Deepening the Sense of Touch in Planetary Exploration with Geometric and Topological Deep Learning

Yuzhou Chen,[1,4] **Yuliya Marchetti,**[2] **Yulia R. Gel**[3, 4]

[1]Department of Statistical Science, Southern Methodist University, Dallas, TX 75275, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[3]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75080, USA
[4]Energy Storage and Distributed Resources Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
yuzhouc@smu.edu, yuliya.marchetti@jpl.nasa.gov, ygl@utdallas.edu

Abstract

Tactile and embedded sensing is a new concept that has recently appeared in the context of rovers and planetary exploration missions. Various sensors such as those measuring pressure and integrated directly on wheels have the potential to add a "sense of touch" to exploratory vehicles. We investigate the utility of deep learning (DL), from conventional Convolutional Neural Networks (CNN) to emerging geometric and topological DL, to terrain classification for planetary exploration based on a novel dataset from an experimental tactile wheel concept. The dataset includes 2D conductivity images from a pressure sensor array, which is wrapped around a rover wheel and is able to read pressure signatures of the ground beneath the wheel. Neither newer nor traditional DL tools have been previously applied to tactile sensing data. We discuss insights into advantages and limitations of these methods for the analysis of non-traditional pressure images and their potential use in planetary surface science.

Introduction

Planetary rovers, like the Mars Exploration Rovers (MER) and the Mars Science Laboratory (MSL) rover, have enabled exploration and collection of invaluable science data across the surface of Mars. Possible future manned and unmanned Lunar rover missions, e.g. Volatiles Investigating Polar Exploration Rover (VIPER), are planned for quick investigation and mapping of the Moon surface, while the Fetch Rover on Mars sample-return mission (MSR) will deliver a sample back to Earth. These planetary rovers are typically equipped with a variety of science instruments, manipulation equipment and traditional imagers. In their operations, they rely heavily on visual images and telemetry, such as motor current or wheel rim thrust. These data are used to assess the health and safety of the vehicle and determine its activities in advance by operations teams, e.g. route planning, hazard avoidance and targeting of onboard science instruments. Currently, such systems lack embedded instruments that have the ability to assess the interaction between the terrain and the rover mobility system in real time.

A tactile wheel is an novel experimental concept for planetary and other exploration rovers that combines a payload of in-situ sensors directly on the wheel, which enables direct measurements of several key contact area parameters and surface properties of potential scientific interest. This new sensing modality has promise to become a primary planetary mission driver, especially for supporting exploration and mapping efforts at cruising rather than crawling speeds for the future rover missions to Mars and Moon and terrestrial mobility applications for ground vehicle support in remote harsh environments. Knowing the current slip and sink rates of each wheel could be used to halt a drive, while knowing about rock composition and size below the wheel could result in a better power distribution to the wheels for avoiding damage and wear. This information can be highly beneficial for scientific endeavors as well. Monitoring terrain change can trigger remote sensing from onboard science instruments, when criteria of interest have been observed. In-situ wheel instruments enable higher science yield, while enabling safer vehicle operations. Combined with the visual odometry, such tactile sensing will add "touch" to the "eyes" of a rover.

A new concept developed in Marchetti et al. (2020) not only incorporates in-situ sensors on a robotic wheel, but also uses machine learning to extract meaningful metrics from the interaction between the wheel and the terrain using the sensors. These include continuous slip estimation, balance, and sharpness for engineering purposes and estimates of hydration, texture, and terrain patterns for science applications. This tactile wheel carries an electrochemical impedance spectrometer (EIS), for hydration detection, and is wrapped in a 2D grid of pressure sensors, which give continual context for the pressure signature of the ground underneath the wheel (see Fig. 1). The pressure sensor supplies a conductivity "image" of the surface/wheel interface. Marchetti et al. (2020) conduct numerous experiments and collect data from the pressure and the EIS sensors, and consequently train classical machine learning models, such as tree-based Stochastic Gradient Boosting (SGB) and random forest (RF), with the derived features to infer engineering conditions and predict various terrain properties.

New and One of a Kind Data for Future Planetary Terrain Exploration. The resulting dataset of raw conductivity images from the pressure sensor mounted on a rolling wheel is a first and unique collection of a new data type that has a

Copyright © 2021, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Figure 1: Tactile wheel mounted on a rover (left) and with a data-taking test cart (right). The pressure sensor array wraps around the wheel and is covered with a yellow protective layer of Kevlar.

likeness to natural images. However, unlike with traditional imaging cameras, each element, called a *taxel*, produces a measurement of pressure rather than light. These images are not in RGB or any other color space, but in Digital Number (DN) units that can be converted to pascals. Moreover, the conductivity images need to be calibrated and further processed so that meaningful information is obtained for successful use with machine learning models.

