Detecting Avalanche Deposits using Variational Autoencoder on Sentinel-1 Satellite Imagery

Saumya Sinha*
University of Colorado
Boulder, USA

Sophie Giffard-Roisin *
ISTerre; Univ. Grenoble Alpes, Univ.
Savoie Mont Blanc, CNRS, IRD, IFSTTAR
Grenoble, France

Fatima Karbou
CNRM-GAME, Météo-France
Grenoble, France

Michael Deschatres
Irstea, Université Grenoble Alpes
France

Anna Karas
CNRM-GAME, Météo-France
Grenoble, France

Nicolas Eckert
Irstea, Université Grenoble Alpes
France

Cécile Coléou
CNRM-GAME, Météo-France
Grenoble, France

Claire Monteleoni
University of Colorado
Boulder, USA

Abstract

Avalanche monitoring is a crucial safety challenge, especially in a changing climate. Remote sensing of avalanche deposits can be very useful to identify avalanche risk zones and time periods, which can in turn provide insights about the effects of climate change. In this work, we use Sentinel-1 SAR (synthetic aperture radar) data on the French Alps for the exceptional winter of 2017-18, with the goal of automatically detecting avalanche deposits. We address our problem with an unsupervised learning technique. We treat an avalanche as a rare event, or an anomaly, and we learn a variational autoencoder, in order to isolate the anomaly. We then evaluate our method on labeled test data, using an independent in-situ avalanche inventory as ground truth. Our empirical results show that our unsupervised method obtains comparable performance to a recent supervised learning approach that trained a convolutional neural network on an artificially balanced version of the same SAR data set along with the corresponding ground-truth labels. Our unsupervised approach outperforms the standard CNN in terms of balanced accuracy (63% as compared to 55%). This is a significant improvement, as it allows our method to be used in-situ by climate scientists, where the data is always very unbalanced (< 2% positives). This is the first application of unsupervised deep learning to detect avalanche deposits.

1 Introduction

Reliable inspection of avalanche debris is important to study the stability of the snowpack and variability of the avalanche activity, which can aid in avalanche forecasting. Note only can such

*Corresponding authors: Saumya Sinha, saumya.sinha@colorado.edu; S. Giffard-Roisin, sophie.giffard-roisin@mines-saint-etienne.org

NeurIPS Tackling Climate Change with Machine Learning workshop (2019), Vancouver, Canada.

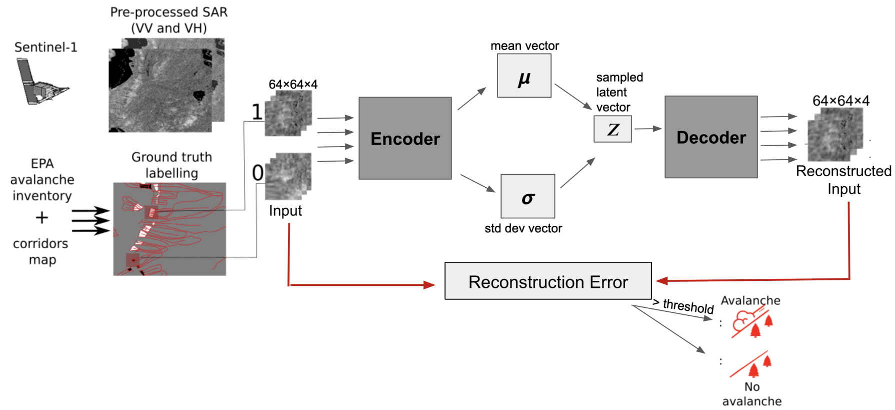

Figure 1: Evaluation Pipeline for the unsupervised learning framework. VAE is first trained only with negative examples. New positive (containing an avalanche) and negative image patches are then passed through the learnt network to obtain reconstruction error which are subject to an optimal threshold for classifying them as *avalanche* or *no avalanche*.

research improve human safety in mountainous regions under a changing climate, but it can also provide insights on climate change. For this work, we have used backscatter coefficients at C-band from the Sentinel-1A and 1B satellites observing the French Alps every 6 days (active microwave imaging). Detection methods which isolate debris-like features by contrasting backscatter between a deposit and surrounding undisturbed snowpack have been used to detect avalanche deposits (Karbou et al. [2018b]).

