Multi-source deep-learning approach for automatic geomorphological mapping: the case of glacial moraines

Isabelle Rocamora, Dino Ienco, Matthieu Ferry

▶ **To cite this version:**

Isabelle Rocamora, Dino Ienco, Matthieu Ferry. Multi-source deep-learning approach for automatic geomorphological mapping: the case of glacial moraines. Geo-spatial Information Science, 2024, 16, pp.1-20. 10.1080/10095020.2023.2292587 . hal-04476128

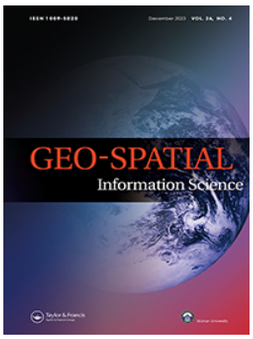
HAL Id: hal-04476128
https://hal.science/hal-04476128v1

Submitted on 27 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License



Geo-spatial Information Science

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgsi20

Multi-source deep-learning approach for automatic geomorphological mapping: the case of glacial moraines

Isabelle Rocamora, Dino Ienco & Matthieu Ferry

To cite this article: Isabelle Rocamora, Dino Ienco & Matthieu Ferry (31 Jan 2024): Multi-source deep-learning approach for automatic geomorphological mapping: the case of glacial moraines, Geo-spatial Information Science, DOI: 10.1080/10095020.2023.2292587

To link to this article: https://doi.org/10.1080/10095020.2023.2292587

© 2024 Wuhan University. Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 31 Jan 2024.

Submit your article to this journal

Article views: 229

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tgsi20

GEO-SPATIAL INFORMATION SCIENCE
https://doi.org/10.1080/10095020.2023.2292587

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Multi-source deep-learning approach for automatic geomorphological mapping: the case of glacial moraines

Isabelle Rocamora [a], Dino Ienco [b] and Matthieu Ferry [a]

[a]Geosciences Montpellier, University of Montpellier - CNRS, Montpellier, France; [b]INRAE, UMR TETIS, University of Montpellier, Montpellier, France

ABSTRACT

Landform mapping is the initial step of many geomorphological analyses (e.g. assessment of natural hazards and natural resources) and requires vast resources to be applied to wide areas at high-resolution. Among geomorphological objects, we focus on glacial moraine mapping, since it is a task relevant to many fields (e.g. paleoclimate and glacial geomorphology). Here we proposed to exploit the potential of Deep Learning-based approaches to map moraine landforms by exploiting multi-source remote sensing imagery. To this end, we propose the first Deep Learning model to map glacial moraines, namely MorNet. As multi-source remote sensing information, we combine together three different sources: Topographic (Pleiades-derived DSM), Multispectral (Sentinel-2), and SAR (Sentinel-1) data. To cope with such heterogeneous information, the proposed model has a dedicated branch for each input source and, a late fusion mechanism is leveraged to combine them with the aim to provide the final mapping. The performance of the MorNet model is evaluated on several glacier valleys in China in the Himalayan range. This area contains minimally eroded moraines, so they are well-defined and of varied morphology. The behavior of the proposed method is compared to models using individual mono-source models in order to highlight the benefit to simultaneously leverage multi-source information. The use of multi-source data allows MorNet to exploit the complementarity of the three input sources and improve its performance from an f1-score of about 41.6 using a single source to 52.8 using three sources. MorNet provides a first-order moraine map through its ability to identify well-defined moraines. Consequently, MorNet can identify areas likely to contain moraines and intends to be used as a tool by experts to facilitate and support large-scale mapping.

ARTICLE HISTORY
Received 22 March 2023
Accepted 4 December 2023

KEYWORDS
Deep learning; data fusion; geomorphology; moraines; mapping

1. Introduction

Geomorphology is the field of geosciences that investigates landforms and their associated formational and evolutional processes, such as erosion and deposition (Evans 2012; Keller et al. 2020; Smith, Paron, and Griffiths 2011). Landforms are geomorphological objects that discontinuously cover the earth's surface and are characterized by their topography. In other words, by their shape (conical, elongated, ridged, etc.), dimensions (height, width, and length), and geomorphometric characteristics such as slope, curvature, roughness, position, and orientation in the landscape (Evans 2012). Geomorphological analyses have both societal and scientific applications, such as the characterization of natural hazards and natural resources or a better understanding of the functioning of natural systems (Bishop et al. 2012; Evans 2012; Keller et al. 2020; Smith, Paron, and Griffiths 2011; Sun et al. 2022). Moreover, these issues are becoming increasingly important with global warming (e.g. Bull 2013; Scherler, Bookhagen, and Strecker 2011; Wobus, Tucker, and Anderson 2010). However, it is necessary before these studies to carry out geomorphological mapping of the objects of interest. These maps are usually drawn manually on the basis of remote sensing images (mostly satellite), then checked and completed by field observations (Chandler et al. 2018; Smith, Paron, and Griffiths 2011). Creating these maps is a complex task that depends on the expertise and skills of the operator and requires a detailed analysis of the landscape and geomorphometric features for accurate delineation of geomorphological objects. As a result, this task is time-consuming, especially when working over large areas. Nowadays, the availability of high ($\sim$ 10 m) and very high-resolution images ($<$ 1 m) permits us to characterize with great precision the shape and nature of geomorphological objects. In addition, the short revisit periods provided by most recent satellite missions make it

CONTACT Matthieu Ferry matthieu.ferry@umontpellier.fr

© 2024 Wuhan University. Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

possible to monitor dynamic phenomena such as bank erosion, landslides, and destabilization of melt lakes. However, due to the time-consuming nature of map production and the limitations of human operators in handling and organizing vast amounts of information, the full potential of this extensive data remains untapped. In response to these problems, many automated mapping systems have been developed using satellite imagery (e.g. Bishop et al. 2012; Han et al. 2022; Smith, Paron, and Griffiths 2011) to generate large-scale maps. These systems primarily rely on user-defined classification rules and thresholds, such as Object-Based Image Analysis (e.g. Anders, Seijmonsbergen, and Bouten 2011; Drăguț and Blaschke 2006; Rastner et al. 2014), band ratio techniques (e.g. Alifu, Ryutaro, and Johnson 2015) and morphometric analysis technique based on pixel (e.g. Alvioli et al. 2016; Wernette, Houser, and Bishop 2016). To a lesser extent, Drăguț and Blaschke (2006) have succeeded in reducing reliance on threshold definitions with fuzzy boundaries, although expert knowledge is still necessary for defining different morphological classes. In comparison, Machine Learning (ML) and Deep Learning (DL) models are entirely data-driven and autonomously acquire knowledge about object characteristics (LeCun, Bengio, and Hinton 2015; Zhang, Zhang, and Du 2016), eliminating the need to develop manual classification rules and threshold.

ML and DL models have been used in various fields, including computer vision, where they were initially employed for classifying natural images (images of ordinary life) and later applied to automate mapping tasks (LeCun, Bengio, and Hinton 2015; Reichstein et al. 2019; Zhang, Zhang, and Du 2016). DL models have demonstrated superior performance compared to ML models, mainly as a result of their capability to handle complex datasets with a large number of variables. Among DL models, Convolutional Neural Networks (CNNs) have emerged as the state-of-the-art technique in computer vision. This strong performance is due to their convolutional filters, which enable them to extract and learn spatial patterns in images, allowing for object identification at various scales (LeCun, Bengio, and Hinton 2015; Zhang, Zhang, and Du 2016). These capabilities make CNNs well-suited for mapping geomorphological objects, particularly given their successful application in automatic mapping for land use and land cover. Additionally, CNNs offer the advantage of combining data acquired from different sensors, and several model architectures and a wide variety of data (multi-spectral, optical, radar, topographic) have been tested in the field of land cover mapping (Chen and Bruzzone 2022; Gbodjo et al. 2021; Hong et al. 2021; Li et al. 2022, 2021; Sainte Fare Garnot, Landrieu, and Chehata 2022). These studies have highlighted that the use of multi-source models (i.e. using several data sources) has often outperformed mono-source ones (i.e. using a unique data type) (Dalla Mura et al. 2015; Gbodjo et al. 2021; Sainte Fare Garnot, Landrieu, and Chehata 2022). These studies mainly use 2D images such as radar and multispectral data which are widely available but do not capture essential characteristics of true 3D landforms (Li et al. 2022). However, in the case of geomorphological mapping, topographic data are of paramount importance for characterizing geomorphological objects from an expert point of view. Recently, several studies have started to use topographic data via a Digital Terrain Model (DTM) or with derived morphometric indices (e.g. slope) for mapping geomorphological objects. These studies focus on particular objects such as landslides (Ghorbanzadeh et al. 2019), mountain summits (Torres et al. 2020), lunar craters (Silburt et al. 2019), sand dunes (Shumack, Hesse, and Farebrother 2020), and debris-covered glaciers (Xie et al. 2020). Other studies work on mapping several geomorphological units within an area (Du et al. 2019; van der Meij et al. 2022). Few studies use only spectral information to map small-scale objects like Martian volcanic and eolian bedforms (Palafox et al. 2017) and rock glaciers (Marcer 2020). Except for Palafox et al. (2017), Marcer (2020) and Ghorbanzadeh et al. (2019), all other studies focus on large-scale shape mapping.

In this study, we focus our efforts on smaller individual objects of the order of kilometers. Our choice is set on glacial moraines because they exhibit a specific geometry resulting from relatively simple formational processes, i.e. cordons with a triangular section formed by the advance of glaciers (Menzies and Ross 2022). Furthermore, precise moraine identification and characterization are relevant to many disciplines. For example, the position and elevation of frontal moraines is used to quantify the equilibrium-line altitude for paleoclimate reconstruction (Hornsey et al. 2022; Saha et al. 2019). Similarly, moraine characteristics, such as width, height, and slope are useful for estimating the ability of a dam moraine to withstand the hydrostatic pressure of the associated meltwater lake in glacial lake outburst flood hazard studies (Prakash and Nagarajan 2017). To the best of our literature review, there are no specific studies on glacial moraines. One of the seven classes defined in the study proposed in van der Meij et al. (2022) corresponds to push moraines. But it is a large object compared to the moraines in our scenario. Indeed, the moraine is almost 2 km wide and several dozen kilometers long. Similarly, several dozen kilometers long. Similarly use ML algorithms to automate geomorphological mapping in the Alps and one of the

classes of their study includes all objects associated with glaciers (moraines, debris-covered glaciers and rock glaciers).

Another objective is to use DL to take advantage of the complementarity of different remote sensing data sources. DL and ML models combining optical or multi-spectral data with radar or DTM data for land cover and geomorphological mapping, have already been used. These studies have shown that multi-source models have better performance than models using only one source of data (Chen and Bruzzone 2022; Du et al. 2019; Gbodjo et al. 2021; Hong et al. 2021; Li et al. 2022, 2020; Sainte Fare Garnot, Landrieu, and Chehata 2022). The specificity of our approach is that we aim to integrate topographic information and evaluate its contribution to the accurate identification of landforms. In this study, we propose a model called MorNet, to automate moraine mapping at the scale of half a dozen glacial valleys and produce a binary map by classifying each pixel. To assess whether topographic information is sufficient to characterize the moraines or needs supplementary information, our model uses three data sources (DTM, multispectral, and radar images) with a three-branch structure to extract complementary information between these sources.

2. Study site and datasets

The study area is located in the Yadong half-graben in China, which intersects the Himalayan range in an N-S orientation at an average elevation of 4500 m. The Yadong normal fault runs along the piedmont of the Jomolhari range and forms spectacular triangular facets whose dip is ~30. Glaciers define accumulation zones at summits of the range and flow westward down very steep slopes to the Yadong basin. The basin fill is mainly composed of Quaternary glacial and fluvio-glacial deposits derived from local granite and metamorphic rocks (Wang et al. 2022; Zuo et al. 2021). Our site comprises two zones of similar size (~55 km^2) separated by 23 km (Figure 1(a)). We use 7.5 km^2 of the study site to train the model and 54.5 km^2 to test the effectiveness of the model. The details of the distribution of the area in these two categories are explained in section 3.3.

We carried out a cartography of the moraines that will be used as ground truth to train the model and test its effectiveness. Moraine mapping was based on a photogrammetric Digital Surface Model (DSM), with a spatial resolution of 1 m, derived by tri-stereo photogrammetry from Pleiades satellites imagery acquired on 20 November 2016. The absence of visible vegetation on the study site allows the DSM to be considered equivalent to a DTM. Pleiades images are composed of five bands, three in the visible spectrum, one in the near-infrared spectrum (RGB and NIR at a spatial resolution of 2 m), and one panchromatic band on the entire visible spectrum (spatial resolution of 0.5 m). We use the triplet of panchromatic images acquired with different incidence angles to produce the DSM using MicMac software (Rupnik, Daakir, and Pierrot Deseilligny 2017). RGB composite images from Sentinel-2 satellite complement the data used to map the moraines.

