Distributed acoustic sensing and machine learning: rockfall detection at Mt. Meager, B.C.

Introduction

A Distributed Acoustic Sensing (DAS) system is utilized at the Mount Meager glacier in British Columbia for an acquisition program to monitor the recently reawakening volcanic region (Klaasen et al., 2021). Thirty days of data were acquired throughout the months of September and October of 2019 to record seismic characteristics related to rockfalls and other strain events. This volcanic region is estimated to have the largest geothermal potential in Canada making it a highly desirable site to monitor, and Klaasen et al. (2021) demonstrates the capabilities of DAS to reveal undiscovered seismicity.

A challenge brought upon the DAS data as a result of the acquisition location is noise. The high elevation glacial environment where the fibre cable is situated is exposed to weather events like wind, rain, and snow (Klaasen et al., 2021). Different weather events occurring at different times and days throughout the acquisition introduce varying noise levels. Fresh snowfall can reduce the wind noise present by acting like an insulator around the cable. These natural noise sources impact the clarity of the DAS trace making it difficult to identify rockfall events. Unnatural noise related to some fibre channel errors also appear in the data, obstructing events. Instrumental noise was also identified by Klaasen et al. (2021), caused by a running generator required to power the battery used by the DAS. Rockfall events vary in the strength they present in the DAS data and low amplitude events have the potential to be overshadowed by these noise sources. The inconsistency of the noise intensity also makes it challenging to identify the same kinds of events at different times.

We explore the effect of training and predicting rockfall events in the presence of noise on variably filtered DAS data to identify the most accurate event detection method, while considering the computational costs associated. Measurement of false positives and false negatives is performed using random sampling and manual classification for each case. Our goal is to identify the best method of efficiently producing high accuracy machine learning event detection systems, using minimal training data and filtering techniques, for application in near real-time large scale monitoring situations.

Theory

A useful workflow of implementing a machine learning model is described by Arief et al. (2021) when processing different microseismic data including the acquisition of raw data, data pre-processing, feature engineering, learning algorithms, and post-processing. We vary from this workflow as we use a technique (transfer learning) that takes advantage of well known deep learning convolutional neural networks and their existing weights (Simonyan and Zisserman, 2015), to control the general task of extracting specific features from data that drive the classification results. We can take these pre-trained models and their knowledge gained solving one problem, and re-purpose / retrain these weights to identify features from the DAS data. Doing so, we avoid tedious manual feature engineering (Heaton, 2016) where image data would need to be extracted and processed in a number of ways, then trained on large datasets to develop fresh weights that would control which extracted / processed images contribute and identify specific features best. We are therefore able to focus on building successful learning algorithms quickly, and on very small training data sets.

Our machine learning workflow for event detection is presented in Table 1. Raw DAS data input is acquired in step one, then in step two it is pre-processed following a given data pre-processing workflow. The third step in Table 1 is developing an initial machine learning model to eventually train with our data using transfer learning. The transferred model is trained to detect over 1000 different classes of images from the ImageNet database (Krizhevsky et al., 2017), and is highly effective at identifying geometric features. To use this model, additional output layers are applied to allow for re-training on new data we provide, while still maintaining and evolving the weights within the network (Convolutional Neural Network (CNN), Arief et al. (2021)) that are already very good at identifying geometric features. Transfer learning re-purposes the models generalizable skill of geometric feature identification, on new input data, dramatically speeding up the training process and improving the results. The next step in

