

AUTOMATED SNOW AVALANCHE MAPPING WITH DEEP LEARNING IN AERIAL IMAGERY FROM THE EXTREME AVALANCHE WINTER OF 1999

Jor Fergus Dal^{1,2,*}, Elisabeth D. Hafner^{1,2,3}, Torben Peters³, Dominik Narnhofer³, Rodrigo Caye Daudt⁴, Holger Heisig⁵, Yves Bühler^{1,2}

¹WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland

²Climate Change, Extremes and Natural Hazards in Alpine Regions Research Centre CERC, Davos, Switzerland

³Photogrammetry and Remote Sensing, ETH Zurich, Switzerland

⁴Sony Advanced Visual Sensing AG, Zurich, Switzerland

⁵Federal Office of Topography swisstopo, Wabern, Switzerland

ABSTRACT: Snow avalanches are a major natural hazard in mountainous areas, posing a significant risk to lives and infrastructure. Knowing the location, frequency, and magnitude of past snow avalanche occurrences is vital to mitigate these risks. Avalanche mapping is therefore critical in risk mitigation. Previous research has explored various techniques to automate snow avalanche mapping from remote sensing sources, including satellite synthetic aperture radar (SAR), optical airborne, and satellite imagery. However, employing historical optical data sources for full snow avalanche delineation has received less attention. These datasets, created in the analog era, could shed light on the patterns of past avalanche activity. The winter of 1999 in the European Alps is notable for its significant avalanche activity with extreme avalanche runouts. It is well documented through aerial images covering over 12,000 square kilometers in the Swiss Alps and neighboring Austria. Our study leverages deep learning techniques to map the release, path, and deposition areas of visible snow avalanches in these historical images. Building upon previous work in deep learning for automated mapping with optical SPOT 6/7 satellite imagery, we have applied an avalanche segmentation model to historical data. The model is based on a convolutional neural network and relies on digital elevation data and orthorectified optical imagery to produce a pixel-level binary snow avalanche segmentation mask. This task requires the adaption of the methods, originally developed and trained on multi-spectral satellite data, for application on grey-scale data, a problem referred to as a domain gap. We do this by augmenting the historical training set with contemporary samples. The contemporary data is transformed to match the target data domain using a Generative Adversarial Network and coupled with the historical data to train the model. We also offer insights into the estimated uncertainty associated with labels and predictions, serving as a resource for future dataset users. Our research shows how cutting-edge mapping techniques can be adapted to historical imagery.

Keywords: Avalanche mapping, remote sensing, deep learning.

1. INTRODUCTION

Snow avalanches (hereafter referred to as avalanches) are recognized as one of the most imposing natural hazards in mountainous environments. They pose significant risks to roads, houses, and other infrastructure in alpine regions. Understanding where and under what conditions they occur is crucial in mitigating the potential damage they may cause. In particular for hazard mapping and the design of permanent mitigation measures, it is important to understand the extent and impact pressures of past events. Mapping avalanches' spatial and temporal distribution is therefore of high importance. Among large historical events that have shaped the development of avalanche planning in the European Alps is the

extreme winter of 1999. Extraordinary amounts of snowfall lead to very high avalanche activity affecting large areas of Switzerland, neighboring France, and Austria. Especially characteristic of this event were the large avalanche runouts which blocked transportation lines and damaged buildings. The most destructive avalanches were documented through field observations, visually identifying the avalanches from their starting points to the deposits and mapping the outlines. Approximately 1200 destructive avalanches were identified. Aerial photography with the Wild RC30 surveying camera, at a 1:30'000 image scale, was used to capture the areas most affected by these avalanches (EISLF, 2000). However, a complete inventory of the avalanches from these images was never produced. This key information could prove invaluable as a reference for future planning. Remote sensing makes otherwise unreachable areas attainable. It avoids sending a field observer into risky terrain and allows large-scale data acquisition. Tested measurement systems for large-scale mapping include optical, multi-spectral, or radar

*Corresponding author address:

Jor Fergus Dal,
WSL Institute for Snow and Avalanche Research SLF
7260 Davos Dorf, Switzerland;
tel: +41 77 529 10 72;
email: jor.fergus.dal@gmail.com

