Automating avalanche detection in ground-based photographs with deep learning

James Fox [a,*], Anna Siebenbrunner [b,c], Sandra Reitinger [a], David Peer [a,d], Antonio Rodríguez-Sánchez [a]

[a] Department of Computer Science, University of Innsbruck, Technikerstraße 21a, Innsbruck 6020, Austria
[b] Lo.La Peak Solutions GmbH, Starres 8, Trins 6152, Austria
[c] Faculty of Geo- and Atmospheric Sciences, University of Innsbruck, Innrain 52f, Innsbruck 6020, Austria
[d] DeepOpinion, Bozner Platz 1/9, Innsbruck 6020, Austria

ARTICLE INFO

ABSTRACT

Automated avalanche detection has previously relied on satellite imagery, which is typically unsuitable for real-time monitoring due to long revisit times. To address this, we propose automating avalanche detection using photographs taken from the ground. This paper introduces image classification and avalanche segmentation tasks on a publicly-released dataset of 4090 photographs annotated by experts. Using the ResNet and YOLO architectures, we achieve avalanche detection F1 scores of 94.4% per image and 65.4% per avalanche region, demonstrating the potential that this method offers for avalanche monitoring.

In contrast with existing approaches, we label images by avalanche type into four distinct categories: glide, loose-snow, slab, and no avalanche. This labelling scheme provides more detail on avalanche events than binary labels and is shown to improve model F1 scores. Moreover, our models do not require a digital elevation model, simplifying application to new areas. Trained models can be used for real-time avalanche monitoring and to gather temporally continuous data for the improvement of existing avalanche forecasting models.

The code and dataset are available at github.com/j-f-ox/avalanche-detection.

1. Introduction

Snow avalanches, hereafter referred to as avalanches, are sudden and rapid downslope movements of snow (Thomas and Goudie, 2013). In addition to infrastructural damage (Fuchs and Bründl, 2005; Jóhannesson and Arnalds, 2001), avalanches cause an average of over 100 fatalities each year in the European Alps (Techel et al., 2016), making avalanche prediction, prevention, and detection an important safety and economic concern.

Despite efforts to mitigate the risks of avalanches, they will never be entirely preventable, which makes continuous monitoring necessary for two main reasons: firstly, rapid avalanche detection is vital for rescue operations and monitoring critical infrastructure. Secondly, detailed information on environmental conditions at the time of an avalanche release can be used to improve avalanche forecasting models (Schweizer and Herwijnen, 2013; Eckerstorfer et al., 2016), which in turn enables more precise risk management.

Avalanche monitoring can be conducted in situ or remotely and with or without human presence. Currently, avalanche monitoring relies in part on field observations which may suffer from biases towards easily accessible avalanches in fair weather conditions and place observers at risk (Eckerstorfer et al., 2016; Schweizer et al., 2015). By contrast, remote sensing can provide data on avalanche events over larger regions, including otherwise inaccessible areas (Eckerstorfer et al., 2016).

Remote sensing methods often utilise satellite data to enable continuous spatial coverage over large regions while lowering maintenance overheads compared to local systems. However, satellites are typically unsuitable for real-time monitoring due to lengthy revisit times of 1–12 days depending on latitude (Eckerstorfer et al., 2016; Hafner et al., 2021). Low temporal resolution poses a severe limitation in time-sensitive scenarios such as rescue operations where responses should be initiated within minutes rather than days. Furthermore, the spatial resolution of satellite images is often insufficient to identify small and medium-sized avalanches (Hafner et al., 2021).

In this paper, we propose a novel method for automated remote avalanche detection in ground-based photographs using convolutional

---

* Corresponding author.
E-mail address: james.fox@student.uibk.ac.at (J. Fox).

neural networks (CNNs). Our contributions are as follows:

1. We release a dataset of 4090 expert-labelled photographs containing 7228 outlined avalanches.
2. We propose training models on data labelled by avalanche type, a marked departure from previous studies which only train models for binary avalanche detection. We then analyse the impact of this additional information on prediction results.
3. We consider two distinct methods of avalanche detection: classifying entire images and segmenting each visible avalanche. Benchmark scores are provided for both approaches to provide a baseline for the research community and to demonstrate the viability of our proposed method.

Our ultimate aim is to automate real-time avalanche monitoring using existing networks of freely accessible mountain web cameras (webcams). By analysing webcam streams with deep learning models, we will be able to gather real-time avalanche data in high temporal resolution without incurring additional installation or maintenance costs. As of March 2023, a case study is ongoing in Tyrol, Austria as a first step towards operational implementation.

2. Current methods of avalanche detection

It is important to detect avalanches *after* they have released for two reasons: firstly, to identify danger to human life or infrastructure as soon as possible. Secondly, data on avalanche releases is essential for avalanche forecasting (EAWS, 2022b), both as an indicator of current snowpack stability and to improve existing models of avalanche activity. Avalanches can be detected either in the field or remotely: we discuss both approaches in this section with particular emphasis on the latter.

2.1. Field observations

Human observers can identify and measure avalanches during fieldwork, and avalanche warning services conduct snowpack stability tests to estimate the probability of additional avalanches. Although it is common practice to rely on field observations for avalanche monitoring and danger assessment, they are impractical over larger regions due to the human effort required and a lack of trained observers. Field observations also lead to biases in gathered data towards easily accessible areas and mild weather conditions (Lato et al., 2012; Eckerstorfer et al., 2016) and are time-consuming and therefore costly. Helicopters enable gathering data in otherwise inaccessible locations but suffer from similar limitations such as low temporal resolution, high costs, and restriction to fair weather conditions.

2.2. Remote monitoring

Avalanches can be detected using ground-based sensors, such as seismic (Bessason et al., 2007; Heck et al., 2018), radar (Meier et al., 2016), and infrasound sensors (Hendrikx et al., 2018). Although these data sources enable weather-independent monitoring, the prerequisite installation of sensors near avalanche paths limits their coverage to relatively small areas (van Herwijnen and Schweizer, 2011; Hendrikx et al., 2018). In contrast, trained experts can use satellite data to manually identify medium and large avalanches over larger regions (Eckerstorfer et al., 2014; Hafner et al., 2021). Synthetic-aperture radar (SAR) satellites such as Sentinel-1 provide freely available data independent of cloud coverage, but also result in the avalanche size being underestimated compared to optical satellites (Eckerstorfer et al., 2017; Hafner et al., 2021).

Previous work has explored two distinct approaches to automating avalanche detection in satellite images. In *classification*-based approaches, models are trained to classify whether an entire image patch contains any avalanches, whereas in *segmentation*-based approaches,

models are trained to outline each individual avalanche. As a result, image classification is considerably simpler than avalanche segmentation; however, classification models are then unable to distinguish between one and multiple avalanches in a single image.

Automated avalanche segmentation in satellite data has been investigated using object-based models (Bühler et al., 2009; Lato et al., 2012; Korzeniowska et al., 2017; Singh et al., 2022), *k*-means image clustering (Larsen et al., 2013; Vickers et al., 2016), thresholding based on radar backscatter (Eckerstorfer et al., 2019; Karas et al., 2021), and regression on additional causative factors (Liu et al., 2021). Convolutional neural networks (CNNs) have also been used for avalanche segmentation. For instance, Bianchi et al. (2021) trained a model based on the U-Net architecture (Ronneberger et al., 2015) to detect avalanches in Sentinel-1 SAR images and Hafner et al. (2022) trained a custom DeepLabV3+ CNN with a ResNet backbone for avalanche identification in SPOT 6/7 images. Moreover, CNNs have been used in classification-based approaches by Kummervold et al. (2018) and Sinha et al. (2019) who both detected avalanches in Sentinel-1 SAR data with VGG models (Simonyan and Zisserman, 2014) pretrained on the ImageNet dataset.

These methods of avalanche detection, which are summarised in Table 1, mostly rely on satellites with revisit times of 1–12 days (Eckerstorfer et al., 2016; Hafner et al., 2021) and are therefore unsuitable for time-sensitive applications where avalanches should be detected within minutes. In the case of airborne cameras and SPOT 6/7, data are not continuously available but must be specifically ordered (Hafner et al., 2021), making these sources unsuitable for ongoing real-time monitoring. Furthermore, most of these approaches require a digital elevation model (DEM) or manual mask to exclude areas where avalanches are unlikely to occur, such as flat terrain and very steep slopes. While this filtering can reduce the search space, it can complicate applying these methods to new areas since a sufficiently detailed DEM or mask must first be obtained for each region of interest.

2.3. Webcams in automated environmental monitoring

Ground-based photography has been employed in automating the monitoring of environmental data such as snow depth (Fromm and Adams, 2016) and snow cover evolution (Valt et al., 2013). In the field of avalanche detection, some avalanche warning services manually review webcam data to obtain an overview of the current avalanche situation (Stucki, 2006). Time-lapse photography has also been used to automatically measure glide crack opening dynamics by identifying snow-free pixels around manually selected regions of interest (Fees et al., 2023). Additionally, Hafner et al. (2023a) recently proposed an

Table 1
A selection of existing methods of automated avalanche detection.

