



# Article Identifying Soil Erosion Processes in Alpine Grasslands on Aerial Imagery with a U-Net Convolutional Neural Network

## Maxim Samarin<sup>1,\*,†</sup>, Lauren Zweifel<sup>2,†</sup>, Volker Roth<sup>1</sup> and Christine Alewell<sup>2</sup>

- <sup>1</sup> Department of Mathematics and Computer Science, University of Basel, Spiegelgasse 1, 4051 Basel, Switzerland; volker.roth@unibas.ch
- <sup>2</sup> Department of Environmental Sciences, University of Basel, Bernoullistrasse 30, 4056 Basel, Switzerland; lauren.zweifel@unibas.ch (L.Z.); christine.alewell@unibas.ch (C.A.)
- \* Correspondence: maxim.samarin@unibas.ch
- + These authors contributed equally to this work.

Received: 25 November 2020; Accepted: 14 December 2020; Published: 18 December 2020



**Abstract:** Erosion in alpine grasslands is a major threat to ecosystem services of alpine soils. Natural causes for the occurrence of soil erosion are steep topography and prevailing climate conditions in combination with soil fragility. To increase our understanding of ongoing erosion processes and support sustainable land-use management, there is a need to acquire detailed information on spatial occurrence and temporal trends. Existing approaches to identify these trends are typically laborious, have lack of transferability to other regions, and are consequently only applicable to smaller regions. In order to overcome these limitations and create a sophisticated erosion monitoring tool capable of large-scale analysis, we developed a model based on U-Net, a fully convolutional neural network, to map different erosion processes on high-resolution aerial images (RGB, 0.25–0.5 m). U-Net was trained on a high-quality data set consisting of labeled erosion sites mapped with object-based image analysis (OBIA) for the Urseren Valley (Central Swiss Alps) for five aerial images (16 year period). We used the U-Net model to map the same study area and conduct quality assessments based on a held-out test region and a temporal transferability test on new images. Erosion classes are assigned according to their type (shallow landslide and sites with reduced vegetation affected by sheet erosion) or land-use impacts (livestock trails and larger management affected areas). We show that results obtained by OBIA and U-Net follow similar linear trends for the 16 year study period, exhibiting increases in total degraded area of 167% and 201%, respectively. Segmentations of eroded sites are generally in good agreement, but also display method-specific differences, which lead to an overall precision of 73%, a recall of 84%, and a  $F_1$ -score of 78%. Our results show that U-Net is transferable to spatially (within our study area) and temporally unseen data (data from new years) and is therefore a method suitable to efficiently and successfully capture the temporal trends and spatial heterogeneity of degradation in alpine grasslands. Additionally, U-Net is a powerful and robust tool to map erosion sites in a predictive manner utilising large amounts of new aerial imagery.

**Keywords:** deep learning; semantic segmentation; remote sensing; object-based image analysis; erosion mapping; landslides; livestock trails; sheet erosion

## 1. Introduction

Soil degradation is a major ecological threat which affects many areas of the world and can be accelerated by land-use management and changing climate parameters, such as precipitation and temperature [1–5]. In Switzerland, some alpine grassland areas are strongly affected by soil erosion

due to the steep terrain and extreme climate conditions. While soil erosion occurs naturally in these environments—in the form of landslides (triggered by snow gliding or heavy precipitation events) or sheet erosion (the process of the removal of topsoil caused by rain drops' impacts and overland flow)—there are also anthropogenic influences (e.g., agricultural activities) which can accelerate erosion rates [3,6,7]. For example, livestock keeping can lead to overgrazing and trampling in favoured grazing areas. Over time, livestock trails develop and trampling and grazing can lead to a reduction in vegetation cover, which in turn is prone to sheet erosion [8,9]. Additionally, livestock keeping can cause instabilities on slopes and ultimately result in landslides [10].

Therefore, erosion processes have strong temporal and spatial dynamic components, which is why large-scale understanding and detailed mapping over time and space is of great importance for long-term sustainable management practices.

Alpine areas are difficult to access and erosion features can affect substantial areas, making a comprehensive understanding of ongoing erosion processes unattainable from the ground. Larger-scale erosion studies for Switzerland have mainly been approached with the help of soil erosion modelling—e.g., the (revised) universal soil loss equation [11-18]. To achieve a thorough understanding of potential soil erosion threats, it is important to combine model outputs with observations for validation purposes [19]. The latter is especially crucial in mountainous and grassland areas, where model suitability has been questioned (see discussion in Alewell et al. [20]). High-resolution aerial imagery offers the opportunity to remotely assess and map the spatial extent of bare soil sites and sites with strongly reduced vegetation cover, allowing certain constraints to be overcome, such as the inaccessibility or extent of a study area. Object-based image analysis (OBIA) is an approach commonly used to identify urban and natural "objects" on satellite and aerial imagery and has been successfully used in the past to map various forms of soil erosion [7,21–30]. OBIA creates image segments by grouping pixels with similar properties together, which can then be classified based on object information (spectral, spatial, textural, and contextual) with expertly developed classification rules and/or various machine learning classifiers. OBIA is a method suitable for smaller study areas, but large-scale studies become difficult to manage. Limitations including processing times, a lack of work-flow transferability to other scenes, and the involvement of manual steps hinder efficient spatial up-scaling of projects. In past years, deep learning methods have progressively been applied in the field of remote sensing for image classification tasks and segmentation tasks [31–34]. In this study, we apply a deep learning method to demonstrate that it is capable of mapping and classifying soil erosion features on aerial images in a fast, objective, reliable, and scalable manner. We apply a fully-convolutional neural network (CNN) framework using the U-Net architecture developed by Ronneberger et al. [35]. In general, the U-Net architecture offers itself to semantic segmentation tasks with limited training data. U-Net and variations of this architecture have become increasingly popular for remote sensing tasks. Many applications focus on urban settings for road [36–39] or building extraction [40–43] from satellite and aerial imagery. Applications in a natural environment are constrained by the limited availability of high-quality labelled training data. Despite this limitation, U-Net has been applied in cloud detection on satellite images [44], mapping of woody vegetation [45], segmentation of plant species [46], forest damage assessment [47], the extraction of Antarctic glacier and ice shelf fronts [48], and archaeological studies [49] to name a few. Our annotated training data has been generated by mapping erosion sites on aerial images using OBIA for a valley in the Central Swiss Alps (Urseren Valley, Canton of Uri). We compare U-Net results to OBIA mapping for a held-out test region (area of 17 km<sup>2</sup>), which was not used for training (9 km<sup>2</sup>) for the years 2000, 2004, 2010, and 2013. Additionally, we investigate both the temporal and the spatial transferability of the U-Net method by mapping a new aerial image not seen during training (2016). Our main objectives of this study are: firstly, to show that the fully automated U-Net approach is capable of reproducing the high-quality soil erosion mapping and the temporal trends as they were attained with OBIA for the same study site; secondly, to show that the U-Net approach generalises well to new aerial images, i.e., can be used in a predictive manner to perform adequate segmentation of previously unseen input

data. In contrast, the OBIA procedure typically eludes such predictive usage and needs to be adjusted for each new aerial image. The capabilities and the fully automated nature of the U-Net approach make it a highly promising tool for efficient large-scale erosion mapping (e.g., alpine-wide analysis of soil erosion in semi-natural ecosystems such as grasslands and bush-land).