While recent work by Karbou et al. [2018b], Waldeland et al. [2018] demonstrates the potential of Sentinel-1 SAR data for avalanche detection, simple change detection methods (Karbou et al. [2018a]) fail to capture the complexity of the interaction between the radar signals and the snow. Few classification-based machine learning techniques have been used to tackle the problem on SAR imagery, e.g., a random forest classifier in Hamar et al. [2016], and a Convolutional Neural Network (CNN) in Waldeland et al. [2018], Sinha et al. [2019]. Both Hamar et al. [2016] and Waldeland et al. [2018] rely on expert labelling and the accuracy of the expert labelling has not been discussed. Unlike the previous supervised works, in this paper, we take an unsupervised learning approach to train our model, allowing our techniques to scale to large unbalanced SAR data sets. In order to evaluate our methods in comparison with related work, we use an avalanche event inventory as ground truth instead of an expert labelled data. This inventory is maintained by forest rangers from ONF (Office National des Forêts) and stored by Irstea research institute. It includes more than 4000 avalanche events in the season 2017-2018 across more than 3000 mountain paths (corridors) in the French Alps. However, the avalanche deposits are only roughly localized.

Avalanche are rare events. As such, there is a need for methods which are able to detect them in unbalanced datasets, in constrast to recently developed methods that use artificially balanced data (Sinha et al. [2019]). Unsupervised learning approaches have recently proven to be efficient for very unbalanced data, i.e. for rare events (Kiran et al. [2018]). We propose training a Variational Autoencoder for this task. We evaluate our method in comparison with the supervised method by Sinha et al. [2019] and a baseline method used in the avalanche detection literature (thresholding by Karbou et al. [2018a]). From the SAR acquisitions of both current time and 6 days earlier acquisition, we first trained our model to only detect new avalanches. Later, we included past avalanches which had occurred in the past one month from the current SAR acquisition and did a comparative study between learning new avalanches vs one month old avalanches.

2 Method

Unsupervised Learning Framework A recent supervised learning CNN-based approach [Sinha 2019] undersamples and throws away a major portion of the dataset in order to keep the positives and negatives balanced. With only 1.4% positives (new avalanches) in our dataset, we used an unsupervised approach using a variational autoencoder (VAE) (Kingma and Welling [2014]) as a rare event classifier. For every SAR acquisition date, we built a label map where we labeled a corridor positive if an avalanche was seen between the last acquisition and the current date (6-day window) and negative if not. We crop 64x64 image patches centered on the lowest elevation part of the corridors with the zone of deposition. We observe the VV and VH SAR polarizations from the descending relative orbit 139 and manipulate these polarizations to give us VV^*, VV^*_{old}, VH^*, VH^*_{old} following Sinha et al. [2019].

Our VAE only observes negative image patches and learns the underlying characteristics of a non-avalanche zone. The model struggles when trying to reconstruct an image containing avalanche (positive). These high differences between the original and reconstructed image give us a measure of anomalous behavior and help label that image as an outlier or *avalanche*. We use 4 feature image channels in our current dataset: VV^*, VV^*_{old}, slope, and angle or orientation (additional measurements provided by SAR). The training set includes 40,551 images with 39,996 as negative samples. While feeding data that is only negative to train our VAE requires the supervision to find negative data, more than 98% of our data is negative. Meanwhile, we have also tried training the VAE on all the data (regardless of label), and found comparable performance.

We observed that vanilla autoencoders struggled to classify the avalanches correctly, thus, we used VAEs following the recent study by Lu and Xu [2018]. The VAE loss can be decomposed into reconstruction error (MSE Loss) and Kullback–Leibler (KL) loss. We use a labeled validation set to tune the threshold on reconstruction error used to classify an avalanche event, i.e., we select the threshold (a hyper-parameter to our algorithm) which gives the best results on the validation data set. Experiments on the validation set showed that the best results were obtained by thresholding only the reconstruction error of an image, as opposed the KL-loss. Fig 1 explains how an image patch fed into a trained VAE outputs a reconstruction score which is used to classify the image as positive or *outlier* based on the optimal threshold. We built a fully convolutional architecture with a latent space dimension of 300 similar to Lu and Xu [2018] and chose a weight of 0.1 for the KL loss term for calculating our VAE loss.