Table 1 is to start creating a training dataset. A training dataset contains a subset of data which is fed into a machine learning model where it is processed and where features are extracted. We utilize a supervised learning scheme (Liu and Wu, 2012) to provide DAS data classified as rockfall events and non-events so that the model is able to identify one or the other after training. Training the CNN consists of exposing the model and weights to the training data in random various batches repetitively, during which the model attempts to predict which images appear similar and what features they share. Each batch is then validated by looking at which binary class the data belongs to. Depending whether the model prediction is correct or incorrect, the model weights are adjusted to increase or decrease weights which contributed to the prediction following the idea of back propagation (Chauvin and Rumelhart, 2013). Back propagation utilizes gradient decent and applies a loss function to decrease or increase weights in accordance to their impact on the result and the results of other predictions. After the training is complete, the output is a predictive neural network model trained to identify rockfall events and non-events, based on how similar a new input image appears to one or the other. The final step in the workflow is applying the model and making predictions on data. New data is fed into the neural network without any prior classification, and the model outputs a prediction claiming what it believes the input matches best. This decision incorporates the gained knowledge from training which altered the weights behind the value of certain extracted features. This process can then be applied quickly on each of the three processed DAS datasets provided for this study.

Machine Learning Event Detection Workflow	
1.	Raw DAS Data Input
2.	Data Pre-Processing
3.	Transfer Learning
4.	Create Training Data
5.	Training Neural Network
6.	Predicting on Data

Table 1: Workflow for machine learning microseismic event detection.

Examples

The raw DAS data was acquired on the Mount Meager glacier over a 30-day period during the months of September & October of 2019 (Klaasen et al., 2021). This DAS fibre consists of 380 channels, with certain intervals laying on rocks, ice, and snow. We focus on one day, September 29th 2019, for processing, analysis, and development. The data is initially processed using a 5-55 Hz bandpass filter to remove high frequency noise. Due to certain channel band errors and the variations in the surfaces on which they are located, we narrow our focus into the channel range of 10-140, which is located on a rocky section. Using this channel range and bandpassed data, the entire day of data is divided into 10 second long traces and stored as image frames. The first 83 minutes, or 500 frames, are then extracted and manually classified to identify the presence of events, recording / marking frames with events to enable the development of a training data-set. The results of the manual classification are seen in Table 2, identifying 79 frames with identified events. In addition to this dataset, two other data sets were created in an identical fashion, apart from using different pre-processing / filtering techniques on the raw data. The two additional datasets are referred to as ***fx*** and ***mix*** data.

Non-Events	Events
421	79

Table 2: Number of events & non-events identified in the first 500 frames (83 minutes)

Applying the machine learning model on each of the datasets produces the results represented by pi-charts in Figure 1. The ***mix*** processed model claims to identify the largest number of events, while the ***fx*** processed model identifies the least. Limiting cases in the results due to the different pre-processing methods are identified in Figures 2 & 3. The ***normal*** and ***fx*** processed model and data results seen in Figures 2a & 2b were unable to identify any event due to persisting channel noise, whereas the ***mix*** processed model and data in Figure 2c was successfully able to identify the event. Alternatively, the

normal & ***fx*** trained models were able to identify a weak event seen in Figures 3a & 3b, while the ***mix*** processed data and model was unable to.

