The Impact of Neural Network Complexity on the Detection of Rockfalls on Mars and the Moon

By Henry Zhao

Author Bio

Henry Zhao grew up in Vancouver, and is currently a grade 11 student at Collingwood school. His favorite subjects are physics and computer science, and he plans on studying mechanical engineering and computer science in university. He is the lead coder in his school robotics club, and often assists other robotics teams with programming. In his spare time, he enjoys working on personal engineering projects, 3D modeling and rendering, origami, and drawing digital art. He is a black belt in martial arts, and has been attending since 2012. He is also actively involved with community service, and helps with tutoring for both his classmates and elementary students.

Abstract

The detection of rockfalls is important to real-time tasks on celestial objects, such as finding potentially hazardous or geologically active areas on Mars. Currently, rockfall processing is done by hand, requiring a large amount of human labor. This study intends to test the effectiveness of machine learning on the detection of rockfalls on Mars and the Moon. Specifically, this study tests the impact of neural network complexity, and thus detection time, on the precision, recall, and F1 score of detection results. 5 different convolutional neural networks were trained and tested on a dataset of Mars and moon rockfalls. It was found that Resnet50 with transfer learning performed the best, with a median detection time and recall, and high precision and F1. There also is a rough pattern corresponding longer detection times with more precision and less recall, and shorter detection time with more recall and less precision. From the data gathered, it appears that it is feasible to use machine learning to quickly and effectively detect rockfalls on Mars, for use in real-time processing and decision making.

Keywords: Machine learning, binary classification, convolutional neural networks, model complexity, rockfall detection, recall, precision, F1 score, training time, classification time

Introduction

Rockfalls are common mass-wasting processes on rocky planets and satellites. They are typically driven by geological and celestial processes, including quakes and meteoritic impacts. By finding the distribution of rockfalls on a planet or satellite, the speed of erosion and geological activity in certain areas can be determined. As well, since erosion acts on the tracks of rockfalls, the density of rockfalls can reveal information about the age and geologic activity of said areas.

Currently, the priority of rockfall studies are on Mars. Such studies could reveal the geologic status of areas on Mars, and provide risk information in the possibility that Mars becomes the next human-occupied planet in our system. The current methods for detecting and studying rockfalls requires human labor to scan through satellite images, marking out all rockfalls.

As of currently, the High Resolution Imaging Science Experiment (HiRISE) imaging system on the Mars Reconnaissance Orbiter has captured only approximately 4% of the Mars surface (Slade, 2021). However, these images have a spatial resolution of up to 0.3m/pixel, meaning the amount and scale of the images gathered are quite significant. Due to this, it takes a significant amount of human labor and time to detect all rockfalls in even a single HiRISE image. It has been previously found by Valentin Tertius Bickel that machine learning, or more specifically, Convolutional Neural Networks (CNNs) were effective in the detection of rockfalls on Mars and the moon. The results showed that CNNs were faster at detecting rockfalls than humans by an order of magnitude, while maintaining an acceptable tradeoff between recall and precision (Bickel et. al, 2020). By training a neural network to identify rockfalls, the processing time and human labor can be greatly reduced, while still resulting in similar accuracy. However, this time reduction is not quantitatively reported, and it is unknown whether it is adequate for real-time processing, which will be important for real-time rover guidance or real-time geologic analysis. Thus, it is necessary to further explore the possibility of optimizing current algorithms, to analyze the potential for processing in near-real-time situations.

The objective of this study is to determine the relationship between the recall or precision of a model compared to the training and detection time. This may eventually lead to a neural network which is capable of near-real-time processing, which would greatly benefit the efficiency and productivity of rovers and similar systems.

Method

A labeled image dataset for deep learning-driven rock fall detection was selected for this study (Bickel 2021). The dataset contains the images with bounding boxes of the rockfalls since it was originally made to train detection systems. As this study is attempting to train a classification network, a new dataset was created by cropping and extracting smaller tiles either containing or not containing rockfalls. Tiles containing rockfalls (Figure 1) were cropped based on the provided bounding boxes, and empty tiles (Figure 2) were extracted from parts of the image not intersected by any boxes.