We focus on latero-frontal moraines which are triangular-shaped mounds formed by the advance of glaciers. They are located near the outlet of a glacial valley and are identified by a prominent crest and two steep flanks (local slope of 20°-30°). Moraines are mapped according to the context of their formation, i.e. their position relative to the paleo-glacier, as well as their morphology. We rely on a DTM and derived morphometric indices such as horizontal curvature and slope to perform the mapping by first locating the glacial valleys and then focusing on the location of moraines near outlets of glacial valleys. DTM, curvature, and slope allow to detect the crests and flanks of the moraines. The slope break at the base of the flanks marks the delineation of the moraine (curvature local extremum).

There are 59 polygons representing the moraines mapped in both zones. The moraines in the northern zone are well-expressed, relatively large, and represent 44% of the mapped moraine surface in the two zones. On the contrary, the moraines in the southern zone are smaller but more numerous; they represent 71% of the 59 polygons localized in the two zones (Figure 1(b)). The morphology of the moraines changes with age, thus old moraines have lower slopes than recent moraines due to erosion. In addition, some moraines can also be very degraded if they have been crossed by drainage networks or covered by debris flows. Some characteristics of the 59 mapped polygons are presented in Figure 2. The length of the moraines varies from 50 m to 4 km, but half are less than 1 km, and a third are less than 500 m long. The median slope ranges from 7° to 33° with a maximum of around 20°, and 70% of the polygons have a median slope between 12° and 25°. The median width varies from 20 m to 270 m, and we observe three peaks, one around 30 m, which corresponds to a quarter of the moraines, a second around 70 m, and a last one around 150 m. The distribution of the area follows a decreasing tendency, with a third of the area less than 0.05 km^2, then almost half of it between 0.05 and 0.2 km^2, and the extremes are between 1000 m^2 and 1.05 km^2. Finally, the median height evolves between 2 m and 69 m with a maximum of around 25 m and a quarter of moraines with a height lower than 11 m.

In our study, we train the DL model by exploiting multi-source remote sensing imagery in order to describe the study area. More in detail, we consider:

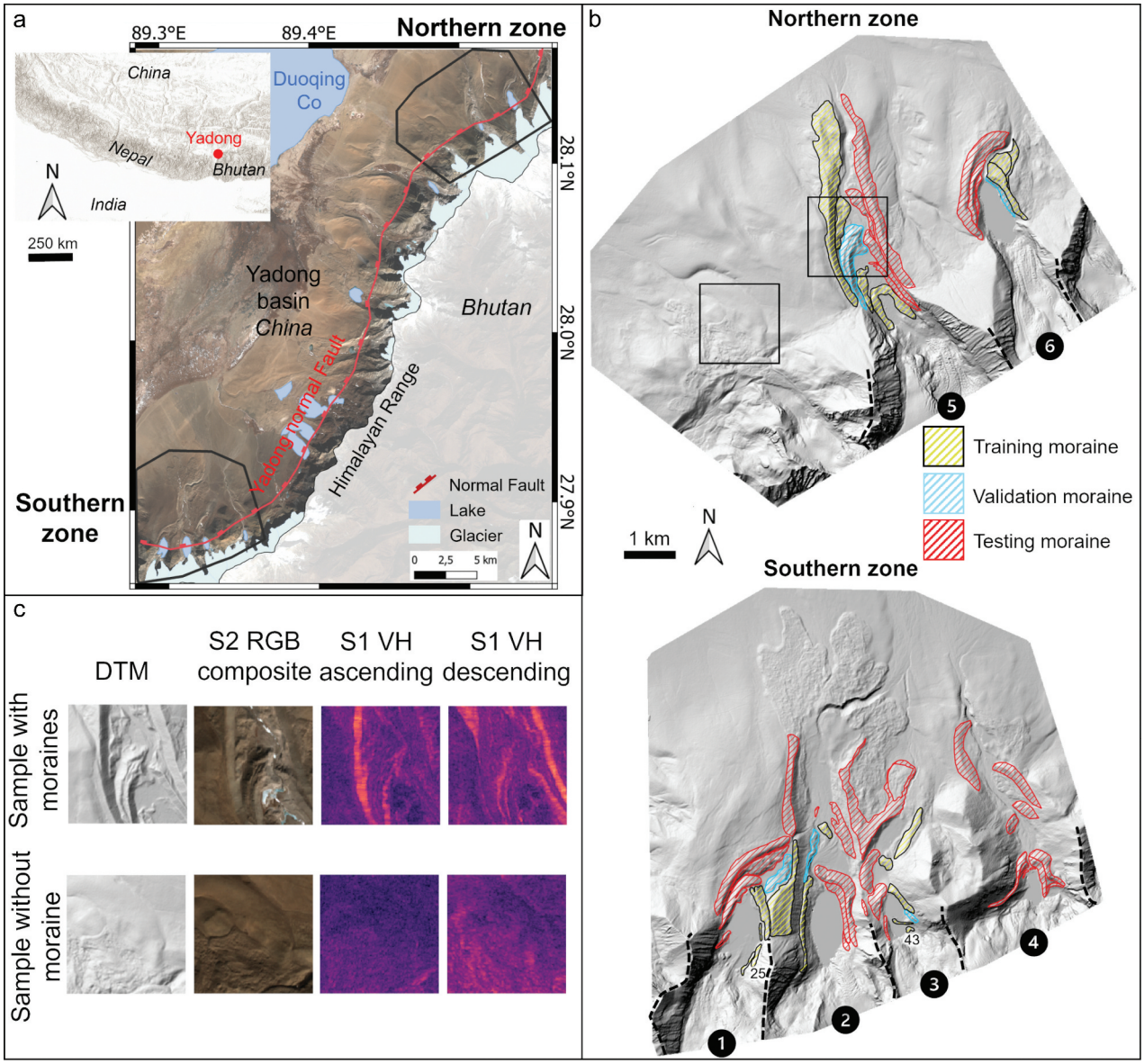


Figure 1. (a) study area in the Yadong half-graben (China) and location of the two zones on RGB composite images provided by Sentinel-2. (b) localization of moraines in the northern and southern zones displayed on the photogrammetric DTM. The numbers 1 to 6 indicate the positions of the glacial valleys. Numbers 25 and 43 correspond to particular moraines whose characteristics are discussed in the results section. Black boxes in solid line correspond to location of the samples of different sources presented in (c).

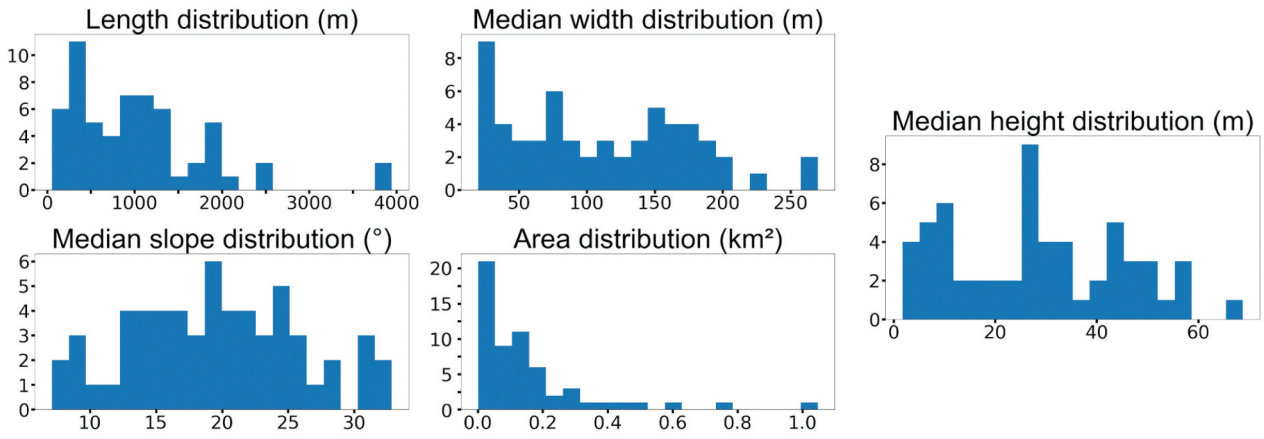


Figure 2. Distribution of length, median width, median slope, area, textcolor black and median height, of the 59 moraines present in the two zones of our study site.

DTM: DTM at a spatial resolution of 10 m, resampled with SAGA-GIS (Conrad et al. 2015) resampling tool from the 1-m-photogrammetric DTM. The bicubic spline parameter is used for downscaling.

Sentinel-1 (S1): Four radar images with a spatial resolution of 10 m from the Sentinel-1 satellite were downloaded via the PEPS platform in the Ground Range Detected format. Images are acquired with the Interferometric Wide acquisition mode, in ascending and descending orbit, and with co- and cross-polarization (VH and VV). The ascending images were collected on 28 December 2020, and the descending images on 24 December 2020. The radar indicates the ground roughness in the direction of incidence. The images are orthorectified and calibrated, and a speckle filter is applied.

Sentinel-2 (S2): Ten images from the Sentinel-2 satellite resampled to a spatial resolution of 10 m. Three bands capture information from the visible spectrum (bands B2, B3, and B4), five in the near-infrared (B5, B6, B7, B8, and B8A), and two images in the short-wave infrared (B11 and B12) and provide information on soil texture, vegetation, and lithology. Images were acquired on 30 December 2020 and downloaded via the PEPS platform.

Sentinel-1 and Sentinel-2 satellites are provided by the Copernicus Programme developed by the European Space Agency (ESA). This programme provides open-source imagery all over the globe with high temporal revisit time. Figure 1(c) shows the different sources on some extracts of the study site.

3. Methodology

In this section, we first present the structure of our three-branch model. Then, we detail the post-processing framework used for filtering MorNet's predictions. Afterward, we introduce the method for the dataset's construction and evaluating the performances of our model. We describe the framework associated with our model in Figure 3(a)). We also describe the sub-models created by removing branches from the three-branch structure of MorNet to compare mono-source and multi-source performances. Finally, we present experimental settings and we report the hyper-parameters of all models.

3.1. Model architecture

3.1.1. Main architecture

The use of multi-source data is getting more and more attention with the unprecedented possibility to access remote sensing data acquired via multiple acquisition modes over a given study area (Chen and Bruzzone 2022; Du et al. 2019; Gbodjo et al. 2021; Hong et al. 2021; Li et al. 2022, 2021, 2021). The two most common ways to combine distinct data sets are: i) Early Fusion and ii) Late Fusion. The former consists in merging the sources before their analysis while the latter involves processing the sources separately and, subsequently, combining them in the middle of the analysis process (Hong et al. 2021; Sebastianelli et al. 2021). For example, Ghorbanzadeh et al. (2019) apply an early merging strategy to optical and topographic data by stacking

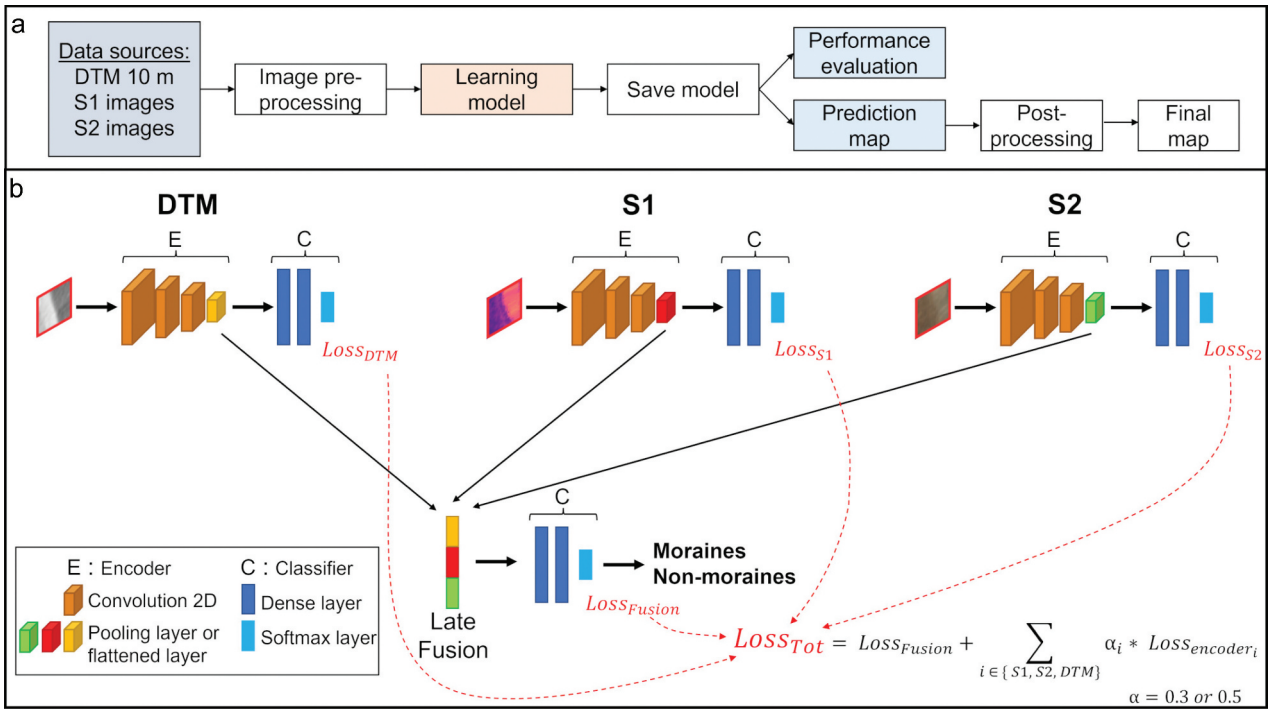


Figure 3. (a) Flowchart of our framework. (b) MorNet's architecture.

