

Article

Not peer-reviewed version

DEDNet: Dual-Encoder DeeplabV3+ Network for Rock Glacier Recognition Based on Multispectral Remote Sensing Image

[Lujun Lin](#) , [Lei Liu](#) , [Ming Liu](#) , Qunjia Zhang , [Min Feng](#) ^{*} , Yasir Shaheen Khalil , [Fang Yin](#)

Posted Date: 27 May 2024

doi: 10.20944/preprints202405.1770.v1

Keywords: Rock glacier; Dual-encoder DeeplabV3+; Multispectral remote sensing images; Spatial-spectral features



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

DEDNet: Dual-Encoder DeeplabV3+ Network for Rock Glacier Recognition Based on Multispectral Remote Sensing Image

Lujun Lin ^{1,†}, Lei Liu ^{1,2,†}, Ming Liu ³, Qunjia Zhang ¹, Min Feng ^{2,*}, Yasir Shaheen Khalil ⁴ and Fang Yin ³

¹ School of Earth Science and Resources, Chang'an University, Xi'an 710054, China; lujun_PSR@chd.edu.cn (L.J.L); liul@chd.edu.cn (L.L); qunjiazhong@chd.edu.cn (Q.J.Z)

² State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources (TPESER), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China

³ Shaanxi Key Laboratory of Land Consolidation, School of Land Engineering, Chang'an University, Xi'an 710054, China; mingliu@chd.edu.cn (M.L); yinf@chd.edu.cn (F.Y)

⁴ Geological Survey of Pakistan, Peshawar 25100, Pakistan; yasirshaheen2189@gmail.com (Y.S.K)

* Correspondence: mfeng@itpcas.ac.cn

† These authors contributed equally to this work.

Abstract: Understanding the distribution of rock glaciers provides key information for investigating and recognizing the status and changes of the periglacial environment. Deep learning algorithms and red-green-blue (RGB) bands from high-resolution satellite images have been extensively employed to map rock glaciers. However, the near-infrared (NIR) band offers rich spectral information and sharp edge features that could significantly contribute to semantic segmentation tasks, but it is rarely utilized in constructing rock glacier identification models due to the limitation of three input bands for classical semantic segmentation networks, like DeeplabV3+. In this study, a dual-encoder DeeplabV3+ network (DEDNet) was designed to overcome the flaws of classical DeeplabV3+ network (CDNet) when identifying rock glaciers using multispectral remote sensing images by extracting spatial and spectral features from RGB and NIR bands, respectively. This network, trained with manually labeled rock glacier samples from QiLian Mountains, established a model with accuracy, precision, recall, specificity, and mIoU (Mean Intersection over Union) of 0.9131, 0.9130, 0.9270, 0.9195, and 0.8601, respectively. The well-trained model was applied to identify new rock glaciers in a test region, achieving a producer's accuracy of 93.68% and a user's accuracy of 94.18%. Furthermore, the model was employed to two study areas in northern Tien Shan (Kazakhstan) and Daxue Shan (Hengduan Shan, China) with high accuracy, which proved that the DEDNet offers an innovative solution to more accurately map rock glaciers on a larger scale due to its robustness across diverse geographic regions.

Keywords: rock glacier; dual-encoder DeeplabV3+; multispectral remote sensing images; spatial-spectral features

1. Introduction

As a predominant type of cryospheric landform [1], rock glaciers play a significant hydrological role because of their high ratio of ice-content, especially in arid and semiarid areas, such as QiLian Mountains (QLMs) [2], where water security issues cause for concern, and rock glaciers situated at higher elevations may function as freshwater reserves in the future [3]. In addition, the rock glaciers may contain paleoclimatic information [4], which is a pivotal factor in semi-quantitatively assessing the meteorological environment of the forming period [5] and climate change on a local or regional scale [3]. Therefore, the recognition and delineation of rock glaciers are meaningful and helpful for evaluating their hydrological contribution [6], studying environment change and rock glacier dynamics [7].

In addition to field measurements [8], the combination of remote sensing images and geographic information system technologies with manual interpretation was used for accurately identifying rock glaciers [9], but these methods are time-consuming [10]. In recent years, studies have attempted to implement self-designed convolutional neural networks [11] along with object-based image analysis [12] to identify rock glaciers in different regions using synthetic aperture radar coherence, multispectral images and digital elevation model, which offered valuable methods to map rock glaciers automatically. Moreover, researchers have trained some mature deep learning networks [10], especially the robust classical DeeplabV3+ network (CDNet) [13], using the RGB bands from high-resolution satellite images as training imagery [14,15], and multi-source rock glacier inventories as training labels [16].

However, the rich spectral information and sharp edge features in near-infrared (NIR) band [17] are helpful for rock glacier recognition because the front and lateral margins with relatively high reflectance could serve as mandatory and general geomorphological criteria for identifying rock glaciers [18,19]. Despite this, NIR band has rarely been utilized in constructing semantic segmentation models for rock glaciers based on CDNet, primarily because CDNet can only accept for inputting three image bands and is limited to processing multispectral images [15]. Therefore, most of the CDNet-based models designed for rock glacier recognition could only use the three-bands natural color image as input datasets [13–15]. The pre-trained model could be fine-tuned to obtain a localized model [20]. Thus, designing a network that simultaneously extracts and fuses spatial and spectral features from three visible bands (red-green-blue, RGB) and the NIR band is helpful for improving image segmentation [21].

In semantic segmentation tasks, powerful backbones are typically employed to extract features [21]. The High-Resolution Net Version 2 (HRNetV2), as a robust backbone [22], has been extended to remote sensing image segmentation tasks [23]. HRNetV2 exhibits strong capability that maintains high-resolution representations through the whole process for extracting and fusing features [24], which is based on the mechanism of parallel multi-resolution convolutions combined with repeated multi-resolution fusions [22]. This capability of HRNetV2 effectively extracts the spatial information, boundary details, and global relationships in images, which are highly beneficial for image segmentation [25].

Therefore, a dual-encoder DeeplabV3+ network (DEDNet) with backbone of HRNetV2-W48 (also called HRNetV2 for simplicity) was designed to optimize the CDNet for processing multispectral remote sensing images and to identify rock glaciers based on GaoFen1/6 (GF1/6) satellite images.

2. Study Area and Materials

2.1. Study Area

The study area is located at the northeast of the QLMs in the northeastern edge of the Tibetan Plateau (**Error! Reference source not found.** a), covering the Qilian county and Menyuan county in Qinghai province, ranging from 99°24'37" to 102°33'27"E and 36°49'25" to 38°41'8"N (**Error! Reference source not found.** b). The primary manifestation of the study area is the multiple parallel mountains, where the developed periglacial landforms continuously supply fresh water to the Heihe and Shiyang Rivers, nourishing the Hexi Corridor and surrounding areas [26]. The QLMs are characterized by extensive distribution of active, transitional, and relict type rock glaciers, particularly concentrated in the northeastern region [27]. To develop a model capable of identifying these diverse types of rock glaciers, it is imperative to compile a dataset that includes samples representing all three categories. Therefore, four sub-areas, referred to as visually interpretation areas (VIAs), were selected for manual interpretation of the rock glaciers (**Error! Reference source not found.** c). Sub-areas A and C exhibit a prevalence of active and transitional rock glaciers, whereas transitional and relict rock glaciers dominate in sub-areas B and D. These areas served as the foundation for constructing a dataset for training the DEDNet. Additionally, another separate sub-region, known as the model test area (MTA), was selected to assess the model's robustness. The

mountain ridges within the MTA are covered by glaciers and snow, creating favorable conditions for rock glacier development.

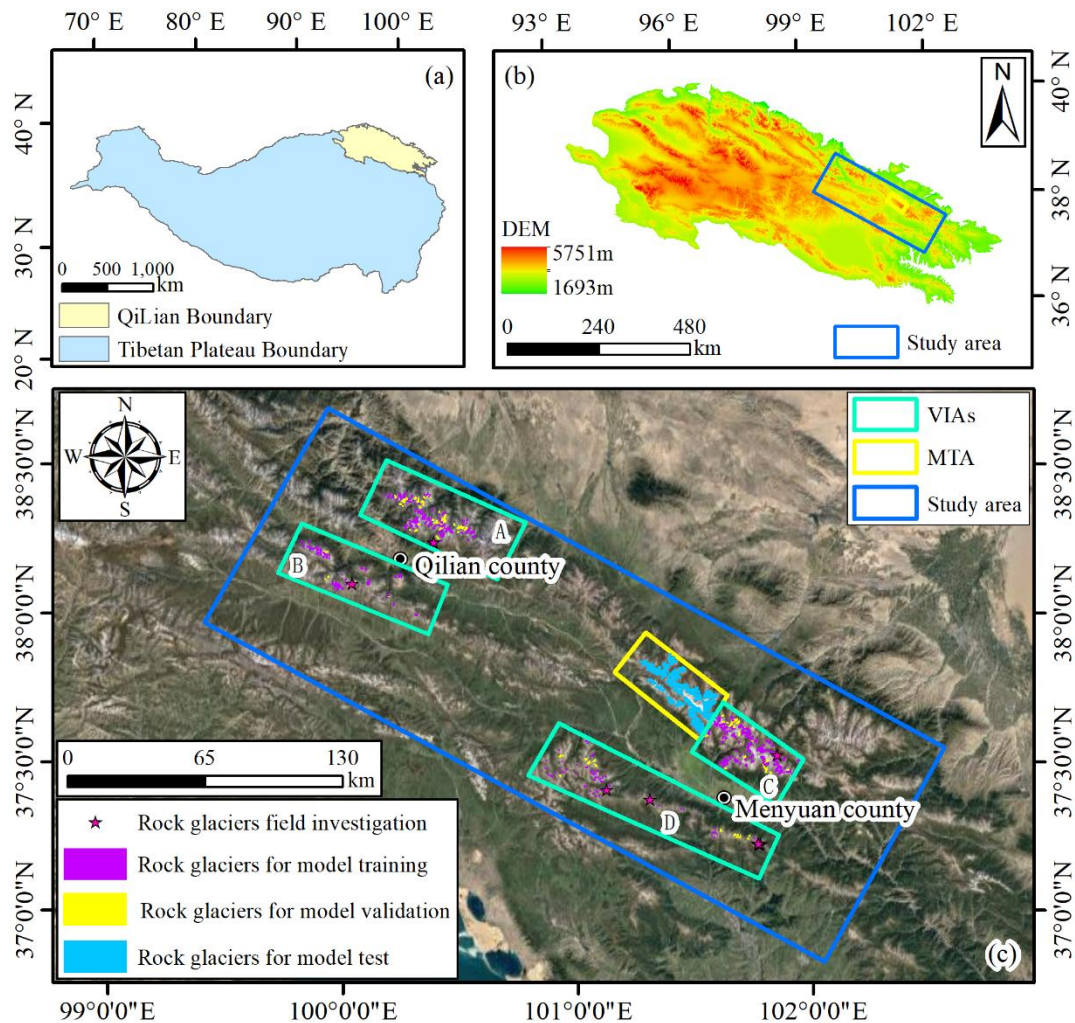


Figure 1. (a) Location of QILMs (b) The location of the study area in QILMs. (c) Location of VIAs (visually interpretation areas), MTA (model test area), field investigation, model training, model validation, and model test rock glaciers in the study area.

