Remote Sensing of Environment 318 (2025) 114577

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Ground surface displacement measurement from SAR imagery using deep learning

Jinwoo Kim [a], Hyung-Sup Jung [b], Zhong Lu [a,c,*]

[a] Roy M. Huffington Department of Earth Sciences, Southern Methodist University, Dallas, TX, USA
[b] Department of Geoinformatics, University of Seoul, Seoul, South Korea
[c] School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou, Jiangsu 221008, China

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
Offset tracking
SAR
Deep learning
Siamese ResNet
Displacement
Glacier
Earthquakes
Volcanoes

ABSTRACT

Offset tracking using synthetic aperture radar (SAR) amplitude imagery is a valuable technique for detecting large ground displacements. However, the traditional offset tracking methods with the SAR datasets are computationally intensive and require significant time for processing. We have developed a novel cross-connection Siamese ResNet (CC-ResSiamNet). The model leverages multi-kernel offset tracking for preprocessing, followed by deep learning architectures that incorporate U-Net, cross-connections, and residual and attention blocks to predict pixel offsets between two SAR amplitude images. It is trained and tested on 200 K pairs of reference and secondary SAR amplitude images, alongside corresponding target offset data from Alaska's glaciers. The comparative analysis with multiple deep learning models confirmed that our designed model is highly generalizable, achieving rapid convergence, minimal overfitting, and high prediction accuracy. Through multi-scenario inference with glacier movements, earthquakes, and volcanic eruptions worldwide, the model demonstrates strong performance, closely matching the accuracy of traditional methods while offering significantly faster processing times through parallel computing. The model's rapid displacement mapping capability shows particular promise for improving disaster response and near real-time surface monitoring. While the approach encounters challenges in accurately capturing small-scale displacements, it opens new possibilities for SAR-based surface displacement prediction using machine learning. This research highlights the advantages of combining deep learning with SAR imagery for advancing geophysical analysis, with future applications anticipated as more commercial and scientific SAR missions launch globally.

1. Introduction

With its day and night imaging capabilities using long-wavelength radar (typically between 1 cm and 1 m), synthetic aperture radar (SAR) datasets have been widely utilized for monitoring the Earth's surface from space and air regardless of weather conditions (Curlander and McDonough, 1991; Hanssen, 2021). Its applications include classifying land use, civilian and military surveillance, mapping areas affected by disasters (e.g., floods, volcanoes, earthquakes, and landslides), and estimating ground deformation caused by anthropogenic activities (e.g., groundwater extraction and transportation construction) and natural events (Handwerger et al., 2019; Qu et al., 2015; Kim and Lu, 2021). In the past, spaceborne SAR missions such as Shuttle Radar Topography Mission (SRTM), TerraSAR-X, European Remote-sensing Satellite (ERS), Environmental satellite (Envisat), and Japanese Earth Resources Satellite (JERS) were conducted primarily by space agencies in only a few countries, including the United States, Germany, Europe, and Japan (Lu and Dzurisin, 2014). However, an increasing number of countries, such as Argentina, India, China, Spain, and South Korea, have recently aimed to launch their own or joint SAR missions for scientific, public, and surveillance purposes. Additionally, more commercial SAR satellites are being operated by companies to profit from selling images to public sectors (Castelletti et al., 2021; Ignatenko et al., 2020). With an expanding availability of SAR images and upcoming missions, SAR has gained popularity among scientists, government agencies, and the general public.

Ground displacement monitoring stands out as a primary objective among SAR's myriad applications (Massonnet and Feigl, 1998), pivotal for enhancing our comprehension of Earth's surface dynamics and bolstering disaster preparedness and response efforts. SAR datasets, each

* Corresponding author.
E-mail address: zhonglu@smu.edu (Z. Lu).

https://doi.org/10.1016/j.rse.2024.114577
Received 28 September 2024; Received in revised form 13 December 2024; Accepted 15 December 2024
Available online 21 December 2024

pixel comprising complex values produced from signal processing of radar returns, yield two key outputs: phase and amplitude, crucial for gauging ground displacement (Rosen et al., 2000). Leveraging interferometric SAR (InSAR), which harnesses phase data of complex values in SAR datasets from multiple acquisitions to generate phase difference images (known as interferograms), yields highly precise estimates of displacement, often reaching centimeter-level accuracy across expansive areas (Rosen et al., 2000; Lu and Dzurisin, 2014). The evolution of InSAR technologies, notably persistent scatterers (PS) and distributed scatterers (DS) methods (Berardino et al., 2002; Hooper et al., 2004; Ferretti et al., 2011), has significantly broadened its utility, enabling comprehensive mapping of ground subsidence in urban locales, tracking surface deformations induced by natural calamities like landslides (Samsonov and Blais-Stevens, 2024; Xia et al., 2024), earthquakes (Fathian et al., 2021), and volcanic activity (Tizzani et al., 2024), assessing infrastructure integrity (Hübinger et al., 2024), and estimating soil moisture fluctuations through non-zero phase triplets (Lu and Kim, 2021; Lazecký et al., 2020; Michaelides et al., 2019). However, effective utilization of phase difference hinges upon the assumption of high InSAR coherence which requires minimal alterations in surface scattering characteristics between acquisitions. This assumption, however, can be easily compromised when surface conditions undergo significant alteration surpassing InSAR's decorrelation limit. In contrast to InSAR, which relies on phase difference, offset tracking (also known as incoherent speckle tracking) draws on amplitude correlation primarily from multiple acquisitions. Its fundamental premise is to trace the peak correlation locations between SAR amplitude images, computed as two-dimensional offsets in both range and azimuth directions for radar coordinates, and east-west and north-south directions for geocoded SAR images. Offset tracking offers a valuable technique for detecting large gradients of displacement induced by seismic activity (Fielding et al., 2020), volcanic eruptions (Bato et al., 2021), landslides (Xu et al., 2020a; Jia et al., 2020), and glacier movements (Strozzi et al., 2002; Feng et al., 2023), with an accuracy of sub-meter level depending on pixel size and correlation condition (Lu and Zhang, 2014). To achieve sub-pixel precision in displacement measurements, oversampling of input SAR amplitude images and cross-correlation via Fourier transform in the frequency domain are typically necessary for offset tracking measurement (Lei et al., 2021). However, both of these processing steps entail significant computational resources and can lead to slower processing speeds. Generating a full-resolution offset map (window size >200) from a pair of a single burst coregistered single look complex Sentinel-1 (CSLC-S1) (dimension: 4800 × 20,000) takes approximately 2 h even with the assistance of graphics processing unit (GPU) and parallel processing.

Deep learning, a subset of machine learning, has revolutionized problem-solving through artificial neural networks and algorithms inspired by the intricate structure and functionality of the human brain. With significant advancements in image recognition, language translation, and medical diagnosis, the scope of deep learning has expanded to include applications in natural language processing (Li, 2018), computer vision (Lemley et al., 2017), healthcare (Miotto et al., 2018), and beyond (Najafabadi et al., 2015). In the domain of machine learning, a diverse array of algorithms and techniques including stochastic gradient descent and backpropagation constitutes the foundational processes that drive these innovations (Goodfellow et al., 2016).

Deep learning has been extensively applied in remote sensing, with significant advancements in key areas such as land use classification, change detection, and disaster assessment (Zhu et al., 2017). In particular, deep learning techniques have been utilized for image enhancement through speckle noise reduction using convolutional neural networks (CNNs) and super-resolution methods (Lattari et al., 2019; Shen et al., 2020), as well as for object detection and classification tasks such as land cover classification and ship detection (Parikh et al., 2019; Li et al., 2022). Other applications include change detection for urban monitoring and disaster impact analysis (Saha et al., 2021; Wu et al., 2023), multi-sensor fusion (Adrian et al., 2021; Hughes and Marcos, 2020), and phase unwrapping (Spoorthi et al., 2020). These applications typically address two main prediction tasks: *1*) classification, which uses supervised learning to categorize input data into predefined classes based on their distinctive features, and *2*) regression, which predicts continuous values by modeling the relationship between input variables and outcomes (Zhang et al., 2023). Many remote sensing tasks have traditionally focused on binary classifications, such as distinguishing water from non-water areas or detecting changes vs. no changes, as well as multi-label classifications, such as identifying roads, lakes, and agricultural fields from optical or radar satellite images, or fused data (Zhu et al., 2017). In contrast, regression tasks, though less common in remote sensing, present unique challenges due to the complexity of predicting continuous values over a wide range, unlike discrete class labels used in classification. While the deep learning approach was applied to estimate sub-pixel ground displacements from earthquake fault ruptures using optical satellite images (Montagnon et al., 2024), an important yet underexplored challenge in regression is estimating ground displacement from SAR imagery—a vital but complex task in SAR data analysis. This study investigates the potential of deep learning techniques to estimate ground surface displacements from SAR imagery, utilizing offset tracking results as the basis.

Traditional methods like InSAR and offset tracking face challenges such as coherence loss and high computational demands, including time-frequency transformation, sub-pixel oversampling, and peak offset detection. To overcome these limitations, we propose an innovative deep learning regression model capable of measuring ground surface displacement instantaneously from SAR imagery. This approach aims to overcome the computational limitations of traditional methods while maintaining accuracy, potentially revolutionizing rapid displacement mapping for disaster response and Earth Science applications.

2. Methodology and implementation

This section outlines a comprehensive workflow for estimating ground displacement using deep learning methods applied to SAR imagery. Our approach integrates four key components: data preparation, segmentation, model architecture, and training and evaluation (Fig. 1). The process begins with the generation of CSLC-S1 bursts. We then apply a multi-kernel offset tracking algorithm to these bursts to estimate easting and northing offsets. These offset estimates, combined with SAR amplitude images, serve as input data for our deep learning models. Labeling the dataset in advance is a fundamental step in deep learning, as it enables the model to learn and predict offsets based on the provided ground truth. In this study, we assumed that the offset tracking results derived from SAR images serve as the true offsets. Although these results are not flawless and may include noise or errors, they provide a practical and consistent proxy for ground truth. This approach allows for dense and reliable offset estimates across large areas, which are essential for effectively training the deep learning model. Central to our methodology is a novel deep learning architecture specifically designed for ground displacement estimation. This model is trained to predict surface displacement using the prepared input data. We employ an iterative training process with multiple sets of input data to refine the model's performance. Model outcomes are evaluated using unseen test sets to assess generalization and prediction accuracy. This evaluation process ensures the robustness and reliability of our approach. In the following subsections, we provide a detailed explanation of the specific methodologies, tools, and techniques applied at each step.

2.1. Data preparation

To ensure an ample and diverse dataset, we have chosen the observational products for end-users from remote sensing analysis (OPERA) CSLC-S1 as the primary data source for implementing our deep learning model. The burst-wise CSLC-S1 provides geocoded in universal

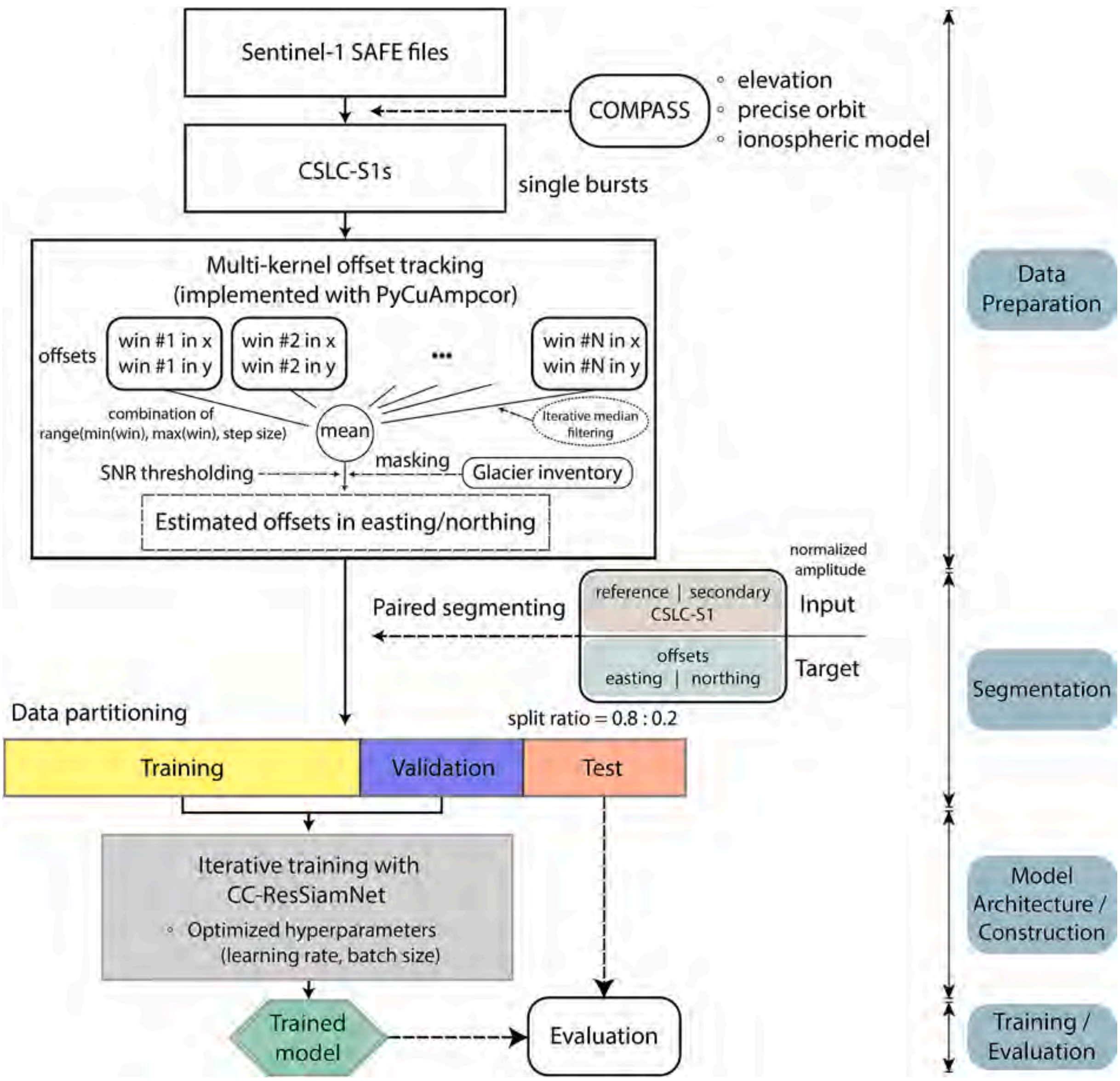

Fig. 1. Overview of the workflow for estimating ground displacement from Sentinel-1 SAR imagery using deep learning. The workflow consists of four main stages, data preparation, segmentation, model construction, and training and evaluation.

transverse Mercator (UTM) coordinates and co-registered SAR images achieved through geometric coregistration utilizing sensor orbit ephemeris, atmospheric models and digital elevation model (DEM) (Brancato, 2023). Sentinel-1's precise orbit determination products obtained from Copernicus Data Space Ecosystem (https://dataspace.copernicus.eu), global ionosphere maps (GIM) from NASA's crustal dynamics data information system (CDDIS; https://cddis.nasa.gov), and Copernicus GLO-30 DEM (https://spacedata.copernicus.eu/collections/copernicus-digital-elevation-model) is used. Additionally, various timing corrections, including Doppler-induced range shift correction, bistatic azimuth delay correction, and atmospheric delay correction, have been applied to achieve adequate geolocation accuracy (Gisinger et al., 2020). Moreover, the mass production of CSLC-S1 from European Space Agency (ESA)'s Sentinel-1 SLC standard archive format for Europe (SAFE) files is facilitated within a short time frame, thanks to the utilization of GPUs embedded in the coregistered multi-temporal SAR SLC (COMPASS) software package (https://github.com/opera-adt/COMPASS).

The research area encompasses Alaska, United States, and the neighboring Yukon Territory of Canada, featuring a multitude of dynamic glaciers (Fig. 2(a)). Furthermore, this region serves as a focal point for NASA's making earth system data records for use in research environments (MEaSUREs) project, ITS_LIVE (https://its-live.jpl.nasa.gov), where Sentinel-1 datasets are being used to track glacier movement. The correlation of SAR amplitude images for offset tracking can generally be highly sustained within a time frame of approximately two weeks. However, this correlation swiftly diminishes as the interval between two acquisitions increases (Kim et al., 2022). Therefore, we opted for CSLC-S1 acquisitions at 12-day intervals (shortest revisits for Sentinel-1 in Alaska), spanning from June to October, to mitigate the impact of snow cover effects in the region.