## 2. Study Area

The Urseren Valley (26 km<sup>2</sup>) is an alpine valley located in Central Switzerland in the southern part of the Canton of Uri (Figure 1). The valley has a NE–SW orientation, and exhibits steep slopes (average angle of 27°) and rough terrain. The valley is geologically divided into two distinct sections and separated by the river Reuss: The northern slope is part of the Aarmassif (granite), and the southern slope belongs to the Gotthard massif (gneiss). Located between these two massifs near the valley floor is the so-called Urseren—Garvera zone (Mesozoic sediments) [50]. The dominant soil types in the catchment are Podzols and Cambisols, with Leptosols commonly found on steep slopes (classified after IUSS Working Group WRB [51]).



**Figure 1.** The Urseren Valley is located in the Central Swiss Alps in the Canton of Uri. The left map contains the topographic map of Switzerland (from low elevations in green to high elevations in brown to white). The right image contains an aerial image of the Urseren Valley overlaid on a hill-shade map of the area.

The 30 year average temperature (1990–2019) of the closest meteorological station in Andermatt (1438 m a.s.l.) is 3.9 °C. The average temperature has increased by 0.7 °C during the last 10 years (compared to the average of 1980–2009). The average rainfall during the last 30 years was 1384 mm with an average maximum 3 day precipitation intensity of 123 mm/3 d. The average seasonal (November–April) snow height is 58 cm with maximum snow heights during February/March (average of 103 cm) (data provided by MeteoSwiss, 2020). The dominant land-covers are grassland (including dwarf-shrubs consisting of Calluna vulgaris, Rhododendron ferrugineum, and Juniperus sibirica), which is mainly used for grazing (i.e., sheep and cattle) and haying, shrubs (mainly Alnus viridis and Sorbus aucuparia), and debris/bare rock areas [3]. Shrub encroachment due to land abandonment and extensification is present in the valley. Avalanches and snow gliding occur frequently in the Urseren Valley, facilitated by the deforested state of the slopes. The dominant erosion processes in this region are (shallow) landslides, sheet erosion, and erosion caused by land-use management (livestock, machinery, and manuring). Additional information on the Urseren Valley and occurring erosion processes can be found in Meusburger and Alewell [3], Zweifel et al. [7], Alewell et al. [52].

## 3. Data Sets

In the following we present the data sets used in our study. Table 1 summarises the data sets used for the mapping procedure conducted with U-Net which were also the basis for the training data set produced with OBIA [7].

**Table 1.** Summary of raster data sets used in this study. All geodata sets<sup>©</sup> Swisstopo. \*: The aerial image of 2016 was only used for validation purposes of the U-Net model.

Data Set	Derivative	Spectral Bands	Spatial Res.	<b>Recording Date</b>	
Aerial Image		Red, Green, Blue	0.5 m	24 August	2000
		Red, Green, Blue	0.5 m	9 September	2004
		Red, Green, Blue	0.25 m	20 July	2010
		Red, Green, Blue	0.25 m	1 August	2013
		Red, Green, Blue	0.25 m	20 July	2016 *
Digital Terrain	Slope		2 m	-	
Model (DTM)	Aspect		2 m		
	Curvature		2 m		

## 3.1. Aerial Imagery

The aerial images of SwissImage are high-resolution georeferenced orthophotos (product of Swisstopo [53]). Five aerial images covering the Urseren Valley were used in the time from 2000 to 2016. These images have a spatial resolution of 0.5 or 0.25 m (Table 1). Spectral information is available in the visible range (red, green, and blue spectral bands). All aerial images have slightly different properties (e.g., spatial resolution, colour distribution, and lighting conditions) but were always recorded during the growing season between late July and early September.

## 3.2. Digital Terrain Model

The digital terrain model (DTM) SwissALTI3D is the surface model of Switzerland without vegetation and development and has a spatial resolution of 2 m (product of Swisstopo [54]). Based on the elevation information of the DTM we derived the slope, aspect, and curvature (plan and profile) using ArcGIS (Version 10.5). The DTM provides valuable information and offers context to the aerial images. Zweifel et al. [7] have shown that for their study using OBIA, the DTM and its derivatives were essential for successful erosion mapping and classification.

## 3.3. Training Data

The data used to train the U-Net model consists of aerial imagery, DTM information, and training labels (see Section 4.3 for the training process). To train our U-Net model, a subsection (9 km<sup>2</sup>) of the Urseren Valley (26 km<sup>2</sup>) was used with the corresponding OBIA-mapped features (Figure 2). Four of the aerial images were used during training, leaving out the year 2016. By separating a subsection for training, we tested the spatial transferability of the model within the larger valley region. In addition, by omitting 2016, we investigated the spatial and temporal transferability when applying U-Net to a different image with properties not known during training.



**Figure 2.** Training (9 km<sup>2</sup>) and testing (17 km<sup>2</sup>) areas are marked on the aerial image with examples of OBIA training labels for 2000 (map on the left). On the right-hand side is an overview of all available years and the sections used for training and testing. All training areas contain OBIA training labels (not shown) for the respective years (2000–2013). Training labels vary for each year due to the continuous evolution of soil erosion sites. The entire area of the image taken in 2016 was used only for testing.

## 3.3.1. Training Labels

The training labels come in the form of mapped erosion sites with attributed erosion classes from a previous study by Zweifel et al. [7]. This data set was created with a semi-automatic method using an OBIA approach described in Section 4.1, which made use of the same aerial imagery and DTM information as used for U-Net. Mapped erosion objects are available for the entire Urseren Valley for all five aerial images (2000, 2004, 2010, 2013, and 2016). Based on random sample evaluation by experts, this data set has an average overall accuracy score of 85.4% [7]. The training labels consist of four different erosion classes: shallow landslides (areas with displaced topsoil layers and clear boundaries to the surrounding vegetation), livestock trails (elongated tracks caused by livestock trampling, mostly perpendicular to the slope), sheet erosion (patches with reduced vegetation cover), and management effects (large areas damaged by heavy machinery, over-fertilisation, or intense grazing in fenced-off areas) (see Figure 3).



**Figure 3.** Examples for the labels used for training the U-Net model. From left to right: shallow landslides, livestock trails, sheet erosion, management effects.

#### 4. Methodology

Our methodology consists of two major parts: the training process and the prediction process, with an overview depicted in Figure 4. To train the U-Net model we use OBIA labels together with the respective aerial image information (RGB) and DTM information for a dedicated training area (9 km<sup>2</sup>). U-Net assigns pixel-wise probability values and thus provides information about the likelihood of pixels belonging to a specific erosion class. Based on these probabilistic assignments, hard segmentations are produced by thresholding. The following sections will describe the methodology in further detail.



**Figure 4.** An overview of the developed workflow on the basis of U-Net showing examples of input files for training and prediction purposes. The output shows one of four erosion classes, namely, shallow landslides, with four different probability thresholds.

## 4.1. Object-Based Image Analysis

Object-based image analysis (OBIA) combines a segmentation algorithm with classification techniques ranging from decision trees to various supervised machine learning algorithms which assign generated segments (or object primitives) to erosion classes. We used the software eCognition Developer (version 9.3.2) implementing a multi-resolution segmentation algorithm for grouping pixels with similar properties to object primitives. Input data consisted of aerial imagery (RGB), the excess green vegetation index, and information from the DTM and its derivatives (slope, aspect, and curvature). The object primitives contained information on their spatial, spectral, textural, and contextual properties based on all input data. Given these extracted feature sets, a random forest classifier was trained on manually selected samples in order to identify bare soil sites or sites with reduced vegetation cover. Subsequently, an additional decision tree was assigned specific erosion classes based on the typical appearances of objects previously identified containing bare soil or reduced vegetation cover. These erosion classes consist of shallow landslides, livestock trails, sheet erosion, and management effects. Note that the entire workflow needed to be performed on every input image to accommodate for varying image properties. Therefore, OBIA-labels for different input images can be considered to be obtained from independent models (i.e., differently calibrated settings). A detailed description of the workflow is presented in Zweifel et al. [7].

