scenario, the feature point pairs obtained by SAR-RL were sparse and unevenly distributed in the images to be registered, affecting the global registration performance of the image pairs. Some of the failure cases involved images with densely distributed point features and line features. This can be attributed to the fact that the SAR-RL in this paper is a global registration method, and the uniform distribution of weights for the globally densely distributed point features and line features can easily result in the failure of the registration.

5. Conclusions

In this paper, we introduced the OptionEM-based reinforcement learning framework to achieve end-to-end SAR image registration. This architecture outputs registered images and affine transformation matrices directly, leading to significant processing time savings compared to multi-step registration algorithms. The proposed framework combines the advantages of Episodic Memory and hierarchical reinforcement learning, which can mitigate the inherent problem of invalid exploration in generalized reinforcement learning algorithms and effectively combine the coarse and fine registration, further improving training efficiency. Furthermore, reinforcement learning can dynamically correct errors, making it more efficient and robust compared to supervised learning mechanisms such as deep learning. The experimental results showed that our method enables sub-second, high-precision SAR image registration. This gives the framework potential for deployment on portable devices.

6. Discussion

Registering SAR images of terrain with undulating regions presents a significant challenge due to inherent noise and geometric distortion in SAR images. The SAR-RL method presented in this paper adopts a global registration approach, but there are instances where registering SAR images of complex terrain regions encounters difficulties, particularly when the training dataset is generated through a self-learning scheme.

Registering two markedly dissimilar images poses a formidable challenge, as elucidated in our examination of the failure instances in Section 4.4.6. When the disparities between the two images slated for registration lie beyond the spectrum of variations encompassed by the training data, it may be necessary to resort to initial registration techniques such as geometric registration [5,50,51], block-based registration [52], or the integration of prior information, such as Digital Elevation Models (DEMs) and Ground Control Points (GCPs). In scenarios where neural-network-based approaches persist in being utilized, it becomes imperative to expand the scope of variations encompassed by the training dataset. Nevertheless, this expansion carries the potential downside of heightened susceptibility to mismatches and a corresponding decrease in the stability of the registration process.

Future research endeavors will be directed towards enhancing the SAR-RL method. This will involve incorporating a priori knowledge of the dataset distribution and optimizing training datasets by introducing diverse distribution patterns and supervision information. These efforts aim to enhance the accuracy and efficiency of the method across various datasets. In addition, methods such as QNN [24] are considered to have the potential to improve the registration performance of SAR images in complex terrain. Simultaneously, there is a consideration for adapting and deploying the trained neural networks to portable devices using TensorFlow Lite. Such an advancement would enable real-time applications by researchers and potentially even non-specialists and enhance its suitability for long-term monitoring and swiftly predicting changes of key areas of concern, such as remote, high-mountain environments [53].

In addition, there are the PolSAR multi-scattering characteristics of different ground objects through polarization features. The degree of polarization highlights the differences

in the incidence angle and surface scattering characteristics, which are key factors affecting the variations in SAR images [54]. Building upon the insights of Usami, N. et al. [54], we aim to explore the feature changes in SAR images more comprehensively, leading to a deeper understanding of our planet.

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

Funding: This research was funded by the National Natural Science Foundation of China Grant Number U2241202.

Data Availability Statement: The sources of "TerraSAR-X Napa 2014" and "Sentinel-1A Napa 2014" are contained within the article.

Acknowledgments: Thanks to Jie Chen from Beihang University for providing a high-performance computer to product the training and test datasets.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; nor in the decision to publish the results.