instruments mounted on satellites, airplanes, and unmanned aerial vehicles (UAVs) (e.g., [Eckerstorfer et al., 2016](#); [Frauenfelder et al., 2012](#)). Avalanches are mapped pixel by pixel from post-processed and georeferenced raster data. This detailed detection and mapping process remains time-consuming and challenging, requiring many hours of manual work from experts. Automating this process in various types of remote sensing data has been attempted with classical machine learning approaches ([Bühler et al., 2009](#); [Lato et al., 2012](#); [Korzeniowska et al., 2017](#)) and more recently using deep learning approaches. Deep learning-based approaches have shown remarkable success in tasks similar to avalanche mapping, such as identifying land cover types, estimating snow cover, and mapping forests ([Zhao et al., 2023](#); [Montginoux et al., 2023](#); [Nguyen et al., 2022](#)). Despite this, its application in avalanche mapping remains relatively new. A key challenge is the limited availability of large, high-fidelity, and accurate datasets. The technique relies on large amounts of data for neural networks to implicitly learn correct classifications. So far, most work in applying deep learning to avalanche mapping has focused on the use of Synthetic Aperture Radar (SAR) ([Waldeland et al., 2018](#); [Eckerstorfer et al., 2019](#); [Bianchi et al., 2021](#)). These methods showcase the feasibility of automated avalanche segmentation, but applications on SAR data do not necessarily translate well to optical imagery. For deep learning approaches using optical imagery we can refer to [Hafner et al. \(2022\)](#). Here they use 1.5 m spatial resolution SPOT 6/7 - Satellite pour l'Observation de la Terre - satellite images coupled with elevation data for large-scale avalanche mapping. They relied on a large set of manually mapped avalanches ([Hafner et al., 2021](#)) for supervised training and model validation. Finally, they compare the model predictions to manual mapping performed by five experts and show that in terms of F1-score, the model predictions are comparable to agreement between experts. Often the limiting factor in training deep learning models is the amount of available labeled data. Recent advances in generative model frameworks have made it possible to produce more data from existing datasets. Among these generative models, there is a class of framework called Generative Adversarial Network (GAN) ([Goodfellow et al., 2014](#)). These models consist of two neural networks that are trained simultaneously through adversarial learning. Synthetic data produced through unpaired image-to-image translation has proven successful in training deep learning models ([Radford et al., 2015](#); [Karras et al., 2019](#)).

We use the model proposed by [Hafner et al. \(2022\)](#) to map avalanches in aerial imagery from the extreme avalanche winter of 1999. The direct appli-

cation is however not feasible, as the SPOT satellite imagery and the 1999 aerial imagery are too dissimilar (domain gap problem). Consequently, the model requires training with manually annotated avalanches from corresponding 1999 images. We increase the amount of available training data by employing a Cycle Generative Adversarial Network (CycleGAN) ([Zhu et al., 2017](#)) to make the SPOT imagery look like the 1999 imagery. We evaluate results, compare them to previous work, examine the reproducibility of the manually mapped avalanches from 1999, and discuss the potential and limitations of using a [CycleGAN](#) to artificially enlarge the training data.

2. DATA AND METHODOLOGY

2.1 Data

2.1.1 Avalanche period

The avalanche winter of 1999 resulted from 30 days of heavy precipitation, strong northwesterly winds, and cold air temperatures between the 27th of January and the 25th of February. New snow accumulation reached 500 cm in some areas, more than the expected amount in a whole season. The snow arrived in three distinct waves, each followed by increased and widespread avalanche activity. For the first time since its introduction in 1993, the highest European avalanche danger level, 'very high', was used for an extended time period (6 days; for details see [EISLF, 2000](#)).

2.1.2 Aerial Images

To document these events, aerial flights were conducted by the Swiss Federal Office of Topography ([swisstopo](#)), on 25.02, 27.02, and 01.03.1999 between 10:30 and 15:30. All images were captured with a Wild RC30 Aerial Film Camera on black and white emulsion ([Geosystems, 2000](#)). Photographs of 23 cm×23 cm were produced with an 80% overlap in the direction of flight at a 1:30'000 scale. [swisstopo](#) later digitized these images with a high-quality aerial image scanner at a scan resolution of 14 μm (1814 dpi; for details see [swisstopo, 2024](#)). The digitized dataset was oriented with automated matching between the images (Tie Points) and toward a digital reference (Ground Control Points). These matches were filtered and used in a Bundle Block Adjustment to produce the global solution. The absolute accuracy of the orthorectified 8-bit mosaic with 0.5 m spatial resolution reaches approximately 1 m and the relative accuracy is in the sub-pixel range (for details see [Heisig and Simmen, 2021](#)).

2.1.3 Digital Terrain Model

The location of an avalanche and its path are dependent and constrained by the terrain. Therefore we use a digital terrain model (DTM) as additional input data. We use the snow-free 2019 [swissALTI3D](#)

2 m resolution DTM (swisstopo, 2019). It offers a 0.5 m vertical accuracy in areas below 2000 m elevation (Light Detection and Ranging (LiDAR) measurements), and a 1 to 3 m vertical accuracy in areas above this elevation (LiDAR or photogrammetry).

2.1.4 Manual Annotations

We use manually mapped avalanches (*labels*) to train the machine learning model and to assess its performance. For manual mapping we follow the annotation procedure described in Hafner et al. (2021). The mapping is performed by an expert from the Institute for Snow and Avalanche Research (SLF), who has extensive experience with using GIS software to map avalanches in optical orthorectified imagery. In total, we map 3597 avalanches in an area of about 2300 km². The mapped avalanches cover 300 km², roughly 13%, of the area captured by aerial imagery. Furthermore, we analyze the reproducibility of the annotations on the 1999 dataset by asking 5 participants to map all visible avalanches in 44.8 km² around the village of Kunkels, GR, Switzerland. We compare the pairwise F1-score between annotators and calculate the mean F1-score to evaluate the general agreement. We evaluate the agreement between the model and the manual mappers by comparing model results to the manually identified avalanches.