the various layers that will serve as input data for its mono-branch model. In contrast, Sainte Fare Garnot, Landrieu, and Chehata (2022) employ both optical and radar data in their study, utilizing a two-branch model that individually processes each data source to extract relevant features. These extracted features are subsequently merged and further processed to generate the final classification. Recent studies suggest that the Late Fusion strategy achieves better performance thanks to the use of recent deep learning methodological developments (Hong et al. 2021; Sainte Fare Garnot, Landrieu, and Chehata 2022; Sebastianelli et al. 2021). Here we adopt a Late Fusion strategy to build a model that takes, as input, a patch of data and predicts the class of the central pixel. To this end, we propose to exploit a model with three branches, one for each of the considered sources (DTM, S1, and S2). More in detail, we adapt the neural network architecture previously introduced in Gbodjo et al. (2021) for land cover mapping from multi-temporal and multi-scale imagery to our task. In this model, each of the three branches is composed of an encoder whose purpose is to extract per-source features. Successively, the per-source features are aggregated together by means of a fusion module and, at the end, the decision is made by a classifier. We call this model MorNet (Figure 3(b)).

3.1.2. Encoder description

Encoders of each branch are identical, but they differ by the number of input bands, the number of filters used, and the dropout rate. Encoders are composed of three 2D convolutional layers and one pooling (Zhang, Zhang, and Du 2016) or one flattened layer at the end. After each convolutional layer, we add a ReLU (Rectified Linear Unit) activation function (Nair and Hinton 2010), batch normalization (Ioffe and Szegedy 2015), and a dropout layer (Srivastava et al. 2014). Batch normalization and dropout layers have been introduced to regularize and stabilize the training of the internal model parameters. The three convolutional layers have a kernel size of 3×3. The first two convolutional layers have a stride value equal to 2. Table 1 supplies information about the difference in the previously mentioned hyper-parameters for each encoder of the multi-branch model.

3.1.3. Training procedure

The model extracts the per-source features through each encoder, whose structure we described in the previous section. Then, the per-source features are fused together via concatenation and, finally, the classifier composed only of fully connected layers can make the decision about the category of the inputs multi-source patch. To train the deep learning multi-source model, the categorical cross-entropy loss function is employed in order to guide the learning process associated to the internal model parameters. Additionally, according to recent literature related to multi-source remote sensing data classification Gbodjo et al. (2021), we adopt a per-source secondary classifier commonly named auxiliary classifier.

The final loss function associated to the training stage of the MorNet model is defined as follows:

$$Loss_{Tot} = Loss_{Fusion} + \sum_{i \in \{S1,S2,DTM\}} \alpha_i * Loss_{encoder_i} \quad (1)$$

where i is the per-source branch spanning the set of sources $\{S1, S2, DTM\}$ and $Loss$ is the loss function for each classifier (i.e. Cross-Entropy loss) (Figure 3(b)). The α parameter weights the influence of the auxiliary classifiers. The technique of auxiliary classifiers improves the learning ability of the model as underlined by recent studies (Gbodjo et al. 2021; Sainte Fare Garnot, Landrieu, and Chehata 2022). All the classifiers have the same structure and parameters. They are composed of three fully connected layers with a dropout rate of 0.3, two first layers with 32 neurons, and a ReLU activation function. The last layer is a fully connected layer with two outputs and a softmax activation function (Table 1). Thus, this layer produces a probability of belonging to the moraine class or not.

Table 1. Details of the model architecture. Numbers (32) and (2) in brackets correspond to the number of neurons.

Encoder's structure	Parameters specific to each branches			
3×3 Conv2D + strides = 2 + *		DTM	S1	S2
3×3 Conv2D + strides = 2 + *	number of input bands	1	4	10
3×3 Conv2D + strides = 1 + *	number of filters	8	32	32
GMP or Flatten layer	dropout rate	0	0,3	0,3
* ReLU + Batch Normalization + Dropout				
GMP = GlobalMaxPooling2D				
Classifier's structure				
Fully connected layer (32) + ReLU + dropout (rate = 0,3)				
Fully connected layer (32) + ReLU + dropout (rate = 0,3)				
Fully connected layer (2) + Softmax				

3.2. Post-processing framework

The objective of the filtering stage is to enhance the prediction map by removing as many false positives (FP) as possible. The first step is to remove aberrant FPs, such as edge artifacts and predictions whose areas are too small. Since we observe that 98% of moraine polygons exhibit areas larger than 20 pixels, we disregard any prediction (positive or negative) smaller than that threshold. The second step is to eliminate predictions that do not fit the context of moraine formation, i.e. located more than 1 km away from a glacial valley. We use the current glacial valleys as indications because the present-day trajectories of main drainages, that usually drain glaciers meltwaters, are strongly similar to the trajectories of the paleo-glaciers that formed the valleys. Each drainage is initiated from the lowest point of the local glacier identified from the Randolph Glacier Inventory 6.0 database (RGI Consortium 2017) and propagated downstream using SAGA-GIS channel networks algorithm (Conrad et al. 2015). The third step consists of reducing FPs along the drainage network by removing the predictions located on the drainage network (1 pixel wide) and taking a 1-pixel buffer on either side. For small drainages, we use the topographic position index (TPI) module of SAGA-GIS (Wilson and Gallant 2000) to provide information on the presence of valleys or crests. This index is calculated based on the DTM with a radius of 100 m which corresponds to the average median width of the moraines. If a pixel has a negative (positive) TPI value then this pixel is located in a valley (crest). The more negative the TPI value is, the more incised the valley is, and conversely, the more positive the value is, the more marked the crest is (De Reu et al. 2013). We, therefore, remove predictions in areas with a TPI between −14 and −30 which corresponds to the talwegs of the small valleys. The fourth and final step aims to smooth the edges of the predictions and remove the small residual elements with the opening operator of mathematical morphology. The MorNet's prediction map obtained after post-processing filtering is hereafter referenced to MorNet_{filt}.

3.3. Assessment

Expert mapping of lateral and frontal moraines along six glacial valleys as well as their associated piedmonts was performed based on the DTM with a spatial resolution of 1 m (Pleiades-derived photogrammetry) and panchromatic images with a spatial resolution of 0.5 m (Pleiades). We observe that moraines are generally well defined and dissected near their lowest point by overflowing melt water and subsequent stream incision. This provides a natural separation between right and left bank moraines (mapped polygons) used hereafter. To account for spatial variability in moraine geomorphology (e.g. width, slope, elevation), the training set is composed of either right or left bank moraines from all valleys, excluding valley #4, where moraines on either side of the banks do not correspond and are too small due to erosion processes to be separated with an acceptable degree of confidence. An additional subset from the training set is reserved for the validation step, while the remaining moraines from opposite banks and both banks of valley #4 constitute the test subset. Furthermore, we decide to take different moraines for training and testing to limit possible spatial autocorrelation issues (Kattenborn et al. 2022). The spatial distribution of the train, validation, and test moraines is presented in Figure 1. From these moraines, we extract remote sensing image patches to constitute three data sets: training, validation and test set. Each set of data is composed of collections of patches extracted from moraines and patches randomly extracted from the rest of the area (areas mapped as non-moraine). Since the maximum median width of moraines observed over the study area is 250 m and our image set is at spatial resolution of 10 m, we decide to set patch size at 25×25pixels.

The label assigned to a given patch corresponds to the class of its central pixel. Train and validation sets are balanced, while the test set is unbalanced to reflect the distribution observed over the entire area between moraine and non-moraine pixels. Details of data sets characteristics (number of patches, number of moraines, etc.) are presented in Table 2. The data is normalized between 0 and 1 by taking the minimum and maximum values over the entire study site. Following recommendations by Filipponi (2019)), a logarithmic transformation is also applied to the SAR images before normalization. To evaluate the performance of the model we use two standard evaluation metrics: the f1-score and the Intersection over Union (IoU) (Garcia-Garcia et al. 2018; Hossin and Sulaiman 2015). F1-score is particularly adapted to our specific data set due to its ability to cope with scenario in which the classes distribution is highly unbalanced. The model is trained over many iterations, called epochs, and we keep the model obtained at the epoch achieving the best f1-score performance on the validation set, then this model is used to provide the final classification on the test set. This strategy is employed to avoid model overfitting if only training data was considered. F1-score is defined as follows:

Table 2. Datasets characteristics.

	Train set	Validation set	Test set
Number of positives patches	27 989	6 021	49 383
Number of negatives patches	27 989	6 021	538 509
Total number of patches	55 978	12 042	587 892
Number of moraines used for patches extraction	19	5	35
Area used for positives patches extraction (km^2)	3.1	0.7	5.4

Table 3. Detailed characteristics of each model.

	Last encoder layer	Fusion layer	Constant α	Sources
MorNet	Flatten	Concatenate	0.3	DTM, S1 and S2
$MorNet_{DTMS2}$	Flatten	Concatenate	0.5	DTM and S2
$MorNet_{DTMS1}$	Flatten	Concatenate	0.5	DTM and S1
$MorNet_{S2S1}$	Flatten	Concatenate	0.5	S1 and S2
$MorNet_{DTM}$	GlobalMaxPooling2D	/	/	DTM
$MorNet_{S2}$	GlobalMaxPooling2D	/	/	S2
$MorNet_{S1}$	GlobalMaxPooling2D	/	/	S1

$$f1\ score = \frac{TP}{TP + 0.5(FP + FN)} \times 100 \qquad (2)$$

with TP = True positive, FP = False Positive and FN = False Negative. In parallel, the IoU compares the overlapping between ground truth and prediction and is defined by Equation 3. This metric is increasingly used in the field and can handle unbalanced datasets; we use it here to facilitate future comparisons with other methods. The values of the f1-score and IoU range from 0 to 100. The closer the value is to 100, the better model performs.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \times 100 \qquad (3)$$

In addition to the full multi-source model called MorNet, we also produced mono-source models (i.e. trained on a single data source) hereafter called $MorNet_{DTM}$, $MorNet_{S1}$, and $MorNet_{S2}$, as well as bi-source models (i.e. trained on two data sources) called $MorNet_{DTMS1}$, $MorNet_{DTMS2}$ and $MorNet_{S1S2}$ (Table 3).

3.4. Experimental settings

There are some variations between the parameters of the mono-source and multi-source models. The last encoder layer of MorNet is a flattened layer (Flatten), and the fusion of the branches is achieved by concatenation. The constant α of the loss function is equal to 0.3. Models involving two input sources (see above), have the same settings as MorNet (flatten layer and fusion by feature concatenation) except for the hyper-parameter α which is equal to 0.5. We choose different α values between bi-source and multi-source so that the contributions of the per-source classifiers are almost equivalent to that of the final multi-source classifier. Mono-source models use a pooling function (GlobalMaxPooling2D) as last operation in the encoder component instead of a flattened layer (Table 1) because, experimentally, mono-source models have exhibited better performance with a pooling layer. This is probably because the GlobalMaxPooling2D layer provides more robust and translation-invariant representations. Table 3 summarizes the characteristics of each model.

