Applied Computing and Geosciences 16 (2022) 100093

Contents lists available at ScienceDirect

Applied Computing and Geosciences

journal homepage: www.sciencedirect.com/journal/applied-computing-and-geosciences

Machine learning – An approach for consistent rock glacier mapping and inventorying – Example of Austria

Georg H. Erharter [a,*], Thomas Wagner [b], Gerfried Winkler [b], Thomas Marcher [a]

[a] Institute of Rock Mechanics and Tunnelling, NAWI Graz Geocenter, Graz University of Technology Rechbauerstraße 12, 8010, Graz, Austria
[b] Institute for Earth Sciences, NAWI Graz Geocenter, University of Graz, Heinrichstrasse 26, 8010, Graz, Austria

ARTICLE INFO

Keywords:
Rock glacier inventory
Permafrost
Hydrological catchment
Digital mapping
Machine learning
Image segmentation

ABSTRACT

Rock glaciers (RG) are landforms that occur in high latitudes or elevations and — in their active state — consist of a mixture of rock debris and ice. Despite serving as a form of groundwater storage, they are an indicator for the occurrence of (former) permafrost and therefore carry significance in the research for the ongoing climate change. For these reasons, the past years have shown rising interest in the establishment of RG inventories to investigate the extent of permafrost and quantify water storages. Creating these inventories, however, usually involves manual, laborious, and subjective mapping of the landforms based on aerial image - and digital elevation model analysis. We propose an approach for RG mapping based on supervised machine learning which can help to increase the mapping efficiency and permits rapid RG mapping in vast and not yet covered areas. We found deep convolutional artificial neural networks (ANN) that are specifically designed for image segmentation (U-Net architecture) to be well suited for this classification problem. The general workflow consists of training the ANNs with orthophotos and slope maps of digital elevation models as input. The output (RG label-maps) is derived from a recently published RG inventory of the Austrian Alps that features 5769 individual RGs and was compiled manually by several scientists. To increase the generalization capabilities, we use live data augmentation during training. Based on this inventory, the ANNs have learned the average expert opinion and the RG map generated by the ANN can be used to increase the consistency and completeness of already existing RG inventories. Moreover, this ANN approach might be valuable for other landform mapping tasks beyond rock glaciers (e.g., other mass movements).

1. Introduction

Rock glaciers (RG) as one of the most-prominent landforms that indicate (former) permafrost conditions were of interest since the early 1960s starting with Wahrhaftig and Cox (1959). An established definition is given by Barsch (1996), more recently by Berthling (2011) and preliminarily standardized in a current project (2018–2023) of the IPA Action Group "Rock glacier inventories and kinematics" (RGIK 2022). Active RGs are tongue-shaped or lobate creeping landforms of perennially frozen debris material and move downslope slowly, thereby often forming transversal ridges and furrows on their lower part and longitudinal ridges on their upper part (e.g., Berthling (2011). Movement may stop due to more gradual terrain, cease of debris-supply or atmospheric warming (ice melt/permafrost thaw) and RGs become inactive. Both active and inactive RGs contain permafrost ice and are historically summarized as intact RGs (e.g. Barsch 1996). When finally, all permafrost ice within the landform completely vanishes, the RG becomes relict. Typical features of a relict RG are extensive vegetation cover and collapse structures. Nevertheless, a clear distinction between relict and intact rock glaciers is not straightforward (e.g., Colucci et al. (2019)). A refinement of the rock glacier activity categories was recently provided by the IPA Action Group "Rock glacier inventories and kinematics" (RGIK 2022).

RGs may serve as palaeo-climate indicators if found in a relict state (e.g., Frauenfelder et al. (2001); Moran et al. (2016)) or may even pose potential natural hazards in their intact state as changes in creeping velocities or disintegration due to permafrost thawing might trigger rock fall or debris flow (e.g., Delaloye et al. (2012); Haeberli et al. (2017)). Climate change and its reflection in a change in mountain permafrost drew upon rock glacier evolution (Haeberli et al., 1993; Kääb et al., 2007). In more recent years the attention was drawn on the importance of RGs for the hydrology of alpine headwaters as highlighted by e.g.,

* Corresponding author.
E-mail address: erharter@tugraz.at (G.H. Erharter).

https://doi.org/10.1016/j.acags.2022.100093
Received 8 September 2021; Received in revised form 10 July 2022; Accepted 21 July 2022
Available online 13 August 2022

Hayashi (2020); Jones et al. (2018a) and Jones et al. (2019) at a global scale. On the one hand, RGs may store water as permafrost ice in intact RGs (e.g., Jones et al. (2018b); Winkler et al. (2018); Pruessner et al. (2021)) and on the other hand they generally store liquid water as groundwater because RGs are considered to be shallow alpine aquifers (e.g., Hayashi (2020); Wagner et al. (2020a); Wagner et al. (2021)). Thus, regional recognition of these landforms yielding into inventories became more and more popular due to available high-resolution digital data. A first globally compiled inventory was established by Jones et al. (2018a). However, the inventories lack on consistent data quality as they are mainly mapped manually (e.g., Guglielmin and Smiraglia (1998); Rangecroft et al. (2014); Wagner et al. (2020b)) and therefore influenced by the personal experience of the mapping scientist. A recent contribution by Brardinoni et al. (2019) showed that, while the front of a RG is usually easily identifiable, the upper side of it – where it transitions into other morphologies (i.e., the rooting zone) – is ambiguous and multiple mappings by different researches would yield different RG boundaries. Therefore, existing approaches of manual delineation of such landforms are subjective (e.g., Brardinoni et al. (2019)) and very often not directly comparable. Furthermore, larger RG inventories usually are broken down into smaller areas which are then mapped by different operators. Although there are ongoing developments that will provide standardized mapping guidelines, individual humans naturally apply their own mapping policies which are different from one another due to different levels of experience and general cognitive differences between human beings. Consequently, human mapping based RG inventories are inherently inconsistent.

To account for some of the above-mentioned challenges of RG mapping we present an RG mapping approach that is based on state-of-the-art machine learning (ML) algorithms. Boosted by easy access to great computational power and the rise in more easily available high-resolution (digital elevation) datasets, the past decade has shown big advances in the field of ML based object detection and image segmentation in general (e.g., Krizhevsky et al. (2012)) and also for general geomorphological topics: Marmion et al. (2008); Stumpf and Kerle (2011); Giaccone et al. (2021). Breaking RG mapping down to a basic problem of semantic image segmentation (i.e., pixelwise classification of an image and thus segmenting the image into semantically coherent parts), we present a supervised ML approach, where we train artificial neural networks (ANN) to map RGs based on an existing inventory (Wagner et al., 2020b) and thus combine multiple expert opinions in one trained algorithm.

Several authors have already presented similar approaches in the context of ML based processing of permafrost related landforms: Deluigi et al. (2017) used non-deep ML classifiers (e.g., logistic regression, support vector machine) to map areas with permafrost; Kofler et al. (2020) also used non-deep ML classifiers to estimate the state of a RG (i.e., intact vs. relict); Marcer (2020) used a simple convolutional neural network (CNN; i.e., a type of deep ANN) to map RGs in a sliding window approach; Robson et al. (2020) also used a CNN to map RGs in an image segmentation approach. While Deluigi et al. (2017) and Kofler et al. (2020) do not address RG mapping directly, our approach differs methodologically from Marcer (2020) and Robson et al. (2020). These two approaches both used simple CNNs with either a binary output (i.e., RG yes or no) for the whole input image (Marcer 2020) or an image segmentation approach without a specialized CNN architecture (Robson et al., 2020).

The overall goal of this study was to develop a framework for RG mapping support that is reproducible and universally applicable and not biased towards specific altitudes or aspects/orientations of the landforms. Therefore, the specific aims of this study are (i) to apply ML based image segmentation for identifying and delineating RGs and validate the ML capability for this approach; (ii) to provide a trained algorithm that might be used on other regions of the world to look for RG suspected landforms. The study provides an approach towards consistent RG mapping in the sense of applying a "consistent mapping policy to an area via one single entity" as opposed to applying multiple different mapping policies to an area via multiple human experts. The recently developed consistent rock galcier inventory of Austria (Wagner et al. 2020b, 2020c) was used as the ANN training data set for this study and based on that the entire Austrian Alps were scanned by the ANN and further compard to the existing inventory. It has to be explicitly mentioned that the ANN algorithm used herein is not considered to replace the manual delineation by experts but to be a helpful assitance.

2. Machine learning pipeline

In ML the process from raw data to final classification or prediction is often referred to as a "pipeline" and an overview of this study's pipelines is given in Fig. 1. Within this pipeline two main phases can be separated: the training phase and the application phase. The training phase starts with selection of suitable raw data consisting of digital elevation models (DEM), orthophotos and an associated RG inventory (section 2.1). In a second step, the raw data needs to be preprocessed into input and output data for an ANN (section 2.2). The ML algorithm (in this case an ANN with the U-Net architecture) must be designed (section 2.3) and trained on the given dataset (section 2.4). The training phase ends with an evaluation of results (section 2.5) and the trained ANN can be used for further application.

In the application phase, new input (DEM data and orthophotos) of not yet mapped areas and/or areas with a previously existing RG inventory that needs revision must be collected and preprocessed in the same way as in the training phase. The previously trained ANN can then be used to generate a probability map that shows areas with potential RGs. This RG probability map can then be used as an assistance to develop a completely new RG inventory of that area or to revise a previous RG inventory. The dashed arrow in Fig. 1 that connects the "Output and Further Processing" of the application phase to the "Raw Data" of the training phase illustrates that the whole ML pipeline can be used in multiple cycles to "fine-tune" an already existing RG inventory (see also the conclusion and outlook in section 5).

2.1. Data base

The used data base consists of digital elevation models (DEM) and orthophotos as input and the existing RG inventory of the Austrian Alps (Wagner et al. 2020b, 2020c) as output, respectively the source of target labels. The DEMs with a ground resolution of 1 × 1 m are based on airborne laser scanning data and fully cover the Austrian federal states of Vorarlberg, Tyrol, Salzburg, Carinthia, Styria and the alpine, southern part of Upper Austria. For this study we used DEMs that were provided directly by the federal governments of the above given Federal states of Austria. References to DEM data can be found in the data availability section at the end of the paper.