2.1.5 Synthetic data

In this section, we aim to leverage the dataset from Hafner et al. (2021) for avalanche detection in SPOT imagery to help us train a model for detecting avalanches in the 1999 images. We'll refer to the SPOT data as the source data and the 1999 imagery as the target data. Our goal is to develop a network or generator that can map the source data to closely resemble the target data, enabling us to use the SPOT labels for training an avalanche detector on the 1999 imagery. This task is made more challenging due to the differences between both modalities: the source data has a 1.5-meter resolution with four channels (RGB-NIR), while the target data has a higher 0.5-meter resolution but only a single grayscale channel. To address this, we upsample the lower-resolution SPOT 2018 data by a factor of 3 using bilinear interpolation. To bridge the gap between these two modalities, we will use CycleGAN, a type of GAN designed for unpaired image translation when matching image pairs are not available for learning the transformation between the two domains. CycleGAN consists of two generators and two discriminators. One generator (G) transforms images from the source domain (SPOT 2018 data) to the target domain (1999 imagery), while the other generator (F) performs the reverse, converting images from the

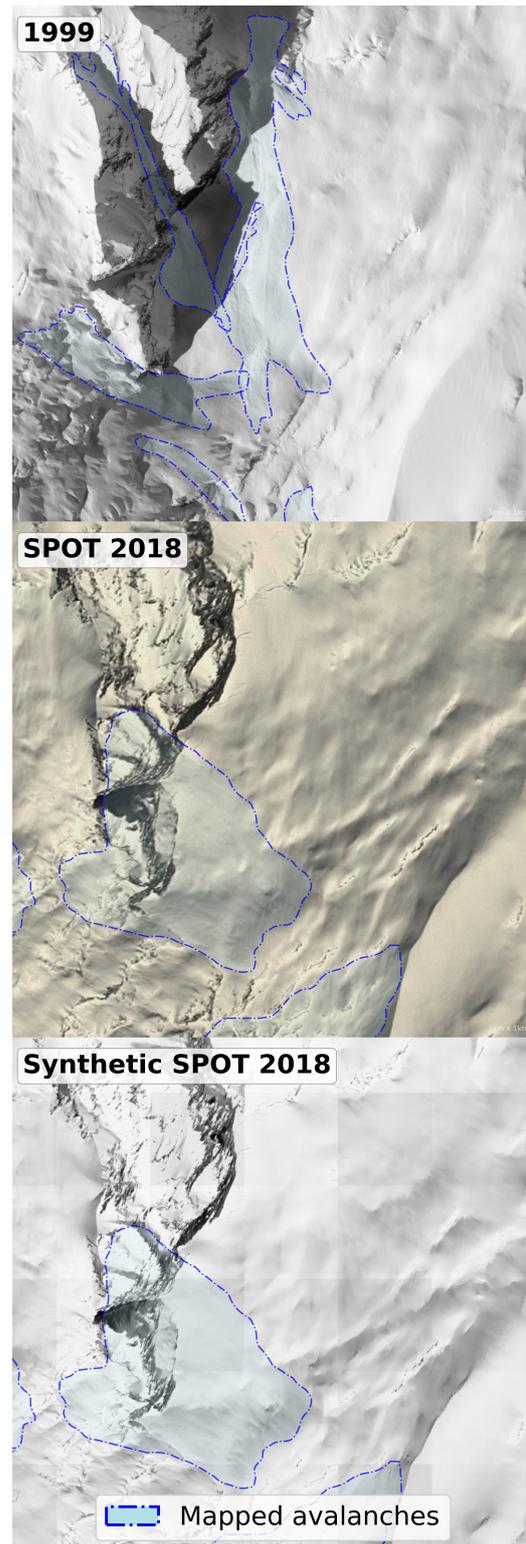


Figure 1: Example of input data from the three different data sources. Each image shows the same 1 km² patch with the overlaid mapped avalanches. The top image stems from the 1999 dataset, the middle is from the SPOT 2018, and the bottom shows the same SPOT 2018 data after transformation with our CycleGAN.

target domain back to the source domain. Correspondingly, there are two discriminators: one (D_T)

that distinguishes between real target domain images and those generated by G , and another (D_S) that differentiates between real source domain images and those generated by F . The cyclic loss in CycleGAN, at a high level, ensures consistency by requiring that after G transforms a source image to the target domain, F can accurately reconstruct the original source image. This loss, calculated by comparing the original and reconstructed images, helps maintain the data’s core characteristics during transformation. The same process is applied in reverse to ensure consistency when converting from the target domain back to the source domain. We create approximately 4800 km² worth of synthetic data covering the Canton of Valais in Switzerland. This adds 4643 manually mapped avalanches to our dataset.