All the models were trained for 500 epochs with a batch size of 128. To learn the neural networks parameters we use the Adam optimizer (Kingma and Ba 2017) with a learning rate equal to 10^{-4}. Models were trained on a workstation with Nvidia GeForce RTX 3090 graphic card, an Intel Xeon Gold 6226 R CPU and 24 GB of RAM. All the models were implemented via the Python Keras/Tensorflow library.

4. Results

In this section, we provide both a quantitative and a qualitative assessment of our deep learning model. To this end, we first present the quantitative results of our model as well as all the per-source ablations. Then, we inspect how the training data homogeneity impacts the behavior of MorNet. Successively, we qualitatively analyze the output of the deep learning framework in order to get additional insight about the characteristics of its predictions and the confidence of the model in its predictions. Finally, we present the improvements made by the filtering process.

4.1. Quantitative results

4.1.1. Models performances

The performance metrics on the test moraine predictions and training time are presented in Table 4.

Table 4. Performances on positive test set for each model (f1-score and IoU) and percentage of false positives (FP) and false negatives (FN) on test set. Time to produce the datasets, train and test the model, and create the prediction map for both areas. $MorNet_{filt}$ corresponds to MorNet's prediction after post-processing.

Model	F1-score on moraine class	IoU	Time	% FP	% FN
$MorNet_{DTM}$	41.6	26.3	~ 1h30	4.3	5.1
$MorNet_{S2}$	40.4	25.3	~ 7h30	**8.7**	4.1
$MorNet_{S1}$	37.7	23.2	~ 5 h	4.8	**5.4**
$MorNet_{DTMS2}$	50,8	34.1	~ 12 h	6.5	3.3
$MorNet_{DTMS1}$	48,4	31.9	~ 10 h	7.8	3.2
$MorNet_{S1S2}$	46,1	30.0	~ 14 h	4.5	4.6
MorNet	**52,8**	**35.9**	~ 18 h	4.8	3.7
$MorNet_{filt}$	**54.2**	**37.3**	/	3.4	4.0

Among the mono-source models, $MorNet_{DTM}$ performs the best, with the highest f1-score and IoU. $MorNet_{DTM}$ has also fewer false positives than $MorNet_{S2}$ and $MorNet_{S1}$ (4.3% versus 8.7% and 4.8% for $MorNet_{S2}$ and $MorNet_{S1}$). Moraines being essentially characterized by their morphology, the DTM source allows better performance than the other mono-source models. Then, we can observe that bi-source models exhibit a better behavior, in terms of performance results, than their mono-source counterparts. $MorNet_{S1S2}$ achieves the lowest performance among all the models exploiting only two input sources. This confirms that the DTM source provides the most relevant information for moraine characterization among all the considered data sources. $MorNet_{DTMS2}$ has a better f1-score and IoU, despite a high percentage of FP, than $MorNet_{S1S2}$. However, $MorNet_{DTMS2}$ makes less FN than the latter (3.3% versus 4.6%). The better performance of $MorNet_{DTMS2}$ compared to $MorNet_{DTMS1}$ can be explained by the results of the mono-source model. $MorNet_{S2}$ achieves better performances than $MorNet_{S1}$, showing that the S2 source is more discriminative than S1. Finally, with an f1-score of 52.8 and an IoU of 36.9, the three-source model, MorNet, slightly outperforms the bi-source models. In addition, with the exception of $MorNet_{S1S2}$, MorNet has a lower percentage of FP than the bi-source models. In conclusion, the multi-source models generalize better than the mono-source models. Thus, the addition of new sources seems to improve the model performance. We can note that the absolute value of the different metrics is not generally high, but this is partly due to the structure of the test set. The test set is highly unbalanced and pulls some of the results down.

4.1.2. Bootstrap analysis on training moraines

To assess whether morphological variability of the moraine stock influences our model, we perform a bootstrap analysis (Boos 2003). Thus, 19 sub-datasets are built from the complete dataset by removing each time one moraine (i.e. all associated polygons) among the total of 19 training moraines. These sub-datasets are used to train all variations of our model with one, two or three sources. Then, for each bootstrap step, we compare the performance of each sub-model to with the complete dataset. If the f1-score obtained with a sub-dataset is better than that obtained with the complete dataset, it indicates that the removed moraine negatively affects the general learning process.

First, in the case of $MorNet_{DTM}$, out of the 19 sub-datasets, 17 outperform the complete dataset (Figure 4), suggesting that the complete dataset contains moraines with highly heterogeneous characteristics. The best f1-score is obtained by removing the lowest and flattest moraine of the training set (moraine #43, indicated in Figure 1). This moraine has a median slope of about 10° and a median height of 8 m and is, therefore, one of the outliers of the training set (Figures 2 and 4). Similarly, the second best result is obtained by removing the steepest moraine with a median slope of 33° (moraine #25, indicated in Figure 1) from the training set. The bootstrap performed on the training moraines causes a variation of the f1-score of 8.2 (from 41,4 to 49,6) (Figure 4).

$MorNet_{S2}$ behavior is similar to $MorNet_{DTM}$, with 17 sub-datasets outperforming the complete dataset. To summarize the characteristics of the ten bands used in the S2 source, we calculated the NDVI (Normalized Difference Vegetation Index) and NDMI (Normalized Difference Moisture Index) as follows:

$$NDVI = \frac{Near\ Infrared\ (Band\ 08) - Red\ (Band\ 04)}{Near\ Infrared + Red} \tag{4}$$

$$NDMI = \frac{Near\ Infrared\ (Band\ 8A) - Short\ wavelength\ infrared\ (Band\ 11)}{Near\ Infrared + Short\ wavelength\ infrared} \tag{5}$$

These two indices are classically used in land cover studies and incorporate information in the visible, near-infrared and mid-infrared ranges, covering the entire spectral range of the ten sentinel-2 bands (Huang et al. 2021; Sahu 2014). NDVI, on the other hand, should enable us to differentiate between snow and water (negative value), rock and bare soil (zero value), and vegetation (positive value). NDMI reflects moisture levels in vegetation, so very negative values correspond to bare soil and very positive values to healthy vegetation. We calculated the median of these indices for each moraine. Similarly, performance is improved when discarding moraines with extreme characteristics such as the highest and lowest median NDVI (moraines #27 and #25, respectively). Overall, the moraines between the first and last decile of the NDVI or NDMI indices are the ones that show the most improvement, with two exceptions. Moraine #30 is outside these deciles but has the 2nd highest reflectance value in band 12, whose spectral values are not

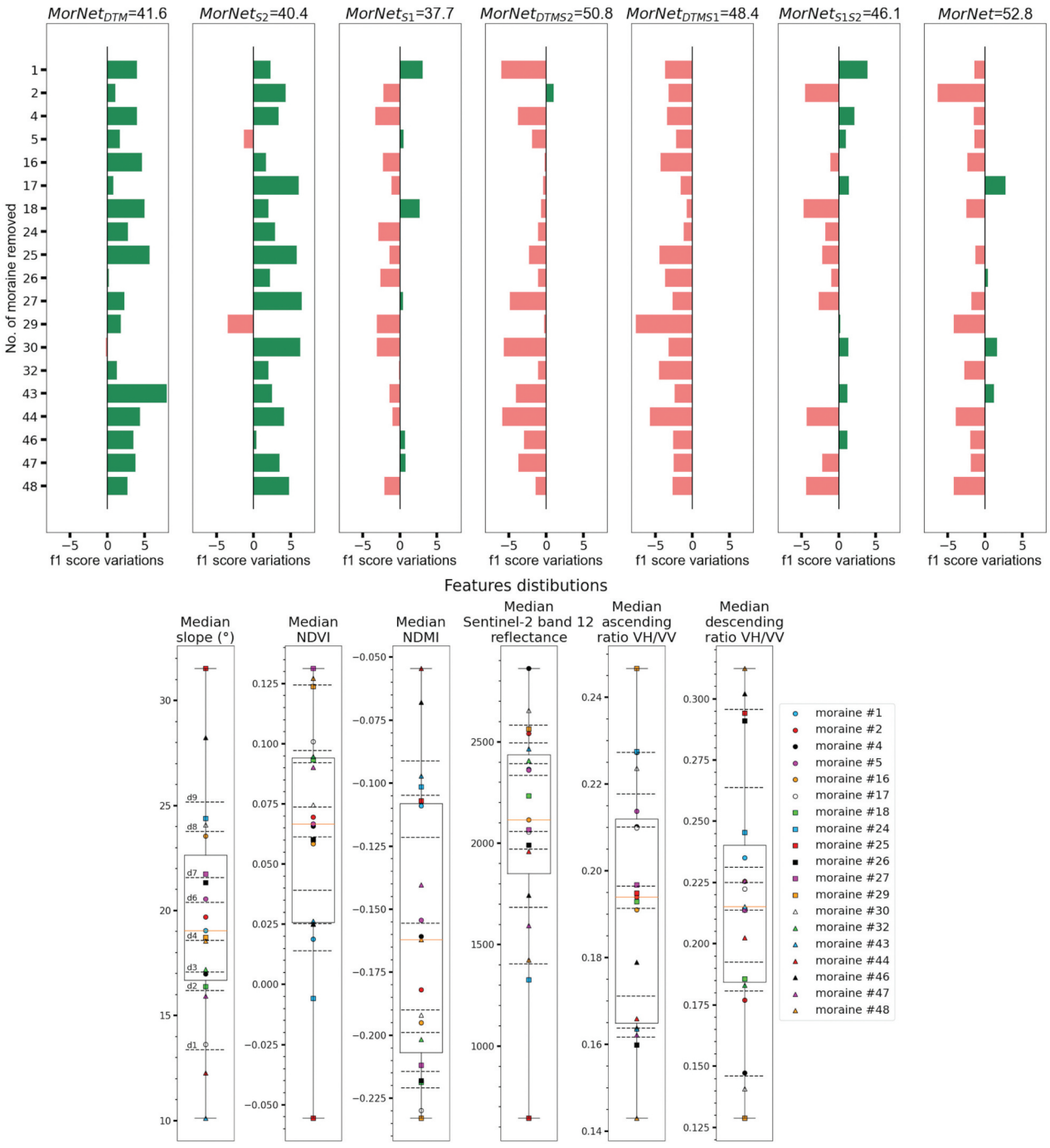


Figure 4. Predictions variations on the 19 modified train sets (sub-datasets) compared to initial train set (complete dataset), whose f1-score value is noted above the graph, and boxplot distributions of moraine features. Outliers are not shown on the boxplot, so the lower and upper whiskers correspond to the minimum and maximum values, and deciles are displayed with dashed lines.

covered by the NDVI and NDMI indices and gives information on lithology (Figure 4). Therefore, eliminating outliers, such as the most vegetated moraines and those with the driest or snowiest soils, improves performance.

Conversely, of the 19 MorNet$_{S1}$ sub-datasets, only 6 exceed the f1-score of the complete dataset, of which 4 are minor improvements ($\leq$ 1.0 variations). The two moraines that were removed, resulting in a significant improvement in the results, are moraines #1 and #18. These moraines have a median descending ratio VH/VV bands between the 1st and 3rd quartile (Figure 4). It appears that removing moraines with a VH/VV ratio between the two quartiles leads to better performance for this particular data source. Presumably, this is because there is a considerable amount of redundant information within the S1 source, and eliminating some of the"classic" moraines allows the model to focus on the more distinct moraine features.

The bootstrap performed on the MorNet$_{DTMS1}$ and MorNet$_{DTMS2}$ bi-source models showed a deterioration in results in almost all cases, which may mean that the information learned from the two sources is contradictory and that the models have difficulty

generalizing. In the case of $MorNet_{DTMS2}$, the moraines that caused little variations (between +1.0 and −1.0) that could be caused by variability intrinsic to the model are moraines with a median NDMI below the 4th decile. Only moraine #27 and #30 corresponds to these characteristics but generates a significant degradation, but they are moraines with extreme features as discussed in the case of $MorNet_{S2}$. The other moraines show significant degradation, except for moraine #24, whose degradation is minor (Figure 4). In the case of $MorNet_{DTMS1}$, the moraines that produce the most degradation are outlier moraines such as #29 (lowest VH/VV descending ratio) and #44 (2nd lowest median slope). Overall, the removal of moraines in the 8th decile and under the 3rd decile of the VH/VV descending ratio caused the highest degradation (Figure 4). Therefore, it seems that, unlike the mono-source models, outlier moraines need to be retained in the dataset for good feature learning in these two bi-source models.