Google Maps satellite images were used as orthophotos and loaded as XYZ tiles in the open geographic information system (GIS) QGIS (QGIS.org, 2021). Although orthophotos of higher quality and resolution are available in Austria, we see Google Maps satellite images as sufficient for this study and they also have the benefit of being a globally available data source.

The inventory consists of 5769 RGs covering a total area of around 303 km^2 and includes polygons that delineate the RGs' geometry as well as a set of selected attributes. The inventory is a compilation and extended version of several previously published RG inventories (Kellerer-Pirklbauer et al., 2012; Lieb et al.; Krainer and Ribis 2012; Stocker and Krainer 2011; Stocker 2012b, 2012a) and is seen as the first consistent RG inventory of entire Austria as potential RG suspected landforms in other federal states are thought to be rare (Upper Austria) or nonexistent (Vienna, Burgenland, Lower Austria).

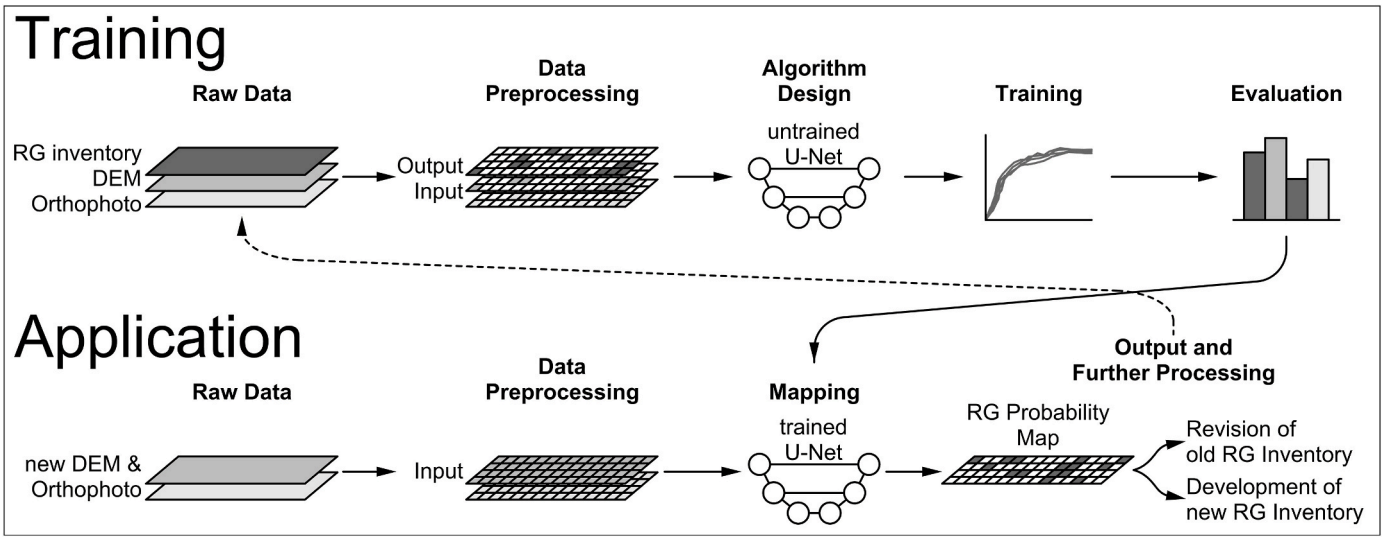

Fig. 1. Schematic overview/flow chart of the Machine Learning pipeline for rock glacier mapping as used in this study. RG = rock glacier, DEM = digital elevation model, U-Net = artificial neural network of this study.

2.2. Dataset preparation

Supervised ML requires dedicated sets of input and output data for the algorithms to be trained on. Input will be denoted as X and output as y in the following. As a first step of dataset preparation these $X - y$ pairs needed to be derived from the above-described data. We therefore placed a 1 × 1 km grid over Austria and each cell with already mapped RGs is taken as one sample for the training dataset. After initially experimenting with 1 × 1 km cells that are centered on individual RGs, we chose a fixed grid as this teaches the algorithm to expect RGs in all positions of the sample and not only at its center. In combination with the data augmentation techniques that are used during training (see section 2.4) we found this approach very promising as it makes the algorithms more robust towards variable positions and orientations of the RGs. Fig. 2 gives a visual summary of the described approach for deriving X and y.

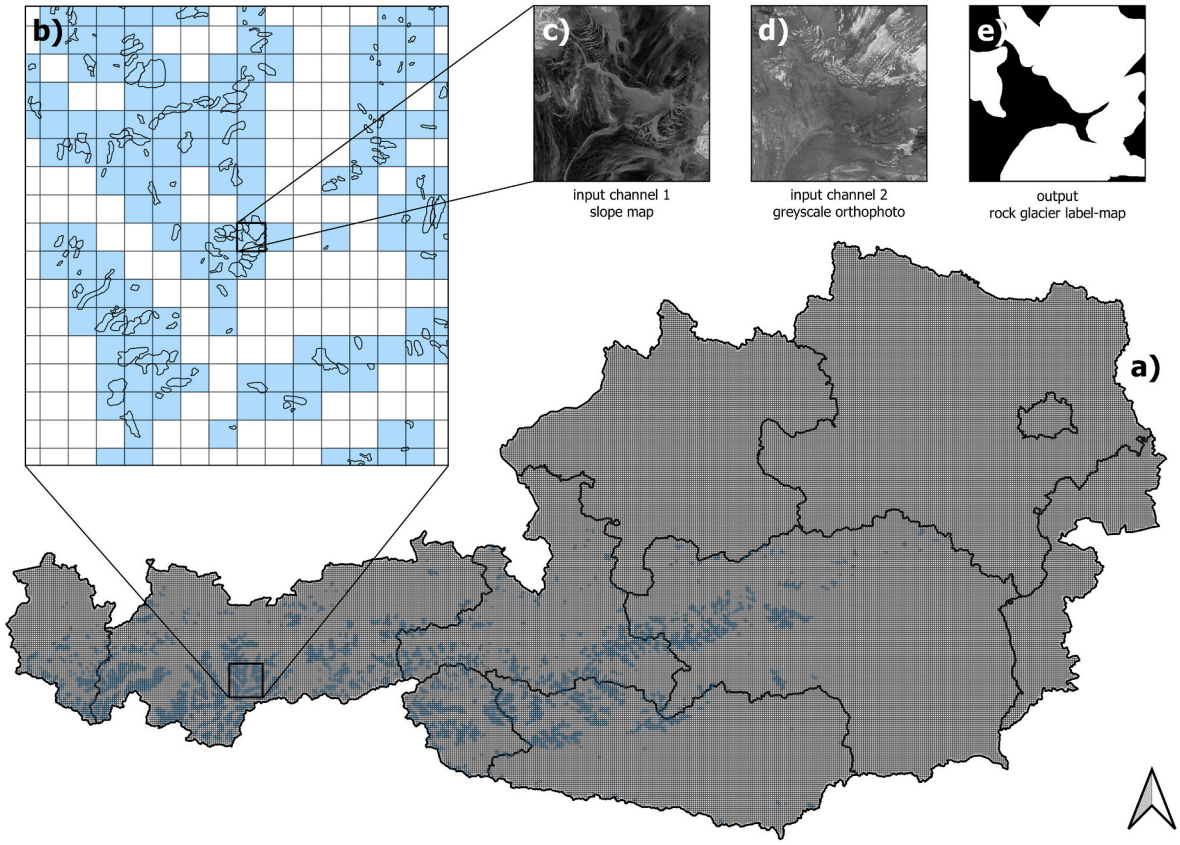

Fig. 2. Map view of the process of dataset preparation: (a). a 1 × 1 km grid is placed over Austria; (b). The cells (1 × 1 km) with already mapped rock glaciers are selected for ANN training and validation (indicated in blue); (c) slope maps of the 1 m resolution digital elevation model and (d) greyscale orthophotos are saved as input for the ANNs and (e) rock glacier label-maps as output/target labels for the training data.

Specifically, each element of the ANN input - X - consists of a 512 × 512 x 2 array. The 512 × 512 pixel size of the arrays was chosen as a trade-off between sufficiently high resolution input/output and computational effort (higher resolution, longer computational time). The two layers/bands/channels of this array can be thought of like a multi band image and will be referred to as "channels" in the following. The first input channel consists of a slope map representation of the DEM and the second input channel of a grayscale orthophoto. Both are resampled to the 512 × 512 pixel image size and thus the two input channels have an approximate resolution of 2 m. The output consists of a label-map which in this case is a binary map showing areas with RGs in white (label = 1) and areas without RGs in black (label = 0). Fig. 2 shows an example of X where the slope map and the greyscale orthophoto of the input are shown as well as the corresponding output – y – which includes multiple large RGs.

The slope map of the input was computed with the program QGIS based on the above mentioned DEMs. In a slope map each raster cell contains information about its slope angle (in degrees) with respect to the neighboring cells. We use the derived slope map as an input, as several studies in the context of slope morphologies (e.g., Liu et al. (2020); Lee et al. (2020)) have shown that the slope inclination is one of the most informative parameters for ML algorithms. Furthermore, using slope data instead of DEMs with absolute heights enables the ML algorithm to identify RGs at all altitudes simply based on their morphological features and there is no risk of the algorithm to become biased towards certain predominant heights of the training dataset. For the same reason we did not use DEM representations like the slope aspect (i.e., the direction into which a slope faces) as the RGs in the training dataset are not equally spread across all orientations (Wagner et al., 2020b) which again would introduce bias in the training.

As given above, we use grayscale orthophotos for the second channel of the input with the idea to give the algorithm additional information about the landscape's surface and vegetation characteristics as they are not well represented in the DEM – slope map representation. We only used one channel of greyscale orthophoto instead of all three R, G, B channels as this is computationally more efficient while still providing the necessary surface information. Building on the above-mentioned studies that also use supervised ML for classification of different morphological features, no tests with various input channels were performed to determine the most informative one. Especially the slope maps provide sufficient morphological information to identify RGs and the orthophotos provide additional information about the land surface.

The output – y – is a 512 × 512 binary raster/label-map containing the information whether a pixel shows a RG or not. The raster was directly derived from the existing RG inventory with the program QGIS (Wagner et al. 2020b, 2020c).