2.2 Model and methodology

For learning avalanche segmentation, we use a model specifically designed for this task by [Hafner et al. \(2022\)](#). This convolutional neural network (CNN) is a modified version of the DeepLabV3+ architecture developed by [\(Chen et al., 2018\)](#), with one of the key modifications being the inclusion of deformable convolutional kernels (for details see [Hafner et al., 2022](#)). Our dataset is separated into exclusive training, validation, and test sets in three different configurations, where the 80-10-10 proportionality between the datasets is maintained. Training, validation, and testing are performed on all three configurations to eliminate regional biases. We report average numbers for each component. The main evaluation of our model performance is conducted on the test dataset spanning approximately 300 km². The area incorporates different terrains and types and sizes of avalanches. To turn the model outputs into binary predictions we choose a threshold value of 0.5. Only the training data is augmented with synthetic data, whereas validation and testing are only performed on real 1999 images. The model is trained on single-channel optical images together with the **DTM**, which are concatenated along the channel dimension. The data is standardized, and samples are selected from “representative points” within known avalanche areas, with a minimum distance between these points to prevent excessive data redundancy. Additionally, we included an equal number of negative samples from areas without avalanches to balance the training data. We minimize the binary cross-entropy (BCE) loss function, measuring the dissimilarity between the predictions and the true labels, to train our model. Weights are updated using an Adam optimizer with a learning rate of $1e^{-4}$ and a multi-step learning rate scheduler. The model is fed samples in batches of 6, and gradients are accumulated over 8 steps before backpropagation to virtually increase the batch size to 48. Additionally, we bilin-

Experiment	F1-score	Precision	Recall
Synthetic 2018	0.303	0.394	0.361
Aerial 1999	0.406	0.525	0.388
Combined	0.451	0.570	0.430

Table 1: Averaged test results for the three experiments. The first row represents training on synthetic data only. The second row results from training without synthetic data. The third row is the combination of the two datasets in training.

early downsample the input images by a factor of two, effectively creating patches with a 1 m resolution. This allows for a larger geographical context per patch while limiting the computational and GPU memory costs. We assess our model’s performance on the test set using precision, recall, and F1-score for the avalanche class. These metrics are based on categorizing pixel-wise predictions into true positives, TP , true negatives, TN , false positives, FP , and false negatives, FN , (for details see [Story and Congalton, 1986](#)). Their definitions are given in equations (1)–(3).

$$\text{precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{F1-score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (3)$$

3. RESULTS

3.1 Test predictions

Our model achieves an average F1-score of 0.451, with a precision of 0.570 and a recall of 0.430 (see Tab. 1). The lower recall value indicates a large number of missed avalanche pixels. We qualitatively identify most false negatives in the avalanche release area (see Fig. 2). These areas are often hard to distinguish visually because of the weak textures in the images. The snow in these parts is mostly smooth. False positives on the other hand mainly occur at high elevations, in patches around summits and cliffs.

3.2 Ablations

To evaluate the effect of including the synthetic data in the training process we run experiments where we train on synthetic data only and without any synthetic data, and compare the performance to the combined dataset. These comparisons are run after hyper-parameter tuning which determines what training parameters provide the best results on the validation dataset. As for the predictions, we run three parallel experiments with different training, validation, and test dataset combinations to reduce bias. The numbers presented here are the average scores on the test set and are summarized in