After generating the CSLC-S1 datasets using COMPASS software, we proceed to estimate two-dimensional offsets between reference and secondary SAR amplitude images. This is accomplished by identifying the location of maximum cross-correlation through a grid search of two patches within the amplitude images. Given that the sensitivity and measurement density of offset tracking results may vary depending on the dimension of template patches (referred to hereafter as kernel size), we employ a multi-kernel offset tracking approach (Chae et al., 2019). This method averages multiple offset tracking results obtained using a

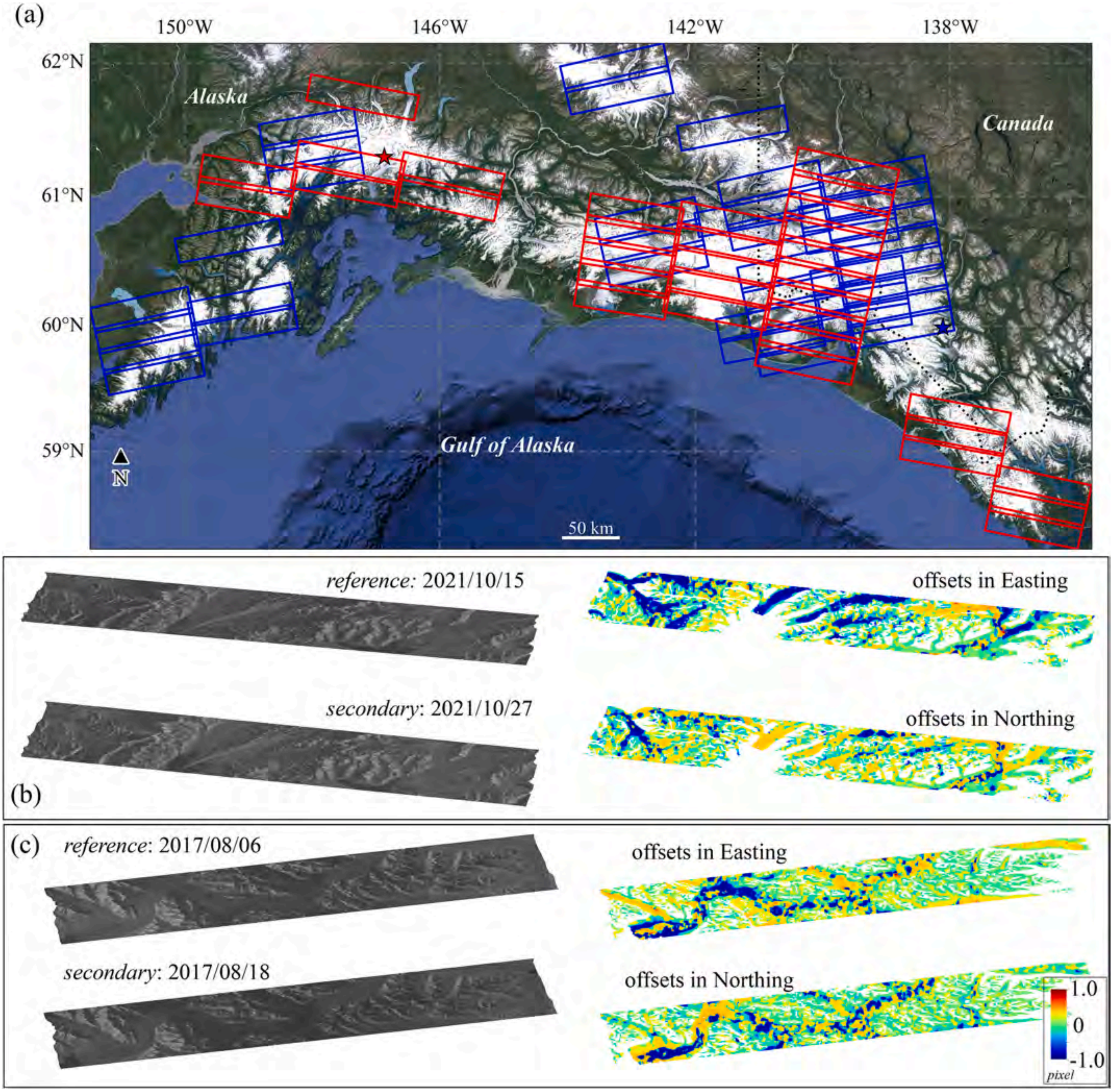

Fig. 2. (a) Burst-wise CSLC-S1 in Alaska and the Yukon Territory used for developing our deep learning model. The CSLC-S1s in red and blue rectangles are from descending and ascending tracks, respectively. (b, c) Exemplary sets of CSLC-1 reference and secondary amplitude images and corresponding offset tracking results in the easting and northing directions from descending and ascending tracks, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

range of kernel sizes, empirically selected from 96 to 230 with 32 intervals in both east-west and north-south directions. For each kernel outcome, iterative median filtering is applied using three adaptive windows (7 × 7, 5 × 5, 3 × 3 in descending order) to smooth the offset results, while signal-to-noise ratio (SNR) thresholding is applied to remove outliers. Utilizing the GPU-powered PyCuAmpcor package (https://github.com/lijun99/cuAmpcor) facilitates efficient computation. With the optimal utilization of GPU cores, generating a full-resolution offset map from a pair of CSLC-S1 images takes between 16 min to 2 h using the package, depending on the kernel size (96 and 230, respectively).

To preserve the original pixel spacings (5 and 10 m in easting and northing, respectively) of CSLC-S1, offset estimation is performed at every pixel with a step size of 1 (e.g., Fig. 2(b, c)). To streamline the training process, we convert the offset estimation unit from meter to pixel. Although the offset tracking results retain the same dimensions as the input CSLC-S1 following this strategy, it is worth noting that the results may look noisier as offset tracking is conducted at every pixel. Maintaining the original pixel spacing from the input CSLC-S1 to the offset tracking results can ultimately aid deep learning methods in efficiently identifying spatial patterns, even if the offsets to be trained exhibit noise. Furthermore, the Randolph glacier inventory (RGI) 7.0 dataset (RGI 7.0 Consortium, 2023), which outlines global glaciers and is compiled by the global land ice measurements from space (GLIMS) initiative, is employed to mask glacier regions from the offset tracking results. The original offset results for the Alaska areas encompass not only glacier movements but also displacements from adjacent topographic features, including mountain slopes, forests, and permafrost areas. While the major glaciers of Alaska, as delineated in RGI dataset, are located on the relatively gradual slopes, narrow valley glaciers interspersed with steep terrain are excluded from further training to reduce potential noise. These surrounding features are prone to geometric distortions inherent in SAR imaging, such as layover, shadow, and foreshortening, and their surface changes are relatively small in size and magnitude. Our deep learning model is specifically designed to learn ground surface displacements, with the consistently flowing glaciers providing the most reliable data for this purpose. By focusing on these large, more stable glacial features, we aim to reduce the impact of noise and SAR imaging artifacts, thereby improving the overall accuracy of our displacement estimation.

2.2. Data partitioning

The datasets, comprising reference and secondary SAR amplitude images along with their corresponding offsets in easting and northing, are partitioned into training/validation and test sets using a ratio of 0.8 to 0.2. This split ratio is maintained for both the training and validation sets. The amplitudes of reference and secondary SAR imagery are converted into decibel unit and normalized to [0,1]. Subsequently, 256 × 256 chips of the amplitudes and offsets are cropped from pixel locations that encompass at least 90 % of glaciers, as determined by the glacier inventory. Because each chip consists predominantly of glacier areas, we can effectively reduce the effects of displacements and noise induced by neighboring surface features such as mountain slopes and permafrost. Our offset tracking method, implemented using windows across tens or hundreds of pixels, may inadvertently include unwanted displacements when estimating offsets in glaciers adjacent to bare earth slopes or other non-glacial features. However, masking the offsets with the RGI dataset remains the best available approach to mitigate such effects. Additionally, to further reduce potential boundary effects during the training process, we apply a masking approach that excludes the outer 16 pixels along all edges (top, bottom, left, right) of the offset chips. Note that 16 pixels is a conservative choice to mitigate artifacts caused by out-of-boundary offsets. Depending on the displacement characteristics, a larger number may be required for fast-moving displacements, while a smaller number may suffice for slow-moving displacements.

To address the imbalance issue arising from different look directions, we employed a balanced dataset comprising an equal number of offset tracking result pairs (248 pairs each) from both ascending (36 CSLC-S1s) and descending (43 CSLC-S1s) tracks (Fig. 2(a)). This approach ensures equitable representation of both orbital configurations in our analysis. Given the substantial volume of data, which exceeded available computational memory for simultaneous processing, we implemented a strategic partitioning protocol. The dataset was systematically divided into four discrete subsets, each structured as 32 K, 8 K, and 10 K samples for training, validation, and testing, respectively. This partitioning scheme adheres to standard deep learning practices, allocating approximately 64 % of the data for training, 16 % for validation, and 20 % for testing. Such distributions ensure robust model development, effective hyperparameter tuning and unbiased performance evaluation. By maintaining consistent proportions across all subsets, we preserve the statistical properties of the original dataset while accommodating hardware limitations. This approach facilitates efficient data handling and promotes reproducibility in our experimental framework.

2.3. Deep learning model construction

In deep learning for remote sensing, an optimized neural network model is crucial for making accurate predictions on unseen data, highlighting the importance of selecting the right model for efficient performance and desired outcomes. The model is composed of multiple interconnected modules that process upstream inputs to generate outputs, which are then used for gradient descent optimization and learning of internal parameters. These modules and their interconnections are strategically designed to produce accurate offset predictions from the input SAR amplitude images.

2.3.1. Modules

Our model comprises three primary modules: residual, attention, and last convolution (Conv) block. Residual neural network (ResNet) can address the vanishing gradient problem with introducing skip connections that directly connect the input of a layer to the output, bypassing one or more intermediate layers (He et al., 2016). Our residual block proposed in this study (Fig. 3(a)), modified from the original ResNet, features a shared flow originating from two inputs (x_1^{in}, x_2^{in}) given f_n (number of filters: varying from 32 to 1024 in our model), followed by a concatenation layer, two Conv blocks (Conv2D with k (kernel size) = (3,3), batch normalization (BN), activation), and one Conv block (Conv2D with k = (3,3), BN) without an activation layer. Each input undergoes a Conv block (with k = (3,3) or (1,1)), with the previous outputs added together to produce two outputs (x_1^{out}, x_2^{out}). Swish function ($\frac{x}{1+e^{-x}}$) is chosen as the activation function in the residual blocks due to its superior performance compared to the rectified linear unit (ReLU). This residual block plays a fundamental role in our deep learning model, facilitating effective training using model inputs.

The attention mechanism in deep learning enables the model to selectively attend to pertinent aspects of the inputs and effectively capture their relative importance. Our attention block (Fig. 3(b)) incorporates both channel and spatial attention mechanisms (Woo et al., 2018). Channel-wise attention focuses on discerning the meaningful channels within an input by applying max-pooing and average-pooling separately, followed by a multi-layer perceptron (Fig. 3(b)). The outputs of these pooing operations are then added and passed through an active layer with a sigmoid function ($\frac{1}{1+e^{-x}}$). For an intermediate feature map $F \in \mathbb{R}^{H \times W \times C}$ (H, W, C are height, width, and channel, respectively. In our model, (H, W, C) = (256,256,32)) as input, the output of the channel attention is $\mathbb{R}^{1 \times 1 \times C}$, which explores the inter-channel relationship of features. Spatial attention, on the other hand, focuses on identifying informative spatial regions within the features. This is achieved by concatenating the results of max-pooling and average-pooling, followed by a convolution layer (Conv2D with f_n = 1, k = (7,7)) and an activation layer with a sigmoid function. For an intermediate feature map $F \in \mathbb{R}^{H \times W \times C}$ as input, the output of spatial attention is $\mathbb{R}^{H \times W \times 1}$, which leverages the inter-spatial relationship of features. Both channel and spatial attention mechanisms complement each other in inferring informative aspects from two distinct dimensions, channel and spatial (Woo et al., 2018), and are integrated into our attention block by multiplication with a convolution layer and a Conv block (including Conv2D with $f_n = 32$, $k = (1,1)$), alongside ReLU (max*(0, x)*) activation functions.

The purpose of our last Conv block is to extract high-level features from the intermediate representations generated by preceding layers and prepare them for the final outputs of our deep learning model (Fig. 3 (c)). Following the application of a Conv block, which includes a Conv2D layers with f_n = 32 and $k = (1,1)$, along with a ReLU activation layer, the inputs undergo further processing. This involves the generation of two-channel outputs representing offsets in both easting and northing directions, achieved through a convolution layer with f_n = 2 and $k = (1,1)$. Given that our model is designed for regression prediction and aims to map learnable features to a desired output space comprising floating-point values, the final output is produced without an activation layer (e.g., ReLU). All convolutional layers within the three primary blocks utilize strides of (1,1) and “same” padding to preserve the dimensionality of the inputs.

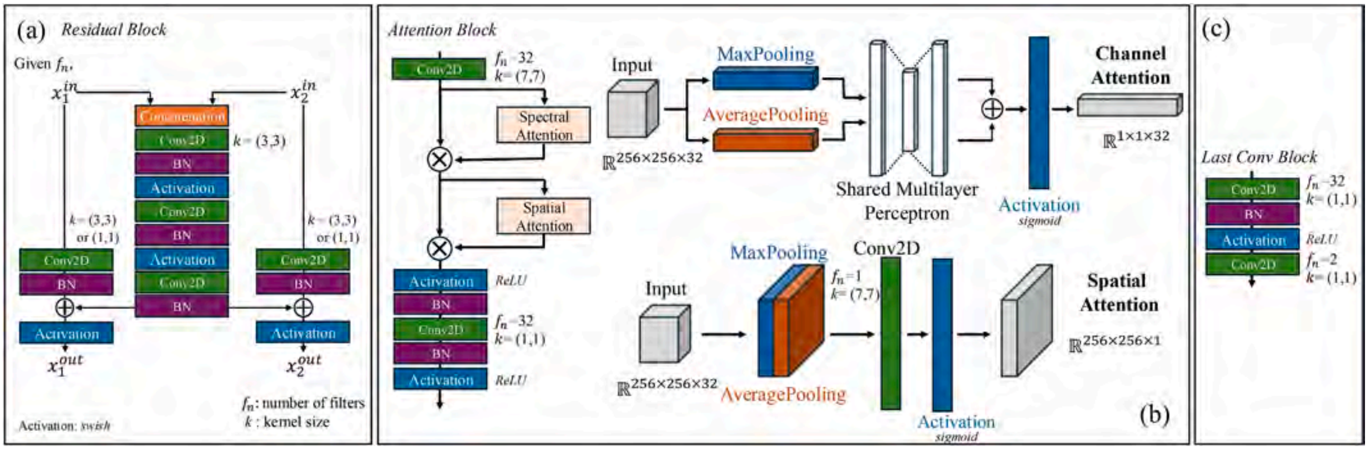

Fig. 3. Three primary modules for our deep learning model: (a) Residual, (b) Attention, and (c) Last Conv block.

2.3.2. Model architecture

We propose an efficient regression model composed of three primary modules: residual, attention, and last Conv blocks, based on the U-Net based Siamese architecture (Ronneberger et al., 2015; Zhan et al., 2017). For this model, we introduce a novel residual block and cross-connection approaches. Hence, we name the proposed model the cross-connection based Siamese ResNet (CC-ResSiamNet). The CC-ResSiamNet has an encoder and a decoder to process input reference and secondary amplitude images (256 × 256 × 1 chips) and produces output 2D offset maps in the easting and northing directions (224 × 224 × 2 chips).

The proposed model consists of six depth levels (*see* Fig. 4), each layer defined by different spatially sized convolution outputs, with two consecutive residual blocks applied at each depth using k = (3,3) and (1,1). In the encoder, the downsampling is applied by the max-pooling with strides of (2,2), resulting in dimensions of (W, H, f_n) = (256,256,32), (128,128,64), (64,64,128), (32,32,256), (16,16,512), and (8,81,024) at each depth, respectively. The encoder extracts features from the two input images. The decoder reverses the encoder's process via the upsampling with strides of (2,2), ultimately restoring the original image dimensions to 256 × 256. Then, 16 pixels along the boundaries of the image are removed to reduce noise effects.