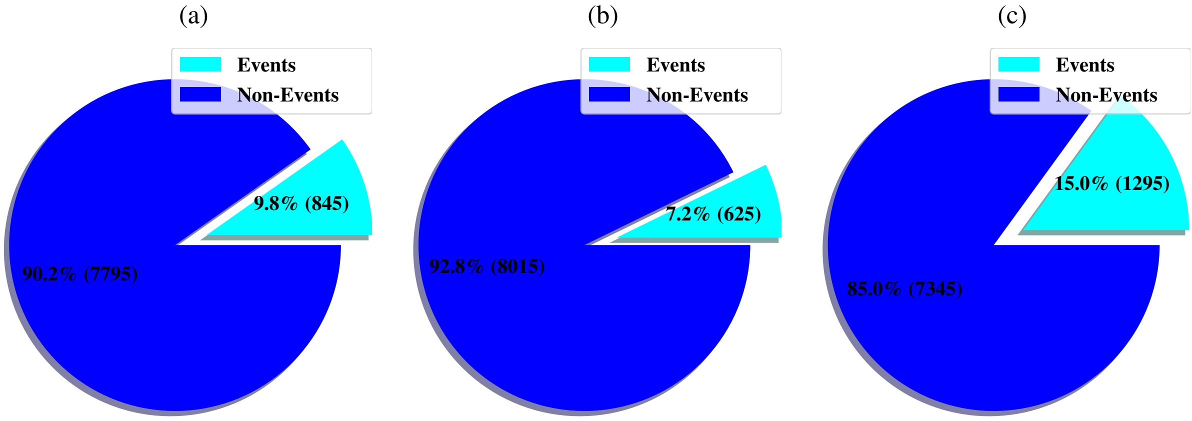

Figure 1: Number of detected events compared to the number of total image frames, a) Events detected with ***normal*** processing, b) Events detected with ***fx*** processing, c) Events detected with ***mix*** processing.

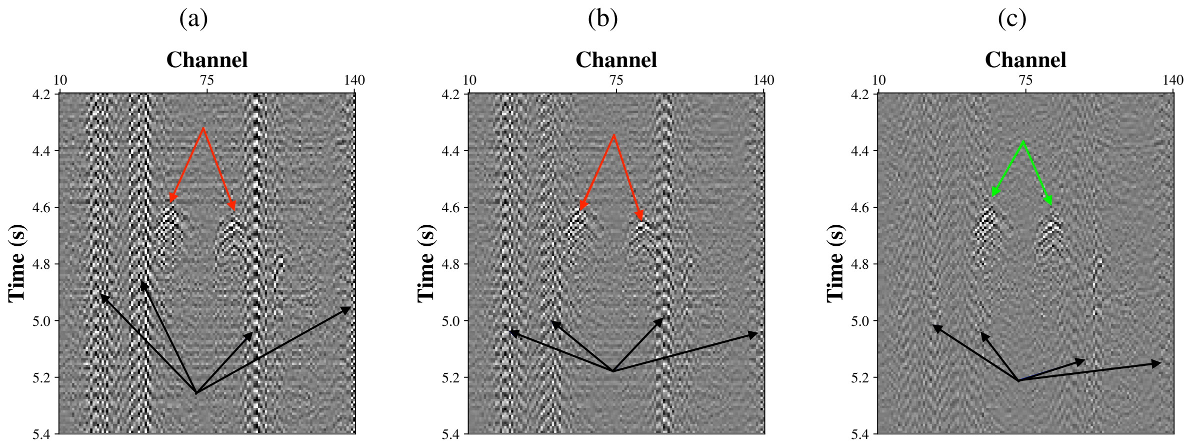

Figure 2: DAS Event only identified by ***mix*** trained data & model. a) Event with ***normal*** processing, black arrow points to channel band noise, red arrow points to undetected event. b) Event with ***fx*** processing, black arrow points to channel band noise, red arrow points to undetected event. c) Event with ***mix*** processing, black arrow points to filtered channel band noise, green arrow points to identified event.