## 4.2. Neural Network Architecture

In this study, we make use of the U-Net architecture [35] illustrated in Figure 5. U-Net is a fully convolutional neural network which consists of a contracting part and an expansive part. See Section S1 in the Supplement for more details on the main components of the neural network which are used in the following description.



**Figure 5.** The employed U-Net architecture: In the first (upper) part, the input is contracted into a compressed representation (right). In the second (lower) part, the compressed representation is expanded into a segmentation map with pixel-wise class probabilities. The input consists of the inpuxt RGB image (three channels) and the DTM derivative maps for the aspect, curvature, and slope (one channel each). The resulting output provides a segmentation map for each considered class: Shallow landslides (indicated by 1 in the output), livestock trail (2), sheet erosion (3), management effects (4), and a class for non-assignable pixels (5).

In the contracting part (upper part in Figure 5), a sequence of two convolutional layers with ReLU activations followed by max pooling layer processes the input. With each max pooling application, the sizes of the resulting feature maps are halved, while the number of features is doubled for the subsequent convolutional layer. In the expansive part (bottom part), a sequence of transposed convolutional layers with ReLU activations followed by two convolutional layers and ReLU activations is applied to restore the original image size. Feature maps from the contracting part are appended to the feature maps obtained through the transposed convolutions to provide fine-detail features in the expansive part. Finally, a  $1 \times 1$  convolutional layer followed by a pixel-wise softmax activation function provides the final segmentation output where each channel represents the segmentation map for the individual classes. The softmax function rescales the activations for each pixel to the [0, 1] interval. More explicitly, for a pixel f in the output map F, the softmax yields a prediction  $p_c(f)$  which can be interpreted as the probability of pixel f to belong to class  $c \in \{1, \ldots, C\}$ . The neural network is trained with the cross entropy loss which penalises incorrect class assignments with

$$-\frac{1}{N}\sum_{f\in F}\sum_{c\in C}y_c(f)\log(p_c(f))\tag{1}$$

where N = |F| is the number of pixels and  $y_c(f)$  is the ground truth class assignment for pixel f, i.e., 1 if c is the correct class and 0 otherwise. For any pixel f in the input image, the softmax prediction  $p(f) = (p_1(f), p_2(f), \dots, p_C(f))$  provides the probabilities for the classes  $c \in \{\text{shallow landslide}, \text{sheet erosion, livestock trail, management effect, non-assignable}\}$ —e.g.,

$$p(f) = \left(\underbrace{0.55, \ 0.1, \ 0.2, \ 0.05,}_{\text{erosion class probabilities}} 0.1\right).$$
(2)

In addition to the four erosion classes, a class for non-assignable pixels is introduced which represents the class for all remaining (potentially ambiguous or vegetation covered and thus stable) objects.

U-Net provides pixel-wise class probabilities like in Equation (2) as the probabilistic output. In the following, for each erosion class we will refer to the full-probability result when only entries for the specific class of this output are considered without applying a threshold (e.g., the first entries for shallow landslides). For the final hard segmentation, we would like to obtain the dominant erosion class and apply different probability thresholds that control to which extent candidate segments are obtained. We only consider the erosion classes and identify the class with the largest probability for pixel *f* as the dominant erosion class. If the selected erosion class probability does not meet the threshold, the respective pixel is considered as a background pixel. For example, in Equation (2), argmax  $\{0.55, 0.1, 0.2, 0.05\}$  implies that shallow landslide is the dominant class and pixel *f* is predicted to be a shallow landslide pixel with a probability of 55%. At a threshold of 0.5, the class probability exceeds the threshold and pixel *f* is assigned to the shallow landslide class, while with a stricter threshold of 0.6 the pixel is considered to be a background pixel. With this kind of threshold segmentation, the final erosion class labels are obtained.

#### 4.3. Training Process

In order to learn how to identify erosion sites, precise boundaries for the different erosion classes are required for training U-Net. Inadequate training labels can deteriorate the spatio-temporal generalisation capability of U-Net. In this study, we used high-quality training labels provided by the OBIA approach (see Section 3.3.1), and we considered the resulting erosion class areas as the ground truth segmentation in this investigation. To process the input images efficiently, we divided the aerial images into tiles of size  $194 \times 176$  m which correspond to  $388 \times 352$  pixels at 0.5 m resolution (2000, 2004) and  $776 \times 704$  pixels at 0.25 m resolution (2010, 2013, 2016). The same is done for the maps of the DTM derivatives aspect, curvature, and slope.

Adjacent tiles overlap such that a 20 m (40 and 80 pixels, respectively) margin of one tile is contained in an adjacent tile. Figure 6 illustrates the resulting tiles for different years. The higher resolution tiles were down-sampled so that all input tiles are of size  $388 \times 352$  pixels. No data augmentation was employed, as we expect object size and orientation (e.g., north/south exposure) to be relevant features. As described previously, U-Net was trained from scratch with tiles extracted from the training area of the years 2000 to 2013, with a total of 1292 training samples. A U-Net of depth 3 with initially 32 (root) filters was used (see Figure 5), resulting in 467,525 network parameters. The network was trained for 300 epochs with a batch size of 20, using the Adam optimizer [55] with a learning rate of 0.001 and a dropout rate of 0.1. We used TensorFlow version 1.10 [56] for our implementation which is based on the U-Net implementation by [57]. The full source code of our analysis pipeline is available under the GNU public license: https://github.com/bmda-unibas/ErosionSegmentation.



**Figure 6.** Example of input RGB images for training for the years 2000, 2004, 2010, and 2013 with a size of  $194 \times 176$  m (corresponding to  $388 \times 352$  pixels at 0.5 m resolution). The images show examples of eroded area on grassland slopes (livestock trails, shallow landslides). Below, the corresponding aspect, curvature, and slope maps are displayed (for all years the same DTM information is used). To obtain the samples, the aerial images of the respective years (Figure 2) and the DTM derivatives were divided into smaller tiles.

#### 4.4. Details on the Evaluation

For the evaluation, only sites with an area of at least 4 m<sup>2</sup> were considered, which we treated as the minimum reasonable object size, and this is in line with the definition used in Zweifel et al. [7]. After choosing an appropriate probability threshold, the quality of the segmentation results was assessed with the precision score (producer's accuracy), recall score (user's accuracy), and their harmonic mean, the  $F_1$  score. We considered objects which overlap in both the OBIA and U-Net results as true positives and weigh true positives, false positives, and false negatives by the areas of the respective segments. Ultimately, our goal was to evaluate the total degraded area on the held-out test area of the training years (2000, 2004, 2010, and 2013) and the validation year 2016 in comparison to the OBIA ground truth results. The emphasis here was to study the temporal trend and relative increase in degraded area as obtained from the different methods. We performed a linear regression to provide the linear trend over the time period from 2000 to 2016.