Similar to the original U-Net, outputs at each depth from the encoder are connected to the corresponding layer in the decoder through skip connections. However, our model introduces a modification: the two outputs from the residual block are interchanged and serve as new inputs to the residual block at the same depth in the decoder, facilitated by a concatenation layer with upsampled inputs from the previous depth. This altered skip connection, termed the "cross-connection", not only facilitates the identification of spatial differences between the two inputs but also helps mitigate the vanishing gradient problem (Fig. 3). The

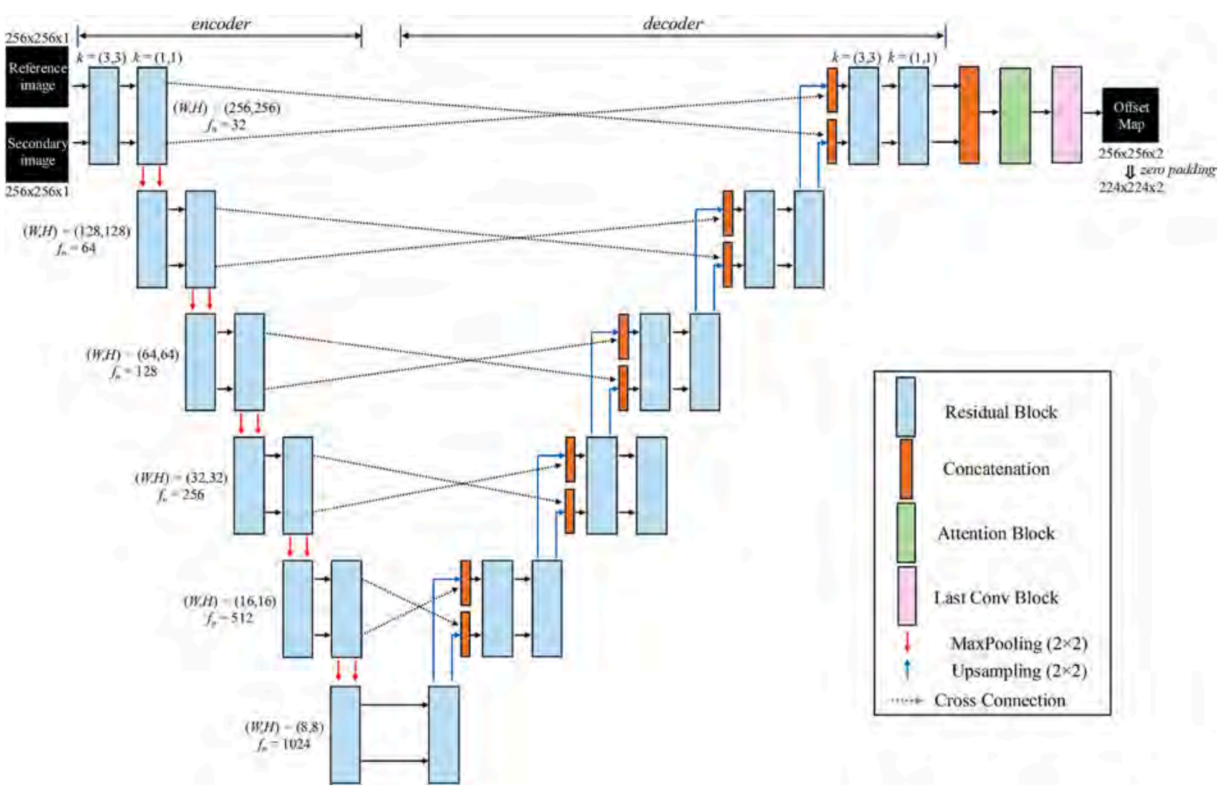

Fig. 4. CC-ResSiamNet proposed in this study. Two reference and secondary images are input to the model, undergo three primary modules (residual, attention, last Conv blocks) with concatenation layers and cross connection, and produce two-channel offset outputs.

outputs from the decoder undergo concatenation before passing through the attention and last Conv blocks to generate 2D offset maps. The final output of our model has dimensions of 224 × 224 × 2, with each channel representing offsets in the easting and northing directions. Our model encompasses 160,876,395 trainable parameters, consisting of weights and biases, with all weights initialized using the He normal initialization method (He et al., 2015).

2.4. Training and evaluation

We configure the hyperparameters, setting the learning rate to 0.0001 and the batch size to 50. For the loss function, we opt for the mean absolute error (MAE), which computes the average absolute difference between actual and predicted values in the dataset (Zhang et al., 2023). This metric provides insight into the proximity of regression predictions to the actual offset tracking values. We employ the adaptive moment estimation (Adam) optimizer to aid in minimizing the loss scalar value through gradient descent and backpropagation across epochs.

During the training process, our CC-ResSiamNet model ingests pairs of reference and secondary images as input. These images are processed through three primary modules: residual blocks, attention mechanisms, and final convolutional layers. The model's architecture incorporates concatenation layers and cross-connections between these modules, allowing for effective feature extraction and comparison between the image pairs. As the model iterates through epochs, it learns to predict two-channel offset outputs, which represent the spatial displacement between the reference and secondary images. The Adam optimizer adjusts the model's parameters based on the computed MAE loss, gradually improving the network's ability to accurately track offsets. This training approach enables the model to learn robust features and relationships between image pairs, preparing it for effective performance on unseen test data.

To address computational constraints and memory limitations, we implement an iterative training strategy. Instead of processing the entire dataset simultaneously, we divide it into four balanced subsets, each containing approximately equal numbers of training, validation, and test sets. The model is trained iteratively on these subsets, with a rotation occurring every 25 epochs. This approach not only circumvents memory constraints but also enhances the model's robustness by exposing it to diverse data distributions throughout the training process. By cycling through different subsets, we ensure that the model learns from a wide range of examples, potentially improving its generalization capabilities and mitigating overfitting risks.

Both training and evaluation procedures employ the same model architecture; however, during evaluation, a feed-forward processing approach is adopted without mini-batch gradient descent and backpropagation. Instead, the trained model parameters, including weights

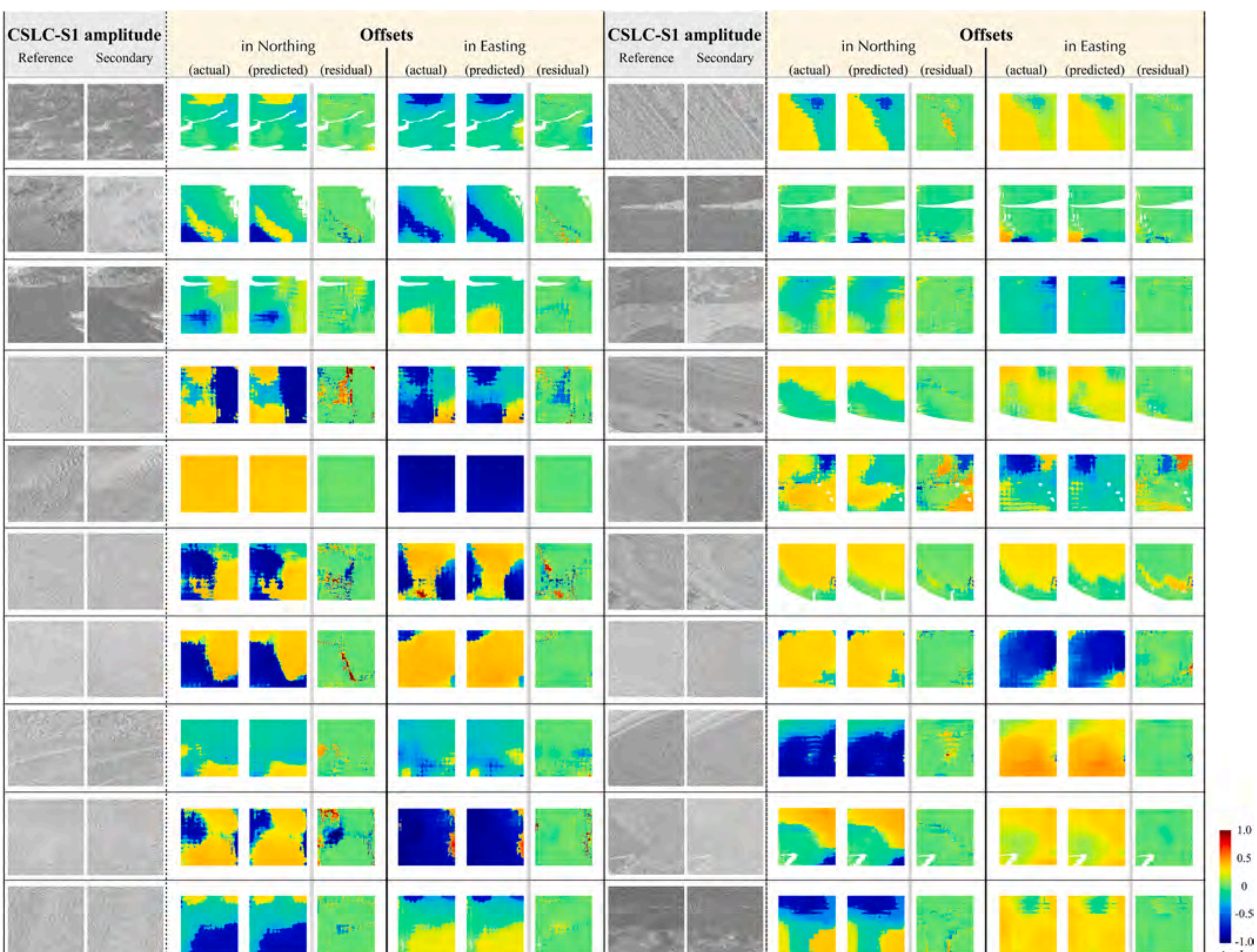

Fig. 5. Evaluation with test sets. Through a feed-forward process that inputs the two CSLC-S1 amplitude images, offsets in the easting and northing directions are produced. The predicted offsets exhibit a high degree of similarity to the actual offsets.

and biases in the residual, attention, and last Conv blocks, are utilized. Consequently, when two randomly assigned reference and secondary amplitude images are inputted into the model, the 256 $\times$ 256-sized images undergo a forward pass through the trained model, resulting in two-channel outputs representing offsets in the easting and northing directions. Evaluation results indicate a high degree of similarity between the predicted offsets generated by the deep learning model and the actual offset tracking results (Fig. 5). Despite the absence of distinct visual "displaced" surface features, our model has managed to produce offset predictions closely aligned with the actual offsets. It is worth noting that the offset tracking results were generated via correlation calculations at a 1 $\times$ 1 step in both the easting and northing directions, potentially introducing more noise. The validation with test sets suggests that our model has successfully learned the intricate relationship between reference and secondary amplitude images, enabling accurate calculation of pixel offsets at the highest resolution. Additionally, using

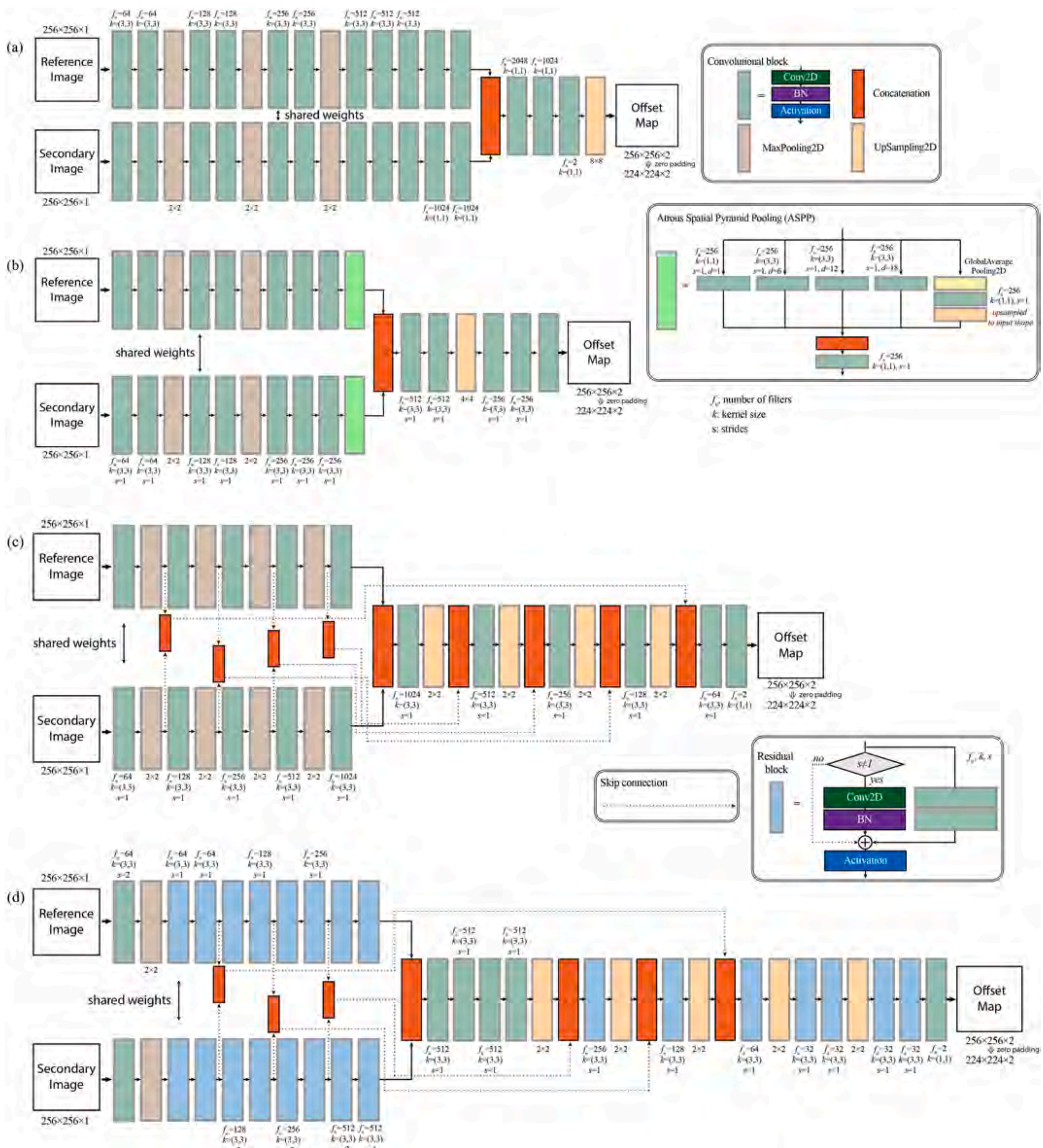

Fig. 6. Deep learning models used for comparative analysis. (a) SiamFCN, (b) SiamDeepLab, (c) SiamUNet, and (d) SiamResNetEncDec.

the proposed and trained deep learning model, generating offset maps from approximately 10 K 256 × 256 chips of CSLC-S1 pairs takes only 1.6 min. This demonstrates that, once the CC-ResSiamNet is trained, we can measure ground surface displacements instantaneously, bypassing the extensive processing time required by traditional offset tracking methods.

3. Comparison with other deep learning models

To assess the effectiveness of our proposed CC-ResSiamNet, we conduct a comparative analysis against multiple deep learning models. It is important to note that the application of deep learning for ground displacement estimation using SAR images represents a novel approach in the field. Given the unprecedented nature of our method, there are no directly comparable deep learning models tailored to our specific purpose. To address this gap and establish a meaningful benchmark, we have designed and implemented four distinct deep learning architectures, including Siamese Fully Convolutional Network (SiamFCN), Siamese DeepLab (SiamDeepLab), Siamese U-Net (SiamUNet), and Siamese ResNet Encoder-Decoder (SiamResNetEncDec). Each of these models is configured to process two SAR amplitude image inputs and generate two-channel offset outputs, mirroring the input-output structure of our CC-ResSiamNet. Consequently, this comparative framework allows us to evaluate the relative performance of our proposed model within the context of contemporary deep learning approaches.