In the case of $MorNet_{S1S2}$, the removal of moraines with VH/VV ascending ratio between the 6th and 9th deciles improve significantly the results (variations $\geq$1.0). Some moraines lead to better performance than the complete dataset when their NDMI and slope characteristics are the highest, without being in the 5% most extreme, like moraine #43 and #46 (Figure 4). Thus, in the case of $MorNet_{S1S2}$, the same phenomenon as in the other bi-source models is observed, although to a lesser degree: having more heterogeneous moraines is preferable. Overall, we note that the relationships between mono-source and multi-source bootstrapping are not linear (removing a moraine that causes an improvement in two mono-source models will not necessarily lead to an improvement when the sources are merged, and vice versa).

In the case of the MorNet model, removing moraine #17 is the only case that leads to a significant improvement. It is the 3rd moraine that brings the most enhancement on $MorNet_{S2}$, and on all the other models it causes little variations (between −1.5 and + 1.5). The highest degradation is caused by removing moraines #2, 29, 44, and 48. Moraine #2 is a moraine with near median characteristics except for VH/VV descending ratio since it is the 4th weakest. Moraines #29, 44, and 48 are moraines with outliers features on at least two sources. Moraine #29 has the lowest NDVI and highest median VH/VV ratio, while moraine #44 has the highest NDMI and 2nd lowest median slope, and moraine #48 has the highest aspect, lowest VH/VV ratio, and second highest NDVI. We note that the degradation caused by removing moraine #2 is present in all models using source S1 (Figure 4). However, there is no such clear relationship for the other moraines. It seems that the removal of moraines with specific extreme characteristics is necessary for the model to generalize well, while other features lead to a loss of model learning. Merging sources and processing them in a non-linear manner makes it difficult to determine the exact characteristics that lead to better performance.

Although the improvements of mono-source models are superior to those of multi-source models, the median variations caused by bootstrapping are lower in the case of the complete MorNet model, which suggests our approach is robust with respect to an heterogeneous training set. In that regard, the end user may consider the effort necessary to build a perfect data set to train the faster $MorNet_{DTM}$ model versus using a readily available unrefined data set to train the slower MorNet. However, these performance variations are not only due to the choice of the training moraines. The non-moraine training patches are randomly selected from the negative class at each new training. It is, therefore, possible that the negative patches were contributing more than the moraine patches to the effective training of the deep learning model. However, these results underline the importance of a good selection of moraines to be included in the training set and ascertain that discarded moraines are not sampled to build the negative class. This also shows that potential errors in ground truth mapping (unidentified or incorrectly delineated moraines) can significantly affect model learning and performance.

4.2. Qualitative results

4.2.1. Mono-source results

First, we focus on the location of the predictions of $MorNet_{DTM}$ shown in Figure 5. $MorNet_{DTM}$ mainly identifies well-defined crests (e.g. moraines #22 and #23), and its predictions describe relatively elongated shapes that resemble moraines. False positives (FPs) cluster into small groups (average area of 750 m^2), and are mostly located around the large debris flow in the northern zone that is at the same elevation as the moraines. The debris flows downstream of the drainages in the southern zone do not show any FPs because they are located at different altitudes compared to the moraines. This is not the case for the debris flow in the northern zone which, in addition to being at the same altitude as the moraines, is located at the outlet of a valley, thus inducing the model to make errors. The rest of the FPs are mainly on crests or areas with textures similar to the debris flow (Figure 5).

$MorNet_{S2}$ identifies almost all the test moraines in the northern area but also makes many false predictions there. A portion of these FPs is located around the training and validation moraines, suggesting that the model has difficulty delineating moraines. However, some of these FPs may be moraines, in

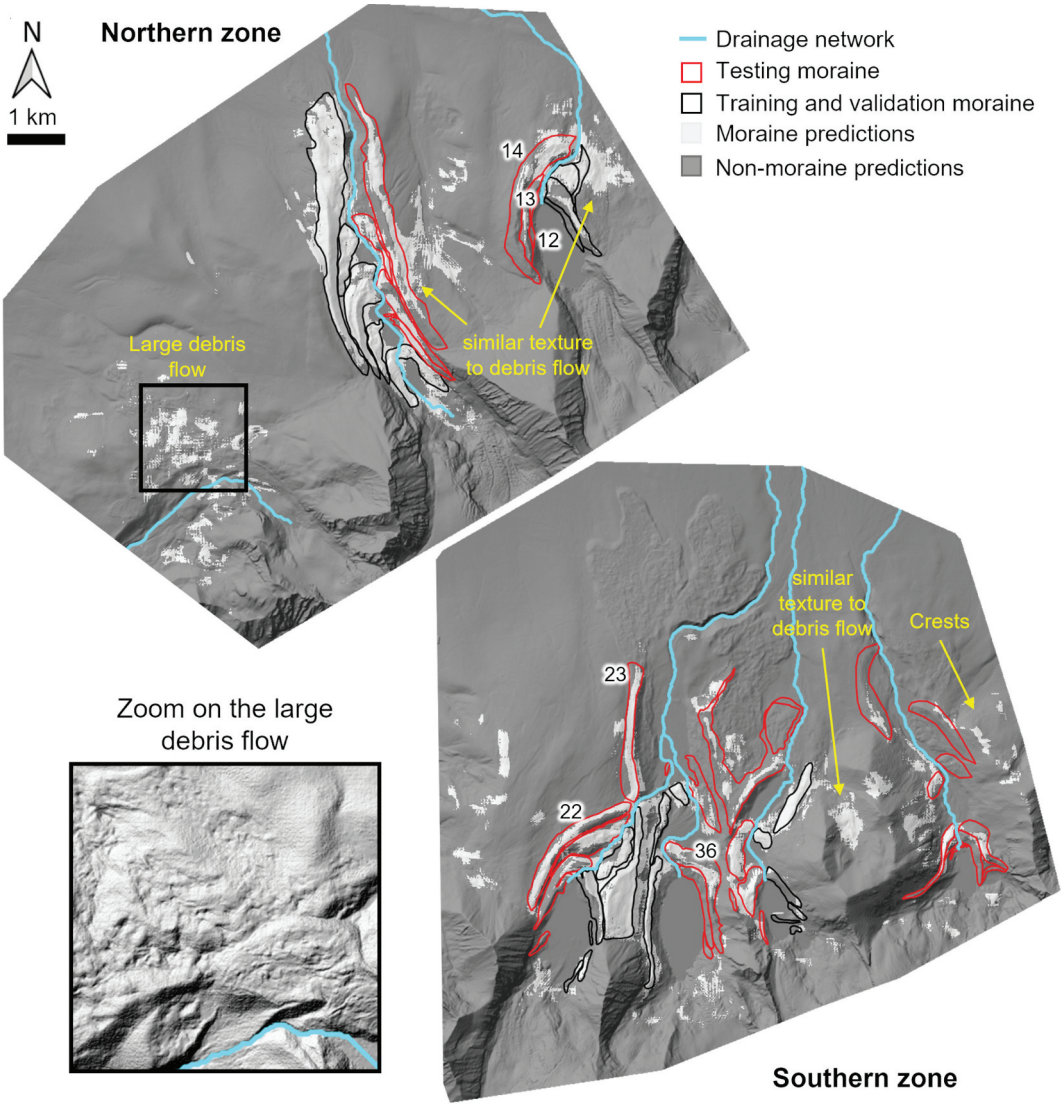


Figure 5. MorNet$_{DTM}$ model predictions at the study site with DTM with a spatial resolution of 10 m in background. The black box is the location of the zoom on the large debris flow with DTM with a spatial resolution of 10 m in the background.

reality. MorNet$_{S2}$ identifies crests less well than MorNet$_{DTM}$, therefore, the predictions of MorNet$_{S2}$ have a less linear shape than MorNet$_{DTM}$. FPs are primarily around drainages and spread over a greater elevation range than moraines, in contrast to MorNet$_{DTM}$ predictions which extend over the same elevation range as moraines (Figure 6(a)). However, the easternmost valley in the southern zone has very few "moraine" predictions (whether right or wrong). In the Sentinel-2 near- and mid-infrared image, this valley has particularly low reflectance. Since the model was trained on this geographical area with only patches depicting non-moraine class, it probably associates all the information from this area to the negative class, resulting in several FNs results. The large areas of FP are rather located in bright areas (Figure 6(b)).

MorNet$_{S1}$ identifies only a part of the crests of the same moraine, such as for moraines #14 and #22. Smaller moraines are better identified especially in the easternmost valley of the southern zone. FPs are located as in MorNet$_{DTM}$ at the large debris flow and glacier edge, as well as along the drainage network (Figure 6(c)). Moreover, east of moraine #23, the prediction follows a small landform with a gradient profile similar to moraines near the drainage network rather than the true moraine (Figure 6(d)).

4.2.2. Multi-source results

We compare the differences between the FPs of multi-source models and the mono-source models from which they are derived. We observe that, in the case of MorNet$_{DTMS2}$, there are fewer FPs near the large debris flow and the glaciers than in the case of MorNet$_{DTMS2}$. We also note that MorNet$_{DTMS2}$ has fewer FPs near the drainage network and moraine boundaries than MorNet$_{S2}$. Similarly, MorNet$_{S1S2}$ predictions have fewer FPs at the drainage network and debris flow than MorNet$_{S1}$ predictions. Thus, it seems that the FPs located at different areas by two mono-source models will be less present in the multi-source model

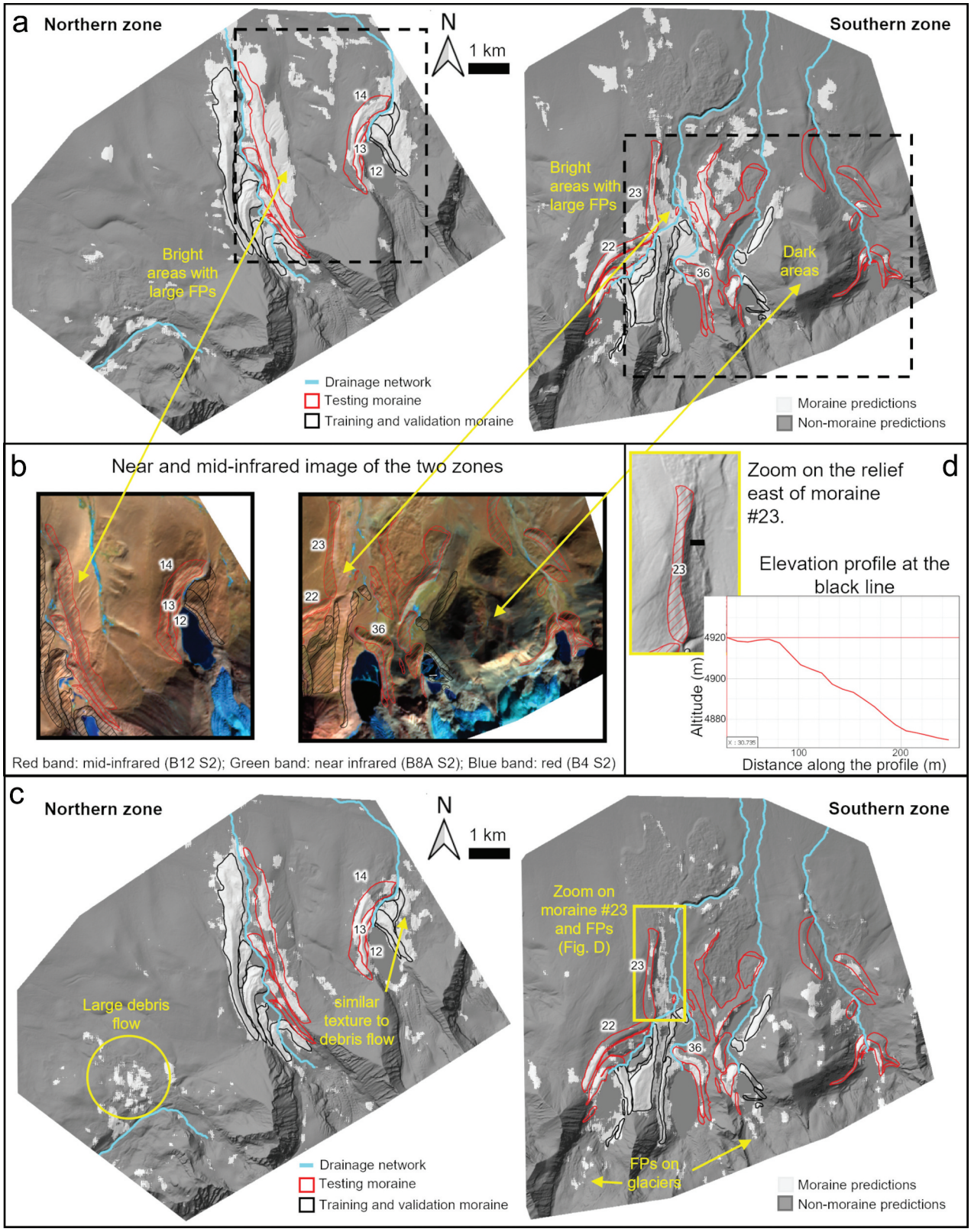


Figure 6. (a) MorNet$_{S2}$ predictions. (b) Zooms in near and mid-infrared Sentinel-2 image. Positions of these zooms are presented in black dashes in Figure 6A. (c) MorNet$_{S1}$ predictions. (d) Zooming in on a false positive and its gradient profile. The location of the zoom is given in Figure 6C. DTM with a spatial resolution of 10 m in background.

using these two sources. On the contrary, we note that MorNet$_{DTMS1}$ has more FPs near the large debris flow and glaciers than MorNet$_{DTM}$ and MorNet$_{S1}$. These FPs, initially present in the two sources, are more extensive in this case. Furthermore, the same FP is found east of moraine #23 between MorNet$_{S1}$ and MorNet$_{S1S2}$. Therefore, mono-source models with FPs in the same areas exhibit many FPs after the fusion of these sources (Figure 7).