To set up the final training dataset, a split of X and y into two subsets is necessary (Raschka and Mirjalili 2019; Chollet 2018). The main subset X_{train} and y_{train} is then used for training the algorithm and a number of random samples is set aside as to be able to objectively validate the trained algorithm's performance: X_{val} and y_{val}. With the above-described approach, we could generate 4522 input-output samples. Since several RGs can appear in the same cell, the number of samples is lower than the total number of RGs in the inventory (i.e., 5769 (see for example the sample of Fig. 2 (c – e) where several RGs appear in identical grid cells). Of the 4522 samples, 800 were randomly selected for validation which leaves 3722 for training. Using 800 samples for validation was determined by trial and error in the experimental phase of the study. While training ANNs multiple times on different subsets of the data (i.e., different splits into training and validation data; also known as "cross validation") we found that a validation dataset size of 800 on the one hand yields a representative sample where most RG-morphologies are well covered while on the other hand the number of samples for training is not too diminished.

Due to the observation that a peculiar number of false positives is associated with several landforms that are not represented in the original dataset (e.g., some anthropogenic structures, hills etc. see section 3) we added 70 manually selected negative samples with no RGs (i.e., y consists of 0 only) to the training data. In summary 3792 samples were used for training and 800 for validation, leading to a ratio of ~79%: 21% and 4592 samples in total. This ratio of training to validation samples is also based on the authors' experience that a share of 20%–30% test data leads to sufficient results, provided that the validation samples are well representative for the whole dataset (Erharter et al. 2019, 2021).

2.3. Artificial neural network

The chosen ANN architecture is based on the "U-Net" ANN architecture after (Ronneberger et al., 2015). U-Net is a type of deep convolutional ANN (Lecun et al., 1998; Hubel and Wiesel 1963; Krizhevsky et al., 2012) with a contracting and expanding path/an encoder and decoder part of the ANN. U-Net has shown state of the art performance on various image segmentation tasks (Wang et al., 2021; Wu et al., 2021; Maxwell et al., 2020) and a graphical representation of our implementation of the U-Net ANN architecture is given in Fig. 3.

The U-Net of our implementation consists of five contracting- and five expanding blocks. Implementation was done with the Python library Keras (Chollet, 2015) with a tensorflow-gpu backend (Abadi et al., 2015). Information on the source code of the used U-Net implementation is given in the Computer code availability section at the end of the paper.

One contracting block consists of two 2D convolution layers (conv2d) with a 3 × 3 kernel size and a stride of 1. Zero padding is applied (i.e., setting: padding = "same", see Keras documentation) and rectified linear unit (ReLU) activation functions (eq. (1)) are used (Hahnloser et al., 2000).

$$f(x) = \max(0, x) \tag{1}$$

Each conv2d layer is followed by a layer of batch normalization. To reduce the input size, a maximum pooling layer with a window size of 2 × 2 and a stride of 2 × 2 is applied in the end of each contracting block. Each contracting block yields two outputs where the residuals of the last conv2d layer are passed on to the corresponding expanding block and the compressed input after the max pooling is passed down to the next contracting block/center of the ANN. In the five contracting blocks, the numbers of filters in the conv2d layers are set to 32, 32, 64, 128 and 256. Due to the max. pooling the dimensions of the input images are reduced by 50% in each block which leads to a size of 32 × 32 pixels after five contracting blocks and initial image sizes of 512 × 512.

Each expanding block receives two inputs, where one is the residual of a corresponding contracting block and one is the output of a previous expanding block or the ANN's center. As a first step, the output of the previous block is upsampled by two (i.e., the input size is doubled) so that the size of the previous block's output matches the size of the incoming residual. Then the residual and the upsampled input are concatenated/merged along the third axis and then fed into a stack of three conv2D layers each followed by batch normalization (identical settings as in the contracting blocks). Going from center to output, the number of filters in the conv2d layers of the expanding blocks is 256, 128, 64, 32 and 32 and the size of the images increases from 32 × 32 to 512 × 512 pixels.

The center part of the U-Net consists of two conv2d layers with 512 filters each, a kernel size of 3 × 3, "same" padding and ReLU activation functions. Each conv2d layer of the center is followed by a layer of batch normalization. The final result of the U-Net is generated in one last conv2d in the end of the ANN with a kernel size of 1 × 1 that applies sigmoid activation functions (eq. (2)), thus creating an output that ranges from 0 (no RG) to 1 (RG).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Inspections of preliminary results during the development phase of

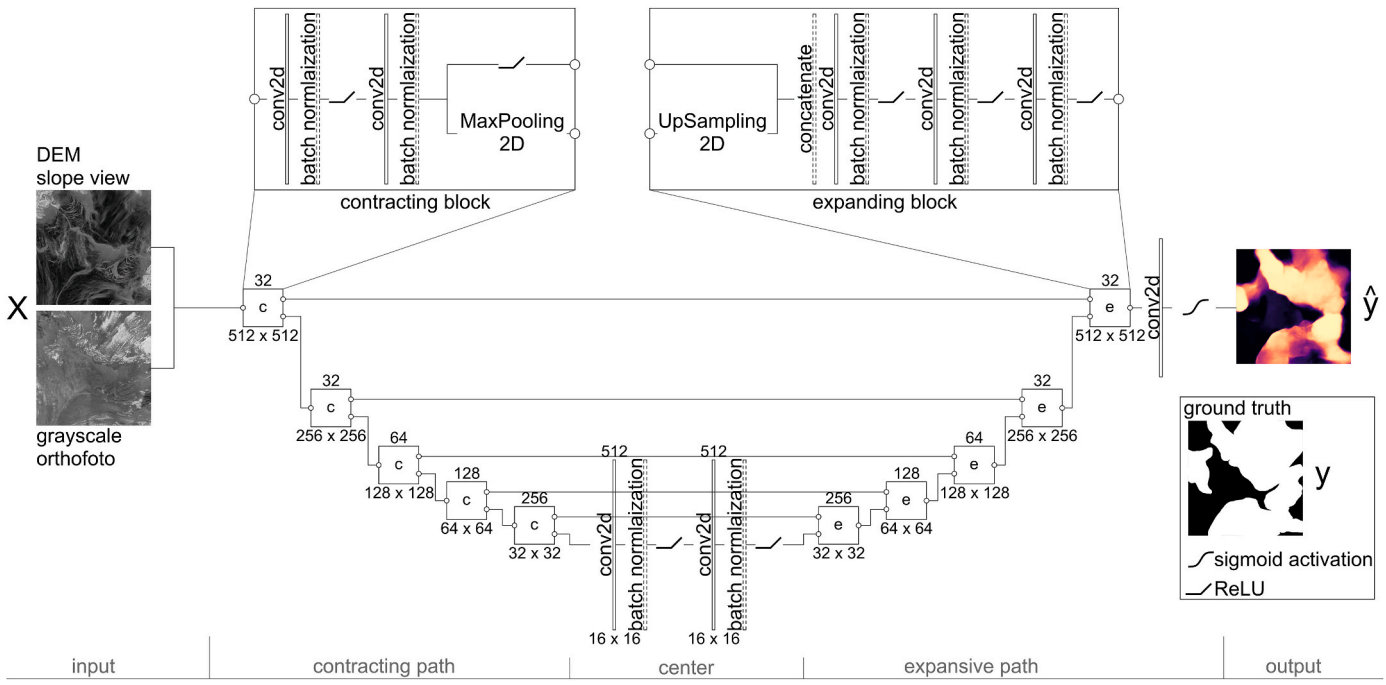

Fig. 3. Graphical representation of the chosen U-Net artificial neural network architecture. ReLU = Rectified linear unit activation function, conv2d = 2D convolution layer; X = input, y = ground truth output, $\hat{y}$ = ANN output.

this study showed considerable variation in individual ANNs' output with respect to how well different morphologies can be detected. As instability is a well-known issue with ANNs (Cunningham et al., 2000; Dietterich 2000) we trained 10 individual ANNs with identical U-Net architecture to compensate for that. The final result is thus created as an unweighted average of all ten individual ANNs, i.e., the arithmetic mean of all ten ANN outputs. Visual result inspection showed no significant improvement when we used a weighted average based on the individual ANNs' performance.

2.4. ANN training

The ten individual ANNs were trained with an "Adam" optimizer (Kingma and Ba 2014) with a learning rate of 0.0001 and no learning rate – decay. To assess the ANNs' performance we chose the Sørensen–Dice coefficient or also called dice similarity coefficient (*DSC*; (Dice 1945; Sørensen 1948)) as this metric has shown good performance for image segmentation tasks before (e.g., Wang et al. (2021); Maxwell et al. (2020)). The *DSC* is computed as follows (eq. (3))

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \tag{3}$$

where A and B are two sets of elements and $|A|$ and $|B|$ are their numbers of elements. The *DSC* therefore divides twice the number of elements that are common in both sets by the total number of elements. A *DSC* of 0 refers to a complete dissimilarity between A and B and a *DSC* of 1 to perfect similarity. We use the *DSC* not only for general performance assessment but also as the loss function for the ANNs by combining the standard binary cross entropy with the "dice loss" (i.e., $1 - DSC$).

During training we did not aim for a dedicated number of epochs but used an "early stopping" mechanism to abort training once the performance did not increase anymore for the test dataset. Specifically, training was aborted once 8 consecutive epochs without an increasing *DSC* of the test-dataset were recorded and only the best performing models are saved. Fig. 4 shows *DSC* records for the training- and testing datasets of all ANNs. Most ANNs reached their peak performance with respect to the test dataset after around 10 epochs and due to the almost

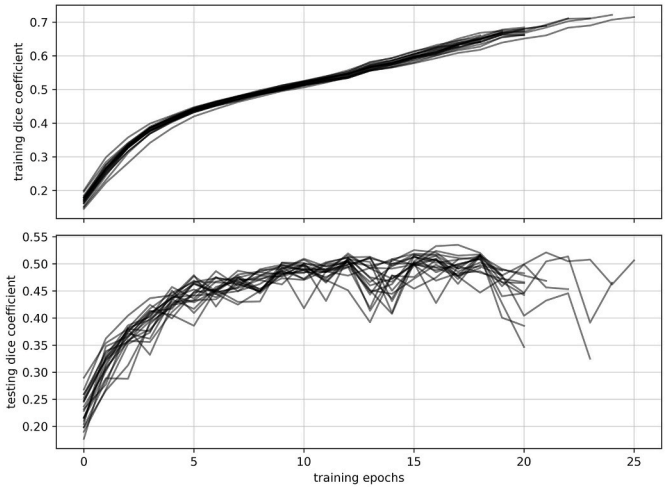

Fig. 4. Recorded dice similarity coefficients for the training and test datasets during ANN training.

linearly increasing *DSC* on the training dataset it can be assumed that further training would lead to overfitting.