Paper	Model/Technique	Sensor
Bühler et al. (2009)	Object-based	Airborne camera (RGB/NIR)
Korzeniowska et al. (2017)	Object-based	
Lato et al. (2012)	Object-based	QuickBird (panchromatic)
Larsen et al. (2013)	Bag-of-textons	
Vickers et al. (2016)	*k*-means	
Kummervold et al. (2018)	VGG-19; AConvNets	
Sinha et al. (2019)	VGG-16	
Eckerstorfer et al. (2019)	Image thresholding	Sentinel-1 (SAR)
Bianchi et al. (2021)	U-Net	
Liu et al. (2021)	SVM	
Karas et al. (2021)	Image thresholding	
Hafner et al. (2022)	DeepLabV3+	SPOT 6/7 (RGB/NIR)
Singh et al. (2022)	Object-based	Sentinel-2 (RGB/NIR/SWIR)

interactive method of avalanche segmentation in webcam images. In their approach, a human annotator initially clicks on an avalanche and then iteratively corrects AI-generated outlines.

However, a thorough review of the literature indicates that, to the best of our knowledge, no prior studies have utilised ground-based photography to fully automate avalanche detection. Our work, therefore, proposes a novel approach to automating avalanche monitoring.

3. Dataset

In conjunction with this paper, we are releasing a dataset of 4090 photographs containing 7228 avalanches which were manually outlined by avalanche experts (University of Innsbruck, 2023).[1] Annotation by domain specialists is commonplace in machine learning and is the standard approach for avalanche monitoring (Hafner et al., 2023b; Bianchi et al., 2021). By releasing this dataset, we aim to encourage further research into automated avalanche detection in ground-based photographs, as we believe that ground-based photography has great potential to improve operational avalanche monitoring.

The studies discussed in section 2 treat automated avalanche detection as a binary classification problem, focusing on determining the presence or absence of avalanches. However, the cause of the avalanche release correlates heavily with the visual features of the resulting avalanche (McClung and Schaerer, 2000). We therefore propose labelling training data by *avalanche type*. We hypothesised that this labelling scheme could improve the performance of avalanche detection models by providing additional contextual information during training.

3.1. Avalanche types

Avalanches can be classified by their release type into three categories: glide avalanches, loose-snow avalanches, and slab avalanches (McClung and Schaerer, 2000; Schweizer et al., 2015). These types of avalanches can be distinguished a posteriori by avalanche experts due to differing visual features which are summarised below:

- **Glide avalanches** are characterised by their gliding mechanism: the entire depth of the snowpack glides downslope, exposing the ground beneath. The avalanche release may be preceded by the formation of full-depth tensile cracks in the snowpack, known as glide cracks (Schweizer et al., 2015), which are visible in Fig. 1.
- **Loose-snow avalanches** release at a single point and spread out in a triangular shape as they move down the slope, entraining more snow. They may appear darker or lighter than the surrounding snow depending on lighting conditions (see Fig. 2).
- **Slab avalanches** occur when a weakness within the snowpack propagates, resulting in the release of a cohesive slab of snow (McClung and Schaerer, 2000; Schweizer et al., 2015). The snowpack above the collapsed weak layer slides down the slope, resulting in a sharp line (the *crown fracture*) where the avalanche released (see Fig. 3). This stands in contrast to glide avalanches in which the entire snowpack moves along the ground below, typically over smooth surfaces like vegetation or rock.

3.2. Dataset photographs

A key contribution of this study is releasing a freely accessible dataset of 4090 photographs[2] containing 7228 labelled avalanches. The dataset contains photographs taken in the field in 2000/2001–2021/

2022 with mean dimension 3431×2431px and standard deviation 1603×1086px. The images each have the date and approximate location at which they were taken available as metadata, with 1276 different locations across the Alps represented.

Mountains and avalanches are depicted at a range of scales in the dataset (see e.g. Fig. 2 or Fig. 3). This heterogeneity reflects the wide range of dimensions inherent to real-world applications of this research, as avalanches may span from a few metres to many kilometres wide (EAWS, 2022a). Most photographs in the dataset were taken from the ground: additional images from low-flying helicopters were included only if the annotators deemed that they could feasibly have been taken by webcams.

3.3. Dataset annotation

Avalanche experts outlined each visible avalanche with a polygonal bounding box and assigned the avalanche to one of three categories: GLIDE, LOOSE, or SLAB referring to glide, loose-snow, and slab avalanches respectively (see Fig. 4). The experts were instructed to draw outlines encompassing the avalanche release area and the runout as these both contain relevant visual features. Images then received an *overall label* corresponding to the avalanche release type with the highest number of annotated pixels, or overall label NONE if no avalanches were annotated.[3]

This multiclass annotation is a novel approach to automating avalanche detection. Our proposed labelling scheme can be viewed as a direct extension of more traditional approaches that consider only the presence or absence of avalanches, providing additional details on the causes of avalanches without losing information. In cases where one wishes to prioritise avalanche detection irrespective of the avalanche type, such as for urgent rescue operations, the GLIDE, LOOSE, and SLAB labels can be combined into a single AVALANCHE label, as shown in Fig. 5.

Of the 3019 dataset images containing avalanches, 311 contain avalanche regions with at least two distinct labels. Slab avalanches have the highest support in the dataset, representing the overall label in 46% of images (see Table 2). In contrast, loose-snow avalanches constitute only 10% of the overall image labels despite representing 25% of the annotated avalanches.

To assess the consistency of annotation between experts, we conducted a reproducibility study on 112 images labelled by two different annotators. Our analysis revealed a pixelwise agreement of 93% and a pairwise IOU of 0.71 for the labelled avalanche regions. The experts also demonstrated a high level of agreement on the avalanche type, predicting the same label for 99.6% of the pixels labelled as avalanches by both annotators. However, the two annotators outlined a different number of avalanches in 47% of the images, suggesting that clearer guidance should be provided about how to delimit individual avalanches during annotation. A similar finding was reported for Radarsat-2 SAR images by Eckerstorfer and Malnes (2015), who found a discrepancy in the number of annotated avalanches due to one expert outlining multiple adjacent avalanches as a single avalanche.

4. Methodology

4.1. Classification versus segmentation

Previous research has explored two distinct approaches to avalanche detection: determining whether an image patch contains *any* avalanches

---

[1] "Experts" here refers to people working in the snow and avalanche industry or conducting research in this field.

[2] 4034 images are from the database of the Avalanche Warning Service Tyrol and other Austrian sources and 56 are from the Avalanche Warning Service Bavaria.

[3] In the unlikely event that two avalanche labels have the same number of visible pixels in a single image, the overall label should reflect the relative frequency of release types, with SLAB having the highest priority, followed by GLIDE, and then LOOSE. However, this was not the case for any images in the dataset.

(a) (b) (c)

Fig. 1. Three images from the dataset depicting glide avalanche at a range of scales. Glide cracks expose the terrain beneath the snowpack in (a), while (b) contains a glide avalanche with textured snow runout visible in the lower half of the photograph. One large glide crack is visible in (c), part of which has released as a glide avalanche.

(a) (b) (c)

Fig. 2. Three images from the dataset containing loose-snow avalanches. Loose-snow avalanches with textured runout at the base are visible in (a) and (b), while (c) contains several loose-snow avalanches visible as lighter regions of perturbed snow. Note that each avalanche originated at a single point and then spread laterally and downwards as it accumulated more snow.

(a) (b) (c)

Fig. 3. Three dataset images showing slab avalanches in a variety of lighting conditions and at different scales. The crown fracture may range from a few metres to many kilometres in width.

(a) (b)
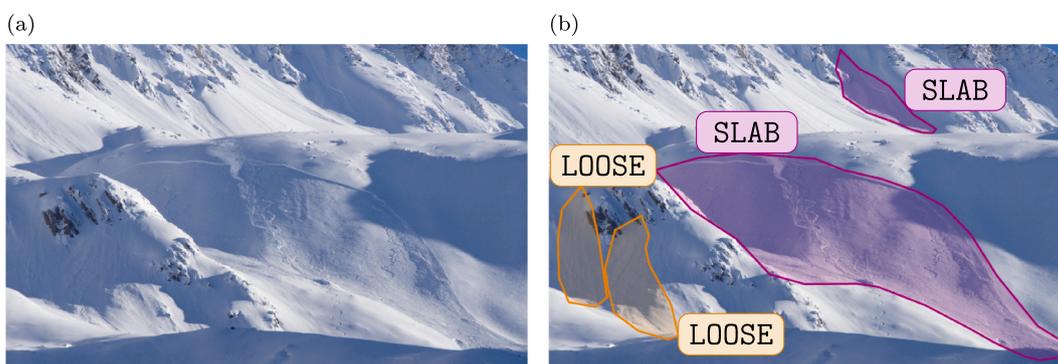
Fig. 4. An image from the dataset containing two loose-snow avalanches and two slab avalanches is shown in (a). Each avalanche was outlined by an expert with a polygonal bounding box, as shown in (b). The image has overall label SLAB as this avalanche type has the most visible pixels.

and identifying each individual avalanche. We defined two tasks on the dataset corresponding to each approach:

– The **image classification** task is to determine whether an entire photograph contains any visible avalanches and classify the dominant avalanche type if so. This corresponds to predicting the overall image labels described in subsection 3.2.

– The **avalanche segmentation** task requires predicting a bounding rectangle and type for each avalanche in a given photograph, corresponding to a typical object detection scenario. Note that, in contrast to the image classification task, there is no need for an explicit NONE label: any image region that is not part of a predicted bounding box is implicitly assumed to be free of avalanches.