2.2. Imagery

Seventeen scenes of GF1/6 satellite images (<http://www.sasclouds.com/chinese/home>) were used in this study (**Error! Reference source not found.**). The GF1/6 images include four multispectral bands (RGB and NIR) with an 8m spatial resolution and a panchromatic band with a 2m spatial resolution. Radiometric calibration, and atmospheric correction using QUAC were employed for GF1/6 images. Subsequently, image fusion based on Gram-Schmidt pan-sharpening were used to preprocess the GF1/6, through Envi5.6 software, to produce four pansharpened multispectral bands with 2m pixels. After that, geographic registration based on Google Earth images was conducted using the ArcGIS 10.8 platform. Due to the significance of vegetation indices in rock glacier recognition [12], the SAVI (Soil Adjusted Vegetation Index) [28], and EVI (Enhanced Vegetation Index) [29] were computed using the NIR, R, and B bands of GF1/6 data. Subsequently, RGB bands were extracted to produce true color images. The NIR band, EVI, and SAVI were then combined into new three-band images, referred to as spectral images, through layer stacking.

Table 1. List of GF1/6 satellite images.

Date	Sensor	Sensor ID	Resolution
26.08.2015	GF1	GF1_PMS1_E101.4_N37.5_20150826_L1A0000999810	2/8m
26.08.2015	GF1	GF1_PMS2_E101.8_N37.4_20150826_L1A0000999891	
26.08.2015	GF1	GF1_PMS2_E101.8_N37.7_20150826_L1A0000999890	
26.08.2015	GF1	GF1_PMS1_E101.5_N37.8_20150826_L1A0000999809	
28.08.2020	GF1	GF1_PMS2_E100.4_N38.2_20200828_L1A0005019914	
28.08.2020	GF1	GF1_PMS1_E100.0_N38.3_20200828_L1A0005019873	
26.07.2020	GF1	GF1_PMS2_E101.4_N37.7_20200726_L1A0004951368	
28.08.2020	GF1	GF1_PMS1_E99.9_N38.0_20200828_L1A0005019874	
28.08.2020	GF1	GF1_PMS2_E100.3_N38.0_20200828_L1A0005019915	
26.07.2020	GF1	GF1_PMS2_E101.4_N37.4_20200726_L1A0004951369	
29.07.2020	GF1	GF1_PMS1_E101.0_N37.5_20200726_L1A0004951209	
07.09.2021	GF6	GF6_PMS_E100.9_N37.3_20210907_L1A1120139417	
26.08.2020	GF6	GF6_PMS_E100.0_N38.7_20200826_L1A1120029769	
03.05.2020	GF6	GF6_PMS_E101.0_N38.0_20200503_L1A1119993834	
01.06.2021	GF6	GF6_PMS_E99.3_N38.0_20210601_L1A1120110250	
01.08.2021	GF6	GF6_PMS_E101.5_N37.3_20210801_L1A1120127842	
26.08.2020	GF6	GF6_PMS_E99.8_N38.0_20200826_L1A1120030072	

3. Methodology

The methodology consists of five crucial steps. First, the rock glaciers in the VIAs and MTA were manually delineated (RG_man) based on the baseline concepts proposed by the International Permafrost Association Action Group [19], and field-investigated geomorphologic features, including developed scale and vegetation coverage, etc. Second, a DEDNet was designed to independently extract spatial and spectral features from true color images and spectral images, and then fuse these features. Third, the labeled images were used to train the DEDNet to obtain a model to identify rock glaciers. Fourth, a workflow was proposed to map and post-process rock glaciers with the well-trained model (RG_mm) based on the true color images and spectral images in VIAs and MTA. Fifth, the area of each RG_mm was calculated and compared to that of RG_man to assess the model's accuracy in the VIAs and its robustness in the MTA.

3.1. Delineating Rock Glaciers with GF1/6 Images and Google Earth

In order to ascertain the geomorphological features of rock glaciers in QLMs, a field investigation was conducted from August 1st to August 15th, 2023 (**Error! Reference source not found.** a1-d1). The main body of rock glaciers is identifiable based on the field-investigated features such as ridges-furrows and heterogeneous vegetation distribution in natural color images, but accurately delineating the boundary is challenging (**Error! Reference source not found.** a2-d2). With the help of a 3D oblique view on Google Earth (<https://earth.google.com>), the delineation of boundaries became more accessible since the elevation difference between the edge of rock glaciers and the valley is obvious (**Error! Reference source not found.** a3-d3). Utilizing the ArcGIS 10.8 platform for assistance, the rock glaciers were meticulously outlined and checked, resulting in 608 polygons in VIAs and 190 polygons in MTA, respectively. The polygons covered mainly by snow and clouds in the natural color images were discarded due to their serious hindering effect for the recognition of rock glaciers [11,13,14]. As a result, 542 snow-free and cloud-free polygons in VIAs and 190 in MTA were ultimately retained.

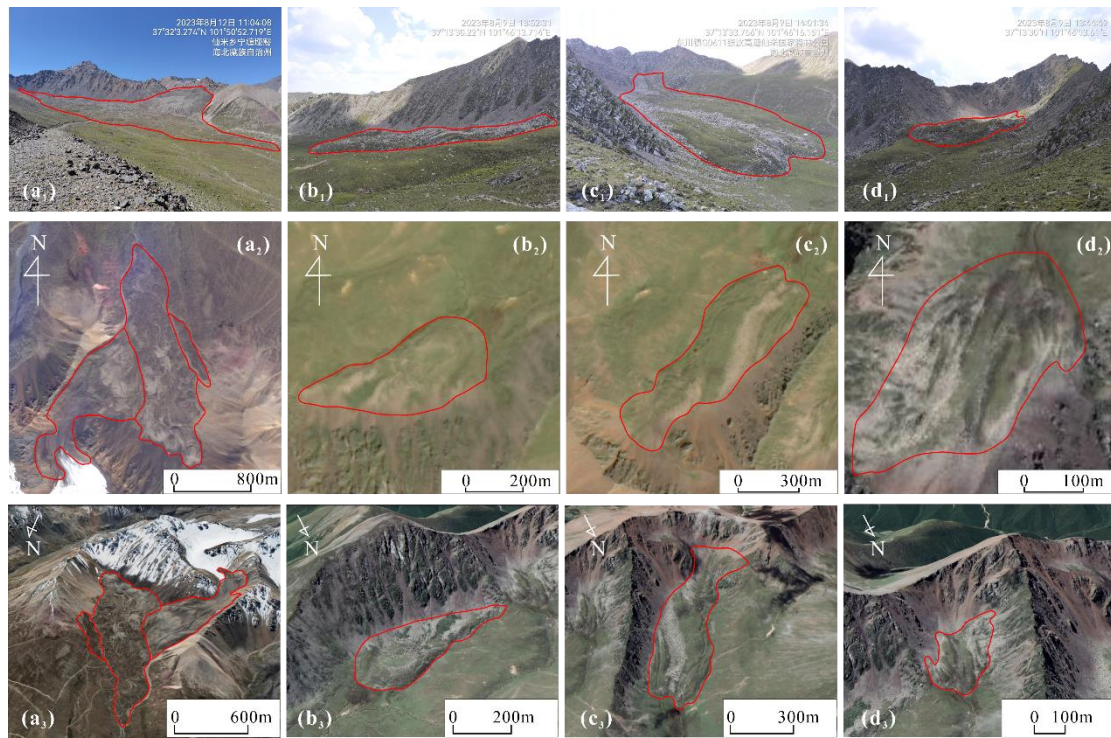


Figure 2. The characteristics of four rock glaciers: (a1-d1) field photos, (a2-d2) GF1/6 images, and (a3-d3) Google Earth 3D oblique view. The red lines represent the boundary of rock glaciers.

3.2. Designing the DEDNet

The DEDNet, based on the CDNet, inherits the classical encoder-decoder structure, where the two encoder modules extract low-level and high-level features by utilizing HRNetV2 and multi-scale atrous convolution, and the decoder module refines segmentation results by fusing these features (**Error! Reference source not found.**). Considering the training efficiency, each encoder designed to process images with three channels, then the pre-trained HRNetV2 could be used through transfer learning. The first encoder processes true color images to extract spatial features like texture, shape, and spatial relationships, while the second encoder handles spectral images to extract spectral features. Two blocks were designed and incorporated into the decoder module to fuse spatial and spectral features. Block 1 was utilized to fuse low-level features, such as local texture, shape, and edge [30], while block 2 was utilized to fuse high-level contextual information in semantic features [31]. In Block 1, low-level feature 1 extracted from true color images and low-level feature 2 extracted from spectral images (both with a shape of $256 \times 88 \times 88$, where 256 represents the channels, and 88 represents the height and width) were initially concatenated to form a new feature map (with a shape of $512 \times 88 \times 88$). The channels of the new feature map are twice that of each low-level feature. Then, dimensionality reduction was performed on the new feature map using 1×1 convolutional operations to match the dimensions of each low-level feature. Inspired by the residual network [32], a shortcut connection was embedded into block 1 to add low-level feature 1 to the new feature map. This addition maximally preserves the low-level information crucial for distinguishing inter-class difference [33], as these features are primarily obtained from the true color images [34]. Apart from the shortcut connection operation, Block 2 mirrors the structure of Block 1. Since high-level semantic information contributes to semantic image segmentation [35], only concatenation and dimensionality reduction operations were carried out in block 2.