3.1. Architectural designs for comparable deep learning models

The following contemporary deep learning models are used for comparative analysis:

- The SiamFCN, primarily composed of a series of convolutional layers, is structured with two identical input branches, each designed to process a SAR amplitude image (Fig. 6(a)). These branches consist of multiple convolutional layers that work to extract features from the input images. Subsequently, after independent feature extraction, the two branches merge, allowing the network to compare and combine the information from both inputs. This merged representation then passes through additional convolutional layers, which further refine and integrate the features. Finally, the network culminates in a final convolutional layer that produces a two-channel output, likely representing the estimated displacement in two dimensions.
- The SiamDeepLab architecture (Fig. 6(b)) builds upon the DeepLab model, which was originally developed for semantic image segmentation (Chen et al., 2018). Like the SiamFCN, it features two identical input branches for processing SAR amplitude images but incorporates specialized components for enhanced feature extraction. Each branch incorporates convolutional layers followed by atrous (dilated) convolutions. Notably, atrous convolutions allow the network to capture wider spatial context without increasing the number of parameters. After the initial feature extraction, the branches merge, and the combined representation undergoes further processing through additional atrous convolutions. A key feature of this model is the inclusion of an atrous spatial pyramid pooling (ASPP) module, which captures multi-scale contextual information. Ultimately, the network concludes with upsampling to restore spatial resolution and a final convolutional layer to produce the two-channel output.
- The SiamUNet model employs two identical U-Net-style branches, one for each SAR amplitude image input (Fig. 6(c)). Each branch follows the characteristic U-shaped architecture of U-Net, consisting of an encoder path for downsampling and a decoder path for upsampling. In this structure, the encoder progressively reduces spatial dimensions while increasing feature depth, capturing hierarchical features. The decoder then gradually restores spatial resolution. A crucial aspect of this architecture is the use of skip connections, which directly connect corresponding levels of the encoder and decoder. These connections allow the decoder to leverage both high-level semantic information and low-level spatial details. Subsequently, after individual processing, the outputs of both branches are merged, and a final convolutional layer generates the two-channel output.
- The SiamResNetEncDec architecture incorporates elements from ResNet into an encoder-decoder framework (Fig. 6(d)). It starts with two identical branches, each processing one of the SAR amplitude images. The initial layers of each branch consist of convolutional operations and max-pooling, reducing spatial dimensions. At the core of the encoder are multiple residual blocks, a hallmark of ResNet architectures (He et al., 2016). These blocks allow for very deep networks by using skip connections to mitigate the vanishing gradient problem. In the decoder path, upsampling operations are used to restore spatial resolution, complemented by skip connections from the encoder. These connections help preserve fine-grained spatial information. After independent processing in each branch, the decoded features are merged. The network then concludes with final convolutional layers to produce the desired two-channel output.

3.2. Comparative analysis

To ensure a fair comparison of multiple deep learning models, we implemented a standardized evaluation framework. This approach includes: *1*) Utilizing identical datasets and data splits for training and testing across all models, *2*) Applying consistent pre-processing techniques to SAR amplitude images, *3*) Maintaining uniform input image sizes for all models, *4*) Standardizing hyperparameters, including learning rate and mini-batch size, across all models, *5*) Training each model for a fixed duration of 200 epochs, and *6*) Employing the same MAE loss function for all models. Based on the training, validation, and testing of all models, their performance is compared by examining the learning curves and error histograms.

3.2.1. Comparison of learning curves

Comparative analysis with learning curves for multiple deep learning models, specifically CC-ResSiamNet, SiamFCN, SiamDeepLab, SiamUNet, and SiamResNetEncDec, presents each model's performance and generalization capabilities (Fig. 7). Among the models analyzed, CC-ResSiamNet demonstrates superior performance. It exhibits the lowest training loss, rapidly converging and maintaining the low loss throughout the training process. Notably, its validation loss is also among the lowest, indicating excellent generalization to unseen data. This combination of low training and validation losses suggest that CC-ResSiamNet achieves the best balance between fitting the training data and generalizing to new instances. SiamResNetEncDec emerges as the second-best performer, closely following CC-ResSiamNet. While its training loss is marginally higher, it maintains a validation loss comparable to CC-ResSiamNet. This performance indicates good generalization capabilities, albeit slightly less efficient in minimizing training error compared to CC-ResSiamNet. SiamFCN and SiamUNet display moderate performance levels. Both models show higher validation losses compared to CC-ResSiamNet and SiamResNetEncDec, with SiamFCN exhibiting slightly higher validation loss than SiamUNet. These results suggest that while both models learn from the training data, they many not generalize as effectively to new, unseen examples. SiamDeepLab presents a unique case in this comparison. Despite competitive performance in terms of training loss, it exhibits the highest validation loss among all models. This substantial discrepancy between training and validation performance is indicative of overfitting, a common challenge in machine learning where a model performs well on training data but fails to generalize to new, unseen data.

The phenomenon of overfitting is particularly evident in the learning

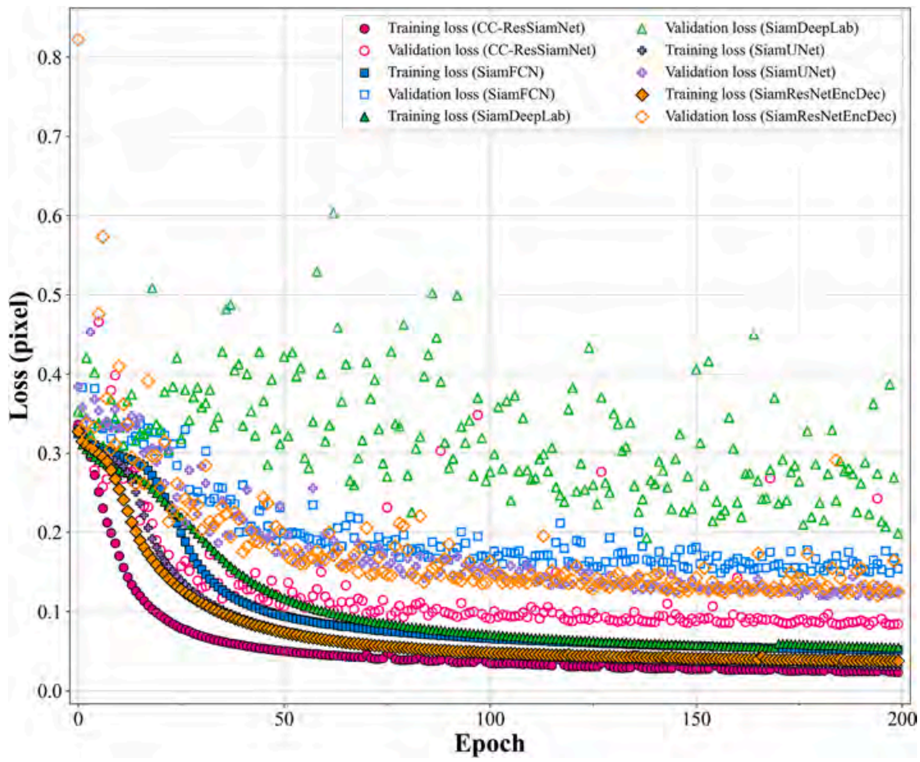

Fig. 7. Learning curve of training and validation losses for CC-ResSiamNet, SiamFCN, SiamDeepLab, SiamUNet, and SiamResNetEncDec.

curves of SiamDeepLab. The large and persistent gap between its decreasing training loss and high, unstable validation loss is a classical sign of a model that has overfit to the training data. In contrast, CC-ResSiamNet and SiamResNetEncDec show minimal signs of overfitting, with small gaps between their training and validation losses, indicating robust generalization. SiamFCN and SiamUNet fall between these extremes, showing moderate overfitting with noticeable, but not severe gaps between training and validation losses.

3.2.2. Comparison of error histograms

The error histogram (Fig. 8) calculated by difference of offset results from both traditional offset tracking and deep learning prediction (distribution of errors in easting and northing directions) can offer a detailed view of the five models' performance on test sets. The CC-ResSiamNet model continues to demonstrate the best performance among all models with the smallest root mean square error (RMSE) less than 0.03 pixels (Fig. 8(a)). Its error distribution is tightly centered around zero for both easting and northing, with the highest peak and narrowest spread. The mean error lines are almost perfectly aligned with zero (both along −0.001), indicating not only high precision but also exceptional accuracy with minimal systematic bias. SiamFCN shows a wider error distribution compared to CC-ResSiamNet (Fig. 8(b)). The spread of errors is larger, particularly in the easting direction. The mean errors reveal a slight bias, especially in the easting direction where the line is visibly offset from zero. This indicates that while the model's predictions are less precise than CC-ResSiamNet, they also tend to have a small systematic error. SiamDeepLab exhibits the widest error distribution among all models (Fig. 8(c)). This is noticeable both in the easting and northing direction. The mean errors show significant offset from zero, indicating a pronounced systematic bias. This aligns with the overfitting behavior observed in the learning curves, suggesting that SiamDeepLab's performance on unseen data is indeed problematic. SiamUNet presents an error distribution that is narrower than SiamDeepLab but wider than CC-ResSiamNet (Fig. 8(d)). Its mean error lines show a slight offset from zero, indicating the presence of a small systematic bias in its predictions. SiamResNetEncDec displays an error distribution that is comparable to SiamUNet in terms of spread (Fig. 8(e)). Its mean errors, particularly in the northing direction, show noticeable offset from zero and it may have a slight systematic bias in the prediction.

Comparative analyses using learning curves and histograms reveal that CC-ResSiamNet is the most effective model among those studied. It exhibits superior performance with rapid convergence in both training and validation, maintaining low losses and minimal overfitting. This suggests a strong ability to generalize to new data, high precision, and low potential biases. The model also shows a well-contained spread of random errors (as indicated by the error distribution) and minimal systematic error (with the mean error close to zero).

4. Multi-scenario inference and validation

The CC-ResSiamNet was trained on data from Alaska's glaciers (Fig. 9), which exhibit distinct characteristics in terms of topography, climate, and ice dynamics. Glacier surface displacement is primarily driven by ice flow and melt, influenced by factors such as temperature, precipitation, and underlying topography. Applying our model to diverse geographic locations and disaster scenarios can be considered extrapolation, as it involves out-of-distribution prediction beyond the parameters learned during training. However, it is crucial to acknowledge the practical limitations of training on an exhaustive global dataset. Our deep learning approach offers a pragmatic solution by focusing on glacier movements derived from pairs of SAR amplitude images in Alaska, which effectively capture ground displacement or pixel shifts between image acquisition times. This workable model opens up possibilities for cross-domain applications, potentially mitigating a degree of domain shift. While not a perfect solution, it may represent a step toward developing more versatile and adaptable Earth observation tools.

To evaluate the model's performance, we applied our trained model to estimate ground displacement across three distinct groups (Fig. 9): *1*) Glaciers in different geographic locations, *2*) Earthquake-affected regions, and *3*) Volcanic areas where the ground surface displaced sufficiently to be measured by SAR images. Importantly, the single scenes

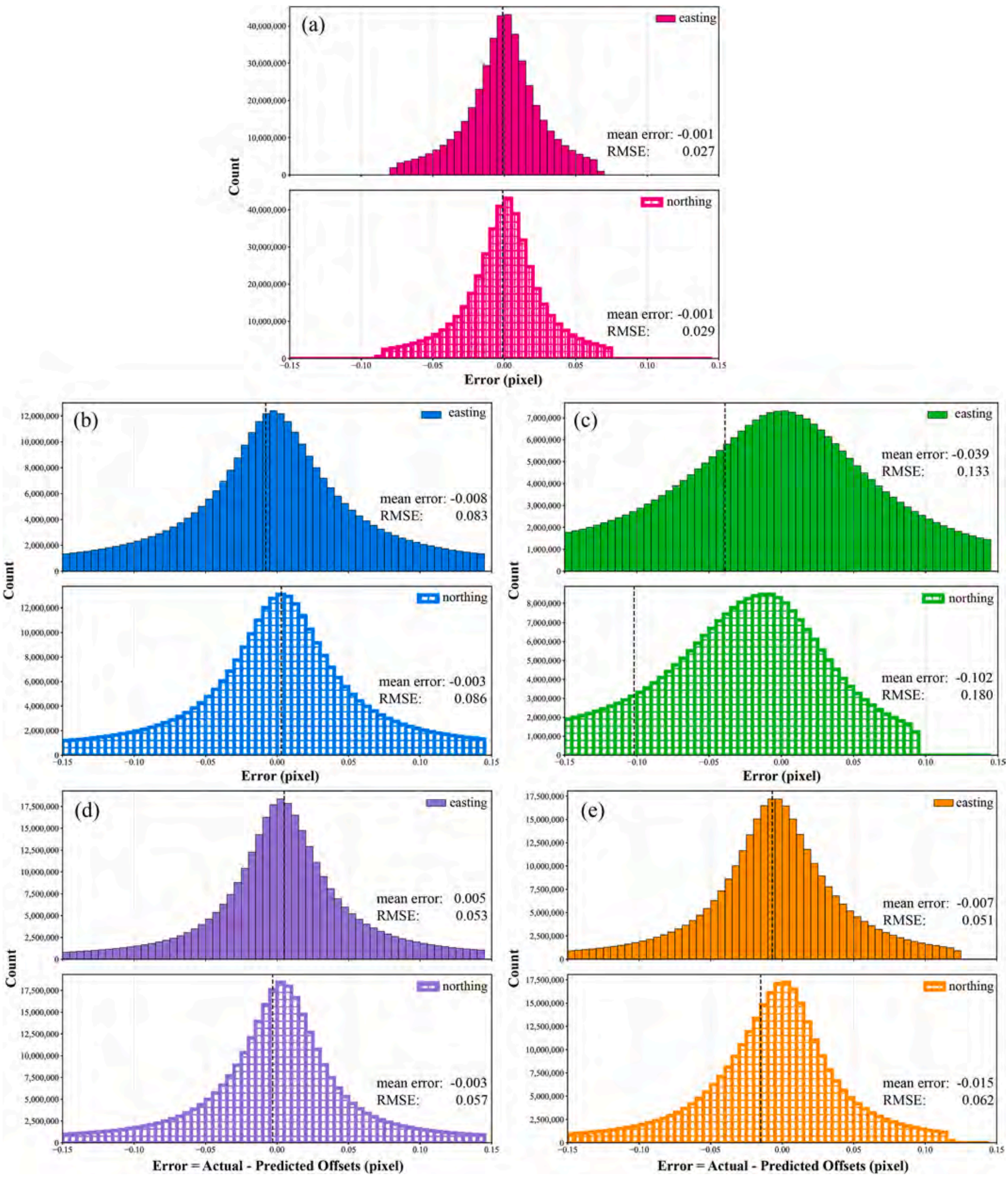

Fig. 8. Error histograms of (a) CC-ResSiamNet, (b) SiamFCN, (c) SiamDeepLab, (d) SiamUNet, and (e) SiamResNetEncDec. Dashed lines represent the mean error values of each distribution.

used for inference in these areas were not included in updating model parameters during training, validation, or testing phases. Consequently, we can assert that the datasets in these areas are truly novel and unseen, providing a rigorous test of the model's generalization capabilities. This multi-scenario validation approach allows us to assess the robustness and versatility of the CC-ResSiamNet in estimating ground displacement across various geological contexts and event types.