The avalanche segmentation task can therefore be considered a more

(a)

(b)
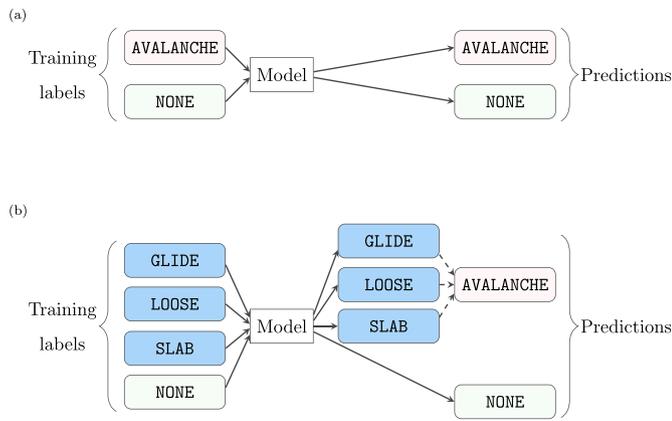
Fig. 5. Labelled data are used to train models which can then predict labels for previously unseen images. Existing approaches to automating avalanche detection use binary labels, as illustrated in (a). By contrast, we propose labelling avalanche data by avalanche type, as shown in (b). The model output can still be reduced to binary labels after prediction if desired (dashed arrows).

Table 2
Label distribution for each task. The distribution of labels reflects the relative occurrence of each avalanche type, with slab avalanches being the most common and loose-snow avalanches being the least common (Schweizer et al., 2015).

Label	Meaning	Number of Images	Number of Avalanches
GLIDE	Glide avalanche	716	2489
LOOSE	Loose-snow avalanche	416	1827
SLAB	Slab avalanche	1887	2912
NONE	No avalanche visible	1071	N/A
Σ		4090	7228

challenging subtask of the image classification task since it requires identifying each individual avalanche rather than the overall image label alone (see Fig. 6).

4.2. Model training

The dataset was split into 3612 training images and 478 test images, with images at the same location grouped into the same split to avoid leaking information due to visual similarities between images at the same locations. In each training run, 10% of the training data were reserved for model validation. We then repeated each training run three times with different validation sets on a GeForce RTX 3090 GPU. The test images were withheld from all training runs and used only to evaluate model performance on unseen data. To improve model generalisation, we applied data augmentation techniques during training including affine transformations, variations in colour, and random horizontal flips.

4.2.1. Classification model training

For the image classification task, we utilised VGG (Simonyan and Zisserman, 2014) and ResNet (He et al., 2016) models which were pretrained on the ImageNet dataset. The models were then finetuned for avalanche detection via transfer learning using the Adam optimiser (Kingma and Ba, 2015) with a learning rate of 2.25×10^{-3} to minimise the cross-entropy loss on the training set. If the validation accuracy had not improved for six consecutive epochs, then training was stopped and the model with the highest validation accuracy from the training run was saved. To reduce model overfitting, we first resized the images so that their shortest side was 105% of the network input size and then cropped a different random square from each image in every training epoch before applying further data augmentation transformations. Unless otherwise stated, the classification models were trained with network input size 704px × 704px.

4.2.2. Segmentation model training

For the avalanche segmentation task, we trained version 3 of the YOLO (You Only Look Once) network (Redmon and Farhadi, 2018) to predict the avalanche bounding boxes. All YOLO models were trained with an Intersection Over Union (IOU) threshold of 0.2 and a confidence threshold of 0.25 using an open-source framework from Ultralytics (Jocher, 2020). The IOU threshold describes the minimum overlap between a predicted bounding box Box_{pred} and a ground truth bounding box Box_{true} for the prediction to be classed as correct, defined as

$$\text{IOU} = \frac{\| \ \text{Box}_{\text{pred}} \ \cap \ \text{Box}_{\text{true}} \ \|}{\| \ \text{Box}_{\text{pred}} \ \cup \ \text{Box}_{\text{true}} \ \|}$$

where $\|\cdot\|$ refers to the number of pixels in a region. Training was then halted if the validation mean average precision (mAP) had not improved for the last 25 epochs, where the mAP for a specific IOU is calculated as

$$\text{mAP}_{\text{IOU}} = \frac{1}{|L||C|} \sum_{\text{label} \in L} \left(\sum_{\text{conf} \in C} \text{precision}_{\text{IOU, label, conf}} \right)$$

for a set of labels L and confidences $C = \{0.001, \ 0.002, \ 0.003, \ldots, \ 1.000\}$.

In addition to the data augmentation transformations mentioned

(a)   (b)
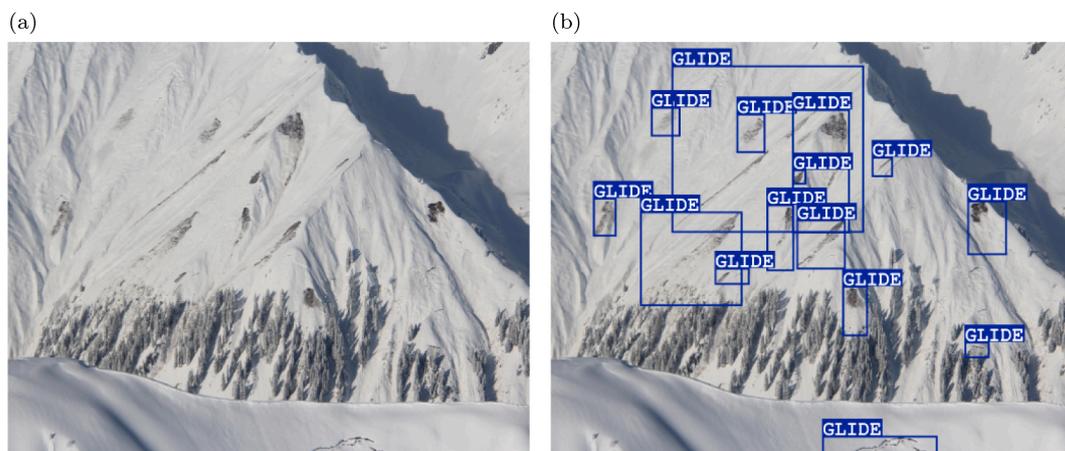
Fig. 6. A dataset image containing 15 annotated glide avalanches. For the image classification task, the overall image label GLIDE should be predicted for (a), whereas for the avalanche segmentation task, a labelled bounding box should be detected for each visible avalanche as shown in (b).

above, we applied mosaic augmentation (Bochkovskiy et al., 2020) to training images to improve model generalisation. This is a technique whereby four images are arranged into a grid before passing this composite image to the network in order to present images to the network in different contexts.

4.3. Experiments

We began by comparing a range of deep learning architectures and depths for both tasks to identify the most suitable architecture for further experiments. For the classification task, we trained ResNet models of depths 18, 34, 50, 101, and 152 (He et al., 2016) and VGG models of depths 13, 16, and 19 (Simonyan and Zisserman, 2014). For the segmentation task, we compared the results for YOLOv3 (Redmon and Farhadi, 2018), YOLOv3-tiny (Adarsh et al., 2020), and YOLOv3-SPP (Huang et al., 2020). YOLOv3-SPP incorporates spatial pyramid pooling, which enables capturing information at multiple scales at the cost of more computational resources (Huang et al., 2020). On the other hand, YOLOv3-tiny has fewer layers than YOLOv3, making training and inference faster (Adarsh et al., 2020).

Following the selection of potential base architectures, we went on to explore the extent to which images can be downscaled without compromising model performance. To investigate this, we trained models on image sizes ranging between 224px and 1120px in step sizes of 224px. Downsizing images before passing them to the network greatly reduces computational requirements but risks removing important details, especially for smaller or low-contrast avalanche regions.

Finally, we investigated the impact of labelling training data by avalanche release mechanism on model performance. We trained networks on the original dataset with avalanche labels GLIDE, LOOSE, and SLAB, and on the same images in which all avalanche regions were mapped to a single AVALANCHE label and compared model results on previously unseen images.

4.4. Metrics

We evaluated model performance using the metrics

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \ , \ \ \text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \ , \ \ \text{F1 score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \ ,$$

where TP refers to the number of true positive classifications, TN to true negatives, FP to false positives, and FN to false negatives. Scores are calculated per image for the image classification task and per bounding box for the avalanche segmentation task. For each prediction, it is then of interest:

 i) whether the predicted release mechanism is correct, and
 ii) whether a predicted avalanche image or bounding box corresponds to any labelled avalanche region, regardless of release mechanism.

For the former, we consider scores on the original dataset labels. In the latter case, we calculate a *binary* avalanche detection score by applying the label mapping

{LOOSE ↦ AVALANCHE, GLIDE ↦ AVALANCHE, SLAB ↦ AVALANCHE}

to the ground truth and predicted labels after inference (see also Fig. 5). These binary scores capture the ability of models to detect avalanches regardless of release type.

The YOLO models output a confidence value for each prediction. The *confidence threshold* describes the minimum confidence required for a predicted bounding box to be accepted. Unless otherwise stated, the scores for the avalanche segmentation task are given at the confidence threshold that maximises the F1 score for each training run.

5. Results

5.1. Architecture comparison

Image classification results for ResNet and VGG networks of various depths are shown in Table 3. Results were compared on the validation set to prevent information about model performance on the test set from influencing the architecture selection, thus ensuring that the test set remains a truly independent measure of the models' generalisation abilities.

The ResNet101 architecture achieved the highest scores on the original dataset labels, with an F1 score of 85.7% and an accuracy of 85.6%, while ResNet152 achieved higher binary scores for avalanche detection regardless of avalanche type. Notably, the ResNet101, ResNet50, and ResNet152 architectures outperformed the VGG networks in all metrics considered. In the following experiments, we utilise ResNet101 and VGG-19 as the base architectures for image classification.