3 Evaluation

We select the model performing best on the validation dataset and report its results on the test set (kept hidden). From Table 1 we can see that, on the "All Alps" data set covering 19 mountain chains, the unsupervised VAE method outperforms the supervised CNN (Sinha et al. [2019]) in terms of balanced accuracy (average of recall on both the classes), while achieving comparable F1 score. Meanwhile it outperforms the baseline method on both metrics. The balanced accuracy score of 63% shows that we are now better at identifying avalanches.

We also show results on the Haute Maurienne chain, a subset which forms the majority of the All Alps data set and is one of the most avalanche-prone mountain ranges. We see similar performance advantages of the VAE method in this region where avalanche monitoring is extremely critical. Our method not only utilizes most of our data but is more generalized and can be extended to locating avalanches as "outliers" even in new mountains.

The F1-score, however, is still low (11% in all Alps, 23% in Haute Maurienne). One hypothesis is that the avalanches might still be visible weeks after, making it difficult to distinguish between past and new avalanches. In order to test this hypothesis, we show in Table 1 the results of Unsupervised - VAE* when we set a corridor 'positive' if an avalanche occurred in the past month and not only 6 days. We can see that the F1-score increases both in all Alps and in the Haute Maurienne.

	All Alps		**Haute Maurienne**	
	Balanced Accuracy	**F1-score**	**Balanced Accuracy**	**F1-score**
Baseline	0.58	0.05	0.58	0.12
Supervised - CNN	0.55	0.12	0.55	0.17
Unsupervised - VAE	0.63	0.11	0.62	0.23
Unsupervised - VAE*	0.62	0.35	0.68	0.66

Table 1: Comparison of our method (Unsupervised - VAE) with the Supervised and Baseline methods on the test set (7,166 samples). Results are shown on the whole mountain range and Haute Maurienne (one of the most susceptible mountain zone). Unsupervised - VAE* is the Unsupervised-VAE method when all avalanches happened in the "past one month" are included (as positive labels).

4 Discussion

This is a first quantitative study exploring the potential of applying unsupervised deep learning methods to detect avalanche deposits and obtain avalanche activity statistics that can be useful in a real-world setting. While work has been done to learn from the original unbalanced dataset in Waldeland et al. [2018], they suffer from relying on a manually labelled dataset. Even though we treated the problem as unsupervised in learning our variational autoencoder, we obtain comparable results to a CNN method which had advantage of supervised training. Moreover, using only the negative labels prevented us from overfitting over the few positive labels of this dataset. We built our model with varying number of feature channels and observed that including additional features such as the slope and orientation of the mountain helped in improving the performance. We also showed that including all avalanches happened in the past month improved the result.

References

Jarle Bauck Hamar, Arnt-Børre Salberg, and Florina Ardelean. Automatic detection and mapping of avalanches in sar images. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 689–692. IEEE, 2016.

Fatima Karbou, Cécile Coléou, Maxime Lefort, Michael Deschatres, Nicolas Eckert, Rémy Martin, Gilles Charvet, and Anne Dufour. Monitoring avalanche debris in the french mountains using sar observations from sentinel-1 satellites. In *International snow science workshop (ISSW)*, 2018a.

Fatima Karbou, Maxime Lefort, Marie Dumont, Nicolas Eckert, Michael Deschrtres, and Rémy Martin. Multi-temporal avalanche debris mapping in the french mountains using synthetic aperture radar observations from sentinel-1. In *EGU General Assembly Conference Abstracts*, volume 20, page 18024, 2018b.

Diederik P Kingma and Max Welling. Stochastic gradient vb and the variational auto-encoder. In *Second International Conference on Learning Representations, ICLR*, 2014.

B Kiran, Dilip Thomas, and Ranjith Parakkal. An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. *Journal of Imaging*, 4(2):36, 2018.

Yuchen Lu and Peng Xu. Anomaly detection for skin disease images using variational autoencoder. *arXiv preprint arXiv:1807.01349*, 2018.

Saumya Sinha, Sophie Giffard-Roisin, Fatima Karbou, Michaël Deschatres, Anna Karas, Nicolas Eckert, Cécile Coléou, and Claire Monteleoni. Can Avalanche Deposits be Effectively Detected by Deep Learning on Sentinel-1 Satellite SAR Images? In *Climate Informatics 2019*, Paris, France, October 2019. URL `https://hal.archives-ouvertes.fr/hal-02278230`.

Anders U Waldeland, Jarle Hamar Reksten, and Arnt-Borre Salberg. Avalanche detection in sar images using deep learning. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 2386–2389. IEEE, 2018.