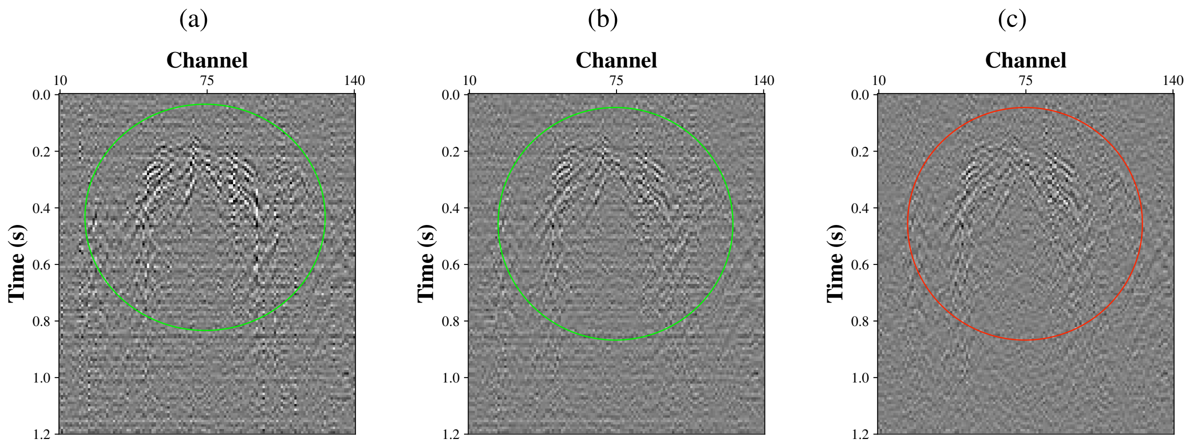

Figure 3: Weak event in the DAS data. a) Event identified with ***normal*** processing, b) Event identified with the ***fx*** processing, c) Event not identified with ***mix*** processing

Results

Building and deploying a machine learning model using CNN's for event image recognition on DAS data, using limited training data, is a highly effective and promising technique. The low quantity of training data allows for a quick adoption of this technique for large scale event detection. We compare applying different pre-processing methods on the dataset used to train and predict results to identify any advantages or improvements.

Considering all of the challenges of working with a small training dataset and overall complex data, these models all perform very well in these early forms. The early success here is extremely valuable since there is significant room for improvement of the models. Future work on these models will consist of increasing the available training data, as well as balancing the number of events and non-events used in it. Increasing training data will improve the model by exposing it to more potential event types. This process of adding events can be done by utilizing a large portion of the currently classified events and then quickly manually classifying them for validation. This means that the early models trained on minimal data can themselves provide additional training data much quicker and be improved. Additionally, we plan to identify the accuracy and improve the statistical understanding of the model in much more detail in order to verify whether there is value in applying the additional pre-processing workflows.

Conclusions

We find that utilization of CNN's and machine learning for image recognition of events in DAS data is an effective and simple approach. Comparisons between this approach and more traditional methods show that the capabilities are the same if not better, even for early underdeveloped models. Implementation of different pre-processing workflows on the data and models appears to have relatively similar performance, but with variation when encountering low amplitude or noise ridden events. Future work will significantly improve the models by using their own predictions to increase training data, and more detailed statistics to measure the accuracy will be implemented.

Acknowledgements

Thanks to Dr. Jan Dettmer, Department of Geoscience, University of Calgary for their assistance and for sharing with us the Mt. Meager data archived at Geoscience BC (Klaasen, 2021).

References

Arief, H.A., Wiktorski, T. and Thomas, P.J. [2021] A Survey on Distributed Fibre Optic Sensor Data Modelling Techniques and Machine Learning Algorithms for Multiphase Fluid Flow Estimation. *Sensors*, **21**(8), 2801.

Chauvin, Y. and Rumelhart, D.E. [2013] Backpropagation. *Psychology Press*.

Heaton, J. [2016] An empirical analysis of feature engineering for Predictive Modeling. *SoutheastCon 2016*.

Klaasen, S. [2021] Data repository for: Distributed Acoustic Sensing in Volcano-Glacial Environments –Mount Meager, British Columbia. This data draws upon research supported by the Government of Canada's New Frontiers in Research Fund (NFRF).

Klaasen, S., Paitz, P., Lindner, N., Dettmer, J. and Fichtner, A. [2021] Distributed acoustic sensing in volcano-glacial environments –Mount Meager, British Columbia. *Journal of Geophysical Research: Solid Earth*, **126**(11).

Krizhevsky, A., Sutskever, I. and Hinton, G.E. [2017] ImageNet classification with deep convolutional Neural Networks. *Communications of the ACM*, **60**(6), 84–90.

Liu, Q. and Wu, Y. [2012] Supervised learning. *Encyclopedia of the Sciences of Learning*, 3243–3245.

Simonyan, K. and Zisserman, A. [2015] Very Deep Convolutional Networks for Large-Scale Image Recognition.