For the avalanche segmentation task, we compared the performance of YOLOv3-tiny, YOLOv3, and YOLOv3-SPP on images of size 448px and 896px. YOLOv3-tiny performed poorly for both input sizes, while the regular and SPP variants achieved comparable results on images of size 448px (see Table 4). YOLOv3-SPP exhibited superior performance on 896px images, so this architecture is used in all subsequent avalanche segmentation experiments.

We next investigated the influence of the IOU threshold used for evaluation on model performance. Decreasing the IOU threshold resulted in fewer false negatives *and* false positives, increasing both recall and precision on the validation set (see Fig. 7). We therefore report mAP values at IOU 0.05 in the remainder of this paper.

5.2. Image size

Following the selection of suitable base architectures, we investigated the impact of image size on network performance. Doubling the image size from 224px to 448px resulted in significant F1 score improvements of 8.1% for the classification models and 10.1% for the segmentation models (see Fig. 8). Further increasing the input size from 448px to 896px resulted in additional improvements of 2.5%, 1.7%, and 5.3% for ResNet101, VGG-19, and YOLOv3-SPP respectively. However, increasing the image size from 896px to 1120px slightly decreased the mean F1 for YOLOv3-SPP by 0.5%.

5.3. Labels

Finally, we investigated the impact of labelling avalanche regions by avalanche type on model performance. We trained models on the original dataset (with avalanche labels GLIDE, LOOSE, and SLAB) and on the same images in which all avalanche regions were assigned a single AVALANCHE label. Note that it is important to consider binary avalanche detection scores in this context to ensure a fair comparison of model performance on an equal number of target labels.

Labelling training images by avalanche type improved performance

Table 3

Benchmark image classification results on the validation set (± standard deviation) sorted by F1 score.

Model	F1	Accuracy	Binary F1	Binary Accuracy
ResNet101	**85.7** ± 0.4	**85.6** ± 0.3	94.5 ± 0.6	92.1 ± 0.9
ResNet50	84.9 ± 0.3	85.0 ± 0.2	94.9 ± 0.2	92.5 ± 0.4
ResNet152	84.1 ± 0.7	84.1 ± 0.7	**95.7** ± 1.0	**93.6** ± 1.3
VGG-19	83.1 ± 0.6	83.1 ± 0.7	94.3 ± 0.4	91.3 ± 0.3
ResNet34	81.4 ± 0.3	81.2 ± 0.1	94.5 ± 1.2	91.7 ± 1.6
ResNet18	80.8 ± 1.8	80.6 ± 1.8	91.9 ± 1.3	88.1 ± 2.0
VGG-13	80.1 ± 0.8	80.2 ± 0.9	92.4 ± 0.5	89.0 ± 0.6
VGG-16	79.1 ± 1.5	79.4 ± 1.0	92.4 ± 0.3	88.6 ± 0.5

Table 4
Results on the validation set for three YOLOv3 variants (± standard deviation).

Size (px)	Model	mAP$_{0.05}$	Bin. mAP$_{0.05}$	F1	Bin. F1
448	YOLOv3-SPP	**58.0** ± 0.4	**65.1** ± 1.1	56.5 ± 2.8	63.7 ± 1.6
	YOLOv3	57.3 ± 1.9	**65.1** ± 2.8	**57.0** ± 3.0	**64.1** ± 2.1
	YOLOv3-tiny	42.6 ± 3.9	53.3 ± 4.4	45.2 ± 4.1	52.3 ± 2.8
896	YOLOv3-SPP	**60.9** ± 0.9	**66.8** ± 1.2	**61.7** ± 2.6	**67.7** ± 1.5
	YOLOv3	58.2 ± 1.1	64.3 ± 2.5	58.9 ± 3.3	65.1 ± 1.7
	YOLOv3-tiny	38.6 ± 5.6	45.3 ± 5.1	38.9 ± 8.5	48.8 ± 3.3

Fig. 7. Validation precision versus recall at a range of IOU thresholds. Results are displayed for YOLOv3-SPP with input size 896px and confidence thresholds ranging between 0.001 and 1.000.

the label LOOSE, misclassifying 10.6% of the loose-snow images as NONE and 8.6% as slab avalanches. In addition, 9.9% of the GLIDE images were misclassified as slab avalanches.

6. Discussion

6.1. Experimental analysis

In our architecture selection experiments, the shallower networks generalised relatively poorly to unseen images. In particular, the ResNet34, ResNet18, VGG-16, VGG-13, and YOLOv3-tiny networks achieved significantly lower F1 scores than their deeper counterparts, as is visible in Table 3 and Table 4. This observation could be attributed to the limited model capacity of these smaller networks.

We further observed a positive correlation between image size and model performance for both image classification and avalanche segmentation (see Fig. 8). These results are noteworthy as we required considerably larger image sizes than those used in other domains to obtain satisfactory results. We believe that this is due to avalanche regions becoming less visible when subjected to excessive downscaling, as shown in Fig. 11. Avalanches tend to be low in contrast relative to the image as a whole, making this domain especially sensitive to image size.

Training deep learning models on larger images requires significantly more computational resources. However, naively cropping training images to make the avalanche regions appear larger would not be an appropriate solution. Avalanches are natural events which can vary from a few metres to several kilometres in width (EAWS, 2022a), so the ability to detect avalanches at a wide range of scales is a challenge that is inherent to the domain. Furthermore, it is impossible to predict exactly where avalanches will release relative to stationary cameras, leading to more variability in apparent size. It is therefore essential that remote avalanche detection systems are robust across a range of apparent scales.

in every metric considered for both classification architectures, leading to F1 score increases of 7.7% and 10.1% for ResNet101 and VGG-19 respectively (see Table 5). For the avalanche segmentation model, labelling training data by type increased the F1 scores by 0.6% and 1.2% for images of size 448px and 896px respectively (see Table 6). In particular, training on data labelled by avalanche type increased the model recall and the confidence threshold at which this maximal F1 score was obtained (see Fig. 9), but led to a decrease in the corresponding precision.

Finally, we measured the labelwise performance on the test set for the ResNet101 and YOLOv3-SPP training runs with the highest validation scores. Both models exhibited poorer performance for loose-snow avalanches than the other two avalanche types, as is visible in Fig. 10, but this discrepancy was significantly more pronounced for the segmentation model. In particular, YOLOv3-SPP achieved F1 scores of around 77% for slab avalanches compared to 47% for loose-snow avalanches. The classification model also displayed significant confusion for

Table 5
Image classification results for ResNet101 and VGG-19 models trained on the original labels (GLIDE, LOOSE, SLAB, NONE) and binary data (AVALANCHE vs. NONE).

Model	Labels	Bin. F1	Bin. Precision	Bin. Recall	Bin. Acc.
ResNet101	Original	**93.8** ± 0.8	**94.3** ± 1.7	**93.3** ± 0.1	**90.9** ± 1.3
	Binary	86.1 ± 0.6	87.5 ± 0.3	84.9 ± 1.0	89.7 ± 0.4
VGG-19	Original	**92.8** ± 0.4	**91.8** ± 0.8	**93.8** ± 1.2	**89.1** ± 0.6
	Binary	82.7 ± 0.6	86.3 ± 3.6	81.3 ± 2.8	87.4 ± 1.1
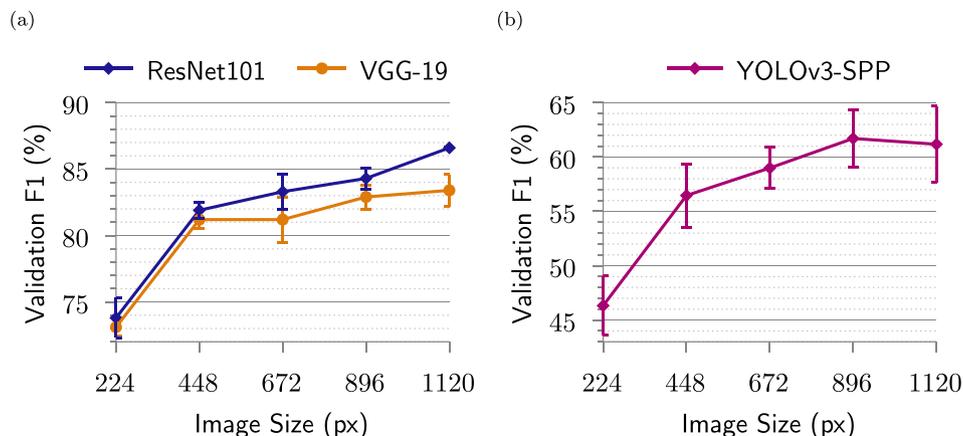
Fig. 8. Validation F1 versus network input size for ResNet101, VGG-19, and YOLOv3-SPP models with error bars indicating the standard deviation.

Table 6

Binary test results for YOLOv3-SPP trained on the original labels (GLIDE, LOOSE, and SLAB) and on the dataset reduced to a single label (AVALANCHE). Training on images labelled by avalanche type increased the mean F1 score and the confidence (Conf.) at which this F1 score was achieved (see also Fig. 9).