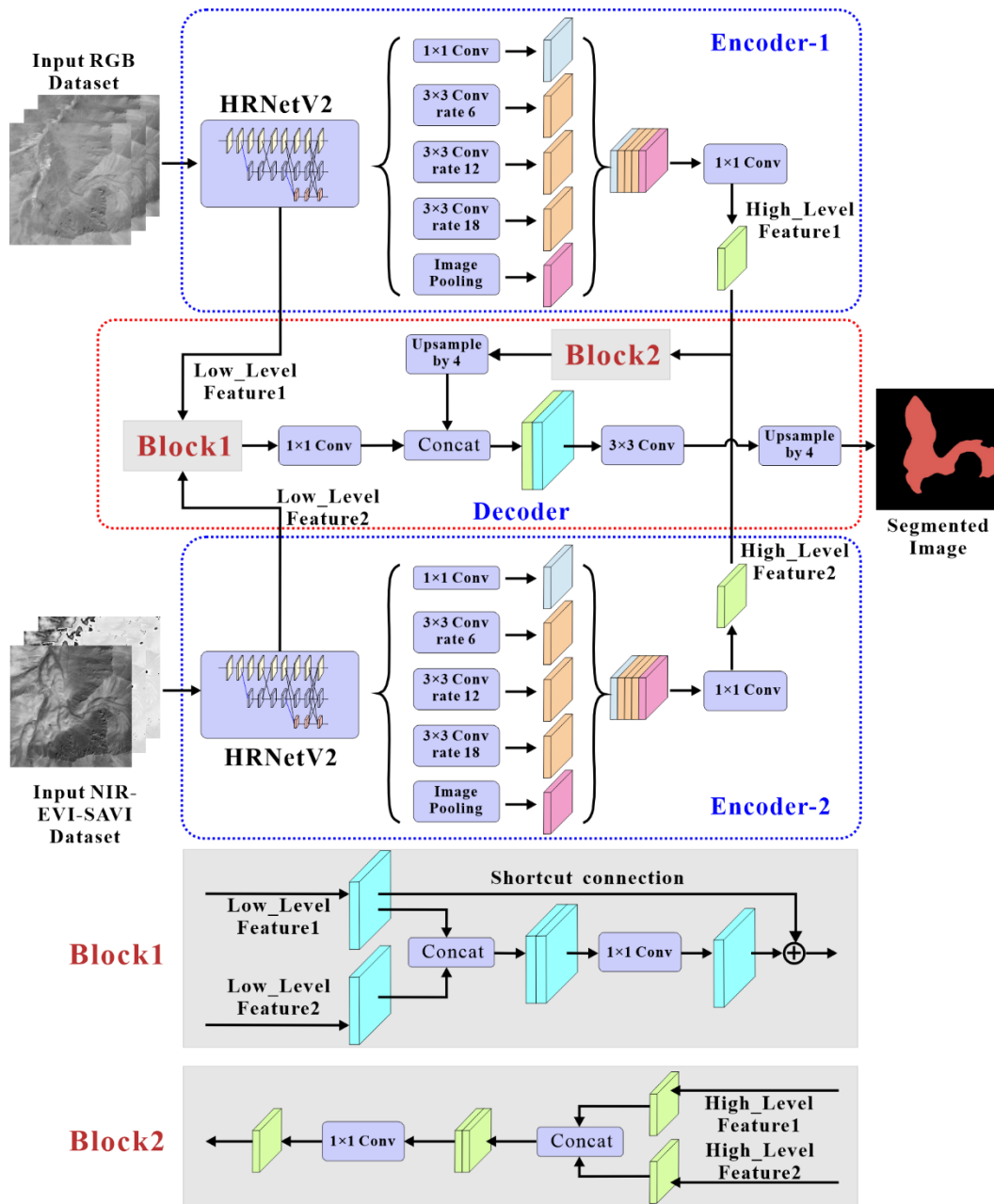


Figure 3. The DEDNet with the backbone of HRNetV2 and the block 1 and block 2 we designed.

3.3. Training and Validating the DEDNet

3.3.1. Preparing the Training and Validation Dataset

The true color images and spectral images were used to create a dataset for training the DEDNet. In this dataset, 542 snow-free and cloud-free RG_man polygons in the VIAs were employed to generate positive samples and their corresponding labels, while 1353 landforms with textures similar to rock glaciers were selected from areas near the rock glaciers to generate negative samples and their labels. To generate positive samples, we initially created $2\text{km} \times 2\text{km}$ rectangles by buffering the centroid of each rock glacier polygon with a distance of 1km. Then, the true color images and spectral images were clipped using these rectangles. For negative sample creation, we generated 1353 rectangles of $2\text{km} \times 2\text{km}$, covering the 1353 landforms, and subsequently clipped the true color images and spectral images. After that, each layer of both positive and negative samples was normalized by the maximum and minimum values for relevant layer of all the samples. Finally, as the samples with a spatial resolution of 2m exceed the memory limit of the GPU we used when

training the DEDNet, they were resampled to a 4m pixel size to accommodate computer performance. The positive labels were generated by creating binary raster images for each positive sample, with pixel values assigned based on the presence of rock glaciers. Pixels within the rock glacier polygons were assigned 1, representing the presence of rock glaciers, while pixels outside the polygons were assigned 0, representing the background. Similarly, the negative labels were generated by creating raster images with a single pixel value of 0 for each negative sample. To avoid data leakage, we picked the positive samples and labels that consist of either a single or 2 to 3 rock glaciers that were not contained in the other positive samples as the validation set. The remaining positive samples and labels were selected as the training set. Following an 8:2 ratio, the 542 positive samples were divided into 433 for training and 109 for validation, while the 1353 negative samples were split into 1082 for training and 271 for validation.

3.3.2. Training and Validating the DEDNet

The DEDNet was trained and validated with some universal hyper-parameters, including: optimizer - Adaptive Moment Estimation, initial learning rate - 3×10^{-5} , learning rate scheduler - Cosine, base size - 500×500 pixels, crop size - 352×352 pixels, batch size - 4, etc. The Dice Loss function was selected to address data imbalance issues (Zhao et al., 2020), given that rock glaciers occupy a small portion in permafrost landforms (Hu et al., 2022). The pre-trained HRNetV2, downloaded from the PyTorch official website, was used for transfer learning to train the DEDNet. The DEDNet was trained on an NVIDIA GeForce RTX 4080 GPU.

3.3.3. Evaluating Metrics

Accuracy, precision, recall, specificity, and mIoU (mean intersection over union) were employed to comprehensively assess model performance. Among them, accuracy is the percentage of true rock glaciers and background in all pixels. Precision is the percentage of true rock glaciers in all detected rock glaciers. Recall is the percentage of the true rock glaciers in all manually-labelled rock glaciers. Specificity is the percentage of true background in all manually-labelled background. The mIoU, providing an overall assessment of segmentation accuracy, is the average of the intersection over union scores for rock glaciers and background.

3.4. Testing the Well-Trained Model

3.4.1. Preparing the Test Dataset

Testing the model robustness using the 190 RG_man polygons in MTA is insufficient to assess the overall performance of the model, such as its propensity to mistakenly identify surrounding landforms with textures similar to rock glaciers. Instead, GF1/6 images covering the MTA should be fed into the model to identify rock glaciers. Therefore, a fishing net covering MTA with each rectangle size of 1km×1km was produced and then used to create the test dataset, following the steps of generating positive samples. In the test dataset, any two adjacent samples overlapped with an area of half size of each sample, which allowed the well-trained model to identify the same rock glacier from different perspectives.

3.4.2. Mapping and Post-Processing Rock Glaciers

Several post-processing steps (**Error! Reference source not found.**) were implemented to improve the quality of recognition results. (1) The lowest area threshold, the smallest area among all manually interpreted rock glacier polygons, was set to remove polygons smaller than 0.022 km². (2) The holes enclosed in the RG_mm polygons were filled for continuity. (3) The RG_mm polygons were extracted from all test samples and were converted into vectors. (4) The single polygon that didn't intersect with other one was discarded, which might be misidentified due to a certain degree of randomness during the recognition process. (5) The union boundary of multiple intersection polygon vectors was calculated. (6) All polygons located within the permafrost range [36] were retained since rock glaciers mainly distributed in the permafrost zone [9].

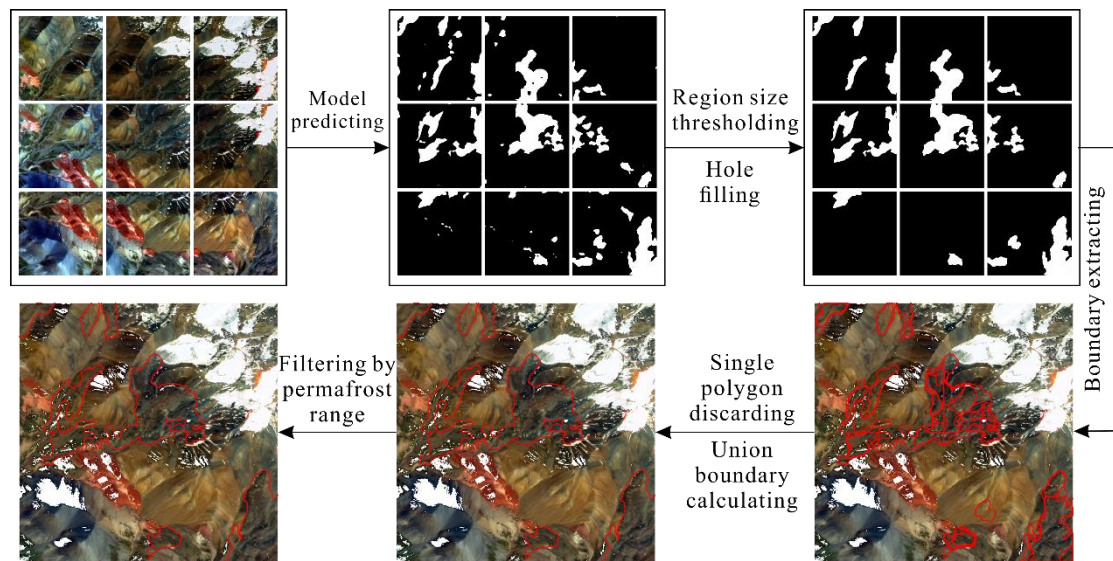


Figure 4. The flowchart of mapping rock glaciers in MTA.

3.4.3. Testing Method and Metrics

The model evaluation metrics selected during the training and validating the DEDNet can assess the overall performance of the model, posing challenges for evaluating the recognition accuracy of individual rock glaciers. Therefore, following the evaluation method in [12], we employ the user's accuracy (the percentage of the model classification that is actually a rock glacier) and the producer's accuracy (the percentage of total rock glaciers that were classified by the model) to evaluate the model performance. In addition, to further explore the identification accuracy of individual rock glacier, we also extracted each polygon's area from the RG_mm and then compared that from RG_man to compute the area deviation. However, during mapping the new rock glaciers, the calculation of the union boundary of multiple intersection vectors may result in multiple adjacent rock glaciers being represented as a single one. Thus, the area-extracting method was designed with three scenarios: the first scenarios is a single RG_man polygon surrounded by a single RG_mm polygon or vice versa, where the polygon area of RG_man corresponds to that of RG_mm. We calculating the area of each rock glacier individually (A in **Error! Reference source not found.**). The second scenario, several RG_man polygons are surrounded by a single RG_mm, where the combined area of these RG_man polygons correspond to the area of the single RG_mm. We calculated the combined area of several RG_man polygons and the area of the single RG_mm polygon (B in **Error! Reference source not found.**). The last scenario is the opposite of the second scenario, where we calculated the combined area of several RG_mm polygons and the area of the single RG_man polygon (C in **Error! Reference source not found.**). After that, we compute the area deviation by taking the absolute value of the difference in polygon area between RG_man and RG_mm. The polygon area of RG_man and RG_mm was plotted in scatter plots, and the area deviations were used to create box plots, facilitating the analysis of recognition accuracy for individual rock glaciers.