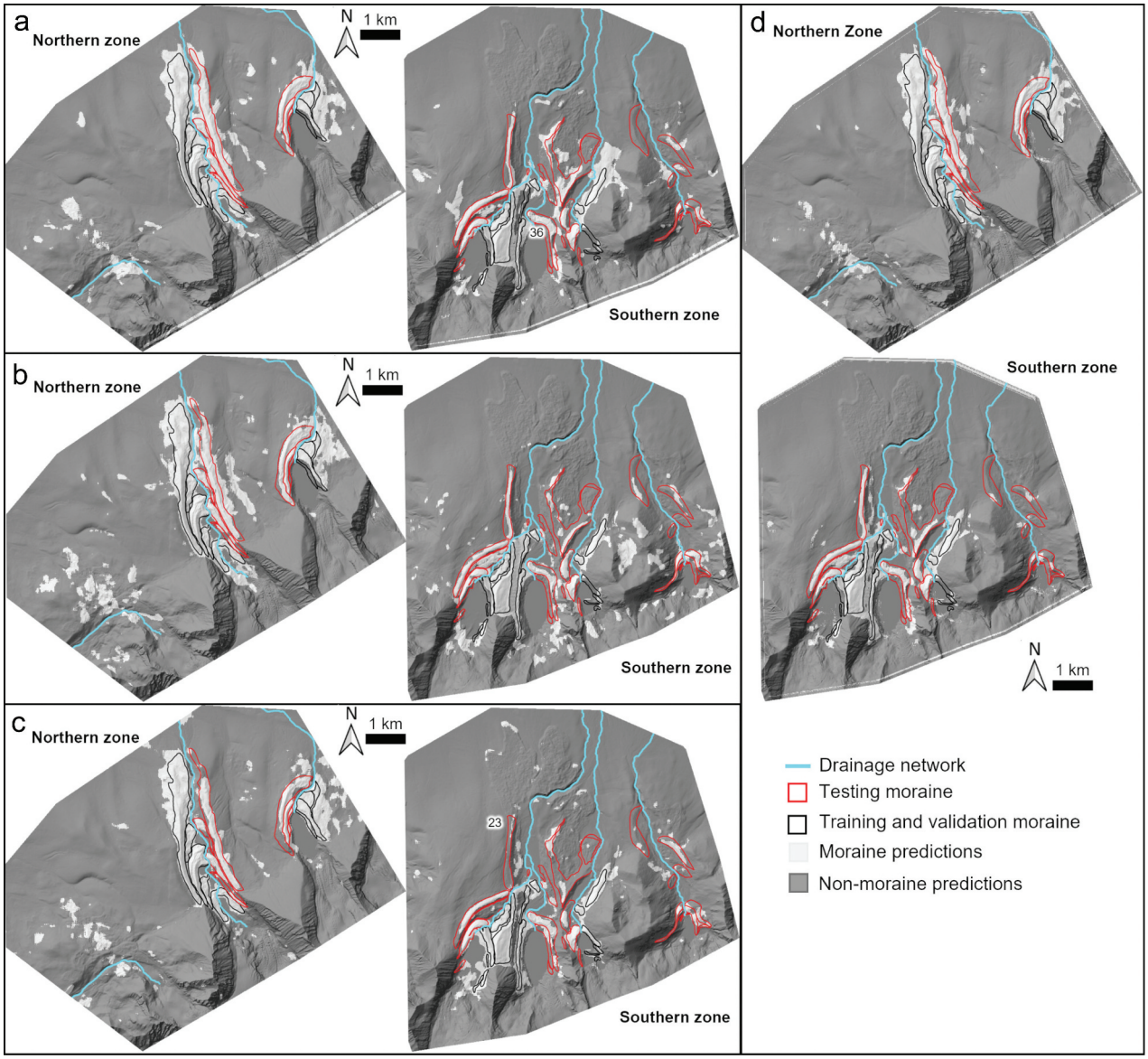


Figure 7. (a) MorNet$_{DTMS2}$ predictions. (b) MorNet$_{DTMS1}$ predictions. (c) MorNet$_{S1S2}$ predictions. predictions. (d) MorNet predictions (three-source fusion). DTM with a spatial resolution of 10 m in background.

This phenomenon also works for correct predictions. Globally, moraines identified by two sources separately will be better identified by the fusion of these two sources (e.g. moraine #36). Multi-source models with DTM source correctly identify the crests and allow to find the large-sized moraines. S2 and S1 sources are the most effective to find the moraines flanks. Finally, the predictions from the fusion of the three sources (MorNet) are a combination of the previous observations. FPs at the debris flow, glacier and drainage network level are present but less frequent, while in some two-source combinations they were absent. However, the model has a good trade-off between errors and correct predictions, as it identifies both the flanks and crests of the most prominent moraines.

4.2.3. Confidence of MorNet

Beyond the traditional boolean representation, we explore the correlation between high probabilities and accurate predictions in detail. We display the probability that a pixel belongs to the moraine class in Figure 8 and observe that most moraine predictions have a probability of belonging to this class higher than 0.9. Particularly, moraine crests exhibit higher probabilities than their flanks, which shows that the model is more confident in identifying a specific shape. FPs do not necessarily exhibit low probability; FPs located near zone boundaries (e.g. surrounding the lakes and borders of the zones) and around moraine boundaries in the northern zone have a high probability of being moraines. There are FPs with a lower probability than 0.8, such as those located on the large debris flow of the northern zone or in the drainage network of the southern zone. However, the correct predictions in the eastern valley of the southern zone have the same probability as the FPs. The high confidence in the other valleys suggests that there may be some autocorrelation effect due to the

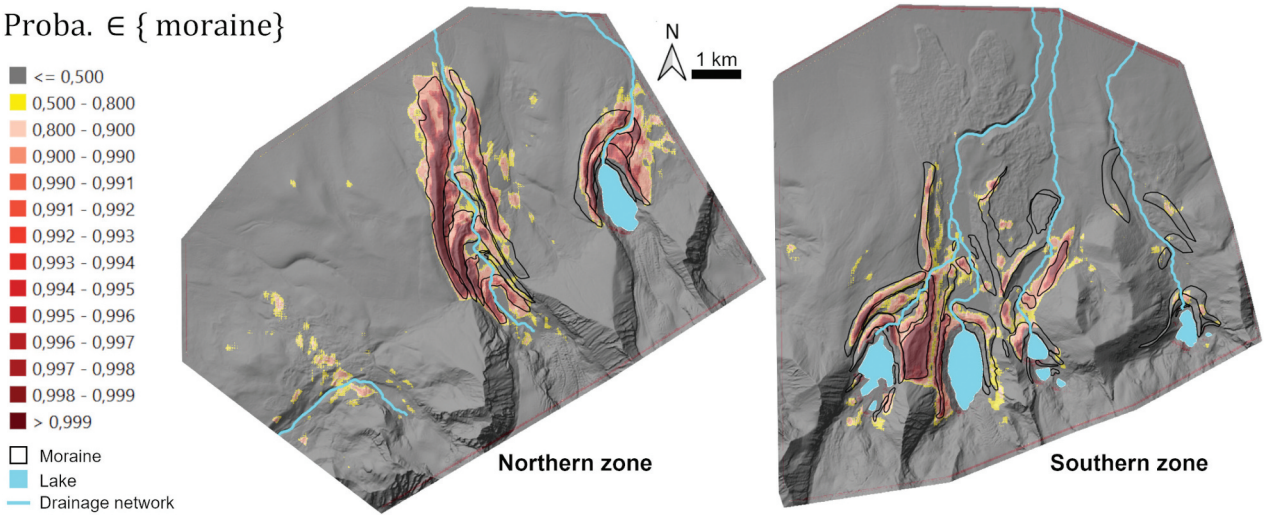


Figure 8. Probabilities that MorNet predictions belong to the"moraine" class.

proximity of the test and train moraines in these valleys. Thus, using the raw model confidence to filter the predictions does not seem suitable since FPs have the same confidence as true moraines. This also highlights the overconfidence of MorNet in its predictions, whether they are right or wrong. It seems necessary to use another indicator than the model confidence of belonging to the moraine class to remove these incorrect predictions and improve the model's performance.

4.3. Post-processing of MorNet predictions

A comparison of the prediction maps before and after the post-process framework described in section 3.2 shows a significant decrease in the number of FPs, especially at the large debris flow in the northern zone where they are reduced by 40% (Figure 9). To a lesser extent, FPs located near glacial valley drainage networks are reduced by about 19%. In general, the contours of the predictions are more homogeneous and easier to distinguish. Indeed, the filtering method removes the FPs in the valleys, some of which connect two moraines. However, this method does not allow for a better delineation of certain moraines, notably around the training moraines to the east of the northern zone, because they are not adjacent to highly incised valleys. Thus, it is not possible to perfectly distinguish each moraine, but moraines separated from others by deep valleys are well delineated. In conclusion, the post-process framework enables a reduction of FPs and allows a better delineation of the moraines near the drainage network.

Overall, MorNet yields a 11.2 increase (26.9%) in f1-score with respect to the best mono-source model MorNet_{DTM} (Table 4). Specifically, the former generates less FNs and more FPs, which are efficiently filtered out by our post-processing framework. Compared to MorNet, MorNet_{filt} (MorNet after filtering) reduces the percentage of FPs by one-third while increasing the number of FNs very slightly (Table 4). Generally, a model that exhibits a high frequency of FPs can be considered as over-predictive, and its performance can be optimized through the implementation of filtering strategies. Conversely, a model with a high incidence of FNs demonstrates under-predictive characteristics and it will be more challenging to enhance its product no matter the post-processing step we will use. Hence, it is more preferable to maintain an over-predictive model.

5. Discussion

To summarize, our study proposes a model to aid in lateral and frontal moraine mapping using multi-source satellite data (DTM, multispectral and radar data).