References

1. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
2. Dellinger, F.; Delon, J.; Gousseau, Y.; Michel, J.; Tupin, F. SAR-SIFT: A SIFT-like algorithm for SAR images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 453–466. [CrossRef]
3. Pan, B.; Jiao, R.; Wang, J.; Han, Y.; Hang, H. SAR image registration based on KECA-SAR-SIFT operator. In Proceedings of the 2022 2nd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Nanjing, China, 23–25 September 2022; pp. 114–119.
4. Hossein-Nejad, Z.; Nasri, M. Image Registration Based on Redundant Keypoint Elimination SARSIFT Algorithm and MROGH Descriptor. In Proceedings of the 2022 International Conference on Machine Vision and Image Processing (MVIP), Ahvaz, Iran, 23–24 February 2022; pp. 1–5.
5. Wang, M.; Zhang, J.; Deng, K.; Hua, F. Combining optimized SAR-SIFT features and RD model for multisource SAR image registration. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–16. [CrossRef]
6. Yu, Q.; Wu, P.; Ni, D.; Hu, H.; Lei, Z.; An, J.; Chen, W. SAR pixelwise registration via multiscale coherent point drift with iterative residual map minimization. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–19. [CrossRef]
7. Wang, S.; Quan, D.; Liang, X.; Ning, M.; Guo, Y.; Jiao, L. A deep learning framework for remote sensing image registration. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 148–164. [CrossRef]
8. Chang, Y.; Xu, Q.; Xiong, X.; Jin, G.; Hou, H.; Man, D. SAR image matching based on rotation-invariant description. *Sci. Rep.* **2023**, *13*, 14510. [CrossRef]
9. Pourfard, M.; Hosseinian, T.; Saeidi, R.; Motamedi, S.A.; Abdollahifard, M.J.; Mansoori, R.; Safabakhsh, R. KAZE-SAR: SAR image registration using KAZE detector and modified SURF descriptor for tackling speckle noise. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 1–12. [CrossRef]
10. Quan, D.; Wang, S.; Ning, M.; Xiong, T.; Jiao, L. Using deep neural networks for synthetic aperture radar image registration. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 2799–2802.
11. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *arXiv* **2017**, arXiv:1706.03762. https://doi.org/10.48550/arXiv.1706.03762
12. Gonçalves, H.; Gonçalves, J.A.; Corte-Real, L. Measures for an objective evaluation of the geometric correction process quality. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 292–296. [CrossRef]
13. Mao, S.; Yang, J.; Gou, S.; Jiao, L.; Xiong, T.; Xiong, L. Multi-Scale Fused SAR Image Registration Based on Deep Forest. *Remote Sens.* **2021**, *13*, 2227. [CrossRef]
14. Jaderberg, M.; Simonyan, K.; Zisserman, A.; Kavukcuoglu, K. Spatial transformer networks. *arXiv* **2015**, arXiv:1506.02025. https://doi.org/10.48550/arXiv.1506.02025
15. Chen, J.; Huang, Z.; Xia, R.; Wu, B.; Sheng, L.; Sun, L.; Yao, B. Large-scale multi-class SAR image target detection dataset-1.0. *J. Radars* **2022**. Available online: https://radars.ac.cn/web/data/getData?dataType=MSAR (accessed on 22 April 2018).
16. Xia, R.; Chen, J.; Huang, Z.; Wan, H.; Wu, B.; Sun, L.; Yao, B.; Xiang, H.; Xing, M. A Visual Transformer Based on Contextual Joint Representation Learning for SAR Ship Detection. *Remote Sens.* **2022**, *14*, 1488. [CrossRef]