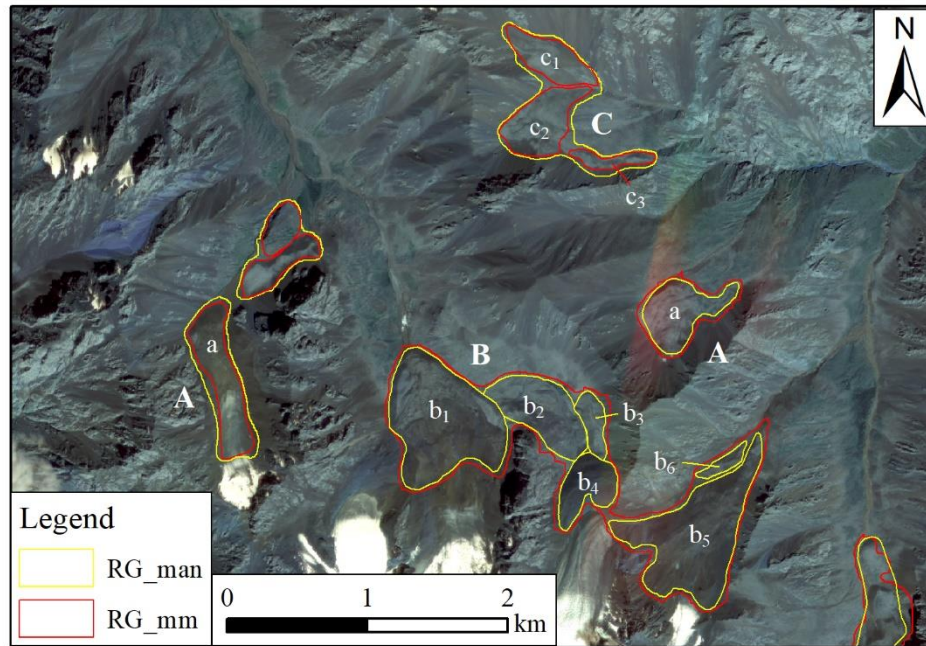


Figure 5. Three area-extracting methods, where the area of A corresponds to that of a, the area of B corresponds to the combine area of b₁, b₂, b₃, b₄, b₅, and b₆, the area of C corresponds to the combine area of c₁, c₂, and c₃.

4. Results

After 100 epochs of training and validating, the model converged with the loss value on the validation set floating around 0.20, and we obtained the best model with an accuracy, precision, recall, specificity, and mIoU of 0.9131, 0.9130, 0.9270, 0.9195, and 0.8601, respectively, on validation set. Subsequently, flowing the steps in the flowchart (**Error! Reference source not found.**), the model was employed to map rock glaciers in VIAs and MTA, respectively.

4.1. Mapping Rock Glaciers in VIAs

Compared to the 542 RG_man polygons, the model identified 536 of them, and only 6 rock glaciers with an area less than 0.11 km² were not recognized. Additionally, the model erroneously identified 17 non-rock glaciers as rock glaciers due to their textural similarity. The total area of 542 RG_man polygons is 155.07 km², while the model delineated an area of 174.27 km², indicating an overestimation of 12.38%. This is mainly because multiple adjacent rock glaciers may be represented as a single one [15], causing the inclusion of non-rock glacier areas between them.

We selected several representative polygons of RG_man and their corresponding RG_mm from four sub-areas to illustrate the strengths and weaknesses of the model (**Error! Reference source not found.**). Overall, the RG_mm polygons on validation dataset are visually accurate, and the boundaries of RG_mm are natural and basically consistent with RG_man, especially in sub-area A, B, and C (**Error! Reference source not found.** c-e), where the rock glaciers are active or transitional and exhibit readily identifiable morphological features. However, in sub-region D, the rock glaciers are mainly relict, featuring indistinct morphological features, and vegetation coverage similarity to surrounding landforms, resulting in insufficient consistency between the boundaries of few RG_mm and RG_man (**Error! Reference source not found.** f). This is consistent with previous studies that rock glaciers with apparent ridges and furrows that could be correctly identified (Marcer, 2020), but smaller rock glaciers with subdued topography and/or evenly distributed vegetation cover were more likely to be missed (Hu et al., 2022).

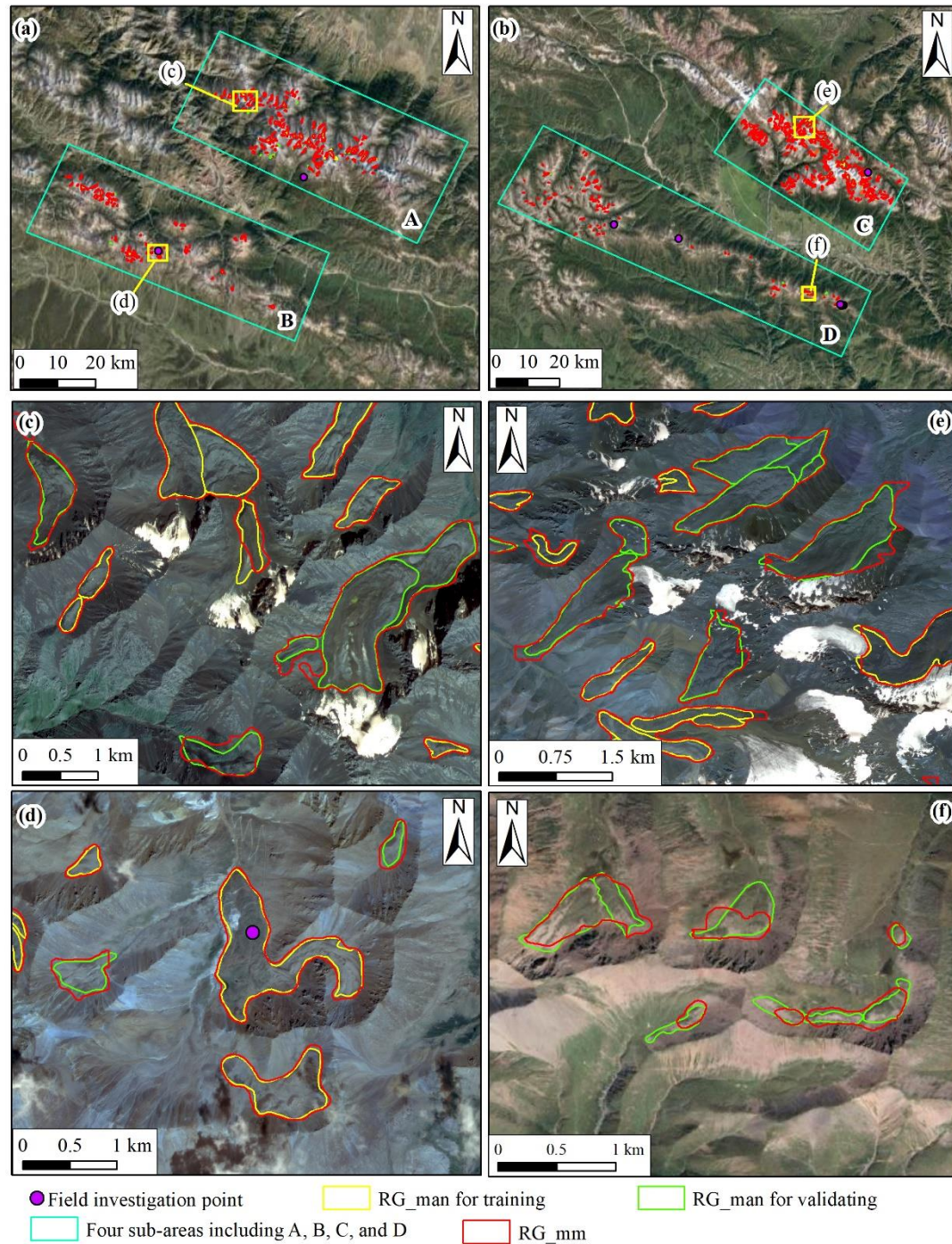


Figure 6. Comparison of boundaries between RG_mm and RG_man in VIAs (a and b). The panels (c-f) correspond to the yellow box in sub-area A, B, C, and D. The background image is a true color image composed of a combination of RGB bands of GF1/6.

To assess the identification accuracy of each individual rock glacier, the areas of each RG_mm polygon and their corresponding RG_man polygon were extracted following the method illustrated in "3.4.3 Testing method and metrics". Nearly all points in sub-areas A, B, C, and D fell near the 1:1 regression line, representing a high accuracy level in identifying each rock glacier on training and validation dataset (**Error! Reference source not found. a**). The range (whiskers) and interquartile range (box length) of the boxplots showed a gradually decreasing trend as the scale of rock glaciers increased, indicating an improvement in the model's recognition accuracy with the increasing size of rock glaciers (**Error! Reference source not found. b**).

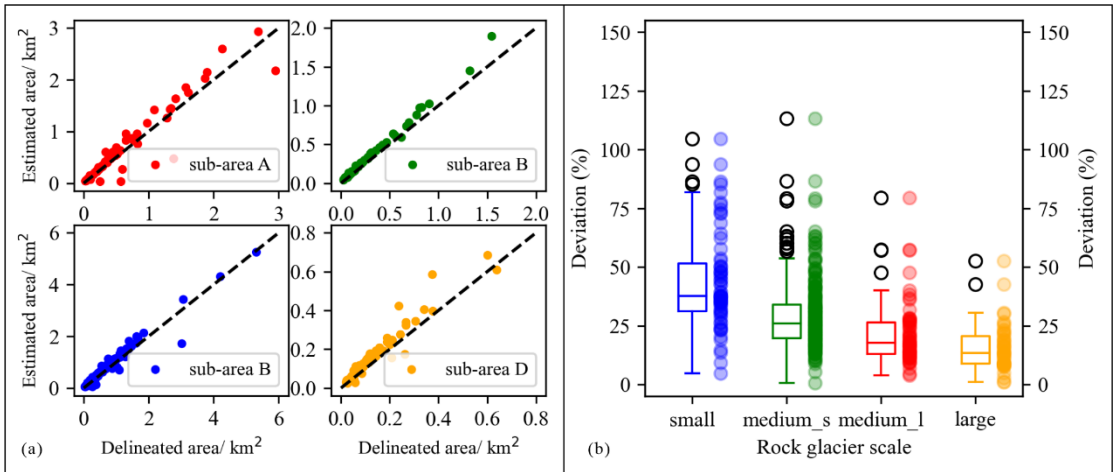


Figure 7. Scatterplots illustrating the areas of RG_mm polygons compared to RG_man polygons on the training and validation datasets (a), and boxplots displaying the deviations in area across various scales on the training and validation datasets (b). In boxplots, "small" denotes areas less than 0.10 km², "medium_s" refers to areas between 0.10 and 0.50 km², "medium_l" signifies areas between 0.50 and 1.00 km², and "large" indicates areas larger than 1.00 km².