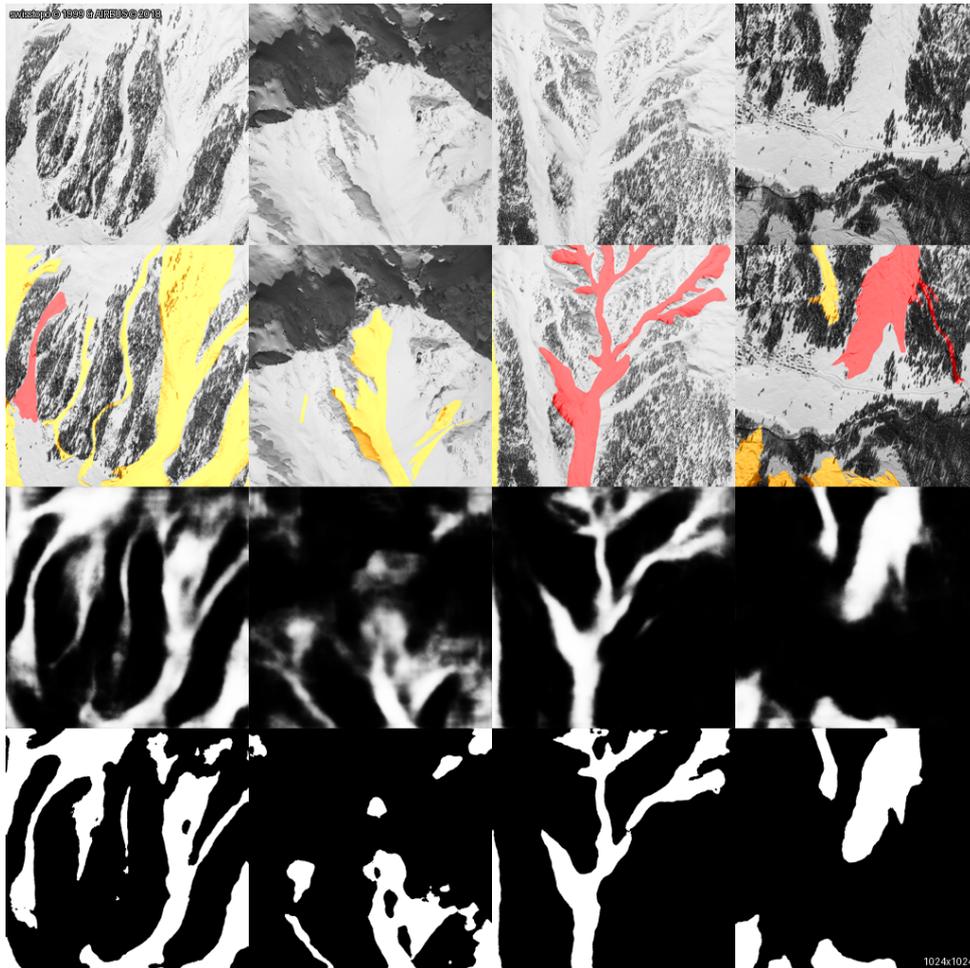


Figure 2: A selection of model predictions on test set. The first column shows the input images. The second column shows the avalanche tracks, with "estimated" avalanches in orange and "created" avalanches in red (for definition see Hafner et al. (2021)). The third column shows raw model outputs. The prediction certainties range from 0 (black), no avalanche, to 1 (white), avalanche. The fourth column shows the predictions thresholded at 0.5.

Tab. 1. Training on synthetic data produces the lowest scores overall. Precision and recall values are similar, suggesting no pronounced bias towards errors of commission or omission. Switching to the 1999 data for training we see a large performance increase. Here, as for the combined experiment, precision is significantly higher than the recall, suggesting significant errors of omission. Combining the two produces the best scores, but the difference is smaller than between the two other experiments.

3.3 Manual mapping study

To assess the consistency among different labelers, we compared their avalanche annotations using pairwise F1-scores. The results reveal significant variability, with F1-scores ranging from 0.29 to 0.58, and an average of 0.42. This indicates that mapping avalanches in our dataset is challenging, with moderate to low agreement among participants. Visually, it is evident that while most participants consistently identify and map avalanches in the same regions (Fig. 3), there is considerable dis-

agreement regarding the precise boundaries. The highest agreement is observed in the lower sections of the identified avalanches, whereas the most significant discrepancies occur at the highest elevations. When comparing the model and the manual labelers by treating the model predictions as a 6th mapper the pairwise F1-scores fall between 0.288 and 0.504, with a mean of 0.406. The calculated agreement is similar but slightly lower than the average between participants. The precision (min 0.184, max 0.629, mean 0.389) and recall (min 0.339, max 0.696, mean 0.531) suggest that the model tends to over-predict avalanches and produces false positives. By visually comparing the predictions, we see that the model typically maps the deposition areas well, whereas the release areas are missed. Additionally, the model maps some smaller avalanches in areas where no participant has mapped any.

4. DISCUSSION

Our study demonstrates that the machine learning algorithm can learn to recognize and segment