## 5. Results and Discussion

U-Net provides pixel-wise probabilities for each erosion class, which allows for assessing the certainty of predictions by studying the resulting heatmaps (see Figure 7 for an example). In practice, this rich information is further post-processed by applying a threshold on the pixel-wise probabilities to form well-delineated segments. In the following, we present both results on the (full-probability) heatmaps and results obtained with a selection of different probability thresholds. The latter enables a more direct comparison to the segmentation results obtained with OBIA. All results were obtained on the held-out test area (see Figure 2). Note that the data from 2016 was not used for training.

#### 5.1. Segmentation of Soil Erosion Sites

The trained U-Net provides satisfying segmentation results which are demonstrated in Figure 7 for exemplary segments of shallow landslides and livestock trails. The heatmaps illustrate the full-probability output of U-Net and display the certainty in the class assignment (upper panel). By selecting different thresholds, hard class assignments can be achieved which lead to slightly different segment shapes depending on the threshold (lower panel). We selected thresholds of 0.2 and

0.8 to display the impacts of a wide range of probability thresholds on the delineation of segments. In general, choosing lower thresholds allows for the identification of a large number of potential erosion sites, while a higher threshold reduces the number of segments and also has an effect on the margins of these object, i.e., shrinks the segments to the most certain area. The probability threshold is a free parameter which can be chosen guided by application requirements or user preferences, or in our case to match baseline results (OBIA).



**Figure 7.** Visualisation of U-Net mapped shallow landslides (**left**) and livestock trails (**right**) for 2016. The lower panel shows segmentation results with different probability thresholds: the lighter colour indicates a lower probability threshold (0.2) and the darker colour indicates a higher probability threshold (0.8). Lower thresholds lead to larger and more numerous segments. For the same region (background omitted for better visualisation), the upper panel shows the full-probability heatmap output of U-Net: darker colours indicate higher probabilities.

In order to evaluate the accuracy of the proposed U-Net approach, we consider the OBIA results for 2016 as the ground truth baseline, which are independent of all other years, as OBIA was separately applied to the aerial image of 2016. For the comparison, we selected a threshold value of 0.3, as this led to the best agreement between U-Net and OBIA segments with respect to the total degraded area (see Section 5.3). OBIA relies on a dedicated, multi-resolution segmentation algorithm which provides clear objects to start with, which can then be classified. In contrast to OBIA, the U-Net approach does not have such a procedure and thus provides less control over segment shapes, as these are determined by pixel-wise thresholding. Consequently, there are cases in which both OBIA and U-Net identify areas as erosion sites but the boundaries of these objects might differ slightly. In that respect, our results for U-Net show that erosion sites with clear, unambiguous boundaries such as shallow landslides (and some very clear cases of livestock trails) generally have better overlaps with the OBIA baseline and contiguous objects are better identified (see Figure 8 on the left). Boundaries of more diffuse erosion sites predicted by U-Net show a slight mismatch with the OBIA baseline (see Figure 8 on the right). In these cases, the correct delineation of sites belonging to management effects or sheet erosion is in general a challenging task which is mirrored in the less accurate matching of the segmentation results from the different methods. Additionally, these erosion classes have similar appearances and are comprised of either bare soil or vegetation areas with strongly reduced vegetation cover which are prone to similar erosion processes (mainly erosion by water run-off), and they differ only in the origin

of the damage. Management affected sites are mostly located near the foot of the slope, are mainly used for the production of hay, and can show signs of heavy machinery usage. Sheet erosion, on the other hand, can be found throughout the entire valley and can be caused not only by livestock trampling and grazing, but also climate-related factors, such as drought, precipitation, and snow-melt [7,58–61]. Still, both methods are able to identify a great majority of overlapping objects. More quantitatively, we obtained scores for a threshold value of 0.3, as presented in Table 2.



**Figure 8.** Comparison of segmentation results of OBIA and U-Net (probability threshold of 0.3) for the aerial image of the year 2016. This aerial image was not used during training of the U-Net model and depicted sections are located in the held-out test area. Lighter colours show OBIA results; darker colours (shaded) are results of U-Net.

**Table 2.** Scores for U-Net with a threshold value of 0.3 for the validation aerial image of 2016. U-Net results are compared to OBIA baseline results.

Scores	U-Net	
Recall	84%	
Precision	73%	
F1	78%	

The precision score indicates that 73% of the predicted U-Net segments have corresponding OBIA segments, and about 27% of predicted U-Net segments do not directly correspond to any OBIA segments. On the other hand, the recall displays that 84% of the OBIA segments are maintained and the remaining 16% of OBIA segments are not identified by U-Net. Both these findings suggest that U-Net successfully identifies a majority of OBIA segments (recall score), but provides more segmented erosion sites than OBIA (false positives). Segments contributing to the 27% false positives still can be valid erosion sites which are not captured by OBIA, as it is known that OBIA tends to give a conservative estimate of the degraded soil [7]. Therefore, it is important to note that these scores mainly highlight the difference between U-Net segmentation with respect to the OBIA segmentation baseline, and it is possible that one method captures valid erosion sites which the other method misses (see example shown in Figure 9 on the right).



**Figure 9.** Examples of two different types of false positives: On the left-hand side, U-Net identifies some rock surfaces as sheet erosion (yellow) and shallow landslides (purple). For both erosion classes, thresholds of 0.2 and 0.8 are shown. Lower threshold choices are linked to more of such false positives. Depicted on the right-hand side are livestock trails with OBIA and U-Net (threshold of 0.2). Here, U-Net is capable of identifying more livestock trails correctly compared to OBIA.

Most cases of false positive predictions can be related to objects which are similar in appearance to the erosion classes, and the reason for misclassification can be recognised in many cases upon manual inspection. False positives are typically patches with rocks located at higher elevations which are classified as shallow landslides (see Figure 9 on the left), or varied classification of sites affected by management and sheet erosion. Nonetheless, singular rocks on grassland areas are successfully left unclassified. These kinds of disagreements are inherent to the U-Net approach, which attempts to identify regularities in the training data and thereby includes objects which share some similarities. In clear cases, such as very small object sizes or predictions at certain altitudes where a particular class of erosion phenomena is not expected, a post-processing step can address these erroneous classifications. Another way of avoiding segmentation ambiguities is to employ pre-processing steps to identify sub-regions of interest for target objects which share some kind of regularity in their appearance, for instance, in the shape of the objects [62]. For the purpose of this study, however, no pre-processing steps were used in the U-Net procedure to ensure objective comparison with OBIA.

#### 5.2. Threshold Selection

In similar studies, the matter of threshold selection is usually not addressed or a fixed threshold value is used. This can be suitable for studies with binary output classes (e.g., [37,48]), but can also be problematic for gradual transitions of classified objects, as discussed by Kattenborn et al. [46]. Other studies employ deep learning approaches for classification of the object primitives in the OBIA framework where object boundaries are already well-defined [63–65]. In our setting, threshold selection can be used to adjust segmentation results in relation to pre-existing knowledge (i.e., segmentation results of other methods such as OBIA), which led to the best fit with a threshold selection of 0.3 for this study (with respect to the total degraded area; see Section 5.3). Additionally, varying thresholds may be applied to make necessary adjustments for different classes with varying appearances. As a standard comparison, a held-out data set of the ground truth segmentation required for training can be used to determine appropriate probability thresholds if necessary. In the absence of appropriate pre-existing knowledge or in cases where visual assessment is not possible, it is advisable to use a range of probability thresholds which capture a variety of segment estimations and assess uncertainty ranges of the estimates.