4.2. Mapping Rock Glaciers in MTA

Using the confusion matrix and metrics values (**Error! Reference source not found.**), we evaluated the model on test dataset. The result shows that the model identified 178 out of 190 RG_man polygons with a Produce’ Accuracy of 93.68%. There are 12 RG_man polygons that the model failed to recognize with the smallest area slightly larger than 0.03 km². Additionally, the model incorrectly identified 11 non-rock glacier landforms as rock glaciers, resulting in User’ Accuracy of 94.18%. Among these 11 misclassified rock glaciers, 10 have areas smaller than 0.10 km². Compared to larger rock glaciers, the model is more likely to misclassify rock glaciers with areas smaller than 0.10 km².

Table 2. Confusion matrix of the model on test dataset and accuracy metrics.

Metrics	Note	Result
True Positive (TP)	Number of correct RG_mm	178
False Positive (FP)	Number of wrong RG_mm	11
False Negative (FN)	Number of missed RG_man	12
Producer's accuracy	TP/(TP+FN)	0.9368
User's accuracy	TP/(TP+FP)	0.9418

When analyzing these 178 RG_mm polygons on the GF1/6 imagery (**Error! Reference source not found.** a), we observed a generally high consistency between the boundaries of RG_mm and RG_man, with superior performance on larger rock glaciers compared to smaller ones (**Error! Reference source not found.** b and c). The polygons were filled by multiple overlapping heatmaps with pixel values close to 1 (**Error! Reference source not found.** d and e), signifying the excellent dependability of RG_mm.

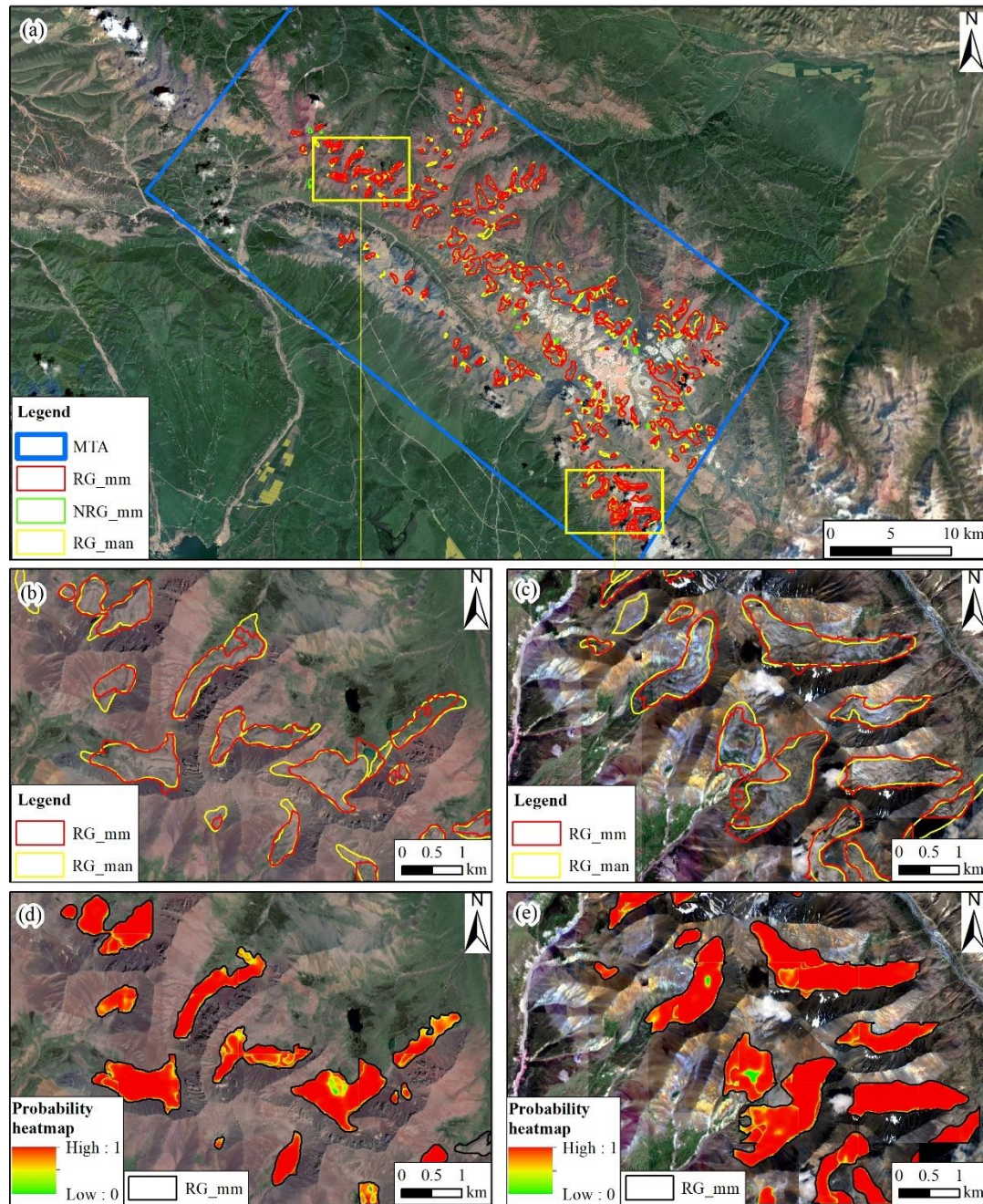


Figure 8. Comparison of RG_man' boundaries with RG_mm' boundaries in MTA (a), along with two typical sub-regions (b and c), and their corresponding probability heatmap (d and e) on GF1/6 imagery. NRG_mm represent non-rock glacier model mapped. Background image is a true color image composed of a combination of RGB bands of GF1/6.

Compared to the total area of 70.14 km² of the 178 RG_man polygons, the model delineated 181 polygons with total area of 67.38 km², indicating a slight underestimation of 3.94%. The vast majority of data points closely aligned with the 1:1 regression line, indicating a high degree of accuracy in identifying individual rock glaciers within the test area (**Error! Reference source not found. a**). As the size of rock glaciers increased, the range (whiskers) and interquartile range (box length) of the boxplots consistently decreased, suggesting an enhancement in the model's recognition accuracy with larger rock glaciers (**Error! Reference source not found. b**). Estimations of individual rock glacier areas exhibit significant variation. Larger rock glaciers were generally identified and delineated with high accuracies, with an average overestimation of 1.81% for glacier areas larger than 1.00 km², and an average underestimation of 2.51% for glacier areas between 0.50 km² and 1.00 km².

Conversely, smaller rock glaciers exhibit lower identification accuracies, with an average underestimation of 13.57% for glacier areas between 0.1 km² and 0.5 km², and an average underestimation of 20.73% for glacier areas smaller than 0.10 km².

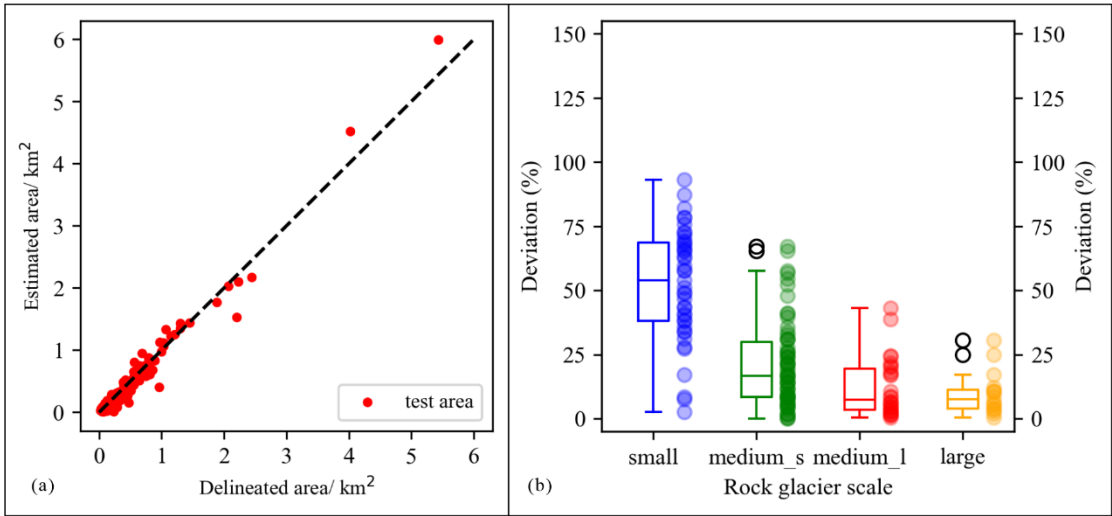


Figure 9. Scatterplots illustrating the areas of RG_mm compared to RG_man on the test datasets (a), and boxplots displaying the deviations in area across various scales on the test datasets (b). "small", "medium_s", "medium_l", and "large" correspond to the same meanings as the terms used in Figure 7.

5. Discussion

5.1. Ablation Experiment

When designing DEDNet based on the CDNet, we proposed various networks, each of which was trained on positive samples (or including negative samples) with the same hyperparameters and evaluated using mIoU, as shown in **Error! Reference source not found..** The CDNet trained on the NIR-EVI-SAVI dataset achieved a comparable mIoU value to that trained on the RGB dataset. The model obtained by exclusively using Block2 for feature fusion in the DEDNet achieved a higher mIoU value compared to the model obtained by exclusively using Block1, indicating that high-level semantic features are more beneficial for rock glacier identification than low-level features. Of course, simultaneously applying both block1 and block2 to the DEDNet results in a superior rock glacier identification model, with an mIoU of 0.8519.

Table 3. The mIoU of different network trained with different dataset.