17. Schwegmann, C.P.; Kleynhans, W.; Salmon, B. The development of deep learning in synthetic aperture radar imagery. In Proceedings of the 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP), Shanghai, China, 18–21 May 2017; pp. 1–2.
18. Jianxu, M. Research on Three-Dimensional Imaging Processing Techniques for Synthetic Aperture Radar Interferometry (InSAR). Ph.D. Thesis, Hunan University, Changsha, China, 2002.
19. Chang, H.H. Remote Sensing Image Registration Based upon Extensive Convolutional Architecture with Transfer Learning and Network Pruning. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [CrossRef]
20. Jie, R. Key Technology Research for Cartographic Applications of Multi-Source Remote Sensing Data. Ph.D. Thesis, University of Chinese Academy of Sciences (Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences), Beijing, China, 2017.
21. Yide, M.; Lian, L.; Yafu, W.; Ruolan, D. The Principles and Applications of Pulse-Coupled Neural Networks. 2006. Available online: https://item.jd.com/10052980.html (accessed on 22 April 2018).
22. Del Frate, F.; Licciardi, G.; Pacifici, F.; Pratola, C.; Solimini, D. Pulse Coupled Neural Network for automatic features extraction from COSMO-Skymed and TerraSAR-X imagery. In Proceedings of the 2009 IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, 12–17 July 2009; Volume 3, pp. III-384–III-387.
23. Zhao, C. SAR Image Registration Method Based on SAR-SIFT and Deep Learning. Master's Thesis, Xidian University, Xi'an, China, 2017.
24. Shang, F.; Hirose, A. Quaternion neural-network-based PolSAR land classification in Poincare-sphere-parameter space. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 5693–5703. [CrossRef]
25. Hu, J.; Lu, J.; Tan, Y.P. Sharable and individual multi-view metric learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 2281–2288. [CrossRef] [PubMed]
26. Ye, Y.; Tang, T.; Zhu, B.; Yang, C.; Li, B.; Hao, S. A multiscale framework with unsupervised learning for remote sensing image registration. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [CrossRef]
27. Quan, D.; Wei, H.; Wang, S.; Li, Y.; Chanussot, J.; Guo, Y.; Hou, B.; Jiao, L. Efficient and Robust: A Cross-modal Registration Deep Wavelet Learning Method for Remote Sensing Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, 4739-4754. [CrossRef]
28. Fan, Y.; Wang, F.; Wang, H. A Transformer-Based Coarse-to-Fine Wide-Swath SAR Image Registration Method under Weak Texture Conditions. *Remote Sens.* **2022**, *14*, 1175. [CrossRef]
29. Li, B.; Guan, D.; Zheng, X.; Chen, Z.; Pan, L. SD-CapsNet: A Siamese Dense Capsule Network for SAR Image Registration with Complex Scenes. *Remote Sens.* **2023**, *15*, 1871. [CrossRef]
30. Deng, X.; Mao, S.; Yang, J.; Lu, S.; Gou, S.; Zhou, Y.; Jiao, L. Multi-Class Double-Transformation Network for SAR Image Registration. *Remote Sens.* **2023**, *15*, 2927. [CrossRef]
31. Mao, S.; Yang, J.; Gou, S.; Lu, K.; Jiao, L.; Xiong, T.; Xiong, L. Adaptive Self-Supervised SAR Image Registration with Modifications of Alignment Transformation. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–15. [CrossRef]
32. Liu, M.; Zhou, G.; Ma, L.; Li, L.; Mei, Q. SIFNet: A self-attention interaction fusion network for multisource satellite imagery template matching. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *118*, 103247. [CrossRef]
33. Chen, J.; Chen, X.; Chen, S.; Liu, Y.; Rao, Y.; Yang, Y.; Wang, H.; Wu, D. Shape-Former: Bridging CNN and Transformer via ShapeConv for multimodal image matching. *Inf. Fusion* **2023**, *91*, 445–457. [CrossRef]
34. Zou, B.; Li, H.; Zhang, L. Self-Supervised SAR Image Registration With SAR-Superpoint and Transformation Aggregation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *61*, 1–15. [CrossRef]
35. Zhao, M.; Zhang, G.; Ding, M. Heterogeneous self-supervised interest point matching for multi-modal remote sensing image registration. *Int. J. Remote Sens.* **2022**, *43*, 915–931. [CrossRef]
36. Quan, D.; Wei, H.; Wang, S.; Gu, Y.; Hou, B.; Jiao, L. A Novel Coarse-to-Fine Deep Learning Registration Framework for Multi-Modal Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 1–16. [CrossRef]
37. Quan, D.; Wang, S.; Gu, Y.; Lei, R.; Yang, B.; Wei, S.; Hou, B.; Jiao, L. Deep feature correlation learning for multi-modal remote sensing image registration. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [CrossRef]
38. Li, L.; Han, L.; Ye, Y. Self-supervised keypoint detection and cross-fusion matching networks for multimodal remote sensing image registration. *Remote Sens.* **2022**, *14*, 3599. [CrossRef]
39. Xiang, D.; Xu, Y.; Cheng, J.; Xie, Y.; Guan, D. Progressive Keypoint Detection with Dense Siamese Network for SAR Image Registration. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *59*, 5847–5858. [CrossRef]
40. Blundell, C.; Uria, B.; Pritzel, A.; Li, Y.; Ruderman, A.; Leibo, J.Z.; Rae, J.; Wierstra, D.; Hassabis, D. Model-free episodic control. *arXiv* **2016**, arXiv:1606.04460.
41. Pritzel, A.; Uria, B.; Srinivasan, S.; Badia, A.P.; Vinyals, O.; Hassabis, D.; Wierstra, D.; Blundell, C. Neural episodic control. In Proceedings of the International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 2827–2836.
42. Savinov, N.; Raichuk, A.; Marinier, R.; Vincent, D.; Pollefeys, M.; Lillicrap, T.; Gelly, S. Episodic curiosity through reachability. *arXiv* **2018**, arXiv:1810.02274.
43. Lin, Z.; Zhao, T.; Yang, G.; Zhang, L. Episodic memory deep q-networks. *arXiv* **2018**, arXiv:1805.07603.
44. Hu, H.; Ye, J.; Zhu, G.; Ren, Z.; Zhang, C. Generalizable Episodic Memory for deep reinforcement learning. *arXiv* **2021**, arXiv:2103.06469.
45. Rocco, I.; Arandjelovic, R.; Sivic, J. Convolutional neural network architecture for geometric matching. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6148–6157.