Firstly, we focused on the contribution of multi-source data to improve the model's ability to identify moraines. In the literature, only a few models applied to the mapping of geomorphological objects exploit multi-source data (e.g. Du et al. 2019; Ghorbanzadeh et al. 2019; Giaccone et al. 2021; Li et al. 2020; Xie et al. 2020), and none of them provide an in-depth study related to the interplay between the different remote sensing information and their impact on the final mapping procedure. Our study highlights the ability of MorNet to exploit the complementarity of remote sensing data from different sources, which increases performance compared to mono-source models. Indeed, DTM data best describes large-dimension moraine crests, Sentinel-2 multi-spectral data highlights moraine flanks, and Sentinel-1 radar data detects small reliefs. Hence, multi-source models detect all moraine parts (crests and flanks) while limiting overestimation. In addition, using open data such

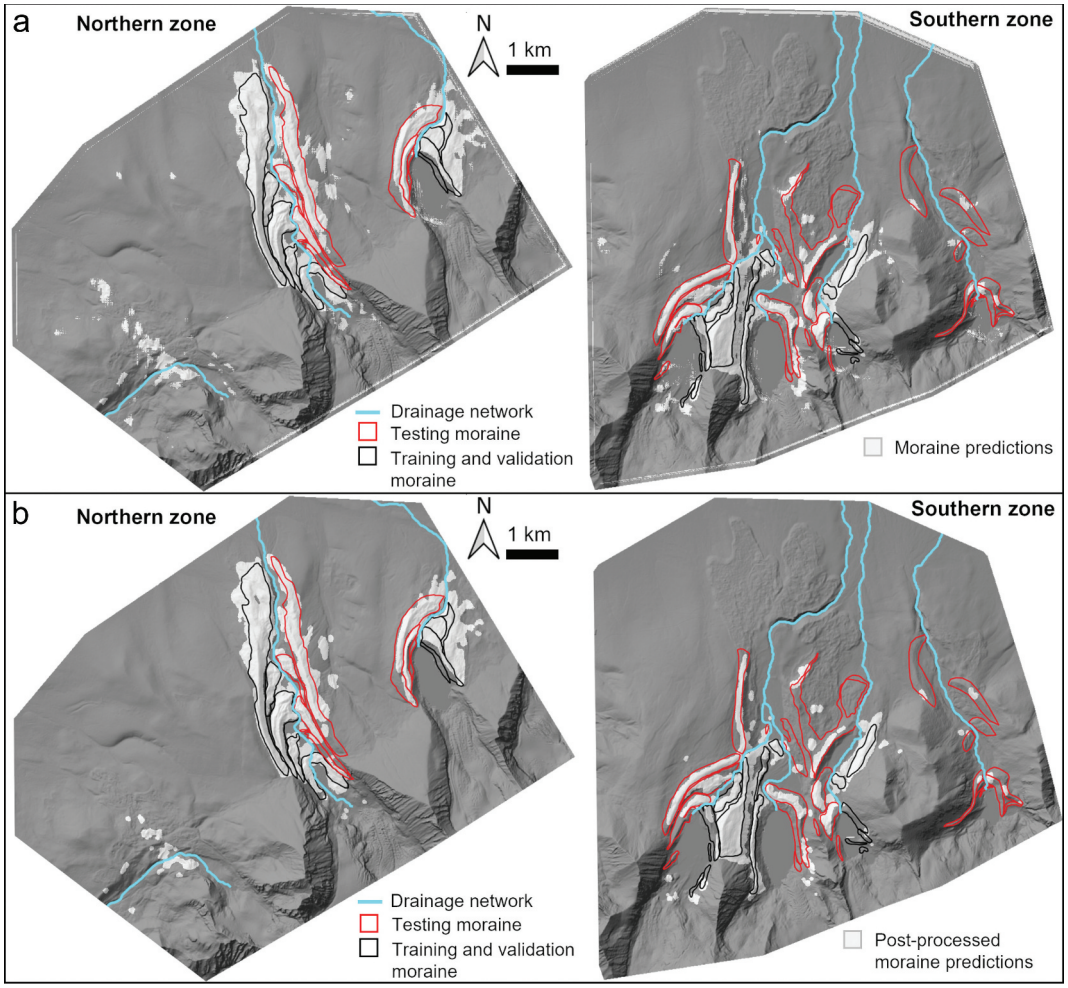


Figure 9. MorNet prediction map before (a) and after (b) post-processing.

as Sentinel-1 and Sentinel-2 data can effectively complement other available information (i.e. topographic data). The use of open data allows to limit the financial effort, and their global spatial coverage makes them actionable to support the mapping process of almost any study area worldwide.

Secondly, MorNet exhibits promising ability to identify well-defined objects associated with pristine moraines. However, there is still room for improvement for second-order features (small and strongly eroded moraines) and tends to produce false positives in debris flow areas and in-between closely grouped moraines. This aspect of our model makes it especially relevant to identify well-preserved landforms with potential to contribute in further studies at more fine-grained level. MorNet can assist the process of large-scale geomorphological mapping over vast areas that would otherwise require extensive resources in terms of human-effort and cost.

Thirdly, the ability of the model to learn moraine characteristics is highly dependent on the training dataset, which should be representative enough to cover a wide range of sizes, geometries, and preservation states. For instance, a selective dataset (i.e, composed of best-defined objects) may permit to identify only pristine moraines with a high performance index, while a wide dataset would enable to detect most moraines but with increased false positives occurrence. It is up to the user to find a trade-off between these two end-members according to the underlying mapping needs.

Finally, we want to point out some limitations and make some recommendations. Firstly, multi-source models provide better performance than mono-source models, but the performance gain remains limited in regard to the longer pre-processing and computation time. In addition, they necessitate dedicated deep learning techniques (multi-branch structure and auxiliary classifiers) that can be challenging to conceive and implement for nonexpert users. Depending on the user's computing resources, skills, and needs, it may be more appropriate to use only the DTM source. However, MorNet_{DTM} is more sensitive to data quality and will require greater attention when building the training dataset, unlike MorNet, which generalizes better even with a heterogeneous training dataset. Though MorNet_{DTM} fails to detect some moraines, the information from the DTM appears to be the most relevant as attested by the relatively few false

positives it produces. Secondly, our model only considers local spatial information, whereas experts take into account the geomorphological context to identify moraines. In the future, it would be interesting to explore the possibility to integrate such contextual information similarly to what we have done at the post-processing stage (distance to the drainage network). Lastly, our study site is located in an semi-arid zone with scarce vegetation and homogeneous lithology, so that satellite photogrammetry may yield a DSM that closely approximates a DTM. However, this also results in rather limited information captured by optical satellite imagery and, therefore, a limited contribution for Sentinel-2 data to multi-source models, which would not be the case in different settings. With the aim to deploy this approach in regions with significant vegetation cover we strongly recommend the use of DTM data (not DSM) in which topography truly reflects geomorphology.

6. Conclusion

We propose the first deep learning model, MorNet, dedicated to glacial moraine mapping. MorNet relies on commonly used SAR and multispectral data combined with topographic information, otherwise rarely used in computer vision approaches. Respectively, these data come from, Sentinel-1, Sentinel-2, and Pleiades (down-sampled photogrammetric DTM derived from panchromatic images) satellite images and are used to build a multi-source model. We assess the contribution of multi-source data in terms of performance by evaluating mono-source and multi-source models. The latter shows a better ability to identify the different parts of the moraines (moraine crest and flanks) by exploiting the complementarity of the data sources and improves overall performance. In addition, it is more robust to the quality of the training dataset, with lower performance variations when applied to very heterogeneous datasets. MorNet exhibits first-order mapping capability by identifying well-preserved moraines, and is destined to be used by experts as a tool to facilitate large-scale mapping efforts by indicating areas that likely contain moraines.

There are several research avenues to explore as possible follow-up of our research work. Assessing the validity of the MorNet model to other morainic landscapes in other parts of the world can provide clues about the comprehension on how well the proposed model can generalize to unseen context and this kind of analysis will also provide some room for investigation related to comparative geomorphological study involving multiple study areas. Contextual information from expert knowledge (such as distance to the drainage network) may be introduced as a data source for the training stage. Furthermore, moraines generally show small-scale spatial variability that would be better characterized by very-high-resolution data (e.g. full-resolution Pleiades images and derived DTM). In that regard, such images generally sample a narrow spectrum (visible and near-infrared), and it would be interesting to assess the evolution of MorNet's behavior considering a trade-off between spatial and spectral resolution.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research is funded by the University of Montpellier (IR's PhD fellowship) and the French Centre National d'Etudes Spatiales (access to Pleiades images).

Notes on contributors

Isabelle Rocamora received her double M.Sc degree in geosciences and engineering from the University of Montpellier and the Ecole des Mines d'Alès, France, in 2021. She is currently preparing her Ph.D. in geosciences at the Geosciences Montpellier laboratory of the University of Montpellier on deep learning approaches for geomorphological analysis based on remote sensing data.

Dino Ienco received the M.Sc. and Ph.D. degrees in computer science both from the University of Torino, Torino, Italy, in 2006 and 2010, respectively. He joined the TETIS Laboratory, IRSTEA, Montpellier, France, in 2011 as a Junior Researcher. His main research interests include machine learning, data science, graph databases, social media analysis, information retrieval and spatio-temporal data analysis with a particular emphasis on remote sensing data and Earth Observation data fusion. Dr. Ienco served in the program committee of many international conferences on data mining, machine learning, and database including IEEE ICDM, ECML PKDD, ACML, IJCAI as well as served as a Reviewer for many international journal in the general field of data science and remote sensing.

Matthieu Ferry is an Associate Professor at the University of Montpellier (France). He received his PhD degree from the Swiss Federal Institute of Technology (ETH-Zurich). His research interests are Earthquake Geology and Remote Sensing.

ORCID

Isabelle Rocamora http://orcid.org/0009-0008-3902-4313
Dino Ienco http://orcid.org/0000-0002-8736-3132
Matthieu Ferry http://orcid.org/0000-0002-6498-7807

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

References

Alifu, H., T. Ryutaro, and B. Johnson. 2015. "A New Band Ratio Technique for Mapping Debris-Covered Glaciers Using Landsat Imagery and a Digital Elevation Model." *International Journal of Remote Sensing* 36 (8): 2063–2075. https://doi.org/10.1080/2150704X.2015.1034886.

Alvioli, M., I. Marchesini, P. Reichenbach, M. Rossi, F. Ardizzone, F. Fiorucci, and F. Guzzetti. 2016. "Automatic Delineation of Geomorphological Slope Units with R.Slopeunits V1.0 and Their Optimization for Landslide Susceptibility Modeling." *Geoscientific Model Development* 9 (11): 3975–3991. https://doi.org/10.5194/gmd-9-3975-2016.

Anders, N. S., A. C. Seijmonsbergen, and W. Bouten. 2011. "Segmentation Optimization and Stratified Object-Based Analysis for Semi-Automated Geomorphological Mapping." *Remote Sensing of Environment* 115 (12): 2976–2985. https://doi.org/10.1016/j.rse.2011.05.007.

Bishop, M. P., L. A. James, J. F. Shroder, and S. J. Walsh. 2012. "Geospatial Technologies and Digital Geomorphological Mapping: Concepts, Issues and Research." *Geomorphology* 137 (1): 5–26. https://doi.org/10.1016/j.geomorph.2011.06.027.

Boos, D. D. 2003. "Introduction to the Bootstrap World." *Statistical Science* 18 (2): 168–174. https://doi.org/10.1214/ss/1063994971.

Bull, W. B. 2013. "Correlation of Fluvial Aggradation Events to Times of Global Climate Change." In *AGU Reference Shelf*, edited by J. S. Noller, M. S. Janet, and R. L. William, 456–464. Washington, D. C: American Geophysical Union. https://doi.org/10.1029/RF004p0456.

Chandler, B. M. P., H. Lovell, C. M. Boston, S. Lukas, I. D. Barr, Í. Ö. Benediktsson, D. I. Benn, et al. 2018. "Glacial Geomorphological Mapping: A Review of Approaches and Frameworks for Best Practice." *Earth Science Review* 185:806–846. https://doi.org/10.1016/j.earscirev.2018.07.015.

Chen, Y., and L. Bruzzone. 2022. "Self-Supervised SAR-Optical Data Fusion of Sentinel-1/-2 Images." *IEEE Transactions on Geoscience and Remote Sensing* 60:1–11. https://doi.org/10.1109/TGRS.2021.3128072.

Conrad, O., B. Bechtel, M. Bock, H. Dietrich, E. Fischer, L. Gerlitz, J. Wehberg, V. Wichmann, and J. Böhner. 2015. "System for Automated Geoscientific Analyses (SAGA) V. 2.1.4." *Geoscientific Model Development* 8 (7): 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015.

Dalla Mura, M., S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson. 2015. "Challenges and Opportunities of Multimodality and Data Fusion in Remote Sensing." *Proceedings of the IEEE* 103 (9): 1585–1601. https://doi.org/10.1109/JPROC.2015.2462751.

De Reu, J., J. Bourgeois, M. Bats, A. Zwertvaegher, V. Gelorini, P. De Smedt, W. Chu, et al. 2013. "Application of the Topographic Position Index to Heterogeneous Landscapes." *Geomorphology* 15 (186): 39–49. https://doi.org/10.1016/j.geomorph.2012.12.015.

Drăguţ, L., and T. Blaschke. 2006. "Automated Classification of Landform Elements Using Object-Based Image Analysis." *Geomorphology* 81 (3–4): 330–344. https://doi.org/10.1016/j.geomorph.2006.04.013.

Du, L., X. You, K. Li, L. Meng, G. Cheng, L. Xiong, and G. Wang. 2019. "Multi-Modal Deep Learning for Landform Recognition." *ISPRS Journal of Photogrammetry and Remote Sensing* 158:63–75. https://doi.org/10.1016/j.isprsjprs.2019.09.018.

Evans, I. S. 2012. "Geomorphometry and Landform Mapping: What is a Landform?" *Geomorphology* 137 (1): 94–106. https://doi.org/10.1016/j.geomorph.2010.09.029.

Filipponi, F. 2019. "Sentinel-1 GRD Preprocessing Workflow." In *3rd International Electronic Conference on Remote Sensing*, MDPI, June 11. https://doi.org/10.3390/ECRS-3-06201.