While utility of DL has been proven in a broad range of application domains, from biomedical imaging to robotics, DL tools have not yet been applied and tested for use with onboard autonomy and in-situ sensors, especially in the context of planetary exploration missions. We take a first step to bringing the power of DL in this context and apply several of these DL tools to infer local terrain characteristics, specifically, for the identification of rock, using a novel dataset of pressure sensor images from the tactile wheel prototype. Terrain typically exhibits a distinct geometric structure, not only on a scale as seen from an airborne instrument but also from under a wheel. Rock will have a distinguishable imprint compared to a sandy dune, for example. In mathematical terms, terrain can intrinsically be represented as a manifold, and thus, it appears natural to also explore utility of DL tools that specifically address analysis of objects that are characterized by geometric structure, including those with non-Euclidean geometry. Hence, the methods of geometric deep learning (GDL) (Bronstein et al. 2017; Monti 2020) such as graph convolutional networks (GCN) (Wu et al. 2020) appear to be especially suited for terrain and surface property analysis. Finally, a number of recent studies (see, e.g. Hofer, Kwitt, and Niethammer 2019; Potts et al. 2019; Gabrielsson et al. 2020) indicate that integration of topological descriptors, i.e. systematic shape analysis, into DL can lead to noticeable gains both in classification and stability performance. This phenomenon can be explained by the important complementary information to many state-of-the-art conventional non-topological descriptors and deeper insight into the intrinsic data shape that topological data summaries are shown to deliver (Pun, Xia, and Lee 2019). Motivated by these results, we also develop and validate a new topological approach to GDL, topological GCN (TOPO-GCN), based on integration of persistent images into the GCN architecture, which could allow us to better account for local and global terrain structure.

While applying the DL machinery to the tactile wheel data, we gained several important insights that will make an impact on utility of any such classification methods if they are implemented with embedded systems like the tactile wheel. Lessons learned are:

- DL tools, in general, and geometric and topological DL, in particular, offer a more powerful alternative to the currently employed ML methods, such as SGB, in application to terrain learning and prediction. However, classification gains (if any) delivered by the DL models vary substantially among models and across surface types.
- Topology and geometry of the the tactile wheel data tend to play an important role in terrain classification. GDL models such as TOPO-GCN appear to be the most competitive approaches for more sophisticated terrain types.
- Not all DL methods integrating topological information of terrain are found to be a feasible choice due to image size limitations.
- Not all terrain types benefit from DL, simpler terrains such as flat surfaces can be successfully classified by simpler ML tools.
- Considerable computational costs, and, as a result, high energy consumption, are a hurdle and might make application of DL challenging for onboard near real-time tasks.

Related Work

Terrain classification for embedded systems. The classification of terrain type is traditionally performed using accelerometers and vibrational analysis, identifying frequency distributions for each of the types in question. Such 1D datasets can infer certain terrain types in fast-moving vehicles (e.g. self-driving cars), but for the cautious speeds typical of explorative robots the vibrational input is weak and unreliable. Further, vibrational analysis cannot alone provide detailed science estimates of surface properties like the presence/thickness of crusts, and rough surface morphology. Past attempts (Arvidson et al. 2017; Senatore et al. 2014) have been made to derive terrain mechanical properties based on such rover telemetry and visual data and applied some ML-based proof-of-principle systems (Rothrock et al. 2016). However, all of these efforts are characterized by inferred interactions that are only available through post-hoc ground analysis. A limited set of embedded pressure sensors have been used to asses wheel contact and load distribution (Shirai and Ishigami 2015; Wu et al. 2016), and estimate drawbar pull (Nagatani et al. 2009), however these do not provide the full contact geometry. Most recently, Marchetti et al. (2020) applied SGB to a set of computed features based on the pressure images and other extractions from the tactile wheel. However, DL has not been attempted with such data. DL methods have previously been applied in planetary terrain classification (Wagstaff et al. 2018; Rothrock et al. 2016; Kerner et al. 2019), but with natural images, taken by rover and orbiter cameras, in the context of traditional image classification tasks. This dataset and pressure "images", however, are unique and their application is for onboard deployment and autonomous driving.

(a) Dense rocks (b) Sparse rocks (c) Small rocks

Figure 2: Select examples of rock experiments. Distribution of rocks and their sizes vary along with the type of material. Composition in the middle plot is much coarser grain size than that on the right and left, which are more sandy.