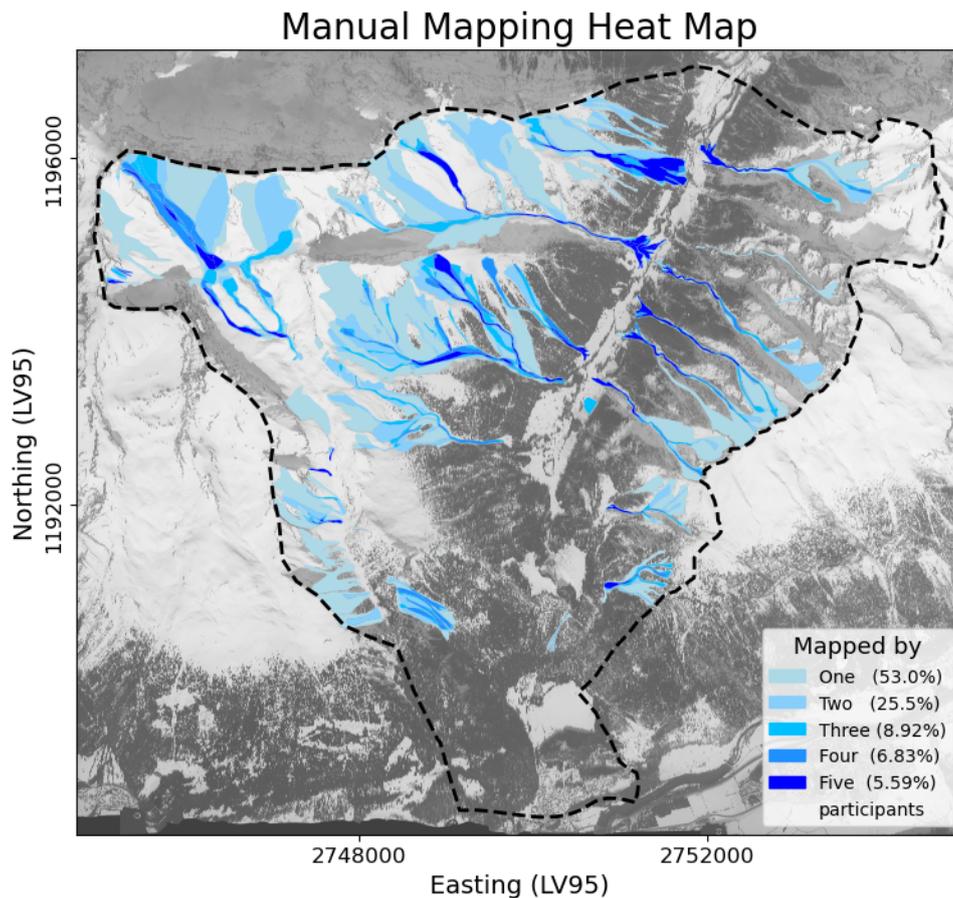


Figure 3: Heat map illustrating agreement on the manual mapping study-area. The darker the hue, the greater the agreement between the five participants (relative area per agreement class in parenthesis).

avalanches in the historical aerial images from 1999. Additionally, we demonstrate that enlarging the training data with synthetic images improves performance. While the model does not perfectly replicate the training mappings, it achieves a level of agreement with human mappers that is comparable to the agreement observed among the mappers themselves. We observe that the model mostly makes mistakes by omitting avalanche release areas. These parts of the avalanches present fewer visual and textural features. Additionally, many avalanches have been snowed on, which would more effectively cover the smoother features of the release areas compared to deposition areas. This is likely the reason why the model performs better on avalanche deposits where the textures are much more visible. Because the input data are single-band images, the model can only learn from differences in pixel values making textures vital. The model also tends to predict avalanches in areas where there are none. This occurs primarily in terrain with the right slope angle, suggesting that the model has learned how to interpret the terrain. Because the release area may look texturally ambiguous it is expected that the model has a harder time

learning to segment this properly. The ambiguity and lack of clear textural cues in many release areas appears to confuse the model, resulting in a higher likelihood of false positives. The input data used to train the model for avalanche segmentation presents several challenges that directly impact the model's performance and learning capabilities. One significant factor are the heavy snowfalls that preceded the image acquisition, making it difficult to clearly identify certain avalanches in the data. This is supported by the lower number of "well-visible" avalanches in the 1999 dataset (for details see [Hafner et al., 2021](#)), compared to the avalanches mapped in the SPOT 2018 data (6% in 1999 versus 33% in 2018), likely due to snow coverage. The reduced visibility of these avalanches limits the availability of clear labels, which affects the model's ability to generalize well, especially in identifying less apparent avalanche features. This issue might be exacerbated by the training data suffering from lower radiometric resolution and is limited to single-spectral imagery. The lack of spectral richness restricts the amount of information the model can extract, making it harder to distinguish between different types of terrain and snow conditions. Higher

spectral diversity and resolution could potentially offer more detailed cues that are critical for segmentation tasks like this one. Another interesting observation is that the model appears to learn useful features from the *DTM*. This is despite the use of data that was collected two decades later and during the summer months, meaning the *DTM* represents summer elevations rather than the winter terrain when avalanches occur. The temporal discrepancy does not seem to pose a significant problem for model learning because the exact elevations are not critical. Instead, the model seems to leverage the general topographical structure, which remains relatively consistent over time, to better infer potential avalanche zones. In the manual mapping study we show how even among manual mappers there is large disagreement. Similar studies comparing manual mappings also find significant disagreement (Hafner et al., 2021, 2023), though our values suggest that the 1999 dataset is particularly prone to disagreement. This further supports the thesis that the avalanches in the 1999 dataset are particularly hard to detect. For example, certain poorly visible avalanches are completely omitted by most participants. Still, most mappers agree on the presence of an avalanche, with most mapped avalanches being mapped by all participants. Yet the strong disagreement concerns the exact outlines. The disagreement is particularly strong in the release area and tends to decrease in the deposition. As discussed above, this likely results from deposits being better visible. This finding is different from the study conducted on manual mappings in SPOT data (Hafner et al., 2021), where the largest differences seemed to stem from lighting conditions. Comparing model predictions to agreement between manual mappers, we see that the model's variability in accuracy mimics agreement between manual mappers. A lot of the uncertainty in avalanche mapping stems from lack of strong agreement between experts, and by consequence a standard way to map avalanches. These uncertainties are propagated into the model because we use manually mapped labels for training. This issue highlights the core challenge of completely automated avalanche mapping. Our ablation study reveals that while the use of synthetic data had a positive effect on model predictions, its impact was somewhat limited. One key limitation stems from the lack of optimized *CycleGAN* training, as our methodology does not ensure that the synthetic data is specifically tailored for the intended task. The quality of the synthetic data was primarily evaluated visually and through the observed improvement in model validation performance. However, training *CycleGAN* models can be challenging, as it requires strategies to align the learned transformations with segmentation improvements. One approach we could have implemented is us-