Size (px)	Labels	mAP$_{0.05}$	F1	Recall	Precision	Conf.
448	Original	63.7 ± 1.4	$\mathbf{63.1} \pm 1.6$	$\mathbf{55.3} \pm 1.8$	73.6 ± 1.2	0.127
	Single-label	$\mathbf{65.2} \pm 0.6$	62.5 ± 1.0	53.6 ± 4.5	$\mathbf{76.0} \pm 5.5$	0.117
896	Original	65.3 ± 1.7	$\mathbf{65.4} \pm 0.2$	$\mathbf{60.1} \pm 1.6$	71.7 ± 1.8	0.131
	Single-label	$\mathbf{66.7} \pm 0.5$	64.2 ± 0.3	56.6 ± 5.5	$\mathbf{75.8} \pm 7.9$	0.065
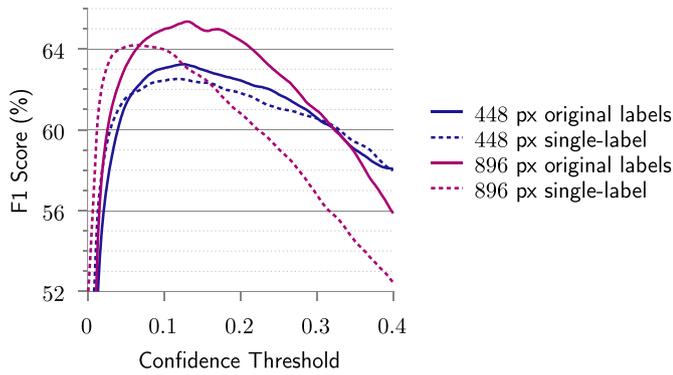
Fig. 9. Mean F1 score versus confidence threshold for YOLOv3-SPP models trained on the original dataset labels (GLIDE, LOOSE, and SLAB) and on the dataset reduced to a single label (AVALANCHE). Labelling training data by avalanche type increased the maximum F1 score and the confidence threshold at which this score was attained.

Finally, we trained models on data with a single AVALANCHE label and on data labelled by avalanche type. Including information on avalanche type increased the F1 scores for both the classification models and the segmentation models (see Table 5 and Table 6). These results support our hypothesis that automated avalanche detection can be improved by incorporating information about avalanche types. We believe that the visual differences between different types of avalanches may have caused confusion when training models solely on binary avalanche labels: a multiclass labelling scheme is therefore more suitable for this domain. A further advantage of this multiclass approach is that trained models can then identify the avalanche type in previously unseen images. This provides a more detailed understanding of the current avalanche situation, which in turn can enable more precise avalanche forecasting.

For the segmentation model, training on data labelled by avalanche type increased the confidence threshold at which the maximal F1 score was obtained (see Fig. 9), suggesting that providing information about avalanche type increased the model's confidence in its predictions. This increase in confidence threshold may then have contributed to the increase in recall and decrease in precision visible in Table 6. However, the high standard deviation for precision and recall of the single-label models may also have played a role and warrants further investigation.

Selecting a confidence threshold presents a natural trade-off between precision and recall, which raises the question of whether one wishes to prioritise reducing false positives or false negatives. Although we consider (unweighted) F1 scores throughout this study, as is standard in the literature, we argue that prioritising recall over precision may be more appropriate in the context of operational avalanche detection. False positive classifications can be easily refuted by experts, whereas a false negative delaying the detection of a new avalanche could have more serious real-world consequences. We therefore propose this optimisation question as an area for future research.
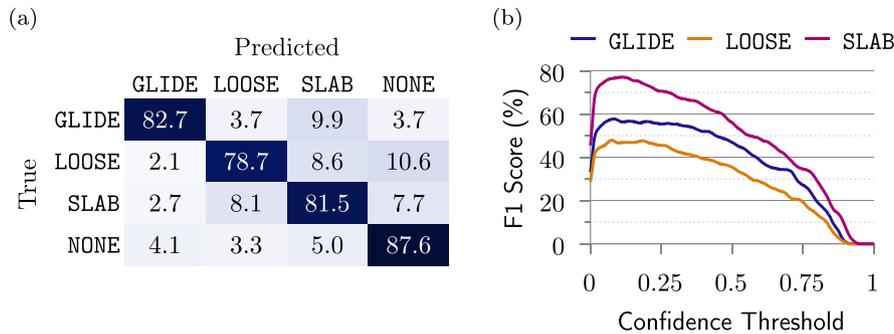
Fig. 10. Labelwise performance on the test set for the models with the highest validation scores. The normalised confusion matrix for ResNet101 is shown in (a) and the labelwise F1 score versus confidence for YOLOv3-SPP with input size 896px is shown in (b).
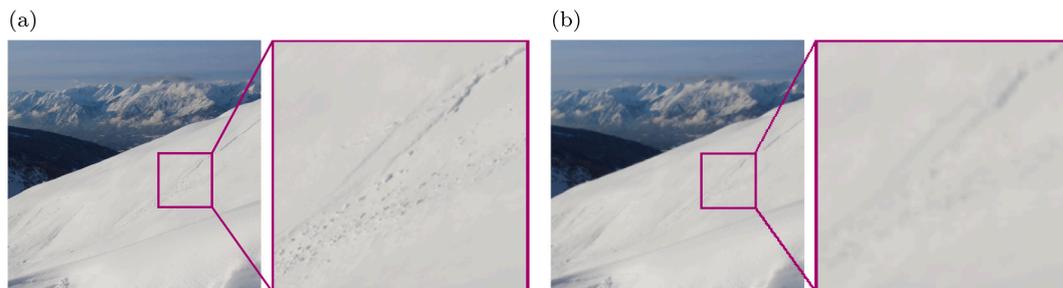
Fig. 11. A dataset image downscaled to a height of 1120px is shown in (a), while (b) depicts the same image downscaled to height 224px. Reducing the image size to 224px results in a significant loss of detail around the avalanche boundary and runout.

6.2. Qualitative analysis

In this subsection, we focus on predictions made by the 896px ResNet101 and YOLOv3-SPP models with the highest validation scores, utilising these as a case study to explore model capabilities and sources of misclassifications. Both models were able to identify slab and glide avalanches across a range of lighting conditions and scales – as is visible in Fig. 12, Fig. 13, and Fig. 16 – and both models exhibited poorer performance for loose-snow avalanches. For real-world applications, one should also consider the relative impact and frequency of each type of avalanche. Slab avalanches were responsible for 97% of recorded avalanche accidents in Switzerland over a 20-year period (Techel et al., 2016), underscoring the importance of detecting slab avalanches in particular.

Initially, we considered that the lower scores for the label LOOSE could be due to the smaller number of loose-snow images in the dataset (as shown in Table 2). However, the classification model attention was consistently highest around the release areas of glide and slab avalanches, while the model typically focused on the entirety of loose-snow avalanches, as can be seen in Fig. 12. This suggests that the lower scores for loose-snow avalanches might also be due to their less clearly defined release zones compared to the sharp lines marking slab crowns or glide cracks.

We further observed that the YOLOv3-SPP model often predicted a single large bounding box surrounding several smaller loose-snow avalanches, as shown in Fig. 14. This was likely due to the segmentation model correctly recognising the visual features associated with loose-snow avalanches but then having difficulty delineating their precise boundaries. This behaviour may in turn have contributed to the improvement in model skill at lower IOU thresholds visible in Fig. 7.

Continuing our analysis, we investigated the causes of misclassifications. In a comprehensive review of the test predictions, we observed that rocks and infrastructure such as buildings and fences often led to false negative classifications for the classification model (as in Fig. 15) and false positive detections for the segmentation model. To address this, one could exclude areas where avalanches can not occur either by aligning each image with a DEM or by manually creating a mask for each camera location. However, this would add significant barriers to deploying models across larger regions due to the overhead required to align every camera with a suitable mask or DEM. We therefore intend to explore other strategies to mitigate this issue in future work, such as unsupervised image preprocessing techniques or the addition of more training data containing rocks and buildings.

A notable finding for operational applications is that trained models were able to identify glide cracks and entire glide avalanches, as shown in Fig. 16. Around half of glide avalanches are preceded by the formation of a glide crack between 15 minutes to several days before the avalanche release (Fees et al., 2023). Trained models could therefore be employed in developing early warning systems to automatically detect glide cracks before they release. Such a system could allow for the timely enactment of safety measures such as road closures to mitigate the potential harm caused by glide avalanches.

6.3. Applications

Automated, ongoing avalanche monitoring has previously not been feasible at scale in real-time. Existing satellite-based approaches are constrained by revisit times on the order of days (Eckerstorfer et al., 2016), making them unsuitable for applications requiring rapid responses such as rescue operations or monitoring critical infrastructure. Although trained experts can identify avalanches within photographs (Eckerstorfer et al., 2016), the vast number of locations where avalanches can occur makes *manual* monitoring via webcams prohibitively time-consuming at scale. However, trained models could be used to *automatically* filter potential avalanche images from webcam streams for confirmation by experts. Automating this initial identification process could thus enable remote avalanche monitoring in near real-time for any areas of interest monitored by webcams.

There are over a thousand freely accessible webcams across the Alps for weather monitoring and tourism[4] which could be used for avalanche monitoring. Using existing webcams has the advantage of not incurring further installation, usage, or maintenance costs. Additional webcams could then be installed around critical infrastructure or higher-risk areas if needed. A workflow whereby models are retrained on corrected predictions to iteratively improve performance over time is described by Fox et al. (2023).

It is not yet clear whether image classification or avalanche segmentation will provide more utility in operational settings. Image classification offers advantages such as ease of annotation and providing a simple overview of the avalanche situation which could serve as a practical solution for the rapid identification of avalanche occurrences at scale. Although the avalanche segmentation task is more challenging than image classification, it provides more details by localising each individual avalanche within photographs. Segmentation models are therefore able to detect multiple avalanches in a single image.