#### 5.3. Trend Analysis of Soil Erosion Sites

In order to study the temporal trend in the extent of soil degradation, we applied U-Net to the series of five aerial images of the Urseren Valley between 2000 and 2016 (see Section 3.3.1). We compare the full-probability U-Net results and the results for the different thresholds to the baseline results of the OBIA approach in Figure 10. In the first case, the heatmap results are added up to form an estimate of degraded area per erosion class. The resulting outcomes of the full-probability U-Net output match the OBIA results closely with respect to the total degraded area. Due to their methodological differences, slight deviations in the segmentation results and the resulting (total) degraded area were expected. The same holds true for the U-Net results with a threshold of 0.3. This threshold was identified to exhibit the most suitable agreement with OBIA segmentation results with respect to the total degraded area. It can be observed that for validation year 2016, the OBIA and U-Net threshold 0.3 results agree very well (in the shaded area in right plot of Figure 10). As expected, the U-Net results display an increase in degraded area for decreasing thresholds. Nevertheless, in all considered U-Net results, the same temporal trends of decrease and increase from one year to another are observed, as in the OBIA baseline. This observation is also supported by the linear regression results, which in all cases provide similar linear temporal trends.



**Figure 10.** Linear trend of the total degraded area in the held-out test region (see Figure 2) as obtained with the OBIA and U-Net approaches. On the left, the results for a range of different threshold values are displayed; on the right the results for the suitable threshold value 0.3 and the full-probability results are given. Qualitatively, a similar increase or decrease of degraded soil in the individual years is retained in all models. The linear interpolation provides a similar temporal trend of increase in degraded soil in all cases. In particular, the full-probability and threshold 0.3 results of the U-Net approach show good agreement with the OBIA baseline. The linear trends with lower and higher thresholds surround the OBIA result. The years 2000 to 2013 provide a result on the spatial generalisation of U-Net (years used for training), while the result for 2016 (shaded column) in addition provides a temporal generalisation result (aerial image of 2016 was not used for training). Note that the OBIA approach needs to be trained on all aerial images.

In order to quantify the relative increase in degraded area, we consider the values for 2000 and 2016 obtained from the linear regression line. Again, the threshold dependency with respect to the total degraded area is observed (top panel in Figure 11). However, for the relative increase in degraded area (quotient of values for 2016 and 2000), the results become mostly independent of the selected threshold (bottom panel in Figure 11). To assess the statistical uncertainty of the linear regression fit and thus the relative increase, one standard deviation each of the fitted parameters

(slope and intercept) is considered to obtain the two most extreme linear trends which are possible within the uncertainty of the fitted parameters. This means the steepest and flattest linear trends with respect to one standard deviation in the parameters are identified, which leads to the error bars for the total degraded area as depicted in Figure 11. As the relative increase considers the ratio of these quantities, the error bars are relatively larger for the relative increase of degraded area. In particular, for a threshold of 0.8, the statistical uncertainty increases due to the comparably small degraded area detected. The obtained U-Net results show similar relative increases of degraded area which fall within the uncertainty range of each other depicting the statistical uncertainty in the linear regression fit (one standard deviation). The U-Net results are in good agreement compared to the baseline method, with an increase of 167% in the test region. This in turn is in line with the increase of  $156 \pm 18\%$  reported in Zweifel et al. [7] for the full Urseren Valley, where  $\pm 18\%$  depicts the estimated propagated error based on expert accuracy assessment (and not the statistical uncertainty in the linear regression fit). Importantly, it has been established that OBIA tends to underestimate the extent of degraded soil [7]. Therefore, the steeper relative increase obtained by the U-Net results is plausible and potentially reflects the increase of degraded area more accurately. Furthermore, the fact that the relative increases for the different probability thresholds coincide with each other within the statistical uncertainty of one standard deviation of the linear regression fit is further evidence for the applicability and robustness of the U-Net approach. Assessing the relative development of aggregated measures, such as the total area of degraded soil, is therefore less sensitive to the choice of threshold. The results on the linear trend (Figure 10) and the relative increase of total degraded area (Figure 11) highlight that the probabilistic output of U-Net aligns with the OBIA results very well, and to study these quantities by choosing a threshold, i.e., hard segmentation, is not required. In our investigation we assess predictions in the held-out test region (see Figure 2) for two validation cases: (i) testing the erosion site prediction of the test region for years for which conditions (colour, shading, vegetation, etc.) were available during training (2000–2013) and (ii) testing the predictions for a new year for which conditions were unknown during training (2016). In the first case, our results provide evidence that the trained U-Net transfers well to adjacent regions with similar conditions, as observed during training, and that shows the spatial generalisation capability of the U-Net approach. Furthermore, the latter validation case gives evidence of suitable erosion site segmentation with the U-Net approach in completely new aerial images with conditions not encountered during training, which in addition highlights the temporal generalisation capability of the approach.



**Figure 11.** Comparison of total degraded area in years 2000 and 2016 for the baseline (OBIA) and the U-Net approach with different thresholds. The total degraded area was obtained from the interpolation results of each year (**top panel**). In all approaches, an increase of degraded area in the Urseren Valley is observed with threshold-specific differences in the total extent. However, the relative increase in degraded area (**bottom panel**) shows that assessing the trend of soil degradation can be done independently of the threshold, as all results fall within the statistical uncertainty of the linear regression fit. Note that the statistical uncertainty for U-Net 0.8 increases due to the comparably small total degraded area detected. The error bars depict the statistical uncertainty of one standard deviation.

For the individual erosion classes, we examine the results for the full U-Net model output and for a threshold of 0.3 (see Figure 12). Especially for sheet erosion and management effects, which contribute to a great amount of the total degraded area, the choice of 0.3 as a threshold for the hard segmentation is appropriate. In the case of livestock trails, the full-probability U-Net results capture the behaviour in the baseline more appropriately. The individual results highlight that an erosion-class-specific choice of the probability threshold can be reasonable in applications such as ours. We provide a result on such a mixture of thresholds for the linear trend for the years 2000 to 2016 exhibits good agreement with the OBIA baseline (similar to Figure 10 on the right). Therefore, although the temporal development of aggregated measures is less dependent on the threshold, choosing different probability thresholds enables flexibility in the number of identified segments and segment boundaries in the U-Net approach. This is especially the case when examining the temporal development with regard to the degraded area per individual erosion class (Figure 12).



**Figure 12.** Mapped degraded area in the test region by erosion class for both the OBIA and U-Net methods (full-probability results and threshold value 0.3). Comparing the two methods, class-specific differences for the yearly degraded area and linear trends can be observed. Moreover, by selecting appropriate thresholds for each erosion class, similar linear trends in both methods can be attained (see Supplementary Figure S2). The years 2000 to 2013 provide a result on the spatial generalisation of U-Net (years used for training), while the result for 2016 (shaded column) in addition provides a temporal generalisation result (aerial image of 2016 was not used for training).

#### 5.4. Deep Learning and OBIA

Deep learning methods for similar applications are predominantly trained with manual labels, and often the objects of interest are precisely defined, such as roads, buildings, or damaged trees in forests [37,43,47]. In our application, the objects are less clearly defined, and some of the segment boundaries concerning both the mapped and omitted areas might be more disputable. The boundaries of objects are often ambiguous due to smooth transitions, especially for erosion sites with reduced vegetation cover. Imprecise delineation of the objects of interest negatively impacts the generalisation capability and applicability of deep learning techniques, and can potentially be a limiting factor for this kind of approach. In particular, it can have a detrimental effect on the accuracy of the U-Net approach if the ground truth misses a great number of relevant objects. Therefore, we do not rely on manual labels of the objects of interest, which might suffer from subjective assessments, require labour-intensive work, and usually are unable to achieve pixel-level precision. Instead, we showcase that any kind of segmentation technique, such as OBIA in our study, can be used as a basis to provide training data to

successfully employ a convolutional neural network for segmentation of natural features, such as the erosion sites in our application.