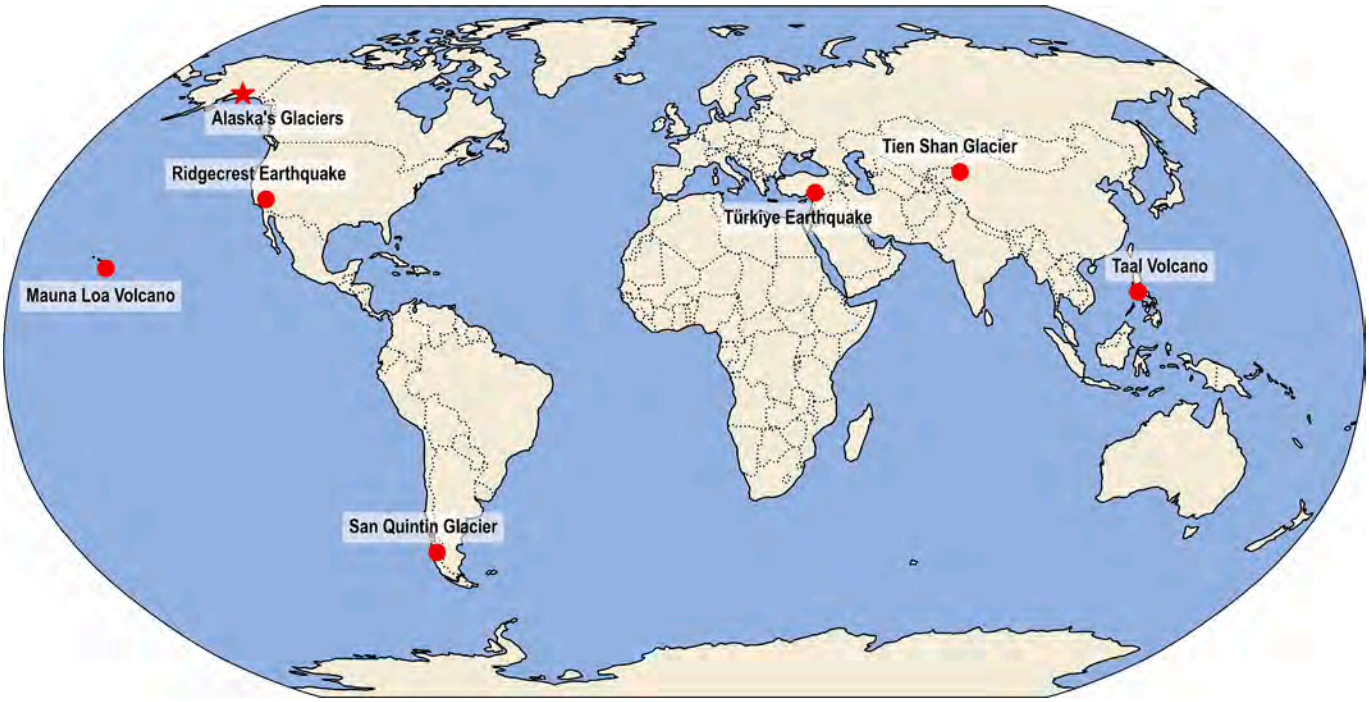

Fig. 9. Locations for multi-scenario inference and validation with our deep learning model. The model was trained on Alaska's glaciers (red star) and used for inferencing ground displacement in glaciers, earthquakes, and volcanoes around the world (red dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

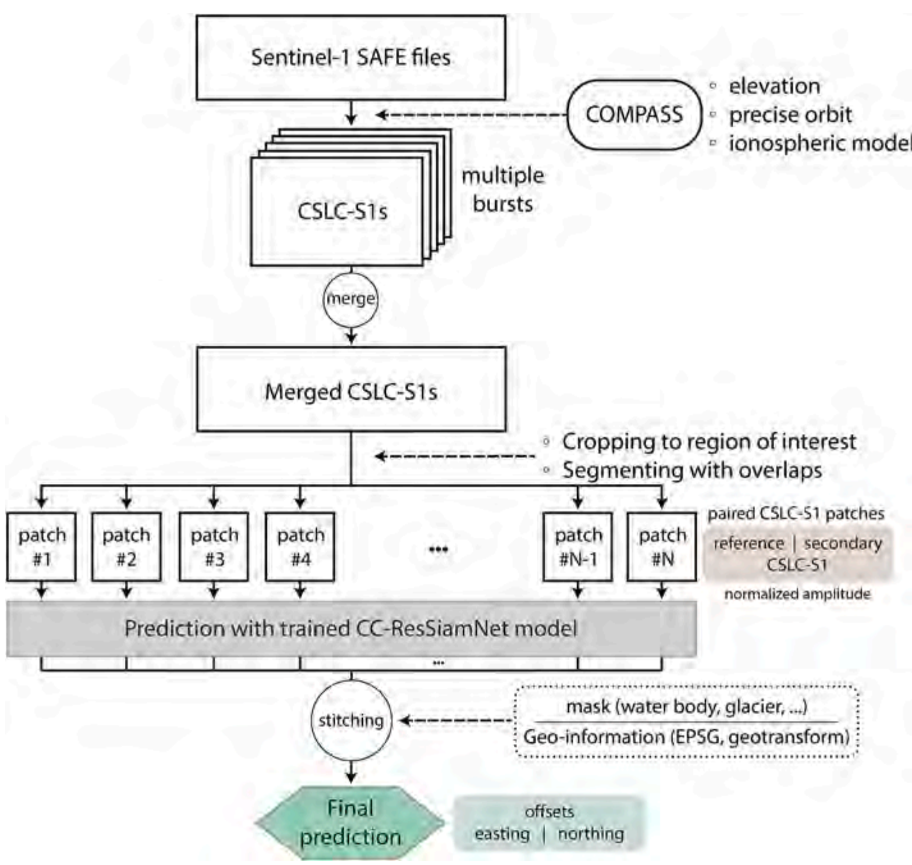

Fig. 10. Workflow of inference processes. Offsets in the easting and northing directions over the region of interest are predicted using reference and secondary CSLC-S1s.

4.1. Inference processes

The data used for inference includes CSLC-S1 files generated by the COMPASS software from Sentinel-1 SAFE files, along with precise orbit data, GLO-30 DEM, and GIM. Unlike the training process, which uses single bursts, multiple bursts are merged to cover the entire area of interest for target glaciers, earthquakes, and volcanoes. This merging ensures that all bursts share the same European Petroleum Survey Group (EPSG) codes. If the bursts are located in different UTM zones, they are reprojected into a common UTM zone, while maintaining the original pixel spacing of 5 and 10 m in the easting and northing directions, respectively (Fig. 10). The coverage of these merged multi-burst input images is slightly larger than the actual region of interest, so the input normalized SAR amplitude images are cropped to the smaller size for computing efficiency.

During the training of our deep learning model, inputs were composed of 256 × 256 chips. However, the real-world data contains irregular shapes. To apply the trained model to SAR amplitude images, we divide the images into multiple 256 × 256-sized patches with overlaps and geoinformation (EPSG code, geotransform composed of origin coordinates and rotation angles), ensuring no image parts are omitted (Fig. 11). Each patch is processed by our trained model, producing two-channel offset maps. After applying 16-pixels zero padding along the edges, these patch offset maps are then stitched into a single two-channel offset map, using each patch's geoinformation in UTM coordinates and applying water or glacier masking, with overlaps weight-averaged. The stitched offset maps represent the final prediction of our deep learning model, which will be compared against the results obtained from traditional offset tracking method.

4.2. Glaciers in different geographic locations

4.2.1. Tien Shan glaciers in China

The Tien Shan, one of the world's largest mountain ranges, extends approximately 2500 km across Central Asia (Fig. 9). Glaciers and ice caps in these high mountains serve as crucial "water towers", providing meltwater to semi-arid regions with seasonal water scarcity. These areas include Kazakhstan, Kyrgyzstan, Uzbekistan, and China's northwestern Xinjiang Uyghur Autonomous Region (Farinotti et al., 2015). From 1960 to 2021, Tien Shan glaciers have undergone significant shrinkage, with an estimated 27 % loss in glacier mass since the 1960s. This retreat has accelerated in recent years, coinciding with a temperature increase of about 0.7 °C per decade in parts of the Tien Shan region (Zhuang et al., 2023). Efforts have been made to assess glacier dynamics, including the measurement of glacier velocity (Millan et al., 2022) and the estimation of mass loss (Farinotti et al., 2015). Our research applies a trained deep learning model to estimate the velocity of a selected glacier in the Tien Shan Mountains.

We utilized both ascending (P136) and descending (P158) track datasets to estimate glacier velocity in the easting and northing directions over the same area (red polygon in Fig. 12). Our prediction is based on CSLC-S1s obtained on 2018/08/07 and 2018/08/19 (P136) and 2019/06/04 and 2019/06/16 (P158), providing a 12-day interval dataset. The results show high similarity with those independently produced by a traditional offset tracking method; specifically multi-kernel offset tracking. The RMSE for P136 is 0.178 pixels in the easting direction and 0.188 pixels in the northing direction, while for P158, it is 0.205 pixels in the easting direction and 0.209 pixels in the northing direction. The glacier is situated in an elongated, narrow valley, and the measured or predicted offsets may be influenced by the steep mountains on the northern and southern sides. Despite these potential confounding factors, our trained deep learning model successfully captured the primary eastward glacier movements and, to a certain degree, the secondary northward glacier velocity.

4.2.2. San Quintin glacier in Chile

San Quintin Glacier, the largest outlet glacier of the Northern Patagonian Ice Field, is situated within Laguna San Rafael National Park in southern Chile (Fig. 9). This glacier flows westward from the Patagonian Glacier Plateau toward the low-lying, swampy Isthmus of Ofqui (Copernicus EU, 2021; Brockmann Consult, 2023). Over the past 30 years, San Quintin has undergone a rapid retreat, with particularly significant changes observed since 1991. As the glacier recedes, it has contributed to the formation of a rapidly expanding proglacial lake at its

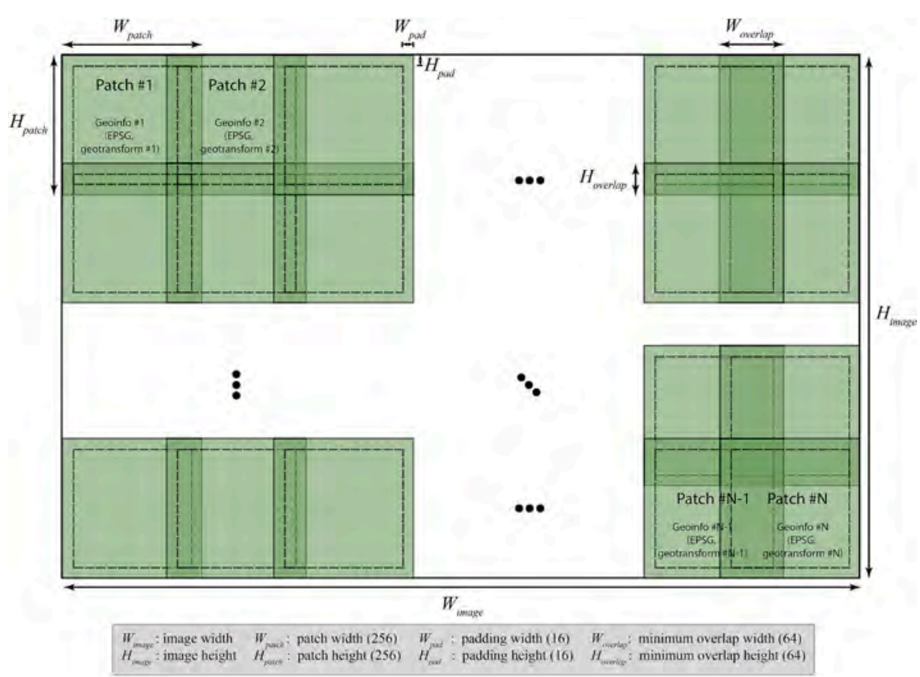

Fig. 11. Strategy for patching and stitching images. The input image is divided into 256 × 256 patches with a minimum overlap of 64 pixels. Each patch includes geoinformation (EPSG code, geotransform) to facilitate stitching in UTM coordinates. The trained deep learning model produces offsets in the easting and northing directions. After applying zero padding along the edges, these offsets are stitched into a single offset map, with overlaps weight-averaged to ensure consistency.

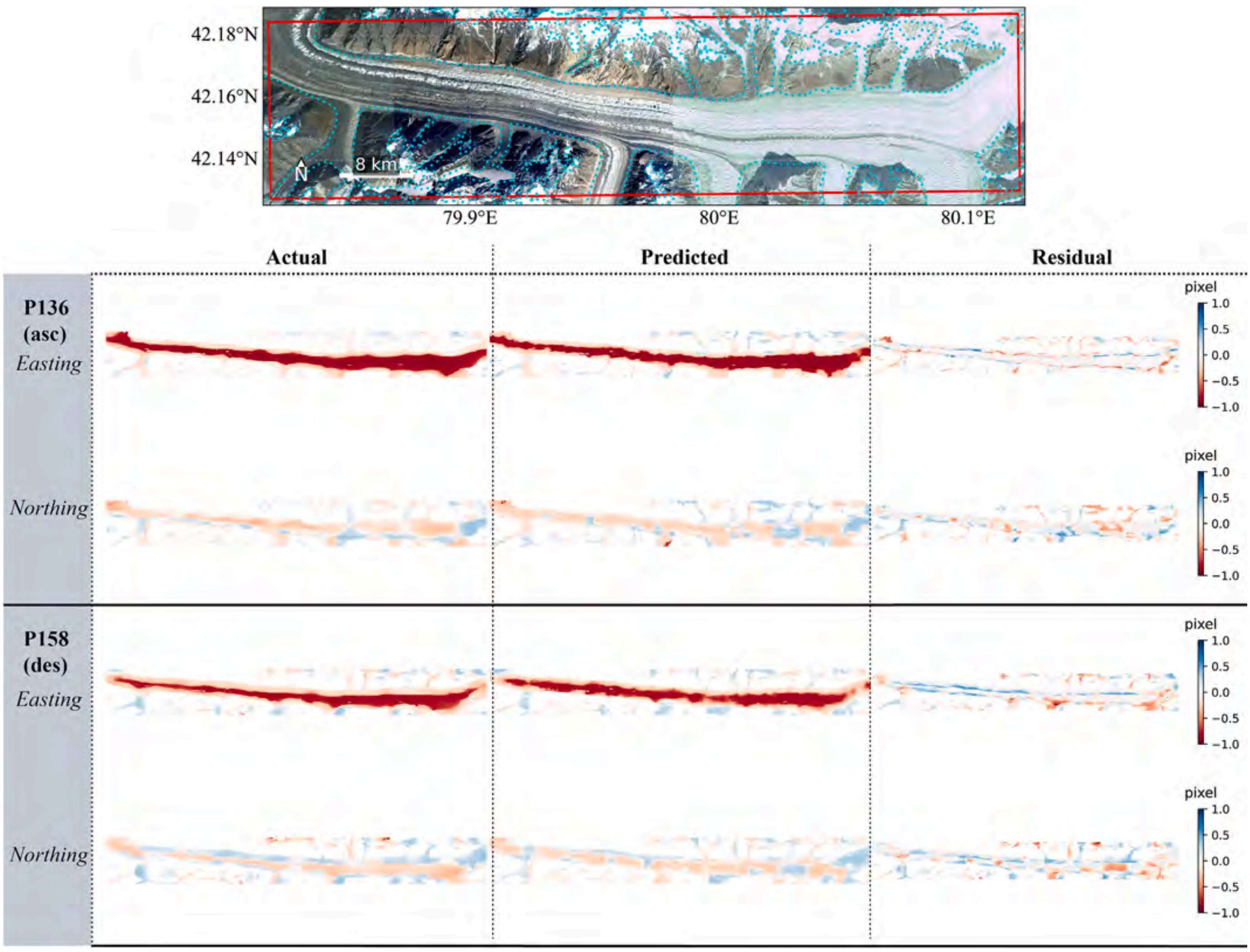

Fig. 12. Tien Shan glacier outlined by RGI datasets (dashed, cyan lines) and the coverage of the datasets (red polygon) used for estimating glacier velocity. This figure presents three columns for each dataset (P136, P158), glacier velocity estimated by traditional offset tracking, velocity predicted by the deep learning model, and residual (actual – predicted offsets). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

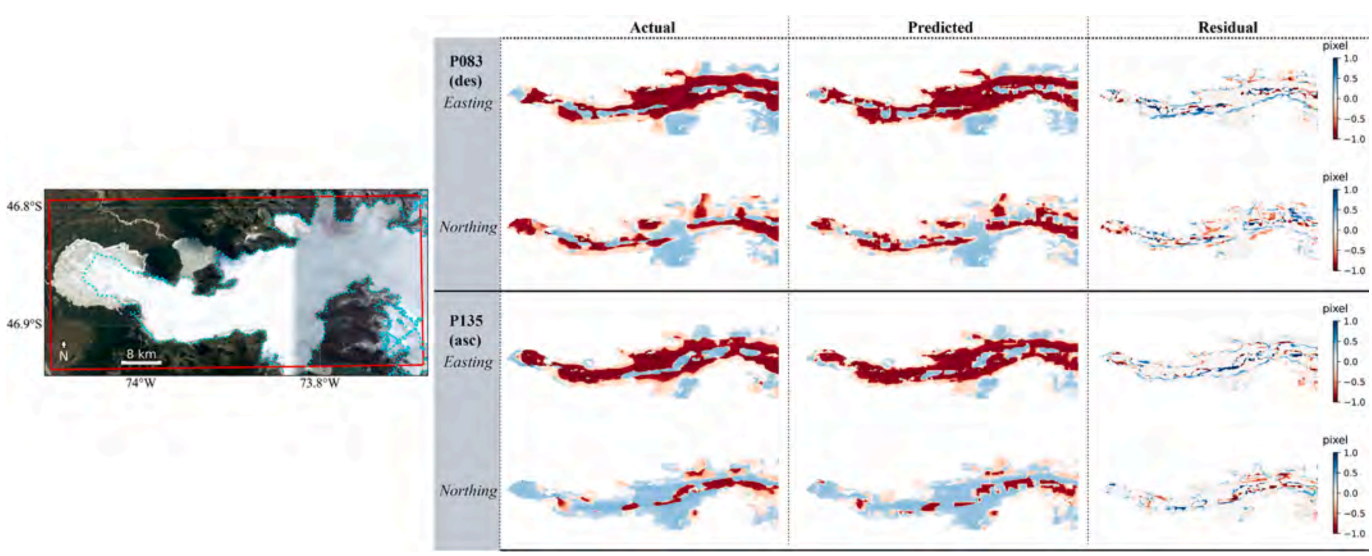

Fig. 13. San Quintin glacier outlined by RGI datasets (dashed, cyan lines) and the coverage of the datasets (red polygon) used for estimating glacier velocity. This figure presents three columns for each dataset (P083, P135), glacier velocity estimated by traditional offset tracking, velocity predicted by the deep learning model, and residual. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

terminus, potentially making San Quintin Glacier Lake the fastest-growing lake in South America during the 21st century (Copernicus EU, 2021). The glacier's retreat was also documented through satellite imagery between 1996 and 2000 (Harrison et al., 2001). As part of global glacier velocity assessments, San Quintin was included in a recent study on ice dynamics (Millan et al., 2022). In this research, we applied a deep learning model to estimate the glacier's velocity.