Network	RGB	NIR-EVI-SAVI	Block1	Block2	Negative Sample	mIoU
CDNet	√					0.8469
CDNet		√				0.8464
DEDNet	√	√	√			0.8348
DEDNet	√	√		√		0.8509
DEDNet	√	√	√	√		0.8519
CDNet	√				√	0.8457
DEDNet	√	√	√	√	√	0.8601

When training the CDNet and DEDNet including negative samples, the mIoU based on CDNet decreased from 0.8469 to 0.8457, while for the DEDNet, it increased from 0.8519 to 0.8601. The four models—training CDNet on positive samples (CDNet_Positive), training DEDNet on positive samples (DEDNet_Positive), training CDNet on positive and negative samples

(CDNet_Positive_Negative), and training DEDNet on positive and negative samples (DEDNet_Positive_Negative)— were utilized to map new rock glaciers in test area (**Error! Reference source not found.**). The CDNet_Positive and DEDNet_Positive models classified some non-rock glacier landforms as rock glaciers, whereas the CDNet_Positive_Negative and DEDNet_Positive_Negative models significantly reduced this error, highlighting the importance of including negative samples during training [11,13]. Comparing the recognition results of models between CDNet_Positive_Negative and DEDNet_Positive_Negative, we observed that both models were able to identify larger rock glaciers (active or transitional), and they both exhibited instances where adjacent rock glaciers were delineated as a single union rock glacier (black arrows in a1 and a3 in **Error! Reference source not found.**), which has also been reported in other rock glacier identification models [15]. However, for delineating rock glacier boundaries, especially the smaller and relict rock glaciers, the recognition performance of DEDNet_Positive_Negative was noticeably better (white arrows in sub-regions a1, a3, b1, b3 and c1, c3 in **Error! Reference source not found.**).

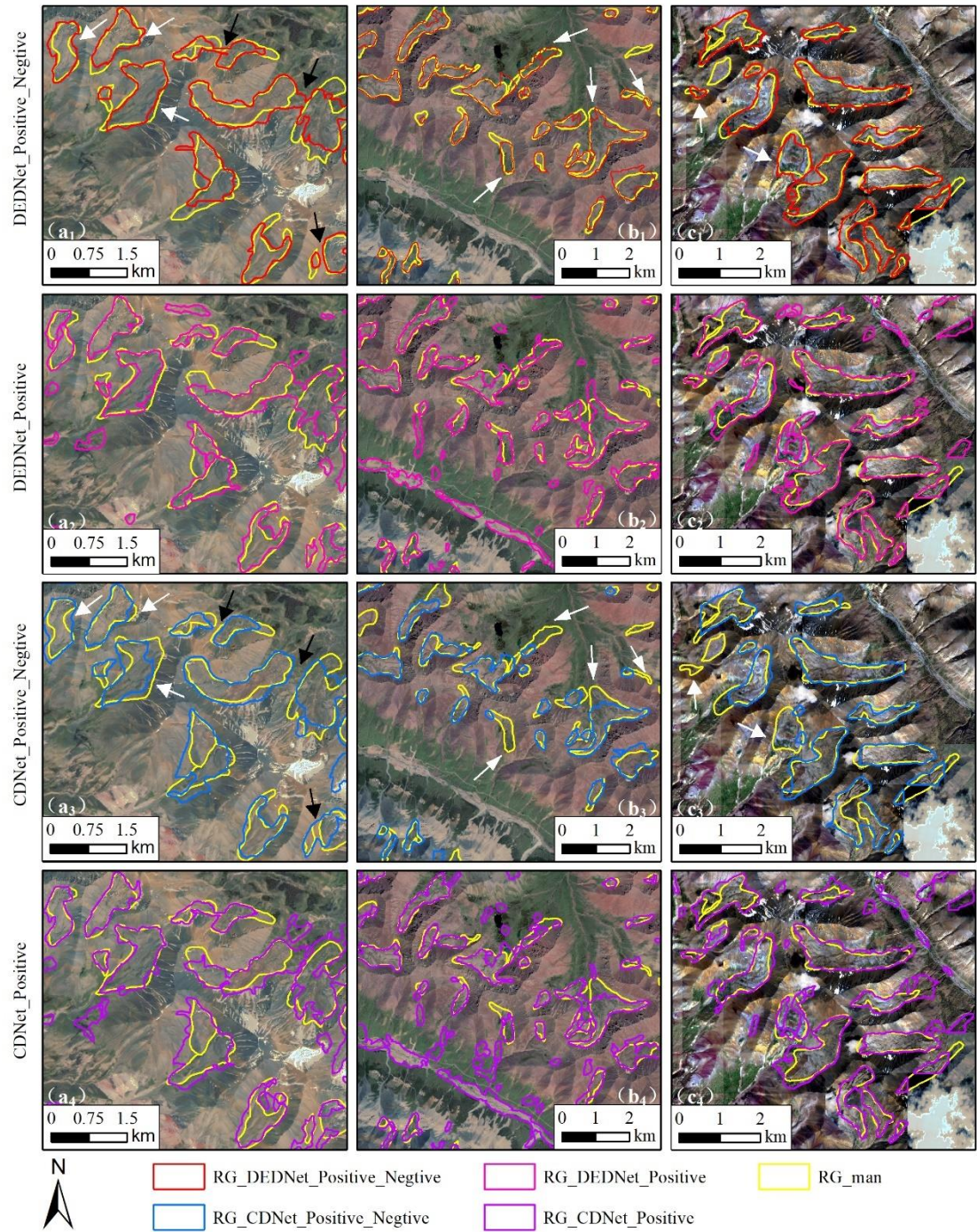


Figure 10. Boundaries of RG_man and four models outlined in three sub-regions in MTA. The four RG_models represent the rock glaciers delineated by corresponding models. For example, RG_CDNet_Positive represent rock glaciers delineated by CDNet_Positive.

The extraction and fusion of spatial and spectral features by DEDNet enhance the accuracy of rock glacier recognition mainly because: (1) For active rock glaciers, the distinct spatial features such as furrows-ridges and steep frontal and lateral margins contribute to the identification of rock glaciers. However, when combining the spatial features with spectral ones, the textural similarity landforms, such as debris-covered glaciers could be excluded readily because the exposed ice in debris-covered glaciers could be distinguished by the lower reflectance in NIR band. In addition, the identification accuracy of the rock glacier boundaries improved because the steep margins exhibit lighter slopes or darker shadows related to the sun position, making them easier to distinguish from the neighboring landforms due to their higher or lower spectral reflectance. (2) For transitional/relict rock glaciers, vegetation has developed along the longitudinal and transverse flow structures, as well as the frontal and lateral margins, forming distinctive and identifiable spatial features and heterogeneous vegetation spectral characteristics. This enables the differentiation of transitional/relict rock glaciers from upstream non-vegetated landforms and downstream homogeneous vegetation landforms.

5.2. Model Performance Comparison

We experimentally compare DEDNet with CDNet, modified CDNet and the mature MSNet [21], using the training and validation datasets, which include both positive and negative samples. The CDNet is presently considered the most advanced and widely adopted method in rock glacier recognition [13–16]. There are two commonly used methods to modify CDNet for processing multi-spectral data: increasing the number of channels in CDNet input layer (MI_CDNet) to match the input data [37], and adding a convolution layer at the beginning of CDNet (AC_CDNet) to transform multispectral images to 3-D features [38]. By the way, The MI_CDNet was trained without the pre-trained HRNetV2 because of the structural mismatch between MI_CDNet and the pre-trained HRNetV2. The significant advantages of MSNet over RTFNet [39], MUFNet [40], and MFNet [41] in processing 4-band images have been verified [21], and the comparison between MSNet and DEDNet can reflect the superiority of DEDNet to fuse spatial and spectral features in the semantic segmentation field. In addition, several powerful backbones, such as ResNet50, ResNet101, Deep Residual Network (DRN), Xception, and Vision Transformer (VIT) was employed to replace the HRNetV2 for training the DEDNet with same hyperparameters, with the results presented in **Error! Reference source not found.**

By training the DEDNet (with backbone of HRNetV2), we obtained the best-performing model, which outperformed CDNet-based model and other comparative models in all evaluation metrics. The models trained with the MI_CDNet and AC_CDNet network achieved low accuracy, with mIoU values of 0.6944 and 0.7112 respectively, representing a 15 percent point decrease compared to the mIoU obtained with the DEDNet model, indicating the challenges of using CDNet for rock glacier identification on multispectral images. Comparing the models trained with MSNet and DEDNet (both using ResNet50 as the backbone) network, we found that the latter achieved an mIoU of 0.8455, slightly higher than the former's 0.8413. This suggests that in rock glacier identification field, the simple yet effective DEDNet is advantageous. Compared to other backbones, HRNetV2's advantages are more pronounced, possibly because HRNetV2 maintains high-resolution representations through the whole process, aiding in distinguishing rock glaciers with spectral similarities from surrounding landforms.

Table 4. The evaluation metrics of different networks on validation datasets.

Network	Backbone	Pretrained	Accuracy	mIOU	Precision	Recall	Specificity
DEDNet	HRNet V2	True	*0.9131	*0.8601	*0.9130	*0.927	*0.9195
CDNet	HRNet V2	True	0.9047	0.8457	0.9045	0.9155	0.9095

MI_CDNet	HRNet V2	False	0.7874	0.6944	0.7875	0.7900	0.7885
AC_CDNet	HRNet V2	True	0.8073	0.7112	0.8075	0.8020	0.8045
MSNet	ResNet 50	True	0.9022	0.8413	0.9025	0.9125	0.9070
DEDNet	ResNet 50	True	0.9056	0.8455	0.9055	0.9145	0.9095
DEDNet	ResNet 101	True	0.9073	0.8490	0.9070	0.9175	0.9112
DEDNet	DRN	True	0.9062	0.8490	0.9060	0.9190	0.9120
DEDNet	Xception	True	0.7563	0.6061	0.7560	0.6660	0.6990
DEDNet	VIT	True	0.8393	0.6405	0.8390	0.690	0.7395

5.3. Generalizability

The well-trained DEDNet model was also applied to two areas with rock glacier inventories published, the northern Tien Shan (Kazakhstan) and Daxue Shan (approximately 1500km in the south of the QLMs), to test the generalizability of the model. The northern Tien Shan, located 2500km in the northwest of QLMs (**Error! Reference source not found.a**), has a rock glacier inventory including only active rock glaciers produced using InSAR kinematics [42]. A region ranging from 76°58'E to 77°13'E and 42°59'N to 43°10'N, with 54 rock glaciers according to the inventory, was selected for the generalizability research. 50 out of the 54 rock glaciers (**Error! Reference source not found.b**) were identified by the well-trained DEDNet model, with a total area of 42.63 km², demonstrating an overestimation compared to the 32.33 km² estimated by the InSAR kinematics, which only extracts the active 'unit' of the rock glacier 'system'.

The Daxue Shan [43] is located to the south of the QLMs (**Error! Reference source not found.d**), in the southeastern edge of the Tibetan Plateau's Hengduan Shan. A rock glacier inventory based on the analysis of Google Earth imagery has been released there. The region, ranging from 101°34'E to 101°40'E and 30°24'N to 30°32', contains 38 rock glaciers with different scale. The DEDNet model identified 35 out of the 38 rock glaciers (**Error! Reference source not found.e**), with a total area of 8.58 km², demonstrating 4.57% underestimation compared to the 8.99 km² of the 38 rock glaciers. Rock glaciers with distinct/vague and identifiable frontal slope and furrows-ridges, not in the inventories, also delineated by the DEDNet model (**Error! Reference source not found.c** and **f**). Overall, our DEDNet-based model exhibits strong robustness and demonstrates great potential for applicability across diverse geographic regions.