Garcia-Garcia, A., S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, and J. Garcia-Rodriguez. 2018. "A Survey on Deep Learning Techniques for Image and Video Semantic Segmentation." *Applied Soft Computing* 70:41–65. https://doi.org/10.1016/j.asoc.2018.05.018.

Gbodjo, Y. J. E., O. Montet, D. Ienco, R. Gaetano, and S. Dupuy. 2021. "Multisensor Land Cover Classification with Sparsely Annotated Data Based on Convolutional Neural Networks and Self-Distillation." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 14:11485–11499. https://doi.org/10.1109/JSTARS.2021.3119191.

Ghorbanzadeh, O., T. Blaschke, K. Gholamnia, S. Meena, D. Tiede, and J. Aryal. 2019. "Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection." *Remote Sensing* 11 (2): 196. https://doi.org/10.3390/rs11020196.

Giaccone, E., F. Oriani, M. Tonini, C. Lambiel, and G. Mariéthoz. 2021. "Using Data-Driven Algorithms for Semi-Automated Geomorphological Mapping." *Stochastic Environmental Research and Risk Assessment* 36 (8): 2115–2131. https://doi.org/10.1007/s00477-021-02062-5.

Han, W., J. Li, S. Wang, X. Zhang, Y. Dong, R. Fan, X. Zhang, and L. Wang. 2022. "Geological Remote Sensing Interpretation Using Deep Learning Feature and an Adaptive Multisource Data Fusion Network." *IEEE Transactions on Geoscience and Remote Sensing* 60:1–14. https://doi.org/10.1109/TGRS.2022.3183080.

Hong, D., L. Gao, N. Yokoya, J. Yao, J. Chanussot, Q. Du, and B. Zhang. 2021. "More Diverse Means Better: Multimodal Deep Learning Meets Remote Sensing Imagery Classification." *IEEE Transactions on Geoscience & Remote Sensing* 59 (5): 4340–4354. https://doi.org/10.1109/TGRS.2020.3016820.

Hornsey, J., A. V. Rowan, M. P. Kirkbride, S. J. Livingstone, D. Fabel, A. Rodes, D. J. Quincey, B. Hubbard, and V. Jomelli. 2022. "Be-10 Dating of Ice-Marginal Moraines in the Khumbu Valley, Nepal, Central Himalaya, Reveals the Response of Monsoon-Influenced Glaciers to Holocene Climate Change." *Journal of Geophysical Research: Earth Surface* 127 (8): e2022JF006645. https://doi.org/10.1029/2022JF006645.

Hossin, M., and M. N. Sulaiman. 2015. "A Review on Evaluation Metrics for Data Classification Evaluations." *International Journal of Data Mining & Knowledge Management Process* 5 (2): 01–11. https://doi.org/10.5121/ijdkp.2015.5201.

Huang, S., L. Tang, J. P. Hupy, Y. Wang, and G. Shao. 2021. "A Commentary Review on the Use of Normalized

Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing." *Journal of Forestry Research* 32 (1): 1–6. https://doi.org/10.1007/s11676-020-01155-1.

Ioffe, S., and C. Szegedy. 2015. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." *arXiv E-Prints*. https://doi.org/10.48550/arXiv.1502.03167.

Kattenborn, T., F. Schiefer, J. Frey, H. Feilhauer, M. D. Mahecha, and C. F. Dormann. 2022. "Spatially Autocorrelated Training and Validation Samples Inflate Performance Assessment of Convolutional Neural Networks." *ISPRS Open Journal of Photogrammetry and Remote Sensing* 5:100018. https://doi.org/10.1016/j.ophoto.2022.100018.

Keller, E., C. Adamaitis, P. Alessio, S. Anderson, E. Goto, S. Gray, L. Gurrola, and K. Morell. 2020. "Applications in Geomorphology." *Geomorphology* 366:106729. https://doi.org/10.1016/j.geomorph.2019.04.001.

Kingma, D. P., and J. Ba. 2017. "Adam: A Method for Stochastic Optimization." *arXiv E-Prints*. https://doi.org/10.48550/arXiv.1412.6980.

LeCun, Y., Y. Bengio, and G. Hinton. 2015. "Deep Learning." *Nature* 521 (7553): 436–444. https://doi.org/10.1038/nature14539.

Li, J., D. Hong, L. Gao, J. Yao, K. Zheng, B. Zhang, and J. Chanussot. 2022. "Deep Learning in Multimodal Remote Sensing Data Fusion: A Comprehensive Review." *International Journal of Applied Earth Observation and Geoinformation* 112:102926. https://doi.org/10.1016/j.jag.2022.102926.

Li, S., L. Xiong, G. Tang, and J. Strobl. 2020. "Deep Learning-Based Approach for Landform Classification from Integrated Data Sources of Digital Elevation Model and Imagery." *Geomorphology* 354:107045. https://doi.org/10.1016/j.geomorph.2020.107045.

Li, X., L. Lei, Y. Sun, M. Li, and G. Kuang. 2021. "Collaborative Attention-Based Heterogeneous Gated Fusion Network for Land Cover Classification." *IEEE Transactions on Geoscience and Remote Sensing* 59 (5): 3829–3845. https://doi.org/10.1109/TGRS.2020.3015389.

Marcer, M. 2020. "Rock Glaciers Automatic Mapping Using Optical Imagery and Convolutional Neural Networks." *Permafrost and Periglacial Processes* 31 (4): 561–566. https://doi.org/10.1002/ppp.2076.

Menzies, J., and M. Ross. 2022. "Glacial Processes and Landforms—Transport and Deposition." In *Treatise on Geomorphology*, 182–202. Elsevier. https://doi.org/10.1016/B978-0-12-818234-5.00027-4.

Nair, V., and G. E. Hinton. 2010. "Rectified Linear Units Improve Restricted Boltzmann Machines." In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, 807–814. Haifa.

Palafox, L. F., C. W. Hamilton, S. P. Scheidt, and A. M. Alvarez. 2017. "Automated Detection of Geological Landforms on Mars Using Convolutional Neural Networks." *Computers & Geosciences* 101:48–56. https://doi.org/10.1016/j.cageo.2016.12.015.

Prakash, C., and R. Nagarajan. 2017. "Outburst Susceptibility Assessment of Moraine-Dammed Lakes in Western Himalaya Using an Analytic Hierarchy Process." *Earth Surface Processes and Landforms* 42 (14): 2306–2321. https://doi.org/10.1002/esp.4185.

Rastner, P., T. Bolch, C. Notarnicola, and F. Paul. 2014. "A Comparison of Pixel- and Object-Based Glacier Classification with Optical Satellite Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (3): 853–862. https://doi.org/10.1109/JSTARS.2013.2274668.

Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and F. Prabhat. 2019. "Deep Learning and Process Understanding for Data-Driven Earth System Science." *Nature* 566 (7743): 195–204. https://doi.org/10.1038/s41586-019-0912-1.

RGI Consortium. 2017. "Randolph Glacier Inventory - a Dataset of Global Glacier Outlines, Version 6." https://doi.org/10.7265/4M1F-GD79.

Rupnik, E., M. Daakir, and M. Pierrot Deseilligny. 2017. "MicMac – a Free, Open-Source Solution for Photogrammetry." *Open Geospatial Data, Software and Standards* 2 (1): 14. https://doi.org/10.1186/s40965-017-0027-2.

Saha, S., L. A. Owen, E. N. Orr, and M. W. Caffee. 2019. "High-Frequency Holocene Glacier Fluctuations in the Himalayan-Tibetan Orogen." *Quaternary Science Reviews* 220:372–400. https://doi.org/10.1016/j.quascirev.2019.07.021.

Sahu, A. S. 2014. "Identification and Mapping of the Water-Logged Areas in Purba Medinipur Part of Keleghai River Basin, India: RS and GIS Methods." *International Journal of Advanced Geosciences* 2. https://doi.org/10.14419/ijag.v2i2.2452.

Sainte Fare Garnot, V., L. Landrieu, and N. Chehata. 2022. "Multi-Modal Temporal Attention Models for Crop Mapping from Satellite Time Series." *ISPRS Journal of Photogrammetry and Remote Sensing* 187:294–305. https://doi.org/10.1016/j.isprsjprs.2022.03.012.

Scherler, D., B. Bookhagen, and M. Strecker. 2011. "Spatially Variable Response of Himalayan Glaciers to Climate Change Affected by Debris Cover." *Nature Geoscience* 4:156–159. https://doi.org/10.1038/ngeo1068.

Sebastianelli, A., M. P. Del Rosso, P. P. Mathieu, and S. L. Ullo. 2021. "Paradigm Selection for Data Fusion of SAR and Multispectral Sentinel Data Applied to Land-Cover Classification." *arXiv E-Prints*. https://doi.org/10.48550/arXiv.2106.11056.

Shumack, S., P. Hesse, and W. Farebrother. 2020. "Deep Learning for Dune Pattern Mapping with the AW3D30 Global Surface Model." *Earth Surface Processes and Landforms* 45 (11): 2417–2431. https://doi.org/10.1002/esp.4888.

Silburt, A., M. Ali-Dib, C. Zhu, A. Jackson, D. Valencia, Y. Kissin, D. Tamayo, and K. Menou. 2019. "Lunar Crater Identification via Deep Learning." *Icarus* 317:27–38. https://doi.org/10.1016/j.icarus.2018.06.022.

Smith, M., P. Paron, and J. Griffiths. 2011. *Geomorphological Mapping Methods and Applications*. Amsterdam: Elsevier Science.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." *Journal of Machine Learning Research* 15 (56): 1929–1958.

Sun, Z., L. Sandoval, R. Crystal-Ornelas, S. M. Mousavi, J. Wang, C. Lin, N. Cristea, et al. 2022. "A Review of Earth Artificial Intelligence." *Computers & Geosciences* 159:105034. https://doi.org/10.1016/j.cageo.2022.105034.

Torres, R. N., P. Fraternali, F. Milani, and D. Frajberg. 2020. "Mountain Summit Detection with Deep Learning: Evaluation and Comparison with Heuristic Methods." *Applied Geomatics* 12 (2): 225–246. https://doi.org/10.1007/s12518-019-00295-2.

van der Meij, W. M., E. W. Meijles, D. Marcos, T. T. L. Harkema, J. H. J. Candel, and G. J. Maas. 2022. "Comparing Geomorphological Maps Made Manually

and by Deep Learning." *Earth Surface Processes and Landforms* 47 (4): 1089–1107. https://doi.org/10.1002/esp.5305.

Wang, S., X. Shen, M.-L. Chevalier, A. Replumaz, Y. Zheng, H. Li, J. Pan, K. Li, and X. Xu. 2022. "Illite K-Ar and (U-Th)/He Low-Temperature Thermochronology Reveal Onset Timing of Yadong-Gulu Rift in Southern Tibetan Plateau." *Frontiers in Earth Science* 10:993796. https://doi.org/10.3389/feart.2022.993796.

Wernette, P., C. Houser, and M. P. Bishop. 2016. "An Automated Approach for Extracting Barrier Island Morphology from Digital Elevation Models." *Geomorphology* 262:1–7. https://doi.org/10.1016/j.geomorph.2016.02.024.

Wilson, J. P., and J. C. Gallant. 2000. "Primary Topographic Attributes." In *Terrain Analysis: Principles and Applications*, edited by J. P. Wilson and J. C. Gallant, 51–85. New York: John Wiley & Sons.

Wobus, C., G. Tucker, and R. Anderson. 2010. "Does Climate Change Create Distinctive Patterns of Landscape Incision." *Journal of Geophysical Research: Atmospheres* 115 (F4): 115. https://doi.org/10.1029/2009JF001562.

Xie, Z., U. K. Haritashya, V. K. Asari, B. W. Young, M. P. Bishop, and J. S. Kargel. 2020. "GlacierNet: A Deep-Learning Approach for Debris-Covered Glacier Mapping." *Institute of Electrical and Electronics Engineers Access* 8:83495–83510. https://doi.org/10.1109/ACCESS.2020.2991187.

Zhang, L., L. Zhang, and B. Du. 2016. "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art." *IEEE Geoscience and Remote Sensing Magazine* 4 (2): 22–40. https://doi.org/10.1109/MGRS.2016.2540798.

Zuo, J., Z. Wu, G. Ha, M. Hu, C. Zhou, and H. Gai. 2021. "Spatial Variation of Nearly NS-Trending Normal Faulting in the Southern Yadong-Gulu Rift, Tibet: New Constraints from the Chongba Yumtso Fault, Duoqing Co Graben." *Journal of Structural Geology* 144:104256. https://doi.org/10.1016/j.jsg.2020.104256.