Similar studies have compared OBIA to deep learning approaches for the detection of landslides on remotely sensed data with the goal of enabling large-scale analysis. In Prakash et al. [66] the comparison was done on the basis of landslide inventories. A study of different machine learning and deep learning methods was conducted by Ghorbanzadeh et al. [67], who used field observations with manual corrections as the ground truth segments. These studies show that deep learning approaches improve segment detection by comparison of the segmentation performances of the different methods. In our study, we leverage the fact that OBIA is a well-suited approach for segmentation tasks on small scales, and thus derive our baseline trends and the ground truth segments from it. Other work like the detection of shrubs on high resolution satellite imagery by Guirado et al. [62] similarly shows that CNN approaches can outperform OBIA in certain cases. That study relied on manually delineated ground truth segments and used dedicated pre-processing steps to identify regions of interest to perform classification of candidate patches. Combining OBIA and CNN approaches was also studied with regard to using CNNs in the classification step of the OBIA framework [63] or using features learned by the CNNs to improve inputs to the OBIA workflow [68]. In our study, OBIA provides the necessary high-quality ground truth segmentation, but our workflow is not bound to OBIA, and any other reliable approach can be used for this too.

The presented results of this study substantiate that the U-Net approach can perform on a par with OBIA. Moreover, the transferability to new data, the insensitivity of trends in aggregated measures to threshold selection, and the flexibility of fitting the U-Net results to existing knowledge or competing segmentation methods—apart from manual inspection of segmentation results—render the proposed approach advantageous for a great variety of applications. Furthermore, large-scale analysis is facilitated by improved running times. For training and prediction, an Nvidia GeForce Titan X Pascal GPU was used. In our study, training required a running time of approximately 6.5 h, while the prediction for the full Urseren Valley took 12 min. This is a significant improvement over the semi-automatic OBIA approach, which takes up to a few days to achieve satisfying results for the Urseren Valley. For large-scale studies (e.g., alpine-wide analysis) the process can efficiently be parallelised using several GPUs, resulting in even faster prediction times.

#### 6. Conclusions

While OBIA is the state-of-the-art approach for mapping objects on remotely sensed images, it suffers from limitations that render this approach unsuitable for larger-scale studies. High-quality segmentation results come at the expense of a lack of transferability of parameter settings from one input image to another, manual adjustments, and a need for expert knowledge in applying the method to the specific task which together lead to long processing times. In particular, the first aspect generally hinders OBIA in a predictive setting for new images. To overcome these shortcomings and enable large-scale analysis, we compared OBIA to a fully convolutional neural network approach which learns relevant features for segmentation by itself and thereby emulates some of the expert knowledge necessary to apply OBIA. We demonstrated that the U-Net approach is capable of performing as well as OBIA with respect to identifying trends in the spatial and temporal development of degraded soil, and can therefore replace OBIA in large-scale studies. Spatial patterns and temporal trends of both methods agree well; nevertheless, some generated segmentation results might partially not overlap  $(F_1 = 78\%)$ . Specifically, we show that U-Net (threshold 0.3) provides a potentially more accurate relative increase of total degraded area in the Urseren Valley than the more conservative estimates of OBIA (201% vs. 167%). This novel approach allows for individual threshold choices for the most successful representation of ongoing soil erosion processes. This is typically possible if some prior knowledge about erosion processes and the spatial extent of degraded soil is available, or if visual assessment is feasible, to which probability thresholds can be calibrated. In our study, we made use of training labels generated with OBIA. However, any kind of (high-)quality training labels can be

used, and the U-Net erosion site segmentation is not limited to combined use with OBIA. In summary, we show that with our approach we can perform erosion site prediction close to similar approaches such as OBIA which provide accurate segmentation results on small scales. A particular strength of the proposed approach is that similar trends are achieved with a more efficient, automatic, and objective method for mapping erosion sites. We require the U-Net approach to be trained only once and obtain much better transferability of the method to new images. Moreover, the approach is insensitive to the threshold choice with respect to trends of aggregated measures, and the improved running times make large-scale analysis of soil erosion is Swiss alpine grasslands feasible.

Still, our model is only as good as the training data; i.e., high-quality training data are important for adequate U-Net performance. Future studies should include a variety of different sample regions to incorporate relevant erosion-type-specific conditions during training (e.g., orientation of erosion sites). Furthermore, U-Net can use as many layers of information as required. A unique feature of fully convolutional neural networks is that inputs of any size and any number of channels can be used, i.e., RGB images with DTM derivatives. Additional maps can be easily incorporated (see Figure 5), which might include more information, such as environmental properties or images with additional spectral information. In that regard, U-Net has the advantage of continual learning; i.e., it can be trained further to incorporate conditions of completely new regions and erosion-type-specific properties. Generally, the U-Net model can be employed in a similar fashion for other segmentation tasks in remote sensing and other inputs, such as UAV or satellite imagery. The requirement for the input data is that the spatial resolution allows for identifying the target objects well enough.

**Supplementary Materials:** The following are available online at http://www.mdpi.com/2072-4292/12/24/ 4149/s1 . Section S1: Details on the neural network architecture, Figure S1: Illustration of the max pooling and convolution operation, Figure S2: Mixed thresholds for trend analysis.

Author Contributions: Conceptualisation, M.S., L.Z., V.R., and C.A.; methodology, M.S. and L.Z.; software, M.S.; validation, M.S. and L.Z.; formal analysis, M.S. and L.Z.; investigation, M.S. and L.Z.; data curation, M.S. and L.Z.; writing—original draft preparation, M.S. and L.Z.; writing—review and editing, M.S., L.Z., V.R., and C.A.; visualization, M.S. and L.Z.; supervision, V.R. and C.A.; project administration, M.S. and L.Z.; funding acquisition, V.R. and C.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Swiss National Science Foundation with the grant number 407540 167333 as part of the Swiss National Research Programme NRP 75 "Big Data".

Acknowledgments: The authors would like to thank the Swiss National Science Foundation for supporting the research. Calculations were performed at sciCORE (http://scicore.unibas.ch/) scientific computing core facility at University of Basel. We also want to acknowledge Swisstopo and MeteoSwiss for providing the data sets we used. Furthermore, we would like to thank the anonymous reviewers for their comments and suggestions which helped us to improve the paper.