ing avalanche labels during training as a consistency check. Additionally, backpropagating the segmentation model's loss to guide the *CycleGAN* training process would have been another potential enhancement. Lastly, there is more annotated SPOT 2018 data available. Transforming this data would likely result in moderate improvement, although we believe the bottle neck lies with optimization. Due to the inherent uncertainties in the labels and model predictions, we recommend manual corrections to ensure the quality of the avalanche segmentation map before making the data available to practitioners.

5. CONCLUSION AND OUTLOOK

We adapted an existing algorithm to map avalanches in panchromatic aerial images from the avalanche winter of 1999. Because deep learning methods need large annotated training datasets, we used synthetically generated samples in combination with 1999 samples to train our model. The study revealed that the use of synthetic data increased model performance. Our experiments show that the approach reaches good performance with an F1-score in the range of agreement between manual mappers. The model has higher accuracy for avalanche deposits but lower reliability in release areas. This is reflected in the results of the manual mapping study. The lack of strong agreement between experts fuelled by the absence of a standard way to map avalanches, causes uncertainty in the model since training relies on manual labels. Further research into this topic could explore using annotations from several experts, and the combination of these to teach the model to focus on features experts agree on. Furthermore, we see potential in further exploring synthetic data to find the best configuration for training our model. We consider efficient and effective strategies to transform existing data into new domains, like the one we proposed, as essential for automatically retrieving information in less explored, less annotated datasets. In conclusion, our study is an important step towards automatic avalanche mapping from historical aerial imagery. We believe the expansion of training sets with synthetic imagery will make automation feasible for further data where the lack of annotations currently impedes automation. Our manual mapping study underscores the complexities of avalanche segmentation and adds another puzzle piece in this critical area. The insights gained from this work have the potential to significantly improve our understanding and management of avalanche hazards, paving the way for safer mountain environments.

=====

ACKNOWLEDGEMENT

We would like to thank all the participants of the manual mapping study for their many hours of dedicated work.