46. Zhou, R.; Zhang, Z.; Wang, Y. Hierarchical Episodic Control. *Preprints* **2023**, 1–18. [CrossRef]
47. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
48. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; pp. 448–456.
49. He, H.; Chen, M.; Chen, T.; Li, D. Matching of remote sensing images with complex background variations via Siamese convolutional neural network. *Remote Sens.* **2018**, *10*, 355. [CrossRef]
50. Xiang, Y.; Jiao, N.; Liu, R.; Wang, F.; You, H.; Qiu, X.; Fu, K. A Geometry-Aware Registration Algorithm for Multiview High-Resolution SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–18. [CrossRef]
51. Li, X.; Wang, T.; Cui, H.; Zhang, G.; Cheng, Q.; Dong, T.; Jiang, B. SARPointNet: An automated feature learning framework for spaceborne SAR image registration. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 6371–6381. [CrossRef]
52. Pallotta, L.; Clemente, C.; Borreca, T.; Giunta, G.; Soraghan, J.J. A joint coregistration of rotated multitemporal SAR images based on the cross-cross-correlation. In Proceedings of the International Conference on Radar Systems (RADAR 2022), Edinburgh, UK, 24–27 October 2022.
53. Shugar, D.H.; Jacquemart, M.; Shean, D.; Bhushan, S.; Upadhyay, K.; Sattar, A.; Schwanghart, W.; McBride, S.; De Vries, M.V.W.; Mergili, M.; et al. A massive rock and ice avalanche caused the 2021 disaster at Chamoli, Indian Himalaya. *Science* **2021**, *373*, 300–306. [CrossRef] [PubMed]
54. Usami, N.; Muhuri, A.; Bhattacharya, A.; Hirose, A. Proposal of wet snowmapping with focus on incident angle influential to depolarization of surface scattering. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 1544–1547.

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.