As given in section 2.2 we used a grid-based approach to create pairs of input and output data as to have the ANNs learn that RGs can appear in all parts of the input. To further enhance this effect, we used live data augmentation during the training to increase the ANNs' general robustness. Training data augmentation (i.e., distortions of input-output pairs of the training data) has shown to successfully improve algorithms' performance as it provides a possibility to artificially increase the number of training samples and increase a dataset's diversity (Shorten and Khoshgoftaar 2019). The augmentation process consists of horizontal and vertical flipping of the samples, random zoom ranges between 100% and 200% magnification of the original size and random ± 90°, 180° and 270° rotations of the samples.

2.5. Evaluation

The averaged output of the ANNs is a continuous range of probability values between 0 (no RG) and 1 (RG) (see section 2.3). While this alone is a valuable information and the direct output is used to detect not yet found RGs, it is not well suited for a direct comparison to the existing Austrian RG inventory. To perform a direct comparison, a probability threshold must be set that allows for a discretization of the output. By testing the range of thresholds between 0 and 1 in steps of 0.05, we found a value of 0.4 as best suited to discretize the results into 2 classes: values ≤ 0.4 = no RG, values > 0.4 = RG.

The diagram of Fig. 5 shows different thresholds plotted against the corresponding *DSC* and the respective values of true positives, false positives, true negatives and false negatives. Lowering the threshold goes along with an increasing amount of true and false positives but also with a decreasing number of true and false negatives. Raising the threshold leads to the opposite trend. A maximum *DSC* of 0.616 is reached at a threshold of 0.4.

In Fig. 6 an example is given where the RG probabilities of the ANN output are shown in a discrete colormap at the example of three RGs in the North Tyrolean "Wurmeskar". It can be seen that a high probability threshold of – for example – 0.8 would lead to small RG outlines and thus a low amount of true and false positives and at the same time a comparably high amount of true and false negatives. A low threshold of 0.2 would lead to very large RG outlines and thus a high amount of true and false positives and a comparably low amount of true and false negatives. The numbers within the green outlines show how much percent of the area of the RGs according to the Austrian RG inventory would have been "properly" detected with the proposed threshold of 0.4. While inspecting these results, one has to keep in mind that also the mapped RG outlines according to the Austrian RG inventory are not the one and only truth (Brardinoni et al., 2019) and might be discussed themselves.

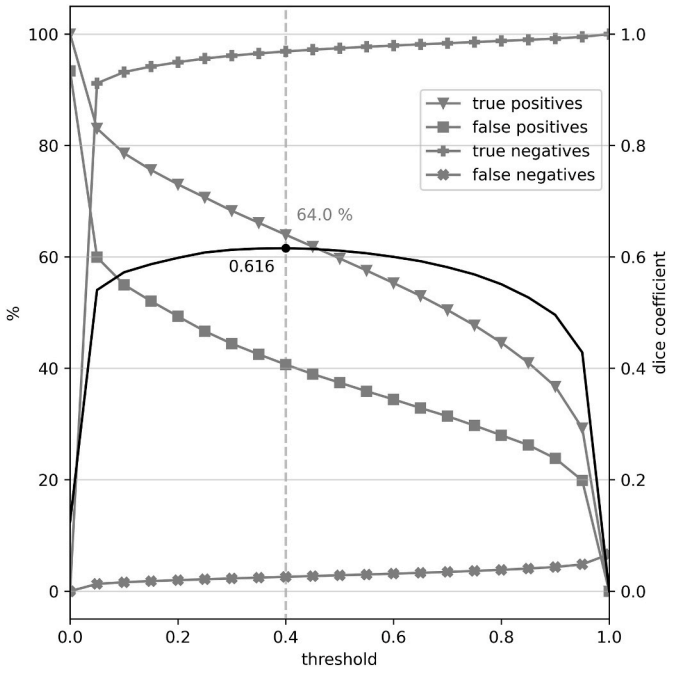

Fig. 5. Different values of probability thresholds plotted against the percentage of: true positives (grey line with triangles), false positives (grey line with squares), true negatives (grey line with pluses) and false negatives (grey line with bold X). The dice coefficient is given as a black line and the maximum of 0.616 at the threshold of 0.4 is marked.

3. Results

3.1. Direct output

As given in sections 2.2 and 2.3 the direct output of the ANNs consists of 1 × 1 km raster images with a resolution of 512 × 512 pixels with scalar values ranging from 0 to 1 where 0 indicated "no RG" and 1 "RG". For the sake of convenience and to facilitate further analysis the individual raster images were georeferenced in the coordinate system "MGI/Austria Lambert" (EPSG code: 31287) and then merged to 105 larger raster images that can be imported with standard GIS software. In Fig. 7, selected samples of not identified RGs (false negatives), falsely identified RGs (false positives), and correctly identified new RGs (true positives) are presented. To access the full RG – probability maps of the ANNs, we refer the reader to the raster images that are provided with this paper and can be retrieved from the repository given in the appendix at the end of this paper. With the probability threshold of 0.4 - as described in the previous chapter - a value of 64% of true positives was achieved. It can be observed that unrecognized RGs are often ones with an overall small area. A more detailed quantitative analysis of the ANNs' performance is given in section 3.2 and section 4.3 and Fig. 11 go into more detail on the relationship between RG size and detection rate.

In the first row of Fig. 7 (a to c), samples of RGs that were not detected by the ANNs are shown. In these images additionally percentages of the correctly detected area of the RGs are additionally given to better highlight the not detected ones (see next section for more information on this quantification). In Fig. 7a, three small RGs were hardly detected, but the big RG in the middle of the figure was well detected. RGs with generally weakly pronounced morphological features (e.g., Fig. 7b) or RGs in areas where the overall morphology has been anthropogenically modified (e.g., Fig. 7c) are also features that tend to remain undetected. In the case of Fig. 7c, the big RG is not detected due to anthropogenic overprint in form of a ski slope, however, parts of the front of that RG are still weakly detected and especially the two neighboring RGs to the South are detected.

In the second row of Fig. 7 (d to f), samples are shown that were falsely classified as RGs. One example would be karstic areas (e.g., Dachstein area, Fig. 7d) where bare rock occurs to be similar to the surface structures of RGs, which is however easily identifiable by experts. A false positive where a debris covered glacier is incorrectly detected as a RG can be found at the southern border of Vorarlberg (Fig. 7e). Unlike the karstic area such a morphology may be similar to "real" RGs and certain ambiguities will remain. In Fig. 7f, another example from southern Vorarlberg is given where a remnant of a lateral moraine deposit was detected as a RG on the right continuing at the top of Fig. 7f (dark colored landform) which is located next to an actual RG according to the current inventory. The moraine is strongly affected by erosion yielding into some gullies at the right top of the image. For the expert it may be clear that there is a wall structure remaining from the former glacier front, and behind this structure a depression is present and not a bulky main body as it would be the case for a RG. However, it needs to be noted that sometimes RGs may have developed out of moraine deposits and the transition is often ambiguous. In addition to the shown examples, further obvious false positives can be observed in connection to anthropogenic structures like quarries, dams, artificial ponds (e.g., in Fig. 7c) or landfills which often resemble the steep and curved frontal slope of RGs.

In the last row of Fig. 7 (g to i) candidates for newly detected RGs are given. In Fig. 7g a potential relict RG that developed out of a moraine/glacial deposit is depicted. It is located North of the existing RG "Futschölferner Ost" with intNr 60342. In Fig. 7h, a new RG is suggested in central East Tyrol in the cirque North of the three RGs with the intNrs 9589, 9587 and 9588. Furthermore, the ANNs indicate a downslope extension of the RG "Ochsenbichl W" (9587) which is also comprehensible from a morphological point of view. In Fig. 7i, yet another downslope extension of the already mapped RG "e Schobertoerl" (intNr 7441)

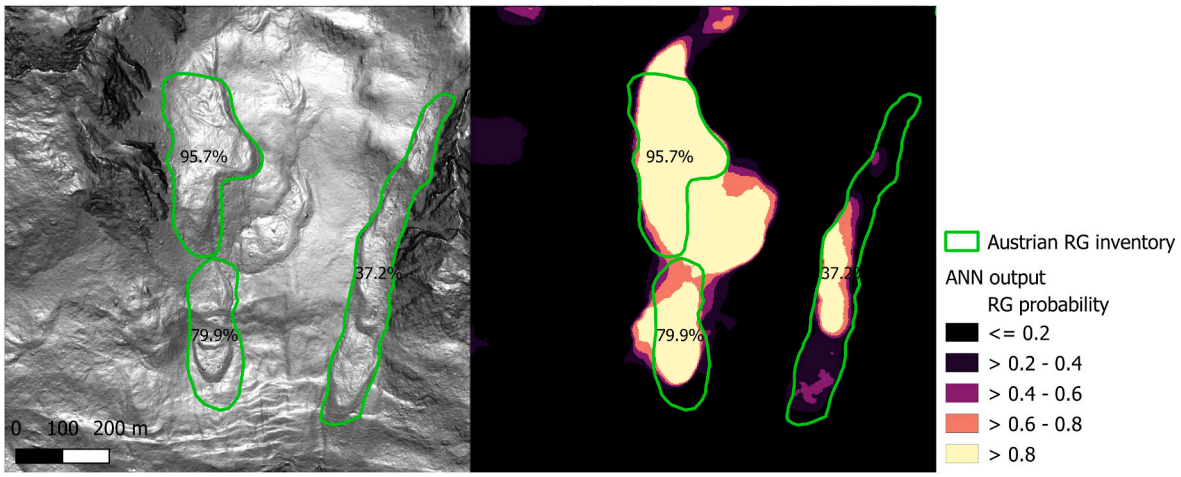

Fig. 6. Three exemplary RGs of the North Tyrolean "Wurmeskar" (left) and the corresponding ANN-based RG probability map with a discrete colormap and an interval of 0.2 (right). It can be seen that a lower threshold would lead to a higher detection rate, but also to more false positives.

is proposed which is also well visible in the slope view and the orthophoto. As opposed to the original Austrian RG inventory, this study also involved the southern part of Upper Austria that borders the Austrian Alps. In this area, there are also a few landforms that were detected as possible RGs and might need to be added to the inventory in a future version.