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

## Abbreviations

The following abbreviations are used in this manuscript:

- OBIA Object-based image analysis
- DTM Digital terrain model
- RGB Red, Green and Blue spectral bands
- CNN Convolutional Neural Networks
- U-Net Name of Convolutional Neural Network architecture
- GPU Graphics Processing Unit
- UAV Unmanned Aerial Vehicle

## References

- 1. EEA. *Regional Climate Change and Adaptation—The Alps Facing The Challenge of Changing Water Resources;* Technical Report 8; European Environmental Agency: Copenhagen, Denmark, 2009.
- 2. Fuhrer, J.; Beniston, M.; Fischlin, A.; Frei, C.; Goyette, S.; Jasper, K.; Pfister, C. Climate risks and their impact on agriculture and forests in Switzerland. *Clim. Chang.* **2006**, *79*, 79–102. [CrossRef]
- Meusburger, K.; Alewell, C. Impacts of anthropogenic and environmental factors on the occurrence of shallow landslides in an alpine catchment (Urseren Valley, Switzerland). *Nat. Hazards Earth Syst. Sci.* 2008, *8*, 509–520. [CrossRef]
- 4. Nearing, M.; Pruski, F.; O'Neal, M. Expected climate change impacts on soil erosion rates: A review. *J. Soil Water Conserv.* **2004**, *59*, 43–50.
- 5. Scheurer, K.; Alewell, C.; Bänninger, D.; Burkhardt-holm, P. Climate and land-use changes affecting river sediment and brown trout in alpine countries—A review. *Environ. Sci. Pollut. Res.* 2009, *16*, 232–242. [CrossRef]
- 6. Tasser, E.; Mader, M.; Tappeiner, U. Effects of land use in alpine grasslands on the probability of landslides. *Basic Appl. Ecol.* **2003**, *4*, 271–280. [CrossRef]
- 7. Zweifel, L.; Meusburger, K.; Alewell, C. Spatio-temporal pattern of soil degradation in a Swiss Alpine grassland catchment. *Remote Sens. Environ.* **2019**, 235, 111441. [CrossRef]
- 8. Apollo, M.; Andreychouk, V.; Bhattarai, S.S. Short-term impacts of livestock grazing on vegetation and track formation in a high mountain environment: A case study from the Himalayan Miyar Valley (India). *Sustainability* **2018**, *10*, 951. [CrossRef]
- 9. Torresani, L.; Wu, J.; Masin, R.; Penasa, M.; Tarolli, P. Estimating soil degradation in montane grasslands of North-eastern Italian Alps (Italy). *Heliyon* **2019**, *5*, e01825. [CrossRef]
- 10. Wiegand, C.; Geitner, C. Shallow erosion in grassland areas in the Alps. What we know and what we need to investigate further. In *Challenges for Mountain Regions: Tackling Complexity*; Boehlau Verlag: Wien, Austria, 2010; p. 76–83.
- 11. Alder, S.; Prasuhn, V.; Liniger, H.; Herweg, K.; Hurni, H.; Candinas, A.; Gujer, H.U. A high-resolution map of direct and indirect connectivity of erosion risk areas to surface waters in Switzerland-A risk assessment tool for planning and policy-making. *Land Use Policy* **2015**, *48*, 236–249. [CrossRef]
- 12. Bircher, P.; Liniger, H.P.; Prasuhn, V. Comparing different multiple flow algorithms to calculate RUSLE factors of slope length (L) and slope steepness (S) in Switzerland. *Geomorphology* **2019**, *346*, 106850. [CrossRef]
- 13. Meusburger, K.; Bänninger, D.; Alewell, C. Estimating vegetation parameter for soil erosion assessment in an alpine catchment by means of QuickBird imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 201–207. [CrossRef]
- 14. Meusburger, K.; Steel, A.; Panagos, P.; Montanarella, L.; Alewell, C. Spatial and temporal variability of rainfall erosivity factor for Switzerland. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 167–177. [CrossRef]
- 15. Prasuhn, V.; Liniger, H.; Gisler, S.; Herweg, K.; Candinas, A.; Clément, J.P. A high-resolution soil erosion risk map of Switzerland as strategic policy support system. *Land Use Policy* **2013**, *32*, 281–291. [CrossRef]
- 16. Schmidt, S.; Alewell, C.; Meusburger, K. Mapping spatio-temporal dynamics of the cover and management factor (C-factor) for grasslands in Switzerland. *Remote Sens. Environ.* **2018**, *211*, 89–104. [CrossRef]
- 17. Schmidt, S.; Alewell, C.; Meusburger, K. Monthly RUSLE soil erosion risk of Swiss grasslands. *J. Maps* **2019**, *15*, 247–256. [CrossRef]
- 18. Schmidt, S.; Tresch, S.; Meusburger, K. Modification of the RUSLE slope length and steepness factor (LS-factor) based on rainfall experiments at steep alpine grasslands. *MethodsX* **2019**, *6*, 219–229. [CrossRef]
- Fischer, F.K.; Kistler, M.; Brandhuber, R.; Maier, H.; Treisch, M.; Auerswald, K. Validation of official erosion modelling based on high-resolution radar rain data by aerial photo erosion classification. *Earth Surf. Proc. Land* 2018, 43, 187–194. [CrossRef]
- 20. Alewell, C.; Borrelli, P.; Meusburger, K.; Panagos, P. Using the USLE: Chances, challenges and limitations of soil erosion modelling. *Int. Soil Water Conserv. Res.* **2019**, *7*, 203–225. [CrossRef]
- 21. D'Oleire-Oltmanns, S.; Marzolff, I.; Peter, K.D.; Ries, J.B. Unmanned aerial vehicle (UAV) for monitoring soil erosion in Morocco. *Remote Sens.* **2012**, *4*, 3390–3416. [CrossRef]
- 22. Eisank, C.; Hölbling, D.; Friedl, B.; Chin, Y. Expert knowledge for object-based landslide mapping in Taiwan. *South-Eastern Eur. J. Earth Observ.* **2014**, *3*, 347–350.