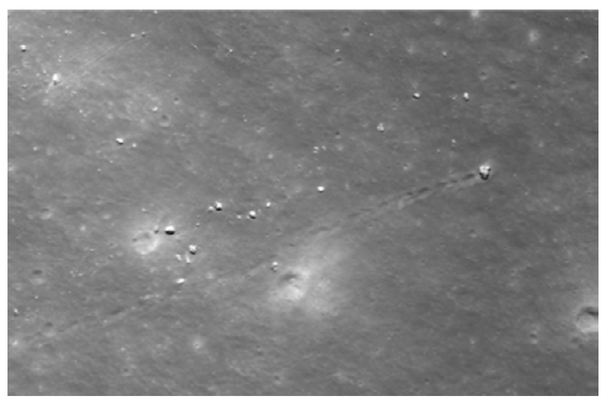

Figure 1: Image tile containing rockfall

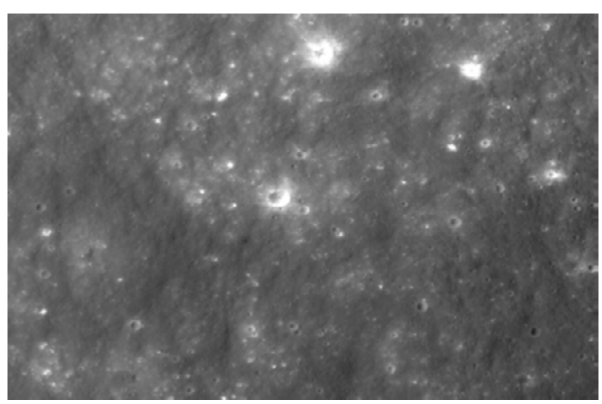

Figure 2: Image tile without any rockfalls

The new dataset was split into 80% training and 20% validation images. A large variety of off-the-shelf classification networks were chosen to be tested in this study. These include 2 unnamed CNNs, and Mobilenet, Resnet, and VGG16 transfer learning models. Each was trained for 100 epochs, with 30 steps, at a learning rate of 5e-5 using the Root Mean Squared Propagation (RMSprop) optimizer. The loss, accuracy, recall, precision, and F1 score, as well as training time per epoch and average classification time for 32 images of each model was recorded in order to make comparisons.

Results

The Resnet50 transfer learning model was found to perform the best, with a median recall, the highest precision, and the highest F1 score. The Mobilenet transfer learning model performed the fastest in both training time per epoch and detection time (Table 1 and Figure 3).

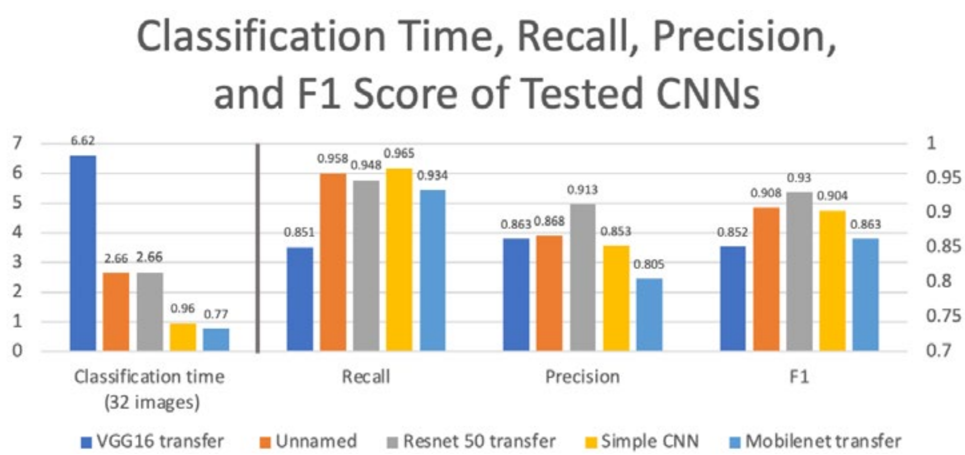

Figure 3: Chart of the classification time, recall, precision, and F1 score of the tested convolutional Neural Networks

Table 1: Performance of models in detecting rockfalls on Mars and the Moon

Model	Layers	Training time per epoch	Classification time (32 images)	Recall	Precision	F1
VGG16 transfer	18	7s	6.62s	0.851	0.863	0.852
Unnamed	50	38s	2.66s	0.958	0.868	0.908
Resnet 50 transfer	176	11s	2.66s	0.948	0.913	0.930
Simple CNN	16	11s	0.96s	0.965	0.853	0.904
Mobilenet transfer	85	6s	0.77s	0.934	0.805	0.863