3.2. Discretized output

As given above, the discretized output enables a direct comparison to the existing Austrian RG inventory. Treating the existing inventory as the ground truth, the intersection of the polygons of the existing inventory with the discretized output enables an analysis of true positive, false positive, true negative and false negative detected areas. Fig. 5 shows how the choice of the probability threshold influences the amount of these areas. An example of the test dataset with the different areas highlighted is given in Fig. 8.

With the identified certainty threshold of 0.4 (see section 2.5) we were able to extract polygons that discretely delineate potential RGs identified by the ANNs (i.e., raster areas with a threshold >0.4). We then intersected the polygons of the existing RG inventory with the polygons derived from the discretized ANN output. By intersecting these two polygon layers, it is possible to determine exactly how many percent of each RG's area has been detected (denoted as A%). To be able to draw objective conclusions from this measure, we only analyzed RGs of the validation dataset that are fully enclosed within one cell, as to avoid possible corner effects. This leaves 538 of the total 800 RGs in the validation dataset. Of these 538 RGs, 119 (22%) have an A% between 0 and 10%, 282 (52%) have an A% of over 70% and 188 (35%) achieve an A% of over 90% and are thus almost completely mapped. A more detailed representation of this measure as well as an example is given in Fig. 9 and an interpretation of these results in section 4.3 where especially Fig. 11 concerns the detection rate of differently sized RGs. We nevertheless want to point out that this result is highly dependent on the choice of the previously described probability threshold. Lowering this threshold will directly influence A%, respectively the number of detected RGs.

4. Discussion

4.1. Comparison with existing approaches

In this study we presented a methodology where we use a specialized ANN architecture that was originally designed for semantic image classification for automatic RG mapping. In comparison to previous, similar studies from Robson et al. (2020) and Marcer (2020) which used simpler ANN frameworks, we also used a much larger RG inventory to train on. The applied U-Net/encoder – decoder type CNN shows state of the art performance in image segmentation and overall better results than standard CNNs for this task (Badrinarayanan et al., 2015; Ronneberger et al., 2015). U-Nets are also increasingly applied on other tasks including geoscientific problems (e.g., Samarin et al. (2020); Dirscherl et al. (2021); Wu et al. (2021)). As there is a broad consensus in the community that U-Nets provide state of the art performance in image segmentation over standard CNN implementations, the approaches from Marcer (2020) and Robson et al. (2020) were not directly replicated for this study.

As the image-segmentation based approach of this study yields a pixelwise RG-probability over the whole mapping area and not a delineation of individual RGs, it is not possible to tell how many new RGs were actually found. It is furthermore problematic to compare absolute numbers of RGs, since some operators delineate individual RGs within bigger clusters of RGs, which always would be detected as one big RG cluster by the presented approach. The analysis with the discretized output of section 3.2 permits however an assessment of how much area was successfully discovered with respect to the testing data. The result analysis furthermore showed that the ANN based mapping shows less certainty in mapping ambiguous parts of RGs like the rooting zone but does well in clearly defined areas like the front and pronounced internal lobes which shows the strong relation between the supervised ML algorithm and the human-made labels that it was trained on.

Considering the used input data of this study in comparison to previous ones, the higher input data resolution and usage of DEM data in the present study helps with this problem and observed difficulties with identification of small landscapes are likely to be related to inconsistencies in the training data. Marcer (2020) used grayscale spot 6 imagery featuring hindering effects such as snow cover, tree cover or excessively shadowed areas and small landforms are therefore discarded.

4.2. Reflection on data processing

In this study we trained ANNs to automatically map RGs based on openly accessible satellite images and digital elevation models. Whereas satellite based orthophotos are already almost globally available, publicly available 1×1 m DEMs are still rare, and this therefore limits the current applicability of the proposed method to countries with available high-resolution DEM data (preferably Lidar based and not satellite based). A follow-up study would be necessary to determine how lower DEM resolutions affect the accuracy of the trained ANNs. For example,

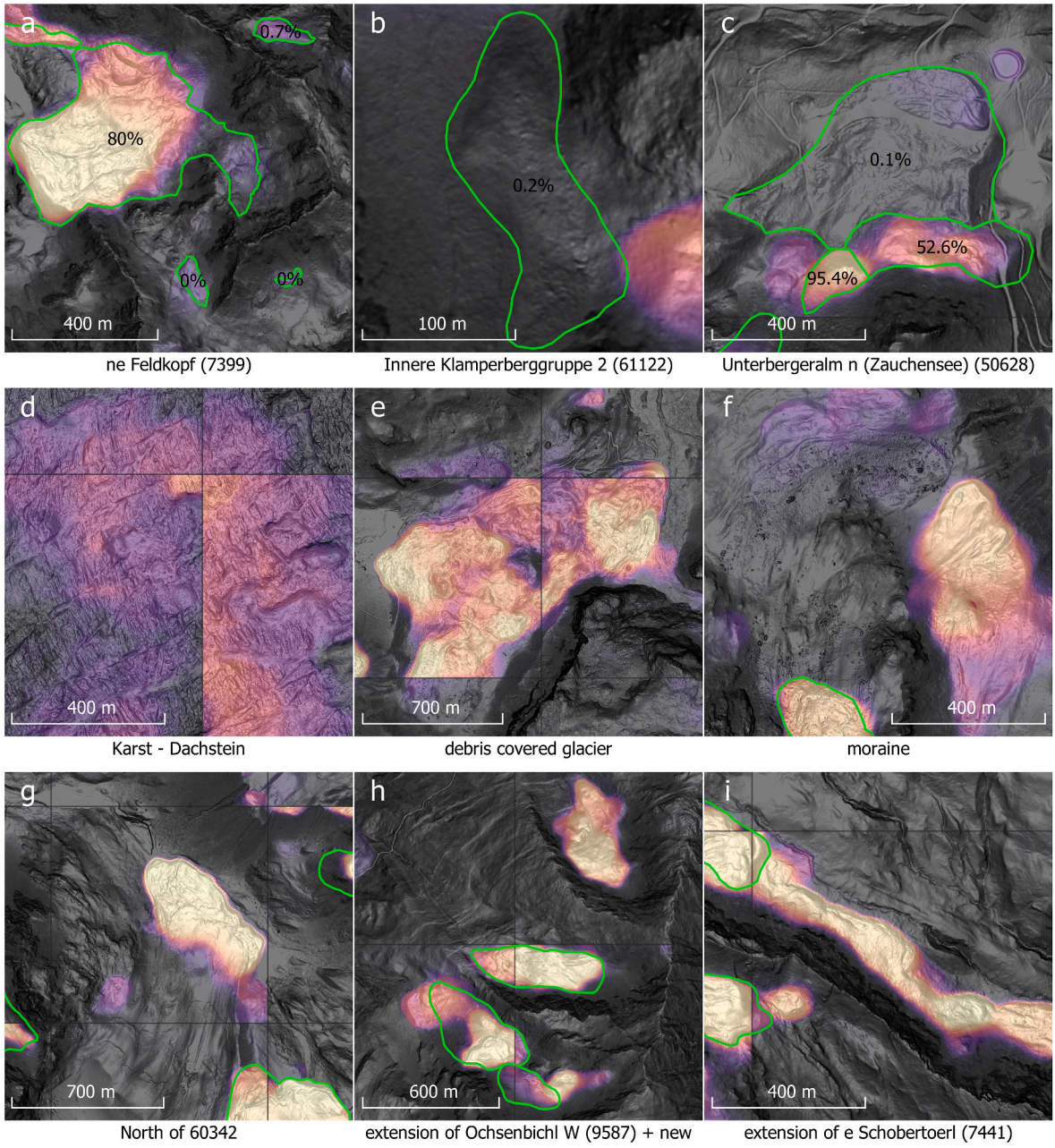

Fig. 7. Samples of the artificial neural network-based rock glacier map with light yellow colors indicating a high rock glacier probability (= 1) and dark colors indicating a low probability (= 0) (same colormap for all images); names and identification number within the Austrian rock glacier inventory (intNr) are given below; already mapped rock glaciers of the Austrian inventory are outlined by green lines and the 1 × 1 km grid is shown by black lines. (a) To (c):examples of rock glaciers from the test dataset that were not detected; (d) to (f): morphologies that were falsely detected as rock glaciers; (g) to (i): morphologies that are seen as newly detected rock glaciers or extensions of already mapped ones. Note the different scales in the sample images.

the average wavelength of furrow-and-ridge structures is around 20–25 m for the active Murtèl RG in Switzerland; a relatively large and obvious rock glacier due to its surface morphology that is representing compressive flow (Frehner et al., 2015). Relict RGs and intact ones with less pronounced morphology are potentially harder to detect (manually and by the ANN) and higher resolution DEMs are likely necessary. In general, it can be assumed that large RGs with pronounced morphological features can still be detected in lower resolution DEMs but not reverse. Therefore, a horizontal resolution of presumably 5–10 m is going to be necessary at least and the use of coarser DEM data (such as the SRTM 90 m DEM Data available for the entire world; e.g., Reuter et al. (2007)) is not going to yield applicable results for this data set. Further investigations due to the proper resolution of the input data set (e.g., DEM) related to the RG size for suitable detectability of these landforms are necessary.

The approach to derive samples from a fixed grid and use data augmentation during training was a success, as the output shows no directional bias or preferred locations of the found RGs within the samples themselves (Fig. 10).

Especially the observation that detected RGs often span over multiple grid cells shows that the ANNs have learned to consider all parts of the RGs and not only well pronounced features like the front of the landform. We also see the U-Net ANN architecture well suited for the given task, although further architecture and parameter tuning could be conducted.