References

- Bianchi, F. M., Grahn, J., Eckerstorfer, M., Malnes, E., and Vickers, H.: 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, doi:10.1109/JSTARS.2020.3036914, URL <https://ieeexplore.ieee.org/document/9254089/>, 2021.
- Bühler, Y., Hüni, A., Christen, M., Meister, R., and Kellenberger, T.: Automated detection and mapping of avalanche deposits using airborne optical remote sensing data, *Cold Regions Science and Technology*, 57, 99–106, doi:10.1016/j.coldregions.2009.02.007, URL <https://linkinghub.elsevier.com/retrieve/pii/S0165232X09000251>, 2009.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H.: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, doi:10.48550/ARXIV.1802.02611, URL <https://arxiv.org/abs/1802.02611>, publisher: arXiv Version Number: 3, 2018.
- Eckerstorfer, M., Bühler, Y., Frauenfelder, R., and Malnes, E.: Remote sensing of snow avalanches: Recent advances, potential, and limitations, *Cold Regions Science and Technology*, 121, 126–140, doi:10.1016/j.coldregions.2015.11.001, URL <https://linkinghub.elsevier.com/retrieve/pii/S0165232X150002591>, 2016.
- Eckerstorfer, M., Vickers, H., Malnes, E., and Grahn, J.: Near-Real Time Automatic Snow Avalanche Activity Monitoring System Using Sentinel-1 SAR Data in Norway, *Remote Sensing*, 11, 2863, doi:10.3390/rs11232863, URL <https://www.mdpi.com/2072-4292/11/23/2863>, 2019.
- EISLF: Der Lawinenwinter 1999. Ereignisanalyse., Tech. rep., Eidg. Institut für Schnee- und Lawinenforschung, Davos, URL <https://www.dora.lib4ri.ch/wsl/islandora/object/wsl:17698/datastream/PDF>, 2000.
- Frauenfelder, R., Solberg, R., Larsen, S. Ø., Salberg, A.-B., and Bjordal, H.: Remote-Sensing Derived Avalanche Inventory Data, in: *Proceedings, 2012 International Snow Science Workshop*, Anchorage, Ak., USA, URL <http://arc.lib.montana.edu/snow-science/item/1708>, 2012.
- Geosystems, L.: Leica RC30 Aerial Film Camera, URL https://geotech.sk/downloads/Fotogrametria/Letecke-kamery/RC30_brochure.pdf, 2000.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative Adversarial Networks, doi:10.48550/ARXIV.1406.2661, URL <https://arxiv.org/abs/1406.2661>, version Number: 1, 2014.
- Hafner, E. D., Techel, F., Leinss, S., and Bühler, Y.: Mapping avalanches with satellites – evaluation of performance and completeness, *The Cryosphere*, 15, 983–1004, doi:10.5194/tc-15-983-2021, URL <https://tc.copernicus.org/articles/15/983/2021/>, 2021.
- Hafner, E. D., Barton, P., Daudt, R. C., Wegner, J. D., Schindler, K., and Bühler, Y.: Automated avalanche mapping from SPOT 6/7 satellite imagery with deep learning: results, evaluation, potential and limitations, *The Cryosphere*, 16, 3517–3530, doi:10.5194/tc-16-3517-2022, URL <https://tc.copernicus.org/articles/16/3517/2022/>, 2022.
- Hafner, E. D., Techel, F., Daudt, R. C., Wegner, J. D., Schindler, K., and Bühler, Y.: Avalanche size estimation and avalanche outline determination by experts: reliability and implications for practice, *Natural Hazards and Earth System Sciences*, 23, 2895–2914, doi:10.5194/nhess-23-2895-2023, URL <https://nhess.copernicus.org/articles/23/2895/2023/>, 2023.
- Heisig, H. and Simmen, J.-L.: Re-engineering the Past: Countrywide Geo-referencing of Archival Aerial Imagery, *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 89, 487–503, doi:10.1007/s41064-021-00162-z, URL <https://link.springer.com/10.1007/s41064-021-00162-z>, 2021.
- Karras, T., Laine, S., and Aila, T.: A Style-Based Generator Architecture for Generative Adversarial Networks, in: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4396–4405, IEEE, Long Beach, CA, USA, doi:10.1109/CVPR.2019.00453, URL <https://ieeexplore.ieee.org/document/8953766/>, 2019.
- Korzeniowska, K., Bühler, Y., Marty, M., and Korup, O.: Regional snow-avalanche detection using object-based image analysis of near-infrared aerial imagery, *Natural Hazards and Earth System Sciences*, 17, 1823–1836, doi:10.5194/nhess-17-1823-2017, URL <https://nhess.copernicus.org/articles/17/1823/2017/>, 2017.
- Lato, M. J., Frauenfelder, R., and Bühler, Y.: Automated detection of snow avalanche deposits: segmentation and classification of optical remote sensing imagery, *Natural Hazards and Earth System Sciences*, 12, 2893–2906, doi:10.5194/nhess-12-2893-2012, URL <https://nhess.copernicus.org/articles/12/2893/2012/>, 2012.
- Montginoux, M., Weissgerber, F., Monteil, C., and Girard, A.: Snow cover estimation by deep-learning segmentic segmentation of radar images based on optical image references, doi:10.5194/egusphere-egu23-7240, URL <https://meetingorganizer.copernicus.org/EGU23/EGU23-7240.html>, 2023.
- Nguyen, T.-A., Kellenberger, B., and Tuia, D.: Mapping forest in the Swiss Alps treeline ecotone with explainable deep learning, *Remote Sensing of Environment*, 281, 113217, doi:10.1016/j.rse.2022.113217, URL <https://linkinghub.elsevier.com/retrieve/pii/S0034425722003248>, 2022.
- Radford, A., Metz, L., and Chintala, S.: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, doi:10.48550/ARXIV.1511.06434, URL <https://arxiv.org/abs/1511.06434>, version Number: 2, 2015.
- Story, M. and Congalton, R. G.: Accuracy assessment: a user's perspective, *Photogrammetric Engineering and remote sensing*, 52, 397–399, 1986.
- swisstopo: swissALTI3D, URL <https://www.swisstopo.admin.ch/de/geodata/height/alti3d.html#dokumente>, 2019.
- swisstopo: Aerial Photographs, URL <https://www.swisstopo.admin.ch/en/analogue-aerial-photographs#Scanned-analogue-images>, 2024.
- Waldeland, A. U., Reksten, J. H., and Salberg, A.-B.: Avalanche Detection in Sar Images Using Deep Learning, in: *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2386–2389, IEEE, Valencia, doi:10.1109/IGARSS.2018.8517536, URL <https://ieeexplore.ieee.org/document/8517536/>, 2018.
- Zhao, S., Tu, K., Ye, S., Tang, H., Hu, Y., and Xie, C.: Land Use and Land Cover Classification Meets Deep Learning: A Review, *Sensors*, 23, 8966, doi:10.3390/s23218966, URL <https://www.mdpi.com/1424-8220/23/21/8966>, 2023.
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A.: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, doi:10.48550/ARXIV.1703.10593, URL <https://arxiv.org/abs/1703.10593>, version Number: 7, 2017.