A further observation from our experiments is illustrated in Figs. 14 and 17: in several cases, the segmentation models detected less clearly visible avalanches which had not been identified by the annotators. A similar finding was reported by Bianchi et al. (2021) for SAR data. This highlights an exciting application of trained avalanche segmentation models for *AI-assisted annotation*. Rather than annotating images from scratch, avalanche experts could utilise avalanche segmentation models to generate initial predictions, which can then be manually refined and corrected. Such an approach has the potential to expedite the annotation process while simultaneously improving the quality of annotation.

6.4. Limitations

An intrinsic limitation of visible-spectrum photography is its reliance on clear visibility. Avalanches that are fully obscured by low clouds, snowfall, or darkness can not be identified in optical data by computer vision or by human observers, whether remotely or in the field. As most photographs in the dataset were captured in clear visibility, further experiments should be conducted on images taken in inclement weather conditions such as heavy precipitation or (partial) cloud occlusion to evaluate the robustness of the trained models.

Using webcams rather than satellites to monitor avalanche avalanches may also introduce data biases towards regions with high levels of tourism such as ski resorts, as these areas are more likely to have preinstalled webcams in mountainous regions. However, ground-based cameras hold a significant advantage over optical satellites in that they are unaffected by high clouds. To enable weather-independent monitoring, one could use ground-based radar sensors either alone or in conjunction with visible-spectrum webcams to enable avalanche monitoring at higher temporal resolutions than achievable via satellite data.

An additional limitation of ground-based photography is the relative difficulty in determining the location and size of detected avalanches. Unlike satellite data, which are generally accompanied by the coordinates of the depicted area, mountain webcams typically provide only the location and viewing direction of the camera. Nevertheless, local avalanche authorities are accustomed to responding to avalanche incidents without precise coordinate data and should be able to infer the approximate location of avalanches from the camera location and surrounding mountain features. Research into estimating avalanche size in oblique photographs has also demonstrated relatively high coherence

---

[4] As of November 2023, there are 276 webcams freely accessible at foto-webcam.eu (https://www.foto-webcam.eu/), 223 more via feratel (https://www.feratel.com/en/webcams.html), and 1366 through Bergfex (https://www.bergfex.at/oesterreich/webcams/).

(a) Image label: GLIDE. (b) Image label: LOOSE. (c) Image label: SLAB.
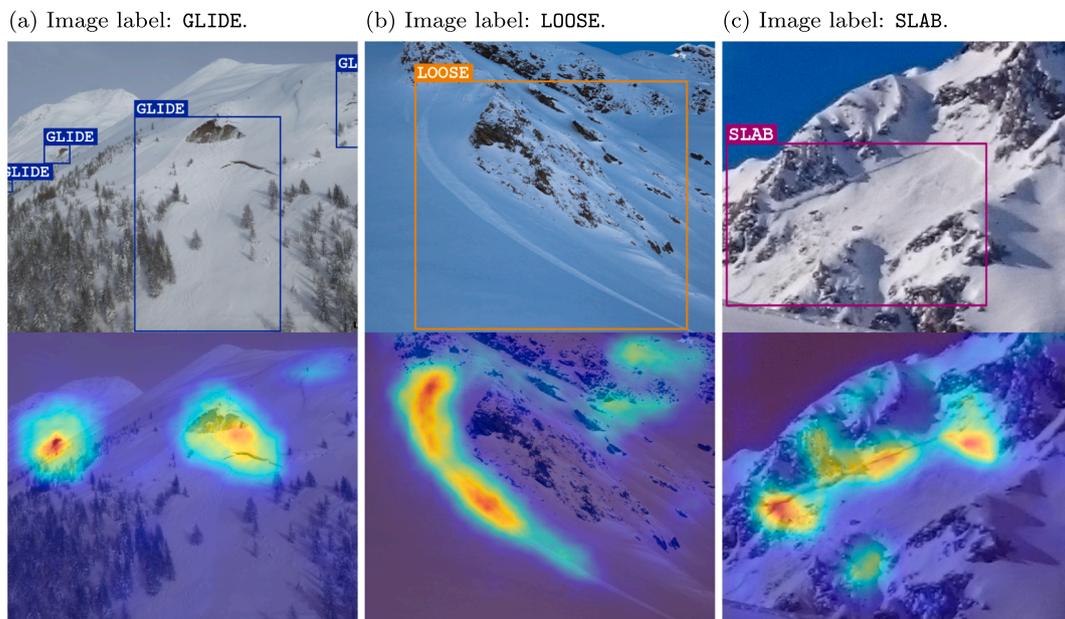
Fig. 12. Three test images that were correctly classified by the ResNet101 model: the ground truth is shown in the top row and the corresponding Gradient-weighted Class Activation Mapping (Grad-CAM) heatmap (Selvaraju et al., 2020) for each prediction is displayed below.
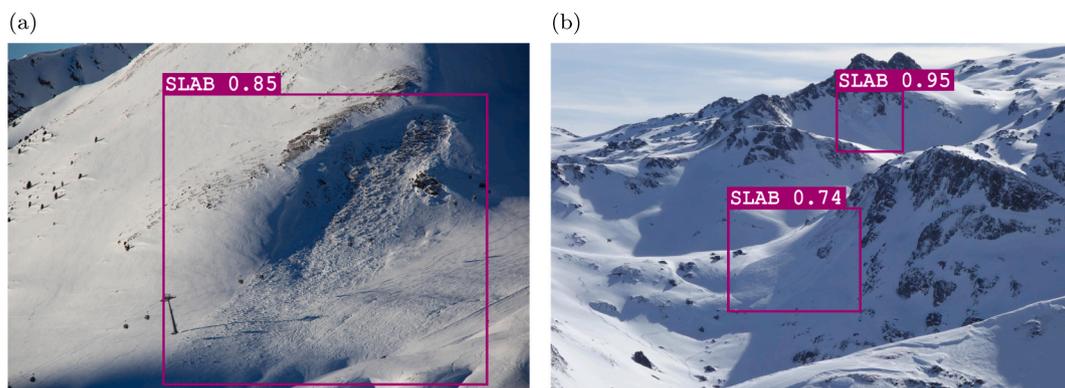
Fig. 13. Example predictions made by the 896px YOLOv3-SPP model. The model was able to identify slab avalanches at a range of scales and in various lighting conditions. The highest Intersection Over Union (IOU) value relative to the ground truth is displayed on each predicted bounding box.
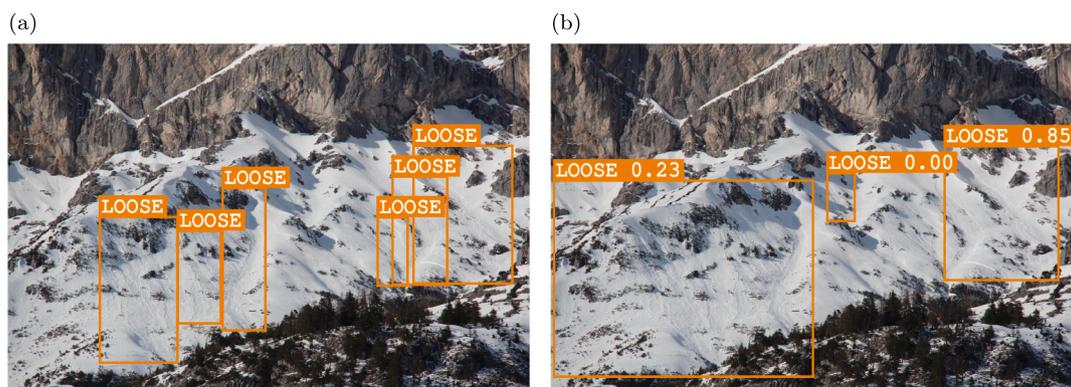
Fig. 14. Ground truth bounding boxes (left) and YOLOv3-SPP predictions with IOU values (right) for a test image containing six loose-snow avalanches. The model detected a very small loose-snow avalanche in the image centre which was not found by the annotators.

between expert estimations of avalanche size (Hafner et al., 2023b). However, the practical implications of estimating avalanche location and size in photographs warrant further investigation. Obtaining more precise measurements would likely necessitate aligning the oblique camera view with a topological model of the depicted area such as a DEM; the disadvantages of which were discussed above.