Discussion

It appears that as classification time decreases, the recall approximately increases and the precision approximately decreases, indicating more correct predictions, but less rockfalls caught. This result matches expectations, since a more complex model should be able to classify more correctly. In comparison, a less complex model should be able to catch more, whether or not a rockfall is actually present. There also does not seem to be a direct relationship between the complexity, or amount of layers a model has, and the classification time. For example, VGG16 consists of 18 layers, and takes approximately 7 seconds to classify 32 images, while resnet50, with 176 layers, only takes approximately 3 seconds. This is likely due to the layer makeup of these networks, as different types of layers compute at differing speeds.

It was unexpected that the transfer learning models, or more specifically, VGG16 and Mobilenet, performed worse than the two unnamed models in both precision and F1. It was expected that the transfer learning models would perform worse in recall, as they should be less sensitive to noise, and thus would be expected to provide more false negatives in return for more correct positives. However, they also lack precision, meaning that they neither classify more, nor do they classify more correctly. There are a few possibilities as to why this could be. Firstly, these models were trained on color images, while the dataset used is in grayscale. Secondly, these models were trained for 1000-class classification, while in this study, they were just used for binary classification. Finally, it could be possible that the dataset these models were initially trained on, and the features they detected and kept, were unfit for detecting rockfalls.

Bickel et al. (2020) have performed a similar study with bounding box object detection networks on the same dataset. They were able to achieve a variety of tradeoffs between recall and precision, in the best case a recall of 0.6 to 0.07 and a precision of 0.39 to 0.95, and a maximum F1 score of 0.48. While comparisons can be made on various metrics, the two methods are fundamentally different, and should not be directly compared.

In conclusion, this study indicates that it is possible, though difficult, to achieve near-real-time processing of rockfalls in martian and lunar images. The HiRISE system collects images at approximately 213 megapixels per second (McEwen et al., 2007), while the fastest method tested could detect rockfalls at approximately 40 megapixels per second on the testing machine. As long as the system used to run these neural networks has approximately 5 times the processing power of the testing machine, super-real-time processing can be obtained. As well, processing speed can be increased by downscaling images. Depending on the requirements of a specific system, different networks or combinations of networks may be used. For example, a faster network with high recall and low precision may generate many potential areas to avoid for rover exploration, or generate safety and risk data for potential habitation candidates, as false positives would be better than false negatives. The fast speed would also be beneficial for full-planet analysis. A lower recall but higher precision network can be combined with a high recall network to improve speed, and would be able to be used in scientific research or targets for rovers, as false negatives would be more desirable than false positives.

Despite the significant contributions offered by the current study, there are some limitations which need to be addressed. The method employed in this study requires the large image to be cut into smaller sized blocks, which is required for these models to train on in a reasonable timeframe. Future studies with more computing resources may be able to train and test models on larger image patches. As well, further studies with more resources may be able to train the transfer learning models without transfer learning, thus computing features from scratch. These can then be compared against the current models, to investigate the effects of transfer learning on recall and precision. Finally, further studies may be able to train and test models that operate on bounding box networks, in order to compare the benefits and drawbacks of each method.

References

1. Bickel, V. T., Mandrake, L., & Doran, G. (2021). A Labeled Image Dataset for Deep Learning-Driven Rockfall Detection on the Moon and Mars, *Frontiers in Remote Sensing*, 2, https://doi.org/10.3389/frsen.2021.640034

2. Bickel, V. T., Conway, S. J., Tesson, P. -A., Manconi, A., Loew, S., & Mall, U. (2020). Deep Learning-Driven Detection and Mapping of Rockfalls on Mars. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 2831-2841, https://doi.org/10.1109/JSTARS.2020.2991588.

3. Slade, S. (2021, March 22). *Stunning images of Mars keep coming 15 years after HiRISE camera first orbited the Red Planet*. The Washington Post. Retrieved January 31, 2022, from https://www.washingtonpost.com/lifestyle/kidspost/stunning-images-of-mars-keep-coming-15-years-since-hirise-camera-first-orbited-the-red-planet/2021/03/16/17af5e30-7c96-11eb-a976-c028a4215c78_story.html

4. McEwen, A. S., et al. (2007), Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE), J. Geophys. Res., 112, https://doi.org/10.1029/2005JE002605.