Converting the continuous output of the ANNs to a discrete one (as

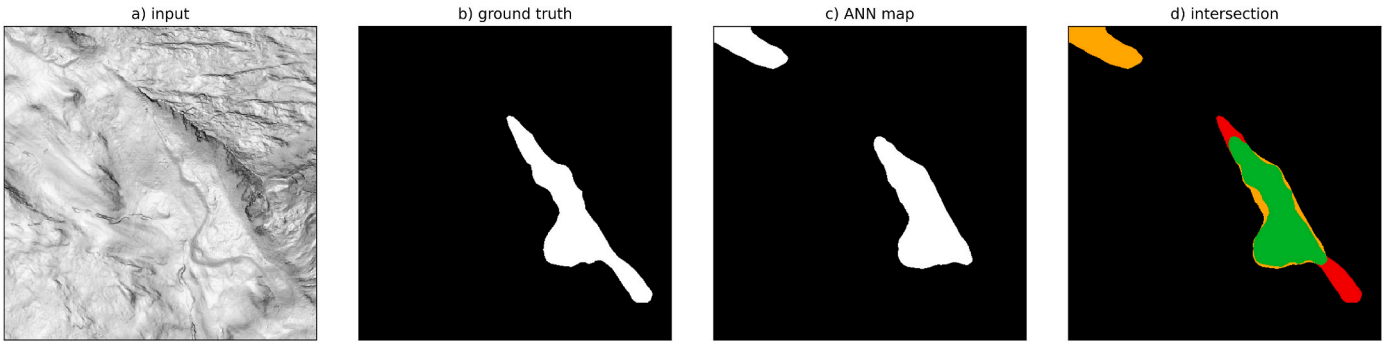

Fig. 8. Discretized sample from the test dataset. From left to right the figure shows: a) The first channel of the respective input sample – i.e., slope view; b) the ground truth label-map derived from the Austrian RG inventory; c) label-map based on a weighted average classification of 10 ANN models and discretized with a threshold of 0.4; d) intersection of the ground truth with the discretized ANN label-map showing true negatives in black, true positives in green, false positives in orange and false negatives in red.

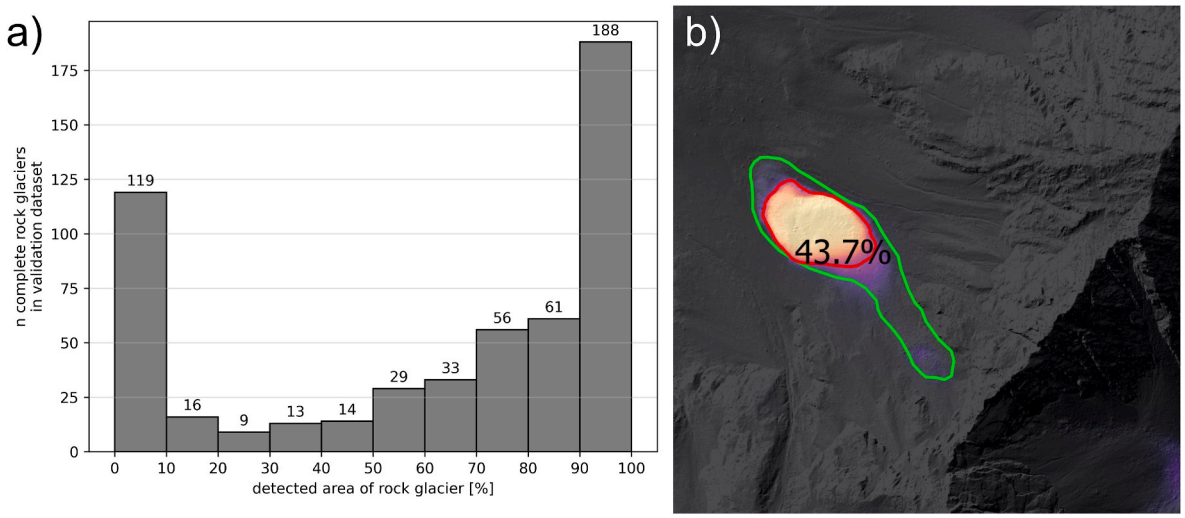

Fig. 9. (a) Barchart showing how many complete rock glaciers of the validation dataset were detected by the ANNs (538 in total). (b) Example of a rock glacier where the originally mapped shape is given in green and the outline of the discretized ANN result in red (A% = 43.7). A hillshade DEM and a colormap that shows the rasterized/direct output of the ANNs is given in the background.

described in section 2.5) enables a direct quantitative comparison to the existing RG inventory on the one hand and also the derivation of vectorized polygons on the other. As given in sections 2.5 and 3.2 however, the choice of the probability threshold that separates RG from non-RG pixels is critical and highly influential for the number of false positives and false negatives. In this study we found a fitting threshold based on the maximum dice coefficient.

4.3. Training data uncertainty and human bias

Given a perfect set of training data, it can be assumed that today's state of the art ML algorithms can achieve and even surpass human level performance for classification tasks. This has already been shown several years ago for with datasets that contain a very low degree of uncertainty in themselves, like handwritten number identification (Ciresan et al., 2012; He et al., 2015). Consequently, the biggest influence on the ANN performance is seen in the training data itself. As in many studies that use ML to do geological/geomorphological classification, the problem arises that during training the data and the labels must be treated as the ground truth. It is however well known that geological labels – especially if they were assigned by human experts – are subject to extensive variability depending on different experiences and conditions of the persons who did the mapping/labelling. This is on the one hand a topic of discussion in the ML community Xu et al. (2021). On the other hand also Wagner et al. (2020b) discuss this for the Austrian RG inventory where the delineation of the individual RGs was done by multiple persons with different levels of experience and it is known that even experts themselves may disagree on the actual delineation of RGs as reported by Brardinoni et al. (2019). Thus, the used training data base – in our case the initial Austrian RG inventory - contains a certain degree of subjectivity, human errors and consequently inconsistency. Particularly the rooting zone delineation is most often done in an ambiguous way that is related to the geomorphological problem of the loose definition of a rock glacier and especially its rooting zone. The IPA Action Group Rock glacier inventories and kinematics is currently trying to come up with the necessary standards, define stricter guidelines for rock glacier mapping and definition to resolve some of these issues in the near future (RGIK 2022). In addition, rock glaciers can also be polymorphic landforms with more than one generation of rock glaciers (e.g., Frauenfelder and Kääb (2000)). The Austrian RG Inventory assigns polymorphic landforms an attribute (LfMorph: m = monomorph; p = polymorph) and list them as individual features (polygons). The ANN might see these individual features as a single one. This might not impact a binary RG mapping (i.e., RG yes or no), but could become an issue for possible multi-class classifications in future studies. Moreover, discrimination of very small RGs with less well pronounced morphologies is problematic even between experts. As shown in Fig. 7 (g – i) there might also be areas where RGs have been overlooked in the initial

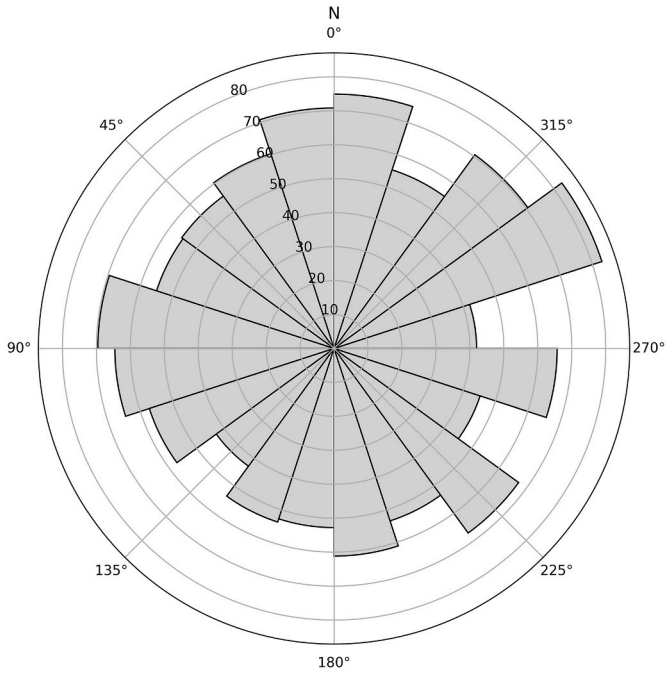

Fig. 10. Polar plot of the average aspect and (former) flow direction [°] (attribute "aspLfDeg" in the Austrian RG Inventory) against the detected area of RGs in percent (18° bars). No directional bias can be detected.

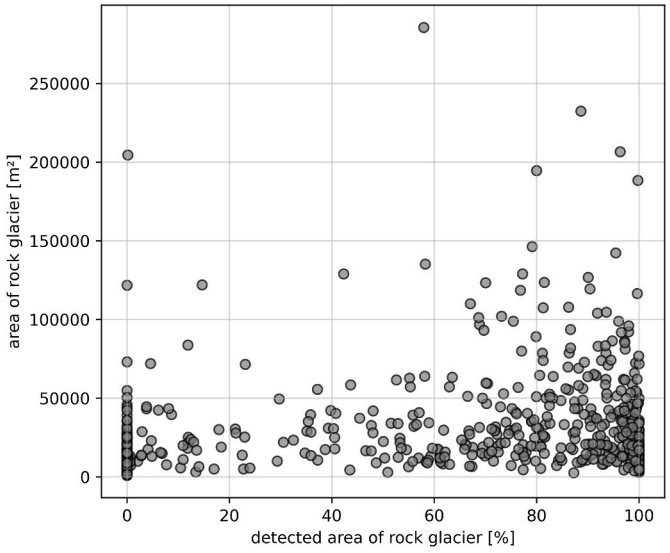

Fig. 11. Scatterplot that shows how the detected area of individual rock glaciers increases with the overall size of the rock glaciers which might be related to inconsistencies in the original data with respect to smaller rock glaciers.

mapping process. We also see the reason for why *A*% (as defined in the previous section) shows an increasing trend with the overall size of the RGs (Fig. 11) in human errors and disagreements in expert opinions with regard to small RGs.

It is not the goal of this study to speculate on the numerous reasons that might introduce inconsistencies and errors into RG inventories (or geological data in general). It can be assumed, however, that these inconsistencies and errors adversely affect training of ML algorithms to a large extent as the algorithm must deal with partially misleading and contradictory training data. To alleviate this, we therefore propose that multiple cycles of ML training and RG inventory revision of the training data set might help to vastly increase an inventory's quality. The proposed ANN integrated in an iterative process of RG inventory refinement might contribute to the IPA Action Group goal.

5. Conclusion and outlook

In conclusion, it has been shown that the chosen methodology for automized RG mapping support generally works and can contribute to the increase in consistency of already established RG inventories (see the introduction for an explanation how "consistency" is used in the context of this study). The increase in consistency of the ANN based approach over the human mapping comes from the fact that once the ANN is trained it applies a fixed and unalterable mapping policy unless it is retrained with more data. In the presented form the proposed ML-framework provides a tool that helps to scan mountain ranges/landscapes efficiently and consistently for specific landforms based on DEMs and orthophotos.