The subjectivity introduced by treating expert annotation as the ground truth may be considered an additional limitation, even though
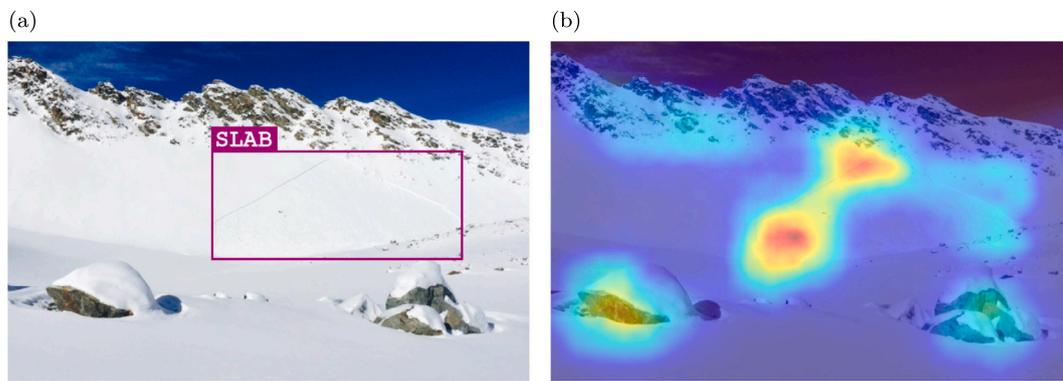
Fig. 15. The `SLAB` image in (a) was misclassified by the ResNet101 model as `NONE`. The model attention is split between the avalanche region and rocks in the foreground and background, as is visible in the Grad-CAM heatmap in (b). Rocks were a common cause of incorrect classifications.
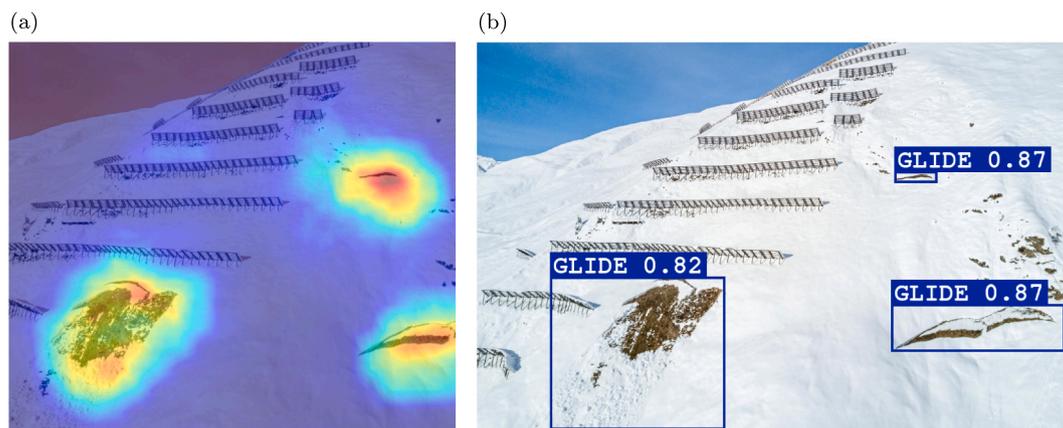
Fig. 16. The Grad-CAM heatmap for ResNet101 is shown in (a) and the bounding boxes predicted by YOLOv3-SPP are shown in (b). Both models were able to identify glide cracks and entire glide avalanches, meaning that they could be used to detect glide avalanches before they release.

Fig. 17. The model detected a small group of loose-snow avalanches (upper left) that the annotators missed. This finding suggests that the model's performance may surpass what is indicated by numerical evaluation metrics.

this is common practice in machine learning. To address this, we provided the avalanche experts with objective guidelines on which parts of the avalanches to outline and observed relatively high coherence between annotators, as discussed in subsection 3.3. An additional measure to minimise subjectivity would be having multiple annotators label all images in the dataset, but this was not feasible for this study due to limited resources.

7. Conclusions

In this paper, we presented a method for automating avalanche detection using ground-based photography. Our models achieved F1 scores of 94.4% per image in the classification task and 65.4% per avalanche region in the segmentation task, demonstrating the potential of our proposed technique to advance operational avalanche monitoring. Furthermore, we found that labelling training data by avalanche type improved model F1 scores. This labelling scheme then enables models to identify the avalanche type in previously unseen images, providing additional details on the current avalanche situation which can be used to enhance forecasting models. In contrast to existing work in this field, our models do not require a DEM of the analysed regions which simplifies application to new regions.

Trained models could be integrated with existing mountain webcams to enable detecting avalanches in near real-time at higher temporal resolutions than previous approaches which have relied on satellite data. To the best of our knowledge, this is the first study on fully automating avalanche detection in photographs that were not obtained from spaceborne or airborne sensors. At the time of writing, we are conducting preliminary experiments using trained models to analyse live webcam data from the Alps. An evaluation of model performance in operational settings will be used to inform further research.

One direction for future work would be to use trained models to continuously gather avalanche data over a designated area of interest. This information could then be cross-referenced with environmental conditions at the time of avalanche releases to improve avalanche forecasting models, for example following recent random forest

approaches proposed by Dkengne Sielenou et al. (2021) and Pérez-Guillén et al. (2022). Such data would also provide valuable insights into the relative frequency of different avalanche types, as there is little data available on this subject.

Another area for future research would be the extension of the models to identify people involved in avalanche situations. Object detection techniques could be used to locate humans in the snow in webcam frames prior to an avalanche, allowing for improved response times in rescue efforts. It is also conceivable that, with sufficient training data, deep learning could be applied to remote monitoring for other environmental hazards such as debris flow, rockfalls, or glacier collapse.

In this study, we have shown that deep learning can be utilised to automate avalanche monitoring in photographs taken from the ground. Our results demonstrate the viability of this method of avalanche detection, which could have significant implications for both avalanche monitoring and forecasting.

CRediT authorship contribution statement

James Fox: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Anna Siebenbrunner:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Data curation, Conceptualization. **Sandra Reitinger:** Writing – original draft, Validation, Software, Investigation, Data curation. **David Peer:** Writing – review & editing, Validation, Supervision, Software, Investigation, Funding acquisition. **Antonio Rodríguez-Sánchez:** Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

Anna Siebenbrunner is an employee of Lo.La Peak Solutions GmbH and David Peer is an employee of Intelligent NLU GmbH. The authors acknowledge that these companies may benefit financially from the commercialisation of models developed during this research. However, the research was conducted objectively and the authors are convinced that the results of the study were not influenced by conflicts of interest. Administrative support for this project was provided by the Avalanche Warning Service Tyrol.

Data availability

Research code and data are available at github.com/j-f-ox/avalanche-detection.

Acknowledgements

References

Adarsh, P., Rathi, P., Kumar, M., 2020. YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 687–694. https://doi.org/10.1109/ICACCS48705.2020.9074315 (iSSN: 2575-7288).

Bessason, B., Eiríksson, G., Thórarinsson, Ó., Thórarinsson, A., Einarsson, S., 2007. Automatic detection of avalanches and debris flows by seismic methods. J. Glaciol.

53, 461–472. https://doi.org/10.3189/002214307783258468 publisher: Cambridge University Press.

Bianchi, F.M., Grahn, J., Eckerstorfer, M., Malnes, E., Vickers, H., 2021. Snow Avalanche Segmentation in SAR Images with fully Convolutional Neural Networks. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14, 75–82. https://doi.org/10.1109/JSTARS.2020.3036914.

Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M., 2020. YOLOv4: Optimal speed and Accuracy of Object Detection. CoRR abs/2004.10934. URL: https://arxiv.org/abs/2004.10934. arXiv: 2004.10934.

Bühler, Y., Hüni, A., Christen, M., Meister, R., Kellenberger, T., 2009. Automated detection and mapping of avalanche deposits using airborne optical remote sensing data. Cold Reg. Sci. Technol. 57, 99–106. https://doi.org/10.1016/j.coldregions.2009.02.007.

Dkengne Sielenou, P., Viallon-Galinier, L., Hagenmuller, P., Naveau, P., Morin, S., Dumont, M., Verfaillie, D., Eckert, N., 2021. Combining random forests and class-balancing to discriminate between three classes of avalanche activity in the French Alps. Cold Reg. Sci. Technol. 187, 103276 https://doi.org/10.1016/j.coldregions.2021.103276.

EAWS, 2022a. Avalanche Size. URL: https://www.avalanches.org/standards/avalanche-size/.

EAWS, 2022b. Determination of the Avalanche Danger Level in Regional Avalanche Forecasting. URL: https://www.avalanches.org/wp-content/uploads/2022/12/EAWS_matrix_definitions_EN.pdf.

Eckerstorfer, M., Malnes, E., 2015. Manual detection of snow avalanche debris using high-resolution Radarsat-2 SAR images. Cold Reg. Sci. Technol. 120, 205–218. https://doi.org/10.1016/j.coldregions.2015.08.016.

Eckerstorfer, M., Malnes, E., Frauenfelder, R., Doomas, U., Brattlien, K., 2014. Avalanche Debris Detection using Satellite-Borne Radar and Optical Remote Sensing. In: International Snow Science Workshop 2014, pp. 131–138. Banff, Canada. URL: https://arc.lib.montana.edu/snow-science/item.php?id=2039.

Eckerstorfer, M., Bühler, Y., Frauenfelder, R., Malnes, E., 2016. Remote sensing of snow avalanches: recent advances, potential, and limitations. Cold Reg. Sci. Technol. 121, 126–140. https://doi.org/10.1016/j.coldregions.2015.11.001.

Eckerstorfer, M., Malnes, E., Müller, K., 2017. A complete snow avalanche activity record from a Norwegian forecasting region using Sentinel-1 satellite-radar data. Cold Reg. Sci. Technol. 144, 39–51. https://doi.org/10.1016/j.coldregions.2017.08.004.

Eckerstorfer, M., Vickers, H., Malnes, E., Grahn, J., 2019. Near-Real Time Automatic Snow Avalanche activity monitoring System using Sentinel-1 SAR Data in Norway. Remote Sens. 11. https://doi.org/10.3390/rs11232863.

Fees, A., Van Herwijnen, A., Altenbach, M., Lombardo, M., Schweizer, J., 2023. Glide-snow avalanche characteristics at different timescales extracted from time-lapse photography. Ann. Glaciol. 1–12. https://doi.org/10.1017/aog.2023.37.

Fox, J., Siebenbrunner, A., Reitinger, S., Peer, D., Rodríguez-Sánchez, A., 2023. Deep learning for real-time avalanche detection in webcam images. In: International Snow Science Workshop Proceedings 2023. Montana State University Library, Bend, Oregon, pp. 1504–1511. URL: https://arc.lib.montana.edu/snow-science/item/3087.

Fromm, R., Adams, M.S., 2016. Spatio-temporal Snow Depth estimates with Time-Lapse Photography. In: International Snow Science Workshop 2016 Proceedings, pp. 628–630. Breckenridge, CO, USA. URL: https://arc.lib.montana.edu/snow-science/item.php?id=2335.