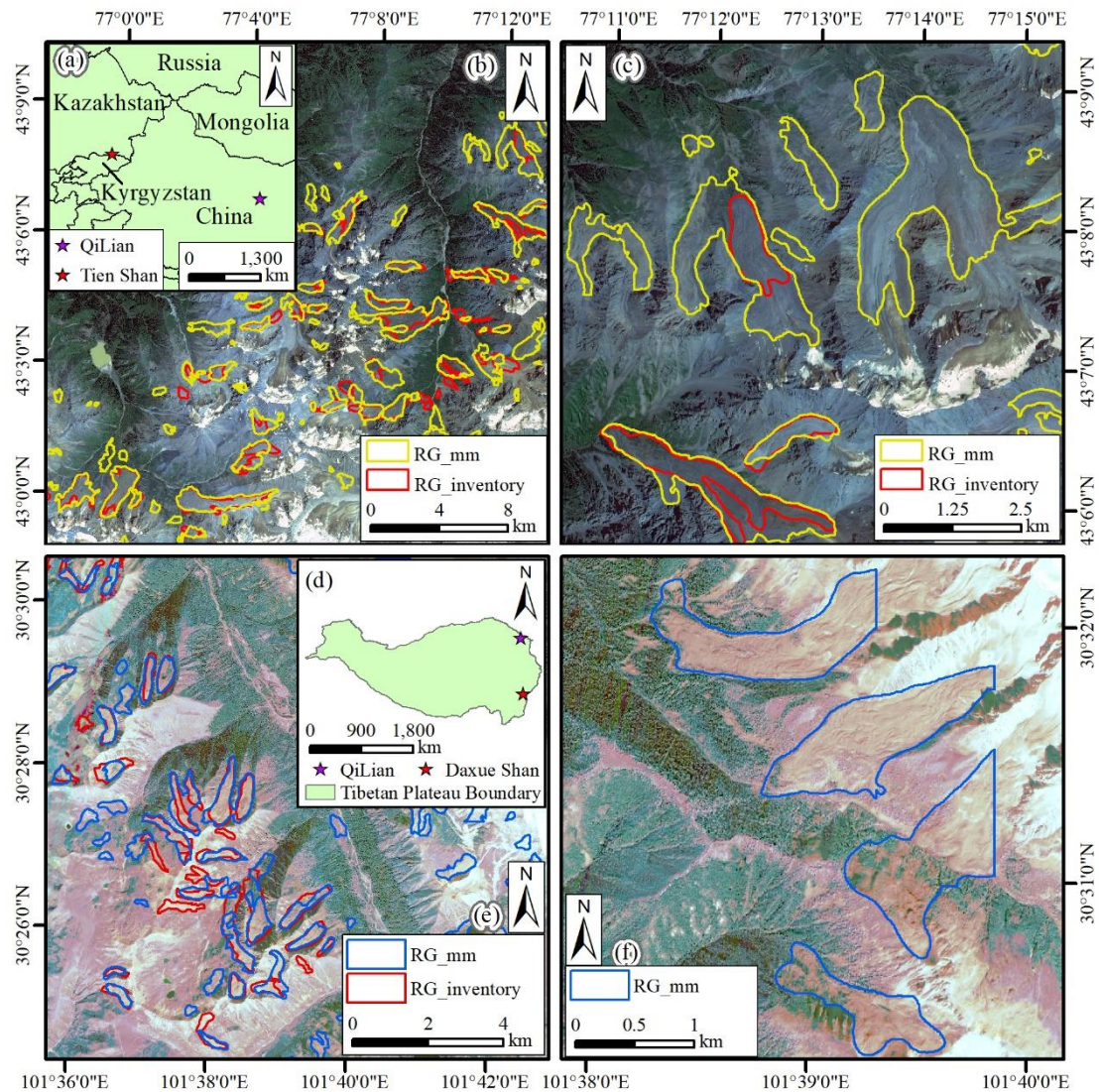


Figure 11. Locations of northern Tien Shan (Kazakhstan) (a) and Daxue Shan (d). Boundaries of RG_man and RG_inventory in northern Tien Shan (b) and Daxue Shan (e). Rock glaciers model mapped but not included in the inventory of northern Tien Shan (c) and Daxue Shan (f).

5.4. Contribution and Limitation

When delineating rock glaciers on a large scale, DEDNet may have two foreseeable potential contributions. Firstly, The DEDNet has showed its robustness when applying the DEDNet-based model to Daxue Shan and Tien Shan, demonstrating the potential capability to map rock glaciers across diverse geographic regions. Therefore, DEDNet can be employed in some alpine regions where inventories remain incomplete to identify and delineate rock glaciers. Secondly, the model trained DEDNet on RGB and NIR-EVI-SAVI datasets outperform those trained solely on RGB or NIR-EVI-SAVI datasets using CDNet, even when supplemented with InSAR images [13], in identifying relict rock glaciers (**Error! Reference source not found.**). Exploring the distribution of relict rock glaciers contributes to a comprehensive understanding of the mountainous periglacial environment because of the cruciality for reconstructing ancient climates [18] and the significance for Hydrogeology research [44].

DEDNet shows promise but still has two limitations. Firstly, the pre-trained HRNetV2 is suitable for processing RGB dataset but not optimized for NIR-EVI-SAVI dataset [21], which affects the extraction of spectral information. This may well explain the phenomenon: compared to the model trained CDNet on 542 positive samples with mIoU of 0.8469, the model trained DEDNet on the same

dataset only achieved a slight increase to 0.8519, an improvement of just 0.005%. While, with the addition of 1353 negative samples, the mIoU increased from 0.8457 to 0.8601, showing a significant improvement of 0.0144%. This is because more training samples result in better fine-tuning of the pre-trained model parameters. Secondly, the number of parameters of the DEDNet is approximately twice that of the CDNet. When training a rock glacier (or other landform) identification model using DEDNet, more computational resources are required. This is the reason why we resampled GF1/6 with a spatial resolution of 2m to 4m during the dataset preparation stage.

6. Conclusions

In this study, we designed a DEDNet with two encoders to simultaneously extract and fuse spatial and spectral features from RGB dataset and NIR-EVI-SAVI dataset, respectively. We trained the DEDNet with positive and negative samples including active, transitional, and relict type rock glacier from VIAs in QLMs and obtained a model with accuracy, precision, recall, specificity, and mIoU of 0.9131, 0.9130, 0.9270, 0.9195, and 0.8601, respectively. Then, we tested the model's robustness in MTA and successfully identified 178 out of 190 RG_man, missing 12 small rock glaciers, and misidentifying 11 other landforms with textural similarity as rock glaciers. Ultimately, we achieved a producer's accuracy of 93.68% and a user's accuracy of 94.18%. Furthermore, Our DEDNet demonstrates its robustness to map rock glaciers with greater accuracy across diverse geographic regions.

Author Contributions: Conceptualization, L.L., M.F., M.L., Y.S.K., and F.Y.; Investigation, L.J.L. and Q.J.Z.; data curation, L.J.L.; funding acquisition, L.L., F.Y. and L.J.L.; methodology, L.J.L., M.L., and Q.J.Z.; supervision, M.L., L.L., M.F., Y.S.K., and F.Y.; project administration, L.L. and M.F.; software, L.J.L.; writing—original draft, L.J.L. and Q.J.Z.; writing—review and editing, M.L., L.L., and M.F.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Open Research Fund of TPESER, grant number TPESER202208; the Central University Basic Scientific Research Business Expenses Special Funds - Chang'an University Excellent Doctoral Dissertation Cultivation Support Project, grant number 300102273720; and the Natural Science Basic Research Program of Shaanxi Province, grant number 2024SF-YBXM-570.

Data Availability Statement: Data used and analyzed in the present study will be made available upon author request.

Acknowledgments: The authors appreciate Google for the free use of the Google Earth Pro software. The authors also highly acknowledge Zhengchao Ren, Jinhao Xu, and Dezhao Yan for their contributions to the rock glacier field investigation in the Qilian Mountains. The permafrost distribution map made by Lin Zhao is available at <https://data.tpdac.cn/zh-hans/data/0231c972-8460-4691-a187-70e4cc356f60/> (accessed date on 18 December 2023).

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

References

1. Azócar, G.F.; Brenning, A. Hydrological and Geomorphological Significance of Rock Glaciers in the Dry Andes, Chile (27°-33°S): Rock Glaciers in the Dry Andes. *Permafr. Periglac. Process.* 2010, 21, 42–53, doi:10.1002/ppp.669.
2. Yan, M.; Tian, X.; Li, Z.; Chen, E.; Li, C.; Fan, W. A Long-Term Simulation of Forest Carbon Fluxes over the Qilian Mountains. *Int. J. Appl. Earth Obs. Geoinformation* 2016, 52, 515–526, doi:10.1016/j.jag.2016.07.009.
3. Sorg, A. Impacts of climate change on glaciers, rock glaciers and water availability in the Tien Shan, Central Asia, Université de Genève, 2013.
4. Humlum, O. The Climatic Significance of Rock Glaciers. *Permafr. Periglac. Process.* 1998, 9, 375–395, doi:10.1002/(SICI)1099-1530(199810/12)9:4<375::AID-PPP301>3.0.CO;2-0.
5. Humlum, O. Rock Glacier Appearance Level and Rock Glacier Initiation Line Altitude: A Methodological Approach to the Study of Rock Glaciers. *Arct. Alp. Res.* 1988, 20, 160–178, doi:10.2307/1551495.
6. Jones, D.B.; Harrison, S.; Anderson, K.; Whalley, W.B. Rock Glaciers and Mountain Hydrology: A Review. *Earth-Sci. Rev.* 2019, 193, 66–90, doi:10.1016/j.earscirev.2019.04.001.
7. Harris, C.; Arenson, L.U.; Christiansen, H.H.; Etzelmüller, B.; Frauenfelder, R.; Gruber, S.; Haeberli, W.; Hauck, C.; Hölzle, M.; Humlum, O.; et al. Permafrost and Climate in Europe: Monitoring and Modelling