Using CSLC-S1 datasets from the descending (P083) and ascending (P135) tracks, we predicted the glacier's velocity in both the easting and northing directions over the same region (depicted by the red polygon in Fig. 13). The data spanned 12-day intervals, covering 2020/09/03 to 2020/09/15 for P083 and 2018/11/10 to 2020/11/22 for P135. Our predictions aligned well with the glacier velocities measured by traditional offset tracking methods (Fig. 13). The RMSE for P083 was 0.343 pixels in the easting direction and 0.324 pixels in the northing direction, while for P135, the RMSE was 0.307 pixels in the easting direction and 0.275 pixels in the northing direction. Although these RMSE values are larger than those observed for the Tien Shan Glaciers, the discrepancy is likely acceptable due to the presence of more pixels with larger glacier velocities approaching −1 and 1 shifts. Similar to the Tien Shan Glaciers, the overall pattern of glacier movements in the easting and northing directions closely matched between our predictions and the traditional offset tracking method.

4.3. Earthquake-affected regions

4.3.1. Ridgecrest earthquake, California, USA

The 2019 Ridgecrest earthquakes in Southern California, including the main M_w 7.1 tremor on July 5, 2019, and numerous aftershocks, significantly affected the town of Ridgecrest. These earthquakes ruptured two conjugate faults in the Airport Lake fault zone and Little Lake fault zone, oriented roughly northwest-southeast and northeast-southwest, respectively. Coseismic displacements were measured using InSAR and offset tracking methods with Sentinel-1 datasets (Xu et al., 2020b). Our deep learning model was applied to the CSLC-S1 datasets from the ascending (P064) and descending (P071) tracks, predicting ground displacement over a 6-day interval (2019/07/04 to 2020/07/10) for P064 and a 12-day interval (2019/07/04 to 2020/07/16) for P071. The model's predictions closely matched traditional offset estimation methods (Fig. 14), with RMSE values of 0.094 pixels (easting) and 0.120 pixels (northing) for P064, and 0.093 pixels (easting) and 0.117 pixels (northing) for P071. While our model slightly underestimates the overall ground displacement and produces less detailed, smoothed displacement compared to the offsets from traditional methods, the earthquake-induced ruptures are clearly identifiable and consistently located across different SAR look directions.

4.3.2. Türkiye earthquake

The devastating 2023 Türkiye earthquake sequence began on February 6, 2023, with an initial M_w 7.8 shock followed by a M_w 7.7 aftershock approximately nine hours later (Fig. 9). Both seismic events occurred along the East Anatolian Fault in southeastern Türkiye, near the Syrian border, with the epicenter located close to the Turkish city of Kahramanmaraş (An et al., 2023). Researchers have employed InSAR and offset tracking techniques using Sentinel-1 and Sentinel-2 data to measure ground surface displacements across the affected regions (He et al., 2023). We applied our deep learning model to CSLC-S1 datasets from ascending (P014) and descending (P021) tracks to predict surface displacements over slightly different regions (red and blue polygons in Fig. 15, respectively). The model successfully captured the overall displacement pattern (Fig. 15) using 12-day interval data from 2023/01/28 to 2023/02/09 (P014) and 2023/01/29 to 2023/02/10 (P021). The RMSE for P014 was 0.241 pixels in the easting direction and 0.284 pixels in the northing direction, while for P021, it was 0.250 pixels and 0.285 pixels, respectively. Our predictions of rupture locations, clearly visible from offsets in the easting direction, closely match those obtained through traditional offset tracking (Fig. 15). However, our model slightly underestimated displacements in the northernmost part of the

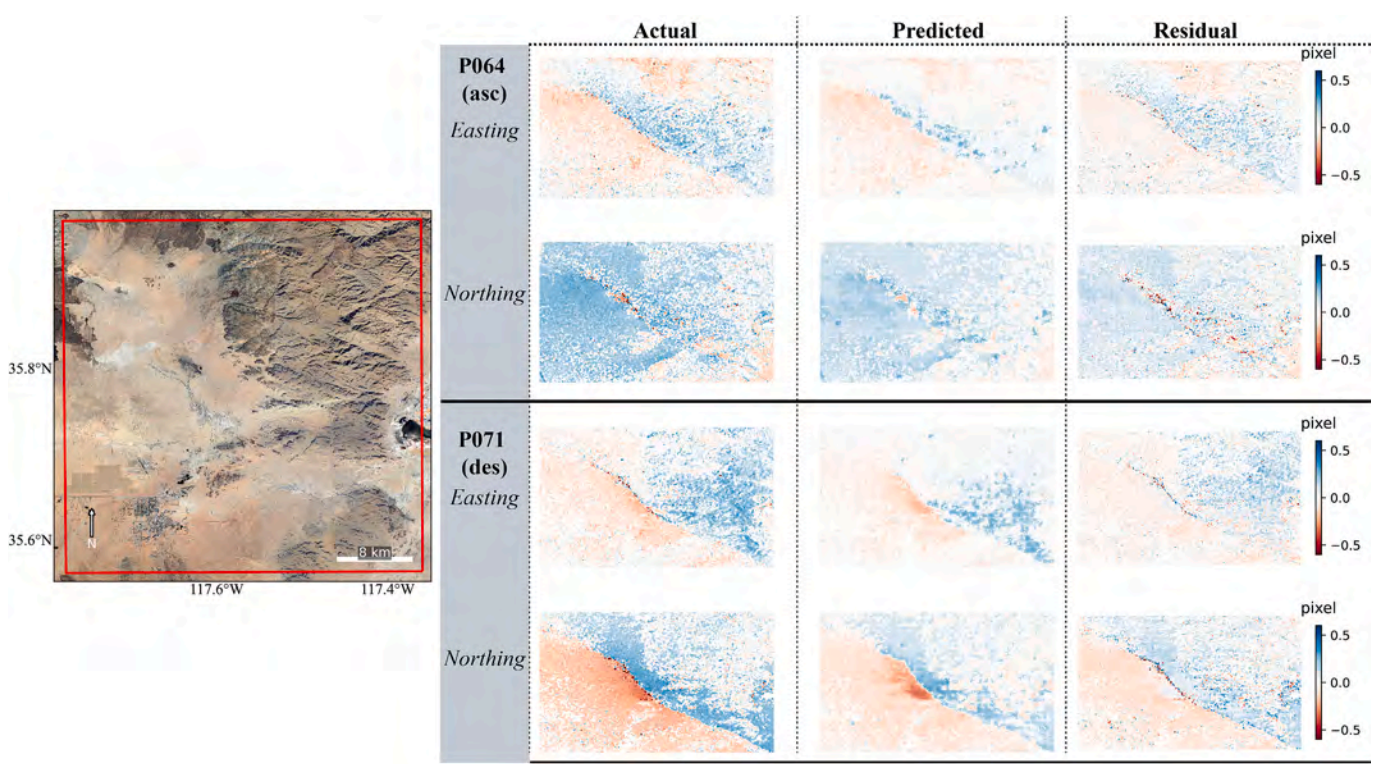

Fig. 14. Ridgecrest area where 2019 earthquakes struck on July 5, 2019, and the coverage of our dataset (red polygon) used for predicting ground displacement. This figure presents three columns for each dataset (P064, P071), ground displacement by traditional offset tracking, velocity predicted by the deep learning model, and residual. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

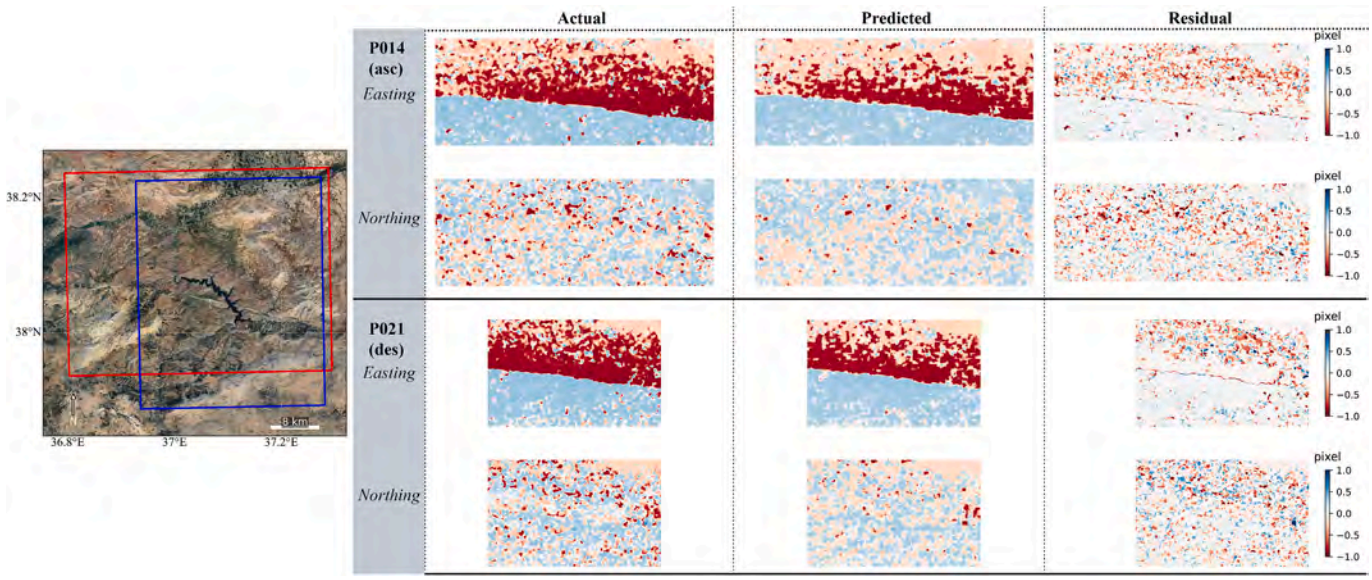

Fig. 15. Regions affected by the 2023 Türkiye-Syria earthquakes on February 6, 2023, with the red and blue polygons representing the dataset coverage area (P014 and P021, respectively) used for ground displacement prediction. The figure consists of three columns for each dataset (P014, P021): ground displacement estimated using traditional offset tracking, velocity predictions from the deep learning model, and the residual between the two methods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

affected area.

4.4. Volcanic areas

4.4.1. Taal volcano, Philippines

Taal Volcano, located in southwestern Luzon, Philippines, approximately 60 km south of Manila (Fig. 9), is nestled within a large caldera filled by Taal Lake. The volcano itself is situated on Volcano Island, positioned in the middle of this lake (Moore et al., 1966). On January 12, 2020, Taal Volcano erupted, generating ash plumes, volcanic lightning, and lava fountains, leading to the evacuation of nearby villages and affecting over 500,000 people. InSAR and offset tracking

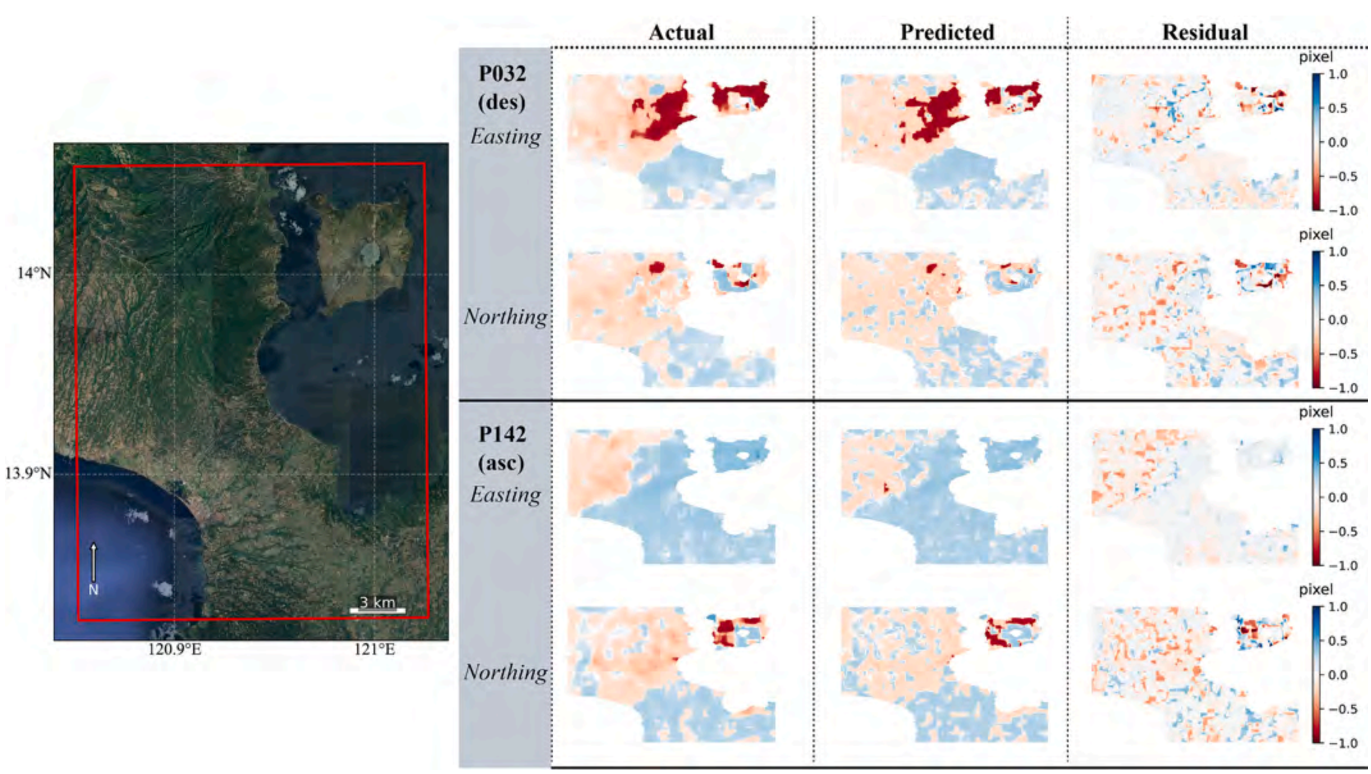

Fig. 16. Taal volcano, Philippines erupted on January 12, 2020, with the red polygon representing the dataset coverage area use for ground displacement prediction. The figure consists of three columns for each dataset (P032, P142): ground displacement estimated using traditional offset tracking, velocity predictions from the deep learning model, and the residual between the two methods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

techniques were employed to analyze the eruption (Bato et al., 2021), while our deep learning model successfully predicted surface displacement using CSLC-S1 datasets over 6-day intervals from 2020/01/09 to 2020/01/15 (P032) and 2020/01/11 to 2020/01/17 (P142). Our predictions closely matched the traditional offset tracking measurements, with RMSE values for P032 being 0.169 pixels in the easting direction and 0.192 pixels in the northing direction, and for P142, 0.110 pixels in the easting and 0.197 pixels in the northing direction. Despite the complex terrain, dense vegetation, and the surrounding water body of Taal Lake (as shown in Fig. 16), our model effectively captured surface displacements in both easting and northing directions, irrespective of SAR viewing geometry.