However, the current ANN output is not directly providing a new RG inventory and should therefore not be used to replace manual RG delineation and assessment by experts is still necessary. Experts might even work with the direct ANN output without applying a threshold to derive vectorized polygons as a base for their manual delineation. The reason for why the ANN output is not directly useable as a new RG inventory is seen in the quality of the data the ML-framework is trained on. Consequently, improving the approach of ANN based RG mapping goes along with improving the database on which the ANNs are trained. To do so, an iterative approach could be chosen where in a first loop ANNs are trained to predict RGs based on a conventionally established RG inventory. With that result, the original RG inventory could be revised by extending it with newly detected RGs and possibly deleting ones which were not detected by the ANNs and where the delineation was questionable in the first place. Having done that, the quality and consistency of the inventory should be improved, and one could use the revised inventory to retrain a new ML framework which should perform better than the initial one. By combining this approach with the work presented by e.g., Reinosch et al. (2021) for automated RG classification, a big step towards more consistent RG inventories as proposed by the IPA is to be expected. With these methods (possibly involving several loops of ML-training and inventory-revising) it should be possible to reach a state where the performance of the ML framework is sufficiently good, and the results can reach or even surpass human level performance as there is no limit to the amount of training a ML algorithm can go through (as opposed to humans). As given before however, the current framework has to be seen as an additional help for the human mapper and ML algorithms today still lack the ability to include geological context and "out of domain experience" in their processing. Only continuous development and simultaneous performance assessment will be able to show if ML based mapping can reach human level performance in the future.

Despite applications to increase the completeness and consistency of current RG inventories, the presented ML based way of RG detection can be used for basic establishment of new RG inventories in mountain ranges that were not yet considered for such an application. The presented framework was designed in a way that it can be applied for all regions where orthophotos and DEMs with a sufficient resolution are available. Although the availability of high-resolution DEM data might still be an obstacle today, we see this as a problem that decreases with ongoing technological progress and digitalization. By using slope maps and orthophotos as an input - as opposed to direct DEM data - we furthermore designed the framework in a way that it only considers the morphological features of the landscape and is independent of regional characteristics like slope orientation or absolute elevation.

Data Availability

Austrian Rock Glacier Inventory based on: https://doi.org/10.1594/PANGAEA.921629

By the time of this writing the following states of Austria provide **1 × 1 m digital elevation models** as openly accessible downloads that can be used to reproduce this study (sorted alphabetically):

- Carinthia, https://www.data.gv.at/katalog/dataset/c4352e0c-f811-4c59-bdac-4ba199684fc2;
- Salzburg, https://www.salzburg.gv.at/api5/datalinq/report/vektorwork@alsdownload@alsdownload;
- Styria, https://data.steiermark.at/cms/beitrag/11822084/97108894/?AppInt_OGD_ID=1211;
- Tyrol, https://tiris.maps.arcgis.com/apps/webappviewer/index.html?id=5e3071044cb44e76843d110baef8b138;
- Upper Austria: https://www.land-oberoesterreich.gv.at/211787.htm;
- Vienna, https://www.data.gv.at/katalog/dataset/1eda4b16-05e3-43c3-b5b0-ef8276df0701;

The following URL was used to retrieve the XYZ tiles for the **Google maps satellite images**: http://www.google.cn/maps/vt?lyrs=s@189&gl=cn&x={x}&y={y}&z={z}

Computer code availability

The source code of the used U-Net implementation for this paper as it is described in section 2.3 is available in the following GitHub repository: https://github.com/geograz/ML_RG_mapping. The coding was done by the main author (contact details as given in the affiliation section of this paper) and it is a modified version of the U-Net model presented in the following GitHub Repository: https://github.com/malhotraa/carvana-image-masking-challenge that was made for a specific Kaggle challenge (https://www.kaggle.com/c/carvana-image-masking-challenge/overview/description). The code for this paper was made available before submission in August 2021 and the ANN was trained on a NVIDA RTX 2080 Ti graphics processing unit. The programming language is Python. The code is licensed under the "MIT license" and further information is given in the repository.

Author statement for paper increasing consistency and completeness of rock glacier inventories with machine learning – the example of Austria

Georg H. Erharter: Conceptualization, Methodology, Software, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, **Thomas Wagner:** Validation; Investigation; Writing - Review & Editing, **Gerfried Winkler:** Conceptualization, Validation, Resources, Writing -Review & Editing, **Thomas Marcher:** Writing -Review & Editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Digital elevation models were provided by the GIS Service of the Federal Governments of Vorarlberg, Tyrol, Salzburg, Styria, and Carinthia which is highly acknowledged.

Appendix

The **105 resulting raster tiles** as described in section 3 can be found under the following link:

https://repository.tugraz.at/records/ckkj2-2me08?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjY2NDQ1YWVlLWE3N2ItNDkzOS05NDdlLTY0N2MxZWFmZjhhZiIsImRhdGEiOnt9LCJyYW5kb20iOiI5NmFmOTgxMGY2OTc2Yzk1MmYwMWI4OGM0ZGZjNmUwYyJ9.FqnfeGcBzd15vpl3NrN7qGUq12-l47dNqN5vGsQGqgaG2dB5uOqxjgmoy-p98dpHY-Pfb128A6WmuxLxqTkOLw

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. https://www.tensorflow.org/.

Badrinarayanan, V., Kendall, A., Cipolla, R., 2015. SegNet: a deep convolutional encoder-decoder architecture for image segmentation, 14pp. URL: http://arxiv.org/pdf/1511.00561v3.

Barsch, D. (Ed.), 1996. Rockglaciers: Indicators for the Present and Former Geoecology in High Mountain Environments. Springer, Berlin, Heidelberg, p. 331pp.

Berthling, I., 2011. Beyond confusion: rock glaciers as cryo-conditioned landforms. Geomorphology 131, 98–106. https://doi.org/10.1016/j.geomorph.2011.05.002.

Brardinoni, F., Scotti, R., Sailer, R., Mair, V., 2019. Evaluating sources of uncertainty and variability in rock glacier inventories. Earth Surf. Process. Landforms 44, 2450–2466. https://doi.org/10.1002/esp.4674.

Chollet, F. (Ed.), 2018. Deep Learning with Python, Manning. Shelter Island, NY, p. 361pp.

Chollet, F., 2015. Keras others.

Ciresan, D., Meier, U., Schmidhuber, J., 2012. Multi-column Deep Neural Networks for Image Classification 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 3642–3649.

Colucci, R.R., Forte, E., Žebre, M., Maset, E., Zanettini, C., Guglielmin, M., 2019. Is that a relict rock glacier? Geomorphology 330, 177–189. https://doi.org/10.1016/j.geomorph.2019.02.002.

Cunningham, P., Carney, J., Jacob, S., 2000. Stability problems with artificial neural networks and the ensemble solution. Artif. Intell. Med. 20, 217–225. https://doi.org/10.1016/S0933-3657(00)00065-8.

Delaloye, R., Morard, S., Barboux, C., Abbet, D., Gruber, V., Riedo, M., Gachet, S., 2012. Rapidly moving rock glaciers in Mattertal. In: Jahrestagung Der Schweizerischen Geomorphologischen Gesellschaft, pp. 21–31.

Deluigi, N., Lambiel, C., Kanevski, M., 2017. Data-driven mapping of the potential mountain permafrost distribution. Sci. Total Environ. 590–591, 370–380. https://doi.org/10.1016/j.scitotenv.2017.02.041 [in eng].

Dice, L.R., 1945. Measures of the amount of ecologic association between species. Ecology 26, 297–302. https://doi.org/10.2307/1932409.

Dietterich, T.G., 2000. Ensemble Methods in Machine Learning Multiple Classifier Systems. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–15.

Dirscherl, M., Dietz, A.J., Kneisel, C., Kuenzer, C., 2021. A novel method for automated supraglacial lake mapping in Antarctica using sentinel-1 SAR imagery and deep learning. Rem. Sens. 13, 197. https://doi.org/10.3390/rs13020197.

Erharter, G.H., Marcher, T., Reinhold, C., 2019. Application of artificial neural networks for Underground construction – chances and challenges – insights from the BBT exploratory tunnel Ahrental Pfons. Geomechanik und Tunnelbau 12, 472–477. https://doi.org/10.1002/geot.201900027.

Erharter, G.H., Oberhollenzer, S., Fankhauser, A., Marte, R., Marcher, T., 2021. Learning decision boundaries for cone penetration test classification. Comput. Aided Civ. Infrastruct. Eng. 36, 489–503. https://doi.org/10.1111/mice.12662.

Frauenfelder, R., Haeberli, W., Hoelzle, M., Maisch, M., 2001. Using relict rockglaciers in GIS-based modelling to reconstruct Younger Dryas permafrost distribution patterns in the Err-Julier area, Swiss Alp. Norsk Geografisk Tidsskrift - Norwegian Journal of Geography 55, 195–202. https://doi.org/10.1080/00291950152746522.

Frauenfelder, R., Kääb, A., 2000. Towards a palaeoclimatic model of rock-glacier formation in the Swiss Alps. Ann. Glaciol. 31, 281–286. https://doi.org/10.3189/172756400781820264.

Frehner, M., Ling, A.H.M., Gärtner-Roer, I., 2015. Furrow-and-ridge morphology on rockglaciers explained by gravity-driven buckle folding: a case study from the Murtèl rockglacier (Switzerland). Permafr. Periglac. Process. 26, 57–66. https://doi.org/10.1002/ppp.1831.

Giaccone, E., Oriani, F., Tonini, M., Lambiel, C., Mariéthoz, G., 2021. Using data-driven algorithms for semi-automated geomorphological mapping. Stoch. Environ. Res. Risk Assess. https://doi.org/10.1007/s00477-021-02062-5.

Guglielmin, M., Smiraglia, C., 1998. The Rock Glacier Inventory of the Italian Alps PERMAFROST - Seventh International Conference (Proceedings).

Haeberli, W., Guodong, C., Gorbunov, A.P., Harris, S.A., 1993. Mountain permafrost and climatic change. Permafr. Periglac. Process. 4, 165–174. https://doi.org/10.1002/ppp.3430040208.