Fuchs, S., Bründl, M., 2005. Damage potential and losses Resulting from Snow Avalanches in Settlements of the Canton of Grisons, Switzerland. Nat. Hazards 34, 53–69. https://doi.org/10.1007/s11069-004-0784-y.

Hafner, E.D., Techel, F., Leinss, S., Bühler, Y., 2021. Mapping avalanches with satellites – evaluation of performance and completeness. Cryosphere 15, 983–1004. https://doi.org/10.5194/tc-15-983-2021. Publisher: Copernicus GmbH.

Hafner, E.D., Barton, P., Daudt, R.C., Wegner, J.D., Schindler, K., Bühler, Y., 2022. Automated avalanche mapping from SPOT 6/7 satellite imagery with deep learning: results, evaluation, potential and limitations. Cryosphere 16, 3517–3530. https://doi.org/10.5194/tc-16-3517-2022.

Hafner, E., Oberson, L., Kontogianni, T., Daudt, R.C., Wegner, J.D., Schindler, K., Bühler, Y., 2023. Using Interactive Object Segmentation to Derive Avalanche Outlines From Webcam Imagery. https://doi.org/10.5194/egusphere-egu23-10867. Conference name: EGU23.

Hafner, E.D., Techel, F., Daudt, R.C., Wegner, J.D., Schindler, K., Bühler, Y., 2023b. Avalanche Size Estimation and Avalanche Outline Determination By Experts: Reliability And Implications for Practice. Preprint. Other Hazards (e.g., Glacial and Snow Hazards, Karst, Wildfires Hazards, and Medical Geo-Hazards). https://doi.org/10.5194/egusphere-2023-586.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning For Image Recognition. In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 770–778.

Heck, M., Hammer, C., Hobiger, M., Herwijnen, A.V., Schweizer, J., Fäh, D., 2018. Automatic classification of continuous seismic data for avalanche monitoring purposes. In: International Snow Science Workshop Proceedings, pp. 631–635. Innsbruck, Austria. URL: https://arc.lib.montana.edu/snow-science/item.php?id=2614.

Hendrikx, J., Dreier, L., Ulivieri, G., Sanderson, J., Jones, A., Steinkogler, W., 2018. Evaluation of an infrasound detection system for avalanches in Rogers Pass, Canada. In: International Snow Science Workshop Proceedings, Innsbruck, Austria, pp. 171–175. URL: https://arc.lib.montana.edu/snow-science/item/2510.

Huang, Z., Wang, J., Fu, X., Yu, T., Guo, Y., Wang, R., 2020. DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection. Inf. Sci. 522, 241–258. https://doi.org/10.1016/j.ins.2020.02.067.

Jocher, G., 2020. YOLOv5 by Ultralytics. URL: https://github.com/ultralytics/yolov5.

Jóhannesson, T., Arnalds, 2001. Accidents and economic damage due to snow avalanches and landslides in Iceland. Jökull 50, 81–94. https://doi.org/10.33799/jokull2001.50.081.

Karas, A., Karbou, F., Giffard-Roisin, S., Durand, P., Eckert, N., 2021. Automatic Color Detection-based Method Applied to Sentinel-1 SAR Images for Snow Avalanche Debris monitoring. IEEE Trans. Geosci. Remote Sens. 60, 1–17. https://doi.org/10.1109/TGRS.2021.3131853.

Kingma, D.P., Ba, J., 2015. Adam: A Method for Stochastic Optimization. In: Bengio, Y., LeCun, Y. (Eds.), 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015. Conference Track Proceedings. URL: http://arxiv.org/abs/1412.6980.

Korzeniowska, K., Bühler, Y., Marty, M., Korup, O., 2017. Regional snow-avalanche detection using object-based image analysis of near-infrared aerial imagery. Nat. Hazards Earth Syst. Sci. 17, 1823–1836. https://doi.org/10.5194/nhess-17-1823-2017.

Kummervold, P., Malnes, E., Eckerstorfer, M., Arntzen, I., Bianchi, F.M., 2018. Avalanche detection in Sentinel-1 radar images using convolutional neural networks. In: International Snow Science Workshop Proceedings. Innsbruck, Austria.

Larsen, S.Ø., Salberg, A.B., Solberg, R., 2013. Automatic Avalanche Mapping Using Texture Classification of Optical Satellite Imagery. EARSeL eProceedings, pp. 399–410.

Lato, M., Frauenfelder, R., Bühler, Y., 2012. Automated detection of snow avalanche deposits: Segmentation and classification of optical remote sensing imagery. Nat. Hazards Earth Syst. Sci. 12, 1–14. https://doi.org/10.5194/nhess-12-2893-2012.

Liu, Y., Chen, X., Qiu, Y., Hao, J., Yang, J., Li, L., 2021. Mapping snow avalanche debris by object-based classification in mountainous regions from Sentinel-1 images and causative indices. CATENA 206, 105559. https://doi.org/10.1016/j.catena.2021.105559.

McClung, D., Schaerer, P., 2000. The Avalanche Handbook, 6th printing ed. The Mountaineers, Seattle, Wash.

Meier, L., Jacquemart, M., Blattmann, B., Arnold, B., 2016. Real-Time Avalanche Detection with Long-Range. Wide-Angle Radars for Road Safety in Zermatt, Switzerland.

Pérez-Guillén, C., Techel, F., Hendrick, M., Volpi, M., Van Herwijnen, A., Olevski, T., Obozinski, G., Pérez-Cruz, F., Schweizer, J., 2022. Data-driven automated predictions of the avalanche danger level for dry-snow conditions in Switzerland. Nat. Hazards Earth Syst. Sci. 22, 2031–2056. https://doi.org/10.5194/nhess-22-2031-2022.

Redmon, J., Farhadi, A., 2018. YOLOv3: an Incremental Improvement. CoRR abs/1804.02767. URL: http://arxiv.org/abs/1804.02767. arXiv:1804.02767.

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.

Schweizer, J., Herwijnen, A.V., 2013. Can near real-time avalanche occurrence data Improve avalanche forecasting?. In: Proceedings of the 2013 International Snow Science Workshop. Grenoble - Chamonix Mont-Blanc, pp. 195–198. URL: https://arc.lib.montana.edu/snow-science/item.php?id=1789.

Schweizer, J., Bartelt, P., van Herwijnen, A., 2015. Snow avalanches, in: Snow and ice-related hazards, risks and disasters. Elsevier, Amsterdam, pp. 395–436. https://doi.org/10.1016/B978-0-12-394849-6.00012-3.

Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D., 2020. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. Int. J. Comput. Vis. 128, 336–359. doi: https://doi.org/10.1007/s11263-019-01228-7.

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition arXiv 1409.1556.

Singh, K.K., Singh, D.K., Thakur, N.K., Dewali, S.K., Negi, H.S., Snehmani, Mishra, V.D., 2022. Detection and mapping of snow avalanche debris from Western Himalaya, India using remote sensing satellite images. Geocarto Int. 37, 2561–2579. https://doi.org/10.1080/10106049.2020.1762762.

Sinha, S., Giffard-Roisin, S., Karbou, F., Deschâtres, M., Karas, A., Eckert, N., Coléou, C., Monteleoni, C., 2019. Can avalanche deposits be effectively detected by deep learning on sentinel-1 satellite SAR images?. In: Climate Informatics Paris, France. URL: https://hal.archives-ouvertes.fr/hal-02278230.

Stucki, T., 2006. Lawinenprognose. Wie entsteht ein Lawinenlagebericht - Möglichkeiten und Grenzen. In: Lawinen und Recht. Proceedings zum Internationalen Seminar vom 6.-9. November 2005 in Davos, pp. 21–34. Davos, Switzerland. URL: https://www.dora.lib4ri.ch/wsl/islandora/object/wsl%3A16752/.

Techel, F., Jarry, F., Kronthaler, G., Mitterer, S., Nairz, P., Pavsek, M., Valt, M., Darms, G., 2016. Avalanche fatalities in the European Alps: long-term trends and statistics. Geographica Helvetica 71, 147–159. https://doi.org/10.5194/gh-71-147-2016.

Thomas, D., Goudie, A., 2013. The Dictionary of Physical Geography. Wiley. URL: https://books.google.at/books?id=aSq9BgAAQBAJ.

University of Innsbruck, . UIBK Avalanche Dataset. URL: https://researchdata.uibk.ac.at//records/h07f4-qzd17 https://doi.org/10.48323/h07f4-qzd17.

Valt, M., Salvatori, R., Plini, P., Salzano, R., Giusto, M., Montagnoli, M., 2013. Climate Change: a New Software to Study the Variations of Snow Images Shot by Web Cam. In: Proceedings of the 2013 International Snow Science Workshop. Grenoble - Chamonix Mont-Blanc, pp. 1004–1008. URL: https://arc.lib.montana.edu/snow-science/item.php?id=1902.

van Herwijnen, A., Schweizer, J., 2011. Monitoring avalanche activity using a seismic sensor. Cold Reg. Sci. Technol. 69, 165–176. https://doi.org/10.1016/j.coldregions.2011.06.008.

Vickers, H., Eckerstorfer, M., Malnes, E., Larsen, Y., Hindberg, H., 2016. A Method for Automated Snow Avalanche Debris Detection through Use of Synthetic Aperture Radar (SAR) Imaging. Earth and Space Science 3. Publisher: Wiley Online Library, pp. 446–462.