4.4.2. Mauna Loa Volcano, Hawaii, USA

Mauna Loa, the Earth's largest active volcano, is located on the Big Island of Hawaii. Its 2022 eruption marked the first activity in 38 years, with the previous eruption occurring in 1984 (Maher et al., 2023). The eruption began on November 27, 2022, and continued until December 13, 2022, initially starting at the summit and later spreading to the Northeast Rift Zone (Moisseeva et al., 2023). To predict the ground displacements caused by the volcanic activity, we employed our trained deep learning model, utilizing 12-day interval CSLC-S1 datasets from two time periods: 2022/11/22 to 2022/12/04 (descending track P087) and 2022/11/25 to 2022/12/07 (ascending track P124). The RMSE for P087 was 0.049 pixels in the easting direction and 0.083 pixels in the northing direction, while for P124, the RMSE values were 0.051 pixels and 0.059 pixels, respectively. Although these small RMSE values reflected relatively small displacements around the summit compared to other cases, the volcanic displacements were not well captured particularly in the northing direction of the P087 track (Fig. 17). The predicted displacement pattern along the rift was similar, but most displacements were underestimated. Predictions for P124 showed better alignment with the results of traditional offset tracking methods, though the displacements appeared smoothed and less detailed.

5. Discussion

Deep learning has become a powerful tool in satellite remote sensing, offering significant advantages such as enhanced accuracy in image classification, object detection, and segmentation (Ma et al., 2019). It also excels in managing large datasets and demonstrates resilience to noise, including atmospheric effects (Wang et al., 2022). However, its applications have primarily been focused on classification tasks and disaster monitoring, such as mapping areas impacted by floods or wildfires (Zhu et al., 2017). This study suggests the new application of deep learning in radar satellite remote sensing by demonstrating its efficacy in estimating ground displacement through regression-based predictions. The combination of U-Net and ResNet has shown strong performance in change detection tasks when used in a Siamese structure with two input images (Zhan et al., 2017). Leveraging the characteristic cross-connections that enhance change detection, we designed our deep learning model, CC-ResSiamNet, based on these two architectures, incorporating an additional attention block and a convolutional block for more accurate regression outputs.

In deep learning, identifying appropriate datasets for training a model is crucial. To predict ground displacements from SAR images, we require datasets over consistently shifting surfaces that offer both sufficient quantity and quality. While earthquakes and volcanic activity displace large areas, the episodic nature of these events makes it difficult to gather sufficient data for effective model training. In contrast, consistently flowing glaciers provide an ideal training dataset. The spatial variations within glaciers, along with their pixel shifts caused by surface displacement, are detectable in SAR amplitude images. Additionally, CSLC-S1 is the best available data source, as it bypasses time-consuming coregistration with Sentinel-1 by utilizing geometry-driven, GPU-accelerated processing. Target offsets can be generated through multi-kernel offset tracking, which uses multiple windows to enhance results from different spatial perspectives. Approximately 200 K pairs of 256 × 256 reference and secondary SAR image chips, along

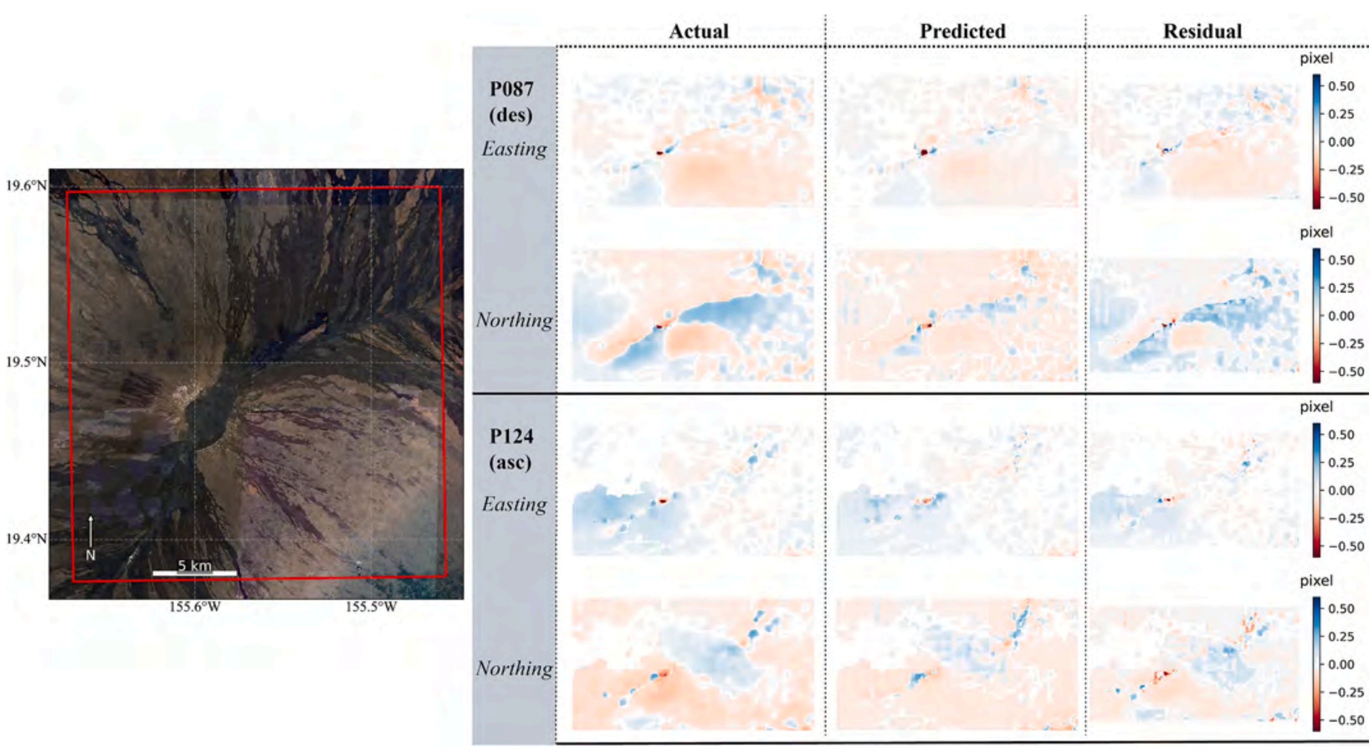

Fig. 17. Mauna Loa volcano, Hawaii, with the red polygon representing the dataset coverage area use for ground displacement prediction. The figure consists of three columns for each dataset (P087, P124): ground displacement estimated using traditional offset tracking, velocity predictions from the deep learning model, and the residual between the two methods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

with target offsets, were input into our deep learning model. Once the model is trained, predicting ground surface displacement from thousands of amplitude image chips takes only minutes using GPU-accelerated parallel inference. Comparative analysis with several contemporary deep learning models revealed that CC-ResSiamNet outperformed them all, exhibiting faster convergence in the learning curve, minimal overfitting, and higher prediction accuracy on the test sets.

Applying our trained model to different geographic regions and disaster scenarios presents a challenge, as deploying the model in entirely new areas or phenomena can result in significant domain shifts or extrapolation issues. Our inference results on glaciers in China and Chile, the Ridgecrest and Türkiye earthquakes, and the Taal and Mauna Loa volcanoes demonstrated that our trained model successfully predicted ground displacement when compared to independently produced traditional offset results. Although the predicted offsets tend to underestimate ground displacement in earthquake and volcanic events, the model accurately captured the overall displacement patterns and earthquake-induced ruptures. Several factors could explain why the model performed well in multi-scenario inferences: *1*) *Feature similarity*: SAR images capture distinct characteristics of the Earth's surface, and despite geographic differences, the key features indicating surface displacement may be similar across these scenarios. The model likely learned to detect these universal indicators rather than location-specific traits; *2*) *Scale invariance*: The model's architecture may have facilitated the development of scale-invariant features, enabling it to recognize patterns regardless of size, which is advantageous when handling different glacier sizes or varying scales of disaster impacts; *3*) *Sufficient training data*: Our training datasets included a wide variety of glacier types, sizes, and movement patterns, which may have inadvertently prepared the model for diverse scenarios. This variety likely contributed to its robustness and generalization capability; and *4*) *Generalization of displacement patterns*: The model may have learned to generalize displacement patterns rather than focusing on specific geographic features, allowing it to perform well across various scenarios where similar displacement dynamics occur. To further validate the applicability and generalization of our trained deep learning model, a cross-regional analysis with more extensive datasets in varying geographic conditions (e.g., earthquakes in tropical areas and volcanoes in arid regions) may be needed.

There remain various challenges in expanding the application of our deep learning model. The key to improving its performance lies in using more diverse training datasets. While training with 32 K samples over 100 epochs can produce valuable outputs, with a validation loss below 0.10 pixels, the limitations become evident during inference. Conventional offset tracking methods have considered accuracies on the order of one-tenth of the SAR pixel spacing to be acceptable (De Zan, 2014). However, the benefits of using more diverse datasets are particularly noticeable during inference. For instance, training with a single 32 K sample set may yield acceptable accuracy for test sets from Alaska's glaciers, but inference on the Tien Shan glaciers reveals issues, especially around boundary predictions. In contrast, iterative training with multiple datasets (32 K $\times$ 4) reduced these artifacts, leading to predictions more consistent with traditional offset results (Fig. 18). Diverse training datasets improve the model's robustness in multi-scenario inferences and help mitigate overfitting. Although our model shows relatively minor overfitting compared to others, its impact still cannot be entirely ignored. Adopting larger or more varied datasets, or refining the model itself, could further reduce overfitting. However, the challenge of gathering more extensive datasets remains, as it is both time-consuming and labor-intensive on a global scale. Additionally, our current computing resources, particularly memory, are insufficient to handle the larger datasets required for further improvements in prediction accuracy. They are also inadequate for implementing more advanced deep learning algorithms, such as transformer-based models, which could potentially enhance our results beyond the capabilities of the current U-Net-based approach.

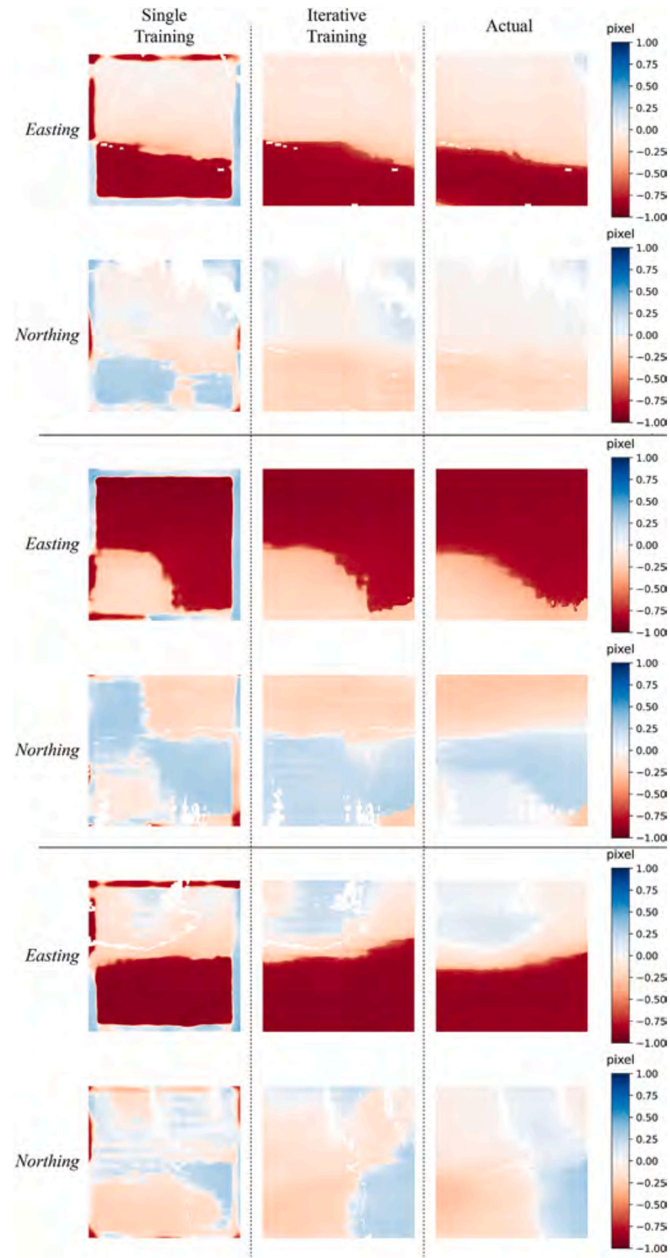

Fig. 18. Comparison of inferred offset results in the easting and northing directions from deep learning models trained on a single dataset (1st column) and multiple datasets (2nd column), alongside results from traditional offset methods (3rd column).

Future improvement in our deep learning model requires the estimation of small-scale displacements. The training data for our model used a minimum window size of 96 pixels, as smaller windows produced low SNR and noisy results over glacial areas. However, this approach has a drawback: the smoothing effect of larger windows prevents our model from learning pixel shifts caused by minor ground displacements. Consequently, our current models and others may struggle to predict the small-scale landslides from SAR amplitude images.

To address this limitation and enable the prediction of landslide-induced displacements, it is essential to expand the training dataset with a significant volume of target offsets that capture smaller-scale movements. Obtaining a sufficient number of training samples to estimate small-scale displacements in diverse geographic settings from SAR imagery using offset tracking can be challenging due to decorrelation limitations. To overcome this, generating synthetic and simulated displacements in different geometric configurations may be necessary to build adequate training datasets for deep learning models and improve prediction accuracy.

Although our model demonstrates strong performance across various scenarios, its accuracy may decrease in regions with complex terrain or noisy datasets, such as densely forested areas or steep mountain slopes. Future efforts could focus on integrating multi-source data, such as SAR from different sensors or fusion with optical imagery, to provide complementary information for mitigating noise and improving prediction accuracy. Additionally, enhancing sample selection strategies by incorporating diverse geographic and surface conditions could further increase the model's adaptability and robustness.

6. Conclusion

We have developed a novel deep learning approach for monitoring ground displacement using CSLC-S1 amplitude images. Our training methodology, which utilizes randomly assigned SAR inputs and corresponding offset outputs, highlights the critical role of a large, diverse dataset in achieving optimal performance and reducing overfitting. Evaluations on unseen test sets have produced promising results, showing a high degree of similarity between actual and predicted offsets. The model's versatility has been demonstrated through multi-scenario applications, including glacier dynamics, earthquake-affected regions, and volcanic zones. These assessments confirm the model's effectiveness in predicting ground surface displacements using a patch-wise prediction and stitching technique. Its capacity to generalize across different geographic locations and disaster events stems from the diversity of SAR datasets focused on glacier movements, allowing the model to adapt effectively to new environments. While there is room for improvement, particularly in handling small-scale displacements, expanding the global dataset, and developing time-series displacement predictions from multi-temporal SAR images using deep learning, our deep learning approach lays the groundwork for further research. With forthcoming spaceborne missions, such as NASA-ISRO SAR (NISAR) and Sentinel-1C/D, the demand for rapid and accurate ground displacement estimation through deep learning will continue to grow. This model offers a valuable framework for future developments in radar remote sensing and can serve as a benchmark for advancements in this field.

CRediT authorship contribution statement

Jinwoo Kim: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Hyung-Sup Jung:** Writing – review & editing, Software, Resources, Methodology, Conceptualization. **Zhong Lu:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no conflicts of interest relevant to this study.

Acknowledgement

This study was funded by Shuler-Foscue Endowment and Hamilton Visiting Scholar Program at Department of Earth Sciences at Southern Methodist University (SMU). This work was also supported by the Institute of Civil Military Technology Cooperation, the Defense Acquisition Program Administration, and the Ministry of Trade, Industry and Energy of Korea (22-CM-EO-02). Deep learning procedures were conducted using TensorFlow's Keras on SMU's NVIDIA DGX SuperPOD, utilizing 8 GPUs equipped with NVIDIA's A100 80GB. Constructive comments from Associate Editor (Mahdi Motagh) and four anonymous reviewers improved the manuscript.