Haeberli, W., Schaub, Y., Huggel, C., 2017. Increasing risks related to landslides from degrading permafrost into new lakes in de-glaciating mountain ranges. Geomorphology 293, 405–417. https://doi.org/10.1016/j.geomorph.2016.02.009.

Hahnloser, R.H.R., Sarpeshkar, R., Mahowald, M.A., Douglas, R.J., Seung, H.S., 2000. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. Nature 405, 947–951. https://doi.org/10.1038/35016072.

Hayashi, M., 2020. Alpine hydrogeology: the critical role of groundwater in sourcing the headwaters of the world. Ground Water 58, 498–510. https://doi.org/10.1111/gwat.12965 [in eng].
He, K., Zhang, X., Ren, S., Sun, J., 2015. In: Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1026–1034. IEEE.
Hubel, D.H., Wiesel, T.N., 1963. Shape and arrangement of columns in cat's striate cortex. J. Physiol. 165, 559–568. https://doi.org/10.1113/jphysiol.1963.sp007079 [in eng].
Jones, D.B., Harrison, S., Anderson, K., Betts, R.A., 2018a. Mountain rock glaciers contain globally significant water stores. Sci. Rep. 8, 2834. https://doi.org/10.1038/s41598-018-21244-w [in eng].
Jones, D.B., Harrison, S., Anderson, K., Selley, H.L., Wood, J.L., Betts, R.A., 2018b. The distribution and hydrological significance of rock glaciers in the Nepalese Himalaya. Global Planet. Change 160, 123–142. https://doi.org/10.1016/j.gloplacha.2017.11.005.
Jones, D.B., Harrison, S., Anderson, K., Whalley, W.B., 2019. Rock glaciers and mountain hydrology: a review. Earth Sci. Rev. 193, 66–90. https://doi.org/10.1016/j.earscirev.2019.04.001.
Kääb, A., Frauenfelder, R., Roer, I., 2007. On the response of rockglacier creep to surface temperature increase. Global Planet. Change 56, 172–187. https://doi.org/10.1016/j.gloplacha.2006.07.005.
Kellerer-Pirklbauer, A., Lieb, G.K., Kleinferchner, H., 2012. A new rock glacier inventory of the eastern European Alps. Aust. J. Earth Sci. 105, 78–93.
Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. URL: http://arxiv.org/pdf/1412.6980v9, 15pp.
Kofler, C., Steger, S., Mair, V., Zebisch, M., Comiti, F., Schneiderbauer, S., 2020. An inventory-driven rock glacier status model (intact vs. relict) for South Tyrol, Eastern Italian Alps. Geomorphology 350, 106887. https://doi.org/10.1016/j.geomorph.2019.106887.
Krainer, K., Ribis, M., 2012. A rock glacier inventory of the tyrolean Alps (Austria). Aust. J. Earth Sci. 32–47.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. Proc. Adv. Neural Inf. Process. Syst. 1090–1098.
Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2324. https://doi.org/10.1109/5.726791.
Lee, D.-H., Kim, Y.-T., Lee, S.-R., 2020. Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various non-linear activation functions. Rem. Sens. 12, 1194. https://doi.org/10.3390/rs12071194.
Lieb, G.K., Kellerer-Pirklbauer, A., Kleinferchner, H. Second Rock glacier inventory (RGI2) of central and eastern Austria, link to shapefile, supplement to: Kellerer-Pirklbauer, Andreas; Lieb, gerhard karl; kleinferchner, harald (2012): a new Rock glacier inventory of the eastern European Alps. Aust. J. Earth Sci., 105(2), 78-93. doi:10.1594/PANGAEA.869805. ([in en]).
Liu, Z., Gilbert, G., Cepeda, J.M., Kydland Lysdahl, A.O., Piciullo, L., Hefre, H., Lacasse, S., 2020. Modelling of shallow landslides with Machine Learning algorithms. Geosci. Front. https://doi.org/10.1016/j.gsf.2020.04.014.
Marcer, M., 2020. Rock glaciers automatic mapping using optical imagery and convolutional neural networks. Permafr. Periglac. Process. https://doi.org/10.1002/ppp.2076.
Marmion, M., Hjort, J., Thuiller, W., Luoto, M., 2008. A comparison of predictive methods in modelling the distribution of periglacial landforms in Finnish Lapland. Earth Surf. Process. Landforms 33, 2241–2254. https://doi.org/10.1002/esp.1695.
Maxwell, A.E., Bester, M.S., Guillen, L.A., Ramezan, C.A., Carpinello, D.J., Fan, Y., Hartley, F.M., Maynard, S.M., Pyron, J.L., 2020. Semantic segmentation deep learning for extracting surface mine extents from historic topographic maps. Rem. Sens. 12, 4145. https://doi.org/10.3390/rs12244145.
Moran, A.P., Ivy Ochs, S., Vockenhuber, C., Kerschner, H., 2016. rock Glacier development in the northern calcareous Alps at the pleistocene-holocene boundary. Geomorphology 273, 178–188. https://doi.org/10.1016/j.geomorph.2016.08.017.
Pruessner, L., Huss, M., Phillips, M., Farinotti, D., 2021. A framework for modeling rock glaciers and permafrost at the basin-scale in high alpine catchments. J. Adv. Model. Earth Syst. 13 https://doi.org/10.1029/2020MS002361.
QGIS.org, 2021. QGIS Geographic Information System. http://www.qgis.org. (Accessed 15 March 2021).
Rangecroft, S., Harrison, S., Anderson, K., Magrath, J., Castel, A.P., Pacheco, P., 2014. A first rock glacier inventory for the Bolivian Andes. Permafr. Periglac. Process. 25, 333–343. https://doi.org/10.1002/ppp.1816.
Raschka, S., Mirjalili, V., 2019. Python Machine Learning - Third Edition: Machine Learning and Deep Learning with python, Scikit… -learn, and Tensorflow 2. PACKT PUBLISHING LIMITED [S.l.].
Reinosch, E., Gerke, M., Riedel, B., Schwalb, A., Ye, Q., Buckel, J., 2021. Rock glacier inventory of the western Nyainqêntanglha Range, Tibetan Plateau, supported by InSAR time series and automated classification. Permafr. Periglac. Process. https://doi.org/10.1002/ppp.2117.
Reuter, H.I., Nelson, A., Jarvis, A., 2007. An evaluation of void-filling interpolation methods for SRTM data. Int. J. Geogr. Inf. Sci. 21, 983–1008. https://doi.org/10.1080/13658810601169899.
RGIK, 2022. Towards Standard Guidelines for Inventorying Rock Glaciers: Baseline Concepts, version 4.2.2.
Robson, B.A., Bolch, T., MacDonell, S., Hölbling, D., Rastner, P., Schaffer, N., 2020. Automated detection of rock glaciers using deep learning and object-based image analysis. Rem. Sens. Environ. 250, 112033 https://doi.org/10.1016/j.rse.2020.112033.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. Springer International Publishing, Cham, pp. 234–241.
Samarin, M., Zweifel, L., Roth, V., Alewell, C., 2020. Identifying soil erosion processes in alpine grasslands on aerial imagery with a U-net convolutional neural network. Rem. Sens. 12, 4149. https://doi.org/10.3390/rs12244149.
Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6. https://doi.org/10.1186/s40537-019-0197-0.
Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Kongelige Danske Videnskabernes Selskab 5, 1–34.
Stocker, K., 2012a. Blockgletscher in Vorarlberg und in der Verwallgruppe. In: Museumsverein Jahrbuch Vorarlberger Landesmuseum, pp. 124–139.
Stocker, K., 2012b. Geologie und Blockgletscher der Madererspitze. Master's Thesis. University of Innsbruck, Innsbruck.
Stocker, K., Krainer, K., 2011. Abschlussbericht Blockgletscherinventar, p. 6p. Unpublished report.
Stumpf, A., Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. Rem. Sens. Environ. 115, 2564–2577. https://doi.org/10.1016/J.RSE.2011.05.013.
Wagner, T., Brodacz, A., Krainer, K., Winkler, G., 2020a. Active rock glaciers as shallow groundwater reservoirs, Austrian Alps. Grundwasser 25, 215–230. https://doi.org/10.1007/s00767-020-00455-x.
Wagner, T., Kainz, S., Krainer, K., Winkler, G., 2021. Storage-discharge characteristics of an active rock glacier catchment in the Innere Ölgrube, Austrian Alps. Hydrol. Process. 35 https://doi.org/10.1002/hyp.14210.
Wagner, T., Pleschberger, R., Kainz, S., Ribis, M., Kellerer-Pirklbauer, A., Krainer, K., Philippitsch, R., Winkler, G., 2020b. The first consistent inventory of rock glaciers and their hydrological catchments of the Austrian Alps. Aust. J. Earth Sci. 113, 1–23. https://doi.org/10.17738/ajes.2020.0001.
Wagner, T., Ribis, M., Kellerer-Pirklbauer, A., Krainer, K., Winkler, G., 2020c. The Austrian Rock Glacier Inventory RGI_1 and the Related Rock Glacier Catchment Inventory RGCI_1 in ArcGis (Shapefile) Format. https://doi.org/10.1594/PANGAEA.921629 [in en].
Wahrhaftig, C., Cox, A., 1959. Rock glaciers in the Alaska range. Geol. Soc. Am. Bull. 70, 383, 10.1130/0016-7606(1959)70[383:RGITAR]2.0.CO;2.
Wang, Z., Fan, C., Xian, M., 2021. Application and evaluation of a deep learning architecture to urban tree canopy mapping. Rem. Sens. 13, 1749. https://doi.org/10.3390/rs13091749.
Winkler, G., Wagner, T., Krainer, K., Ribis, M., Hergarten, S., 2018. HYDROGEOLOGY OF ROCK GLACIERS - STORAGE CAPACITY AND DRAINAGE DYNAMICS - AN OVERVIEW, pp. 329–334. https://doi.org/10.25680/8822.2018.82.81.168 [in en].
Wu, C., Zhang, F., Xia, J., Xu, Y., Li, G., Xie, J., Du, Z., Liu, R., 2021. Building damage detection using U-net with attention mechanism from pre- and post-disaster remote sensing datasets. Rem. Sens. 13, 905. https://doi.org/10.3390/rs13050905.
Xu, Q., Walder, C., Xu, C., 2021. Humanly certifying superhuman classifiers, 11pp. URL: http://arxiv.org/pdf/2109.07867v1.