- 23. Guzzetti, F.; Mondini, A.C.; Cardinali, M.; Fiorucci, F.; Santangelo, M.; Chang, K.T. Landslide inventory maps: New tools for an old problem. *Earth-Sci. Rev.* **2012**, *112*, 42–66. [CrossRef]
- 24. Hölbling, D.; Friedl, B.; Eisank, C. An object-based approach for semi-automated landslide change detection and attribution of changes to landslide classes in northern Taiwan. *Earth Sci. Inform.* **2015**, *8*, 327–335. [CrossRef]
- 25. Hölbling, D.; Betts, H.; Spiekermann, R.; Phillips, C. Identifying spatio-temporal landslide hotspots on North Island, New Zealand, by analyzing historical and recent aerial photography. *Geosciences* **2016**, *6*, 48. [CrossRef]
- 26. Hölbling, D.; Abad, L.; Dabiri, Z.; Prasicek, G.; Tsai, T.t.; Argentin, A.l. Mapping and analyzing the evolution of the Butangbunasi landslide using Landsat time series with respect to heavy rainfall events during Typhoons. *Appl. Sci.* **2020**, *10*, 630. [CrossRef]
- Martha, T.R.; Kerle, N.; van Westen, C.J.; Jetten, V.; Vinod Kumar, K. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. *ISPRS J. Photogramm. Remote Sens.* 2012, 67, 105–119. [CrossRef]
- 28. Shruthi, R.B.V.; Kerle, N.; Jetten, V. Object-based gully feature extraction using high spatial resolution imagery. *Geomorphology* **2011**, *134*, 260–268. [CrossRef]
- 29. Wang, B.; Zhang, Z.; Wang, X.; Zhao, X.; Yi, L.; Hu, S. Object-based mapping of gullies using optical images: A case study in the black soil region, Northeast of China. *Remote Sens.* **2020**, *12*, 487. [CrossRef]
- 30. Wiegand, C.; Rutzinger, M.; Heinrich, K.; Geitner, C. Automated extraction of shallow erosion areas based on multi-temporal ortho-imagery. *Remote Sens.* **2013**, *5*, 2292–2307. [CrossRef]
- 31. Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **2019**, 152, 166–177. [CrossRef]
- 32. Heydari, S.S.; Mountrakis, G. Meta-analysis of deep neural networks in remote sensing: A comparative study of mono-temporal classification to support vector machines. *ISPRS J. Photogramm. Remote Sens.* **2019**, 152, 192–210. [CrossRef]
- 33. Huang, B.; Zhao, B.; Song, Y. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* **2018**, *214*, 73–86. [CrossRef]
- Yuan, Q.; Shen, H.; Li, T.; Li, Z.; Li, S.; Jiang, Y.; Xu, H.; Tan, W.; Yang, Q.; Wang, J.; et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* 2020, 241, 111716. [CrossRef]
- 35. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-assisted intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015; Volume 9351, pp. 234–241.
- Yuan, M.; Liu, Z.; Wang, F. Using the wide-range attention u-net for road segmentation. *Remote Sens. Lett.* 2019, 10, 506–515. [CrossRef]
- Zhang, Z.; Liu, Q.; Wang, Y. Road Extraction by Deep Residual U-Net. *IEEE Geosci. Remote Sens. Lett.* 2018, 15, 749–753. [CrossRef]
- 38. Alshaikhli, T.; Liu, W.; Maruyama, Y. Automated method of road extraction from aerial images using a deep convolutional neural network. *Appl. Sci.* **2019**, *9*, 4825. [CrossRef]
- 39. Wulamu, A.; Shi, Z.; Zhang, D.; He, Z. Multiscale Road Extraction in Remote Sensing Images. *Comput. Intel. Neurosci.* **2019**, 2019, 1–9. [CrossRef]
- 40. Xu, Y.; Wu, L.; Xie, Z.; Chen, Z. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. *Remote Sens.* **2018**, *10*, 144. [CrossRef]
- 41. Yi, Y.; Zhang, Z.; Zhang, W.; Zhang, C.; Li, W.; Zhao, T. Semantic segmentation of urban buildings from VHR remote sensing imagery using a deep convolutional neural network. *Remote Sens.* **2019**, *11*, 1774. [CrossRef]
- 42. Ivanovsky, L.; Khryashchev, V.; Pavlov, V.; Ostrovskaya, A. Building detection on aerial images using U-NET neural networks. In Proceedings of the Conference of Open Innovation Association, FRUCT, Moscow, Russia, 8–12 April 2019; pp. 116–122.
- 43. Mboga, N.; Georganos, S.; Grippa, T.; Lennert, M.; Vanhuysse, S.; Wolff, E. Fully convolutional networks and geographic object-based image analysis for the classification of VHR imagery. *Remote Sens.* **2019**, *11*, 597. [CrossRef]
- 44. Yang, J.; Guo, J.; Yue, H.; Liu, Z.; Hu, H.; Li, K. CDnet: CNN-based cloud detection for remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6195–6211. [CrossRef]

- Flood, N.; Watson, F.; Collett, L. Using a U-net convolutional neural network to map woody vegetation extent from high resolution satellite imagery across Queensland, Australia. *Int. J. Appl. Earth Obs.* 2019, *82*, 101897. [CrossRef]
- Kattenborn, T.; Eichel, J.; Fassnacht, F.E. Convolutional Neural Networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution UAV imagery. *Sci. Rep.* 2019, 9, 17656. [CrossRef] [PubMed]
- 47. Hamdi, Z.M.; Brandmeier, M.; Straub, C. Forest damage assessment using deep learning on high resolution remote sensing data. *Remote Sens.* **2019**, *11*, 1976. [CrossRef]
- 48. Baumhoer, C.A.; Dietz, A.J.; Kneisel, C.; Kuenzer, C. Automated extraction of antarctic glacier and ice shelf fronts from Sentinel-1 imagery using deep learning. *Remote Sens.* **2019**, *11*, 2529. [CrossRef]
- 49. Bundzel, M.; Jaščur, M.; Kováč, M.; Lieskovský, T.; Sinčák, P.; Tkáčik, T. Semantic segmentation of airborne lidar data in maya archaeology. *Remote Sens.* **2020**, *12*, 1–22. [CrossRef]
- 50. Wyss, R. Die Urseren-Zone—Lithostratigraphie und Tektonik. Eclogae Geol. Hel. 1986, 79, 731–767.
- 51. IUSS Working Group WRB. *World Reference Base for Soil Resources;* IUSS Working Group WRB: Wageningen, The Netherlands, 2006; pp. 1–128.
- 52. Alewell, C.; Egli, M.; Meusburger, K. An attempt to estimate tolerable soil erosion rates by matching soil formation with denudation in Alpine grasslands. *J. Soils Sediments* **2015**, *15*, 1383–1399. [CrossRef]
- 53. Swisstopo. Swissimage. Das Digitale Farborthophotomosaik der Schweiz; Swisstopo: Wabern, Switzerland, 2010.
- 54. Swisstopo. SwissALTI3D. Das hoch aufgelöste Terrainmodell der Schweiz; Swisstopo: Wabern, Switzerland, 2014.
- 55. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015; pp. 1–15.
- 56. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. TensorFlow: A System for Large-Scale Machine Learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), Savannah, GA, USA, 1 November 2016; pp. 265–283.
- 57. Akeret, J.; Chang, C.; Lucchi, A.; Refregier, A. Radio frequency interference mitigation using deep convolutional neural networks. *Astron. Comput.* **2017**, *18*, 35–39. [CrossRef]
- Alewell, C.; Meusburger, K.; Brodbeck, M.; Bänninger, D. Methods to describe and predict soil erosion in mountain regions. *Landsc. Urban Plan.* 2008, *88*, 46–53. [CrossRef]
- 59. Meusburger, K.; Alewell, C. *Soil Erosion in the Alps*; Federal Office for the Environment FOEN: Bern, Switzerland, 2014; p. 118.
- 60. Konz, N.; Baenninger, D.; Konz, M.; Nearing, M.; Alewell, C. Process identification of soil erosion in steep mountain regions. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 675–686. [CrossRef]
- 61. Konz, N.; Prasuhn, V.; Alewell, C. On the measurement of alpine soil erosion. *Catena* **2012**, *91*, 63–71. [CrossRef]
- Guirado, E.; Tabik, S.; Alcaraz-Segura, D.; Cabello, J.; Herrera, F. Deep-learning Versus OBIA for scattered shrub detection with Google Earth Imagery: Ziziphus lotus as case study. *Remote Sens.* 2017, *9*, 1220. [CrossRef]
- Fu, Y.; Liu, K.; Shen, Z.; Deng, J.; Gan, M.; Liu, X.; Lu, D.; Wang, K. Mapping impervious surfaces in town-rural transition belts using China's GF-2 imagery and object-based deep CNNs. *Remote Sens.* 2019, 11, 280. [CrossRef]
- 64. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sens. Environ.* **2018**, *216*, 57–70. [CrossRef]
- Lu, H.; Ma, L.; Fu, X.; Liu, C.; Wang, Z.; Tang, M.; Li, N. Landslides information extraction using Object-Oriented Image Analysis paradigm based on Deep Learning and Transfer Learning. *Remote Sens.* 2020, 12, 752. [CrossRef]
- 66. Prakash, N.; Manconi, A.; Loew, S. Mapping landslides on EO data: Performance of deep learning models vs. Traditional machine learning models. *Remote Sens.* **2020**, *12*, 346. [CrossRef]
- 67. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.R.; Tiede, D.; Aryal, J. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sens.* **2019**, *11*, 196. [CrossRef]

68. Pan, X.; Zhao, J.; Xu, J. An object-based and heterogeneous segment filter convolutional neural network for high-resolution remote sensing image classification. *Int. J. Remote Sens.* **2019**, *40*, 5892–5916. [CrossRef]

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



 $\odot$  2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).