Data availability

Datasets and code are available at https://zenodo.org/records/11358571.

References

Adrian, J., Sagan, V., Maimaitijiang, M., 2021. Sentinel SAR-optical fusion for crop type mapping using deep learning and Google earth engine. ISPRS J. Photogramm. Remote Sens. 175, 215–235.

An, Q., Feng, G., He, L., Xiong, Z., Lu, H., Wang, X., Wei, J., 2023. Three-dimensional deformation of the 2023 Turkey mw 7.8 and mw 7.7 earthquakes sequence obtained by fusing optical and SAR images. Remote Sens. 15 (10), 2656.

Bato, M.G., Lundgren, P., Pinel, V., Solidum Jr., R., Daag, A., Cahulogan, M., 2021. The 2020 eruption and large lateral dike emplacement at Taal volcano, Philippines: insights from satellite radar data. Geophys. Res. Lett. 48 (7), e2021GL092803.

Berardino, P., Fornaro, G., Lanari, R., Sansosti, E., 2002. A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. IEEE Trans. Geosci. Remote Sens. 40 (11), 2375–2383. https://doi.org/10.1109/TGRS.2002.803792.

Brancato, V., 2023. OPERA Level-2 Coregistered Single Look Complex (CSLC) from Sentinel-1 a/B Product Specification (Technical Report D-108278). Jet Propulsion Laboratory, Pasadena, CA.

Brockmann Consult, 2023. San Quintín Glacier. https://www.brockmann-consult.de/san-quintin-glacier/ (accessed 16 September 2024).

Castelletti, D., Farquharson, G., Stringham, C., Duersch, M., Eddy, D., 2021. Capella space first operational SAR satellite. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 11–16 July 2021. https://doi.org/10.1109/IGARSS47720.2021.9554100.

Chae, S.-H., Lee, W.-J., Baek, W.-K., Jung, H.-S., 2019. An improvement of the performance of SAR offset tracking approach to measure optimal surface displacements. IEEE Access 7, 131627–131637.

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 40 (4), 834–848.

Copernicus EU, 2021. San Quintín Glacier in Chile. https://www.copernicus.eu/en/media/image-day-gallery/san-quintin-glacier-chile (accessed 16 September 2024).

Curlander, J.C., McDonough, R.N., 1991. Synthetic Aperture Radar: Systems and Signal Processing, 11. Wiley, New York.

De Zan, F., 2014. Accuracy of incoherent speckle tracking for circular gaussian signals. IEEE Geosci. Remote Sens. Lett. 11 (1), 264–267.

Farinotti, D., Longuevergne, L., Moholdt, G., Duethmann, D., Mölg, T., Bolch, T., Vorogushyn, S., Güntner, A., 2015. Substantial glacier mass loss in the Tien Shan over the past 50 years. Nat. Geosci. 8, 716–722.

Fathian, A., Atzori, S., Nazari, H., Reicherter, K., Salvi, S., Svigkas, N., Tatar, M., Tolomei, C., Yaminifard, F., 2021. Complex co- and postseismic faulting of the 2017-2018 seismic sequence in western Iran revealed by InSAR and seismic data. Remote Sens. Environ. 253, 112224.

Feng, X., Chen, Z., Li, G., Ju, Q., Yang, Z., Cheng, X., 2023. Improving the capability of D-InSAR combined with offset-tracking for monitoring glacier velocity. Remote Sens. Environ. 285, 113394.

Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., Rucci, A., 2011. A new algorithm for processing interferometric data-stacks: SqueeSAR. IEEE Trans. Geosci. Remote Sens. 49 (9), 3460–3470.

Fielding, E.J., Liu, Z., Stephenson, O.L., Zhong, M., Liang, C., Moore, A., Yun, S.-H., Mark Simons, M., 2020. Surface deformation related to the 2019 M_w 7.1 and 6.4 Ridgecrest earthquakes in California from GPS, SAR interferometry, and SAR pixel offsets. Seismol. Res. Lett. 91 (4), 2035–2046.

Gisinger, C., Schubert, A., Breit, H., Garthwaite, M., Balss, U., Willberg, M., Small, D., Eineder, M., Miranda, N., 2020. In-depth verification of Sentinel-1 and TerraSAR-X geolocation accuracy using the Australian corner reflector array. IEEE Trans. Geosci. Remote Sens. 59 (2), 1154–1181.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Cambridge, MA.

Handwerger, A.L., Fielding, E.J., Huang, M.-H., Bennet, G.L., Liang, C., Schulz, W.H., 2019. Widespread initiation, reactivation, and acceleration of landslides in the northern California coast ranges due to extreme rainfall. J. Geophys. Res. Earth Surf. 124 (7), 1782–1797. https://doi.org/10.1029/2019JF005035.

Hanssen, R.F., 2021. Radar interferometry: data interpretation and error analysis, 2. Springer Science & Business Media. https://doi.org/10.1007/0-306-47633-9.

Harrison, S., Warren, C.R., Winchester, V., Aniya, M., 2001. Onset of rapid calving and retreat of glacier san Quintin, Hielo Patagónico Norte, southern Chile. Polar Geogr. 25 (1), 54–61.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: 2015 IEEE International Conference on Computer Vision, Santiago, Chile.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA.

He, L., Feng, G., Xu, W., Wang, Y., Xiong, Z., Gao, H., Liu, X., 2023. Coseismic kinematics of the 2023 Kahramanmaras, Turkey earthquake sequence from InSAR and optical data. Geophys. Res. Lett. 50 (17), e2023GL104693.

Hooper, A., Zebker, H., Segall, P., Kampes, B., 2004. A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers. Geophys. Res. Lett. 31 (23). https://doi.org/10.1029/2004GL021737.

Hübinger, C., Fluet-Chouinard, E., Hugelius, G., Peña, F.J., Jaramillo, F., 2024. Automating the detection of hydrological barriers and fragmentation in wetlands using deep learning and InSAR. Remote Sens. Environ. 311, 114314.

Hughes, L.H., Marcos, Lobry S., Tuia, D., Schmitt, M., 2020. A deep learning framework for matching of SAR and optical imagery. ISPRS J. Photogramm. Remote Sens. 169, 166–179.

Ignatenko, V., Laurila, P., Radius, A., Lamentowski, L., Antropov, O., Muff, D., 2020. ICEEYE microsatellite SAR constellation status update: Evaluation of first commercial imaging modes. In: IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, 26 Sep – 02 Oct, 2020. https://doi.org/10.1109/IGARSS39084.2020.9324531.

Jia, H., Wang, Y., Ge, D., Deng, Y., Wang, R., 2020. Improved offset tracking for predisaster deformation monitoring of the 2018 Jinsha River landslide (Tibet, China). Remote Sens. Environ. 247, 11899.

Kim, J.W., Lu, Z., 2021. Taking the pulse of aging oil wells in the Permian Basin from space. Sci. Total Environ. 757, 143932. https://doi.org/10.1016/j.scitotenv.2020.143932.

Kim, J.W., Coe, J.A., Lu, Z., Avdievitch, N.N., Hults, C.P., 2022. Spaceborne InSAR mapping of landslides and subsidence in rapidly deglaciating terrain, Glacier Bay National Park and preserve and vicinity, Alaska and British Columbia. Remote Sens. Environ. 281, 113231. https://doi.org/10.1016/j.rse.2022.113231.

Lattari, F., Leon, B.G., Asaro, F., Rucci, A., Prati, C., Matteucci, M., 2019. Deep learning for SAR image despeckling. Remote Sens. 11 (13), 1532.

Lazecký, M., Spaans, K., González, P.J., Maghsoudi, Y., Morishita, Y., Albino, F., Elliott, J., Greenall, N., Hatton, E., Hooper, A., et al., 2020. LiCSAR: An automatic InSAR tool for measuring and monitoring tectonic and volcanic activity. Remote Sens. 12 (15), 2430. https://doi.org/10.3390/rs12152430.

Lei, Y., Gardner, A., Agram, P., 2021. Autonomous repeat image feature tracking (autoRIFT) and its application for tracking ice displacement. Remote Sens. 13 (4), 749.

Lemley, J., Bazrafkan, S., Corcoran, P., 2017. Deep learning for consumer devices and services: Pusing the limits for machine learning, artificial intelligence, and computer vision. IEEE Consumer Electronics Magazine 6 (2), 48–56.

Li, H., 2018. Deep learning for natural language processing: advantages and challenges. Natl. Sci. Rev. 5 (1), 24–26.

Li, J., Xu, C., Su, H., Gao, L., Wang, T., 2022. Deep learning for SAR ship detection: past, present and future. Remote Sens. 14 (11), 2712.

Lu, Z., Dzurisin, D., 2014. InSAR Imaging of Aleutian Volcanoes: Monitoring a Volcanic Arc from Space. Springer Praxis Books: Geophysical Sciences, New York, USA.

Lu, Z., Kim, J.W., 2021. A framework for studying hydrology-driven landslide hazards in northwestern US using satellite InSAR, precipitation and soil moisture observations: early results and future directions. Geohazards 187, 49–61. https://doi.org/10.3390/geohazards2020002.

Lu, Z., Zhang, L., 2014. Frontiers of radar remote sensing. Photogramm. Eng. Rem. S. 80 (1), 5–13.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., Johnson, B.A., 2019. Deep learning in remote sensing applications: a meta-analysis and review. ISPRS J. Photogramm. Remote Sens. 152, 166–177.

Maher, S.P., Dawson, P., Hotovec-Ellis, A., Thelen, W.A., Jolly, A., Bennington, N., Chang, J.C., Dotray, P., 2023. Characterizing and locating seismic tremor during the 2022 eruption of Mauna Loa volcano, Hawaii, with network covariance. Seism. Record 3 (3), 228–238.

Massonnet, D., Feigl, K.L., 1998. Radar interferometry and its application to changes in the Earth's surface. Rev. Geophys. 36, 441–500.

Michaelides, R.J., Zebker, H.A., Zheng, Y., 2019. An algorithm for estimating and correcting decorrelation phase from InSAR data using closure phase triplets. IEEE Trans. Geosci. Remote Sens. 57 (12), 10390–10397.

Millan, R., Mouginot, J., Rabatel, A., Morlighem, M., 2022. Ice velocity and thickness of the world's glaciers. Nat. Geosci. 15, 124–129.

Miotto, R., Wang, F., Wang, S., Jiang, X., Dudley, J.T., 2018. Deep learning for healthcare: review, opportunities, and challenges. Brief. Bioinform. 19 (6), 1236–1246.

Moisseeva, N., Businger, S., Elias, T., 2023. VogCast: a framework for modeling volcanic air pollution and its application to the 2022 eruption of Mauna Loa volcano, Hawaii. J. Geophys. Res. Atmos. 128 (22), e2023JD039281.

Montagnon, T., Giffard-Roisin, S., Mura, M.D., Marchandon, M., Pathier, E., Hollingsworth, J., 2024. Sub-pixel displacement estimation with deep learning: application to optical satellite images containing sharp displacements. J. Geophys. Res.: Mach. Learn. Computat. 1, e2024JH000174.

Moore, J.G., Nakamura, K., Alcaraz, A., 1966. The 1965 eruption of Taal volcano. Science 151 (3713), 955–960.

Najafabadi, M.M., Villanustre, F., Khoshgoftaar, T.M., Seliya, N., Wald, R., Muharemagic, E., 2015. Deep learning applications and challenges in big data analytics. J. Big Data 2 (1).

Parikh, H., Patel, S., Patel, V., 2019. Classification of SAR and PolSAR images using deep learning: a review. Int. J. Image Data Fusion 11 (1), 1–32.

Qu, F.F., Lu, Z., Zhang, Q., Bawden, G.W., Kim, J.W., Zhao, C.Y., Qu, W., 2015. Mapping ground deformation over Houston-Galveston, Texas using multi-temporal InSAR. Remote Sens. Environ. 169, 290–306. https://doi.org/10.1016/j.rse.2015.08.027.

RGI 7.0 Consortium, 2023. Randolph Glacier Inventory – A dataset of global glacier outlines, Version 7.0. Boulder. NSIDC: National snow and Ice Data Center, Colorado, USA. https://doi.org/10.5067/f6jmovy5navz. Online access.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany.

Rosen, P.A., Hensley, S., Joughin, I.R., Li, F.K., Madsen, S.N., Rodriguez, E., Goldstein, R. M., 2000. Synthetic aperture radar interferometry. Proc. IEEE 88 (3), 333–382.

Saha, S., Bovolo, F., Bruzzone, L., 2021. Building change detection in VHR SAR images via unsupervised deep transcoding. IEEE Trans. Geosci. Remote Sens. 59 (3), 1917–1929.

Samsonov, S.V., Blais-Stevens, A., 2024. Estimating volume of large slow-moving deep-seated landslides in northern Canada from DInSAR-derived 2D and constrained 3D deformation rates. Remote Sens. Environ. 305, 114049.

Shen, H., Lin, L., Li, J., Yuan, Q., Zhao, L., 2020. A residual convolutional neural network for polarimetric SAR image super-resolution. ISPRS J. Photogramm. Remote Sens. 161, 90–108.

Spoorthi, G.E., Gorthi, R.K.S.S., Gorthi, S., 2020. PhaseNet 2.0: phase unwrapping of noisy data based on deep learning approach. IEEE Trans. Image Process. 29, 4862–4872.

Strozzi, T., Luckman, A., Murray, T., Wegmuller, U., Werner, C.L., 2002. Glacier motion estimation using SAR offset-tracking procedures. IEEE Trans. Geosci. Remote Sens. 40 (11), 2384–2391.

Tizzani, P., Fernández, J., Vitale, A., Escayo, J., Barone, A., Castaldo, R., Pepe, S., De Novellis, V., Solaro, G., Pepe, A., Tramelli, A., Hu, Z., Samsonov, S.V., Vigo, I., Tiampo, K.F., Comacho, A.G., 2024. 4D imaging of the volcano feeding system beneath the urban area of the Campi Flegrei caldera. Remote Sens. Environ. 315, 114480.

Wang, P., Bayram, B., Sertel, E., 2022. A comprehensive review on deep learning based remote sensing image super-resolution methods. Earth Sci. Rev. 232, 104110.

Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. CBAM: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), Computer Vision – ECCV 2018. Lecture Notes in computer science, Vol. 11211. Springer Cham, pp. 3–19.

Wu, X., Zhang, Z., Xiong, S., Zhang, W., Tang, J., Li, Z., An, B., Li, R., 2023. A near-real-time flood detection method based on deep learning and SAR images. Remote Sens. 15 (8), 2046.

Xia, Z., Motagh, M., Wang, W., Li, T., Peng, M., Zhou, C., Karimzadeh, S., 2024. Modeling slope instabilities with multi-temporal InSAR considering hydrogeological triggering factors: a case study across Badong County in the three gorges area. Remote Sens. Environ. 309, 114212.

Xu, Y., Lu, Z., Schulz, W.H., Kim, J.W., 2020a. Twelve-year dynamics and rainfall thresholds for alternating creep and rapid movement of the Hooskanaden landslide from integrating InSAR, pixel offset tracking, and borehole and hydrological measurements. J. Geophys. Res. Earth Surf. 125, e2020JF005640.

Xu, X., Sandwell, D.T., Smith-Konter, B., 2020b. Coseismic displacements and surface fractures from Sentinel-1 InSAR: 2019 Ridgecrest earthquakes. Seismol. Res. Lett. 91 (4), 1979–1985.

Zhan, Y., Fu, K., Yan, M., Sun, X., Wang, H., Qiu, X., 2017. Change detection based on deep Siamese convolutional network for optical aerial images. IEEE Geosci. Remote Sens. Lett. 14 (10), 1845–1849.

Zhang, A., Lipton, Z.C., Li, M., Smola, A.J., 2023. Dive into Deep Learning. Cambridge University Press.

Zhu, X.X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci. Remote Sens. Mag. 5 (4), 8–36.

Zhuang, L., Ke, C., Cai, Y., Nourani, V., 2023. Measuring glacier changes in the Tianshan Mountains over the past 20 years using Google earth engine and machine learning. J. Geogr. Sci. 33, 1939–1964.