- Thermal, Geomorphological and Geotechnical Responses. *Earth-Sci. Rev.* 2009, 92, 117–171, doi:10.1016/j.earscirev.2008.12.002.
8. Barsch, D.; Fierz, H.; Haeblerli, W. Shallow Core Drilling and Bore-Hole Measurements in the Permafrost of an Active Rock Glacier near the Grubengletscher, Wallis, Swiss Alps. *Arct. Alp. Res.* 1979, 11, 215–228, doi:10.2307/1550646.
 9. Bolch, T.; Gorbunov, A.P. Characteristics and Origin of Rock Glaciers in Northern Tien Shan (Kazakhstan/Kyrgyzstan). *Permafr. Periglac. Process.* 2014, 25, 320–332, doi:10.1002/ppp.1825.
 10. Feng, M.; Xu, J.; Wang, J.; Ran, Y.; Li, X. Identifying Rock Glacier in Western China Using Deep Learning and Satellite Data. In *Proceedings of the AGU Fall Meeting Abstracts*; December 2019; Vol. 2019, pp. GC53G-1249.
 11. Marcer, M. Rock Glaciers Automatic Mapping Using Optical Imagery and Convolutional Neural Networks. *Permafr. Periglac. Process* 2020, 31, pp.561-566.
 12. Robson, B.A.; Bolch, T.; MacDonell, S.; Höblbling, D.; Rastner, P.; Schaffer, N. Automated Detection of Rock Glaciers Using Deep Learning and Object-Based Image Analysis. *Remote Sens. Environ.* 2020, 250, 112033, doi:10.1016/j.rse.2020.112033.
 13. Hu, Y.; Liu, L.; Huang, L.; Zhao, L.; Wu, T.; Wang, X.; Cai, J. Mapping and Characterizing Rock Glaciers in the Arid West Kunlun of China. *ESS Open Arch.* 2022, 1–37, doi:10.1002/essoar.10512700.1.
 14. Sun, Z.; Hu, Y.; Liu, L.; Racoviteanu, A.; Harrison, S. Mapping Rock Glaciers on the Tibetan Plateau from Planet Basemaps Using Deep Learning. In *Proceedings of the AGU Fall Meeting Abstracts*; December 2022; Vol. Vol. 2022, pp. C42E-1078.
 15. Sun, Z.; Hu, Y.; Racoviteanu, A.; Liu, L.; Harrison, S.; Wang, X.; Cai, J.; Guo, X.; He, Y.; Yuan, H. TPRoGI: A Comprehensive Rock Glacier Inventory for the Tibetan Plateau Using Deep Learning; ESSD – Ice/Permafrost, 2024;
 16. Sun, Z.; Hu, Y.; Liu, L.; Racoviteanu, A.; Harrison, S. Mapping and Inventorying Rock Glaciers on the Tibetan Plateau from Planet Basemaps Using Deep Learning. In *Proceedings of the EGU General Assembly Conference Abstracts*; April 23 2023; p. EGU-6816.
 17. Jiang, J.; Feng, X.; Liu, F.; Xu, Y.; Huang, H. Multi-Spectral RGB-NIR Image Classification Using Double-Channel CNN. *IEEE Access* 2019, 7, 20607–20613, doi:10.1109/ACCESS.2019.2896128.
 18. Barsch, D. Permafrost Creep and Rockglaciers. *Permafr. Periglac. Process.* 1992, 3, 175–188, doi:10.1002/ppp.3430030303.
 19. RGIK Towards Standard Guidelines for Inventorying Rock Glaciers: Baseline Concepts (Version 4.2.2). IPA Action Group Rock Glacier Invent. Kinemat. 2022, 13, doi:https://bigweb.unifr.ch/Science/Geosciences/Geomorphology/Pub/Website/IPA/Guidelines/V4/220331_Baseline_Concepts_Inventorying_Rock_Glaciers_V4.2.2.pdf.
 20. Pan, B.; Shi, Z.; Xu, X.; Shi, T.; Zhang, N.; Zhu, X. CoinNet: Copy Initialization Network for Multispectral Imagery Semantic Segmentation. *IEEE Geosci. Remote Sens. Lett.* 2019, 16, 816–820, doi:10.1109/LGRS.2018.2880756.
 21. Tao, C.; Meng, Y.; Li, J.; Yang, B.; Hu, F.; Li, Y.; Cui, C.; Zhang, W. MSNet: Multispectral Semantic Segmentation Network for Remote Sensing Images. *GIScience Remote Sens.* 2022, 59, 1177–1198, doi:10.1080/15481603.2022.2101728.
 22. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep High-Resolution Representation Learning for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 2021, 43, 3349–3364, doi:10.1109/TPAMI.2020.2983686.
 23. Yang, X.; Fan, X.; Peng, M.; Guan, Q.; Tang, L. Semantic Segmentation for Remote Sensing Images Based on an AD-HRNet Model. *Int. J. Digit. Earth* 2022, 15, 2376–2399, doi:10.1080/17538947.2022.2159080.
 24. Wu, H.; Liang, C.; Liu, M.; Wen, Z. Optimized HRNet for Image Semantic Segmentation. *Expert Syst. Appl.* 2021, 174, 114532, doi:10.1016/j.eswa.2020.114532.
 25. Xu, Z.; Zhang, W.; Zhang, T.; Li, J. HRCNet: High-Resolution Context Extraction Network for Semantic Segmentation of Remote Sensing Images. *Remote Sens.* 2020, 13, 71, doi:10.3390/rs13010071.
 26. Lou, P.; Wu, T.; Chen, J.; Fu, B.; Zhu, X.; Chen, J.; Wu, X.; Yang, S.; Li, R.; Lin, X.; et al. Recognition of Thaw Slumps Based on Machine Learning and UAVs: A Case Study in the Qilian Mountains, Northeastern Qinghai-Tibet Plateau. *Int. J. Appl. Earth Obs. Geoinformation* 2023, 116, 103163, doi:10.1016/j.jag.2022.103163.
 27. Hu, Z.; Yan, D.; Feng, M.; Xu, J.; Liang, S.; Sheng, Y. Enhancing Mountainous Permafrost Mapping by Leveraging a Rock Glacier Inventory in Northeastern Tibetan Plateau. *Int. J. Digit. Earth* 2024, 17, 2304077, doi:10.1080/17538947.2024.2304077.
 28. Gilabert, M.A.; González-Piqueras, J.; García-Haro, F.J.; Meliá, J. A Generalized Soil-Adjusted Vegetation Index. *Remote Sens. Environ.* 2002, 82, 303–310, doi:10.1016/S0034-4257(02)00048-2.
 29. Jiang, Z.; Huete, A.; Didan, K.; Miura, T. Development of a Two-Band Enhanced Vegetation Index without a Blue Band. *Remote Sens. Environ.* 2008, 112, 3833–3845, doi:10.1016/j.rse.2008.06.006.

30. Zhu, L.; Ji, D.; Zhu, S.; Gan, W.; Wu, W.; Yan, J. Learning Statistical Texture for Semantic Segmentation.; 2020; pp. 12975–12984.
31. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation.; 2015; pp. 3431–3440.
32. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); IEEE: Las Vegas, NV, USA, June 2016; pp. 770–778.
33. Su, R.; Xu, D.; Sheng, L.; Ouyang, W. PCG-TAL: Progressive Cross-Granularity Cooperation for Temporal Action Localization. *IEEE Trans. Image Process.* 2021, 30, 2103–2113, doi:10.1109/TIP.2020.3044218.
34. Yan, J.; Liu, J.; Liang, D.; Wang, Y.; Li, J.; Wang, L. Semantic Segmentation of Land Cover in Urban Areas by Fusing Multisource Satellite Image Time Series. *IEEE Trans. Geosci. Remote Sens.* 2023, 61, 1–15, doi:10.1109/TGRS.2023.3329709.
35. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* 2018, 40, 834–848, doi:10.1109/TPAMI.2017.2699184.
36. Li, J.; Wang, Q.; Zhang, Y.; Yang, S.; Gao, G. An Improved Active Layer Thickness Retrieval Method over Qinghai-Tibet Permafrost Using InSAR Technology: With Emphasis on Two-Dimensional Deformation and Unfrozen Water. *Int. J. Appl. Earth Obs. Geoinformation* 2023, 124, 103530, doi:10.1016/j.jag.2023.103530.
37. Carvalho, O.L.F.D.; De Carvalho Júnior, O.A.; Albuquerque, A.O.D.; Bem, P.P.D.; Silva, C.R.; Ferreira, P.H.G.; Moura, R.D.S.D.; Gomes, R.A.T.; Guimarães, R.F.; Borges, D.L. Instance Segmentation for Large, Multi-Channel Remote Sensing Imagery Using Mask-RCNN and a Mosaicking Approach. *Remote Sens.* 2020, 13, 39, doi:10.3390/rs13010039.
38. Han, W.; Li, J.; Wang, S.; Zhang, X.; Dong, Y.; Fan, R.; Zhang, X.; Wang, L. Geological Remote Sensing Interpretation Using Deep Learning Feature and an Adaptive Multisource Data Fusion Network. *IEEE Trans. Geosci. Remote Sens.* 2022, 60, 1–14, doi:10.1109/TGRS.2022.3183080.
39. Sun, Y.; Zuo, W.; Liu, M. RTFNet: RGB-Thermal Fusion Network for Semantic Segmentation of Urban Scenes. *IEEE Robot. Autom. Lett.* 2019, 4, 2576–2583, doi:10.1109/LRA.2019.2904733.
40. Xu, F.; Shang, Z.; Wu, Q.; Zhang, X.; Lin, Z.; Shao, S. MUFNet: Toward Semantic Segmentation of Multi-Spectral Remote Sensing Images. In Proceedings of the 2021 4th Artificial Intelligence and Cloud Computing Conference; ACM: Kyoto Japan, December 17 2021; pp. 39–46.
41. Ha, Q.; Watanabe, K.; Karasawa, T.; Ushiku, Y.; Harada, T. MFNet: Towards Real-Time Semantic Segmentation for Autonomous Vehicles with Multi-Spectral Scenes. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); IEEE: Vancouver, BC, September 2017; pp. 5108–5115.
42. Bertone, A.; Barboux, C.; Bodin, X.; Bolch, T.; Brardinoni, F.; Caduff, R.; Christiansen, H.H.; Darrow, M.M.; Delaloye, R.; Etzelmueller, B.; et al. Incorporating InSAR Kinematics into Rock Glacier Inventories: Insights from 11 Regions Worldwide. *The Cryosphere* 2022, 16, 2769–2792, doi:10.5194/tc-16-2769-2022.
43. Ran, Z.; Liu, G. Rock Glaciers in Daxue Shan, South-Eastern Tibetan Plateau: An Inventory, Their Distribution, and Their Environmental Controls. *The Cryosphere* 2018, 12, 2327–2340, doi:10.5194/tc-12-2327-2018.
44. Colucci, R.R.; Forte, E.; Žebre, M.; Maset, E.; Zanettini, C.; Guglielmin, M. Is That a Relict Rock Glacier? *Geomorphology* 2019, 330, 177–189, doi:10.1016/j.geomorph.2019.02.002.

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