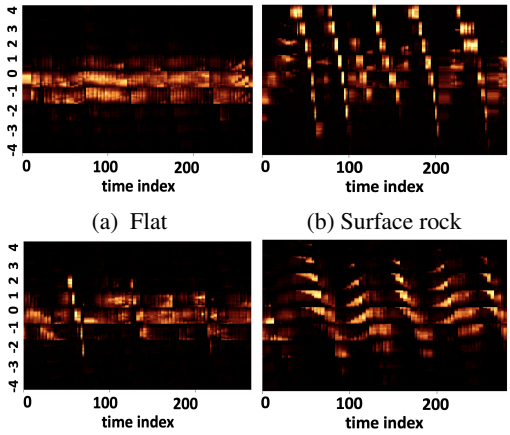
Data Description

Tactile wheel and data collection. The structure of the tactile wheel is comprised of the base cylinder that has several sensors mounted on the inside and that is wrapped circumferentially with an inductive pressure sensor array as shown previously in Fig. 1, that measures the contact pressure between the wheel and the media underneath. The pressure pad consists of 1,920 distinct 16mm square pressure nodes or taxels, arranged in an 20×96-array and measured in digital number (DN) units. Steel grousers, i.e. flat bars intended to increase the traction with the ground, are placed over every other row of the pressure sensor array. These grousers can alter the wheel-terrain interaction and can leave a unique informative pressure signature on the images.

To collect the data, the tactile wheel is mounted on a test cart that is used to power and control the wheel. The cart is then operated in large metal troughs over seven types of materials, with grains of varying sizes and on various configurations of rocks and terrain patterns. Examples of experiments with rocks are shown in Fig. 2. As the wheel is slowly moved along, the measurements are taken by the pressure sensor and other instruments embedded on the wheel and the cart.

Pressure sensor images. Pressure sensor images carry information about the state of the sensor and what the wheel is touching on the ground at each time point. One of the mounted sensors inside the tactile wheel is inertial measurement unit (IMU) that measures orientation of the wheel. At any given point in time, IMU reads the "gravity down" position of the wheel, which should fall within the contact area as measured by the pressure sensor. We use the IMU sensor to stack the columns of the pressure sensor image that correspond to the gravity down column and columns adjacent to it. The resulting "unwrap" image represents the pressure sensor in the space and time dimension. The unwrap images capture the pressure signature as the wheel moves forward in time, with the middle row representing the position of the IMU's gravity column down, the row above it representing the front of the wheel and the row below showing the back of the wheel (see Fig. 3 for examples).

Rock detection. Marchetti et al. (2020) collected 163 experiments across different materials and distributions, resulting in over 90,000 images, each labeled with a corresponding binary class, indicating the presence of rock. We use binary classification with these images as the tactile dataset does not have rocks explicitly labeled in various categories. Rock classification, however, can be extended to multiple types, e.g. small or large, sharp or flat, surface or buried. Several types of rock could also be present in the same pressure image.



(a) Flat (b) Surface rock

(c) Below-surface rock (d) Sharp landform

Figure 3: Pressure "unwrap" images for the full experimental runs. The x-axis represents time, and the y-axis represents a row on the pressure sensor, with 0 indicating the row touching the ground according to IMU.

We focus on rock identification as it greatly impacts both mobility and safety of a vehicle as well as its science yield. The monitoring of rock can reduce damage risk to the wheels and make energy consumption of the vehicle more efficient, resulting in longer traverses and more science data collected. The identification of below-the-surface rock is even more crucial since it cannot be performed from visual odometry. Hidden rocks not only present potential hazard, but can also indicate that a vehicle is on firm ground rather than on loose sand. Importantly, below-surface rock has not yet been exposed to the atmosphere or radiation and can be sampled for science analysis.

Fig. 3 shows images of flat, rock and dune patterns under the pressure sensor. Flat terrain in Fig. 3a has mostly constant pressure signature, with slight fluctuations due to the sandy nature of the material, centered on the gravity vector in the middle of the image. In contrast, rock signature in Fig. 3b has a characteristic pattern that creates diagonal lines across the image. The hidden rock in Fig. 3c leaves a similar pattern, however, it is much more subtle with increasing depth. Finally, Fig. 3d displays a pressure imprint that is representative of a dune instead of a rock. These images show that distinct pressure signatures are characterized by topological features for different terrain types.

Topological Layer for Geometric Deep Learning

Graph-based semi-supervised learning on images. As a part of GDL, graph-based semi-supervised learning (G-SSL) is one of the most popular approaches for supervised learning on images, allowing to capture important local,

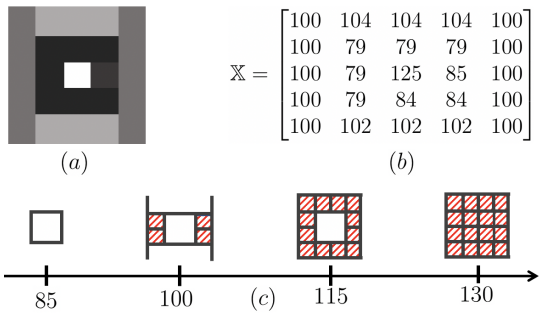


Figure 4: PH for a grayscale digital image. (a) a grayscale image; (b) pixel values of the grayscale image; (c) a filtration of cubical complexes induced by the grayscale image.

stationary, and compositional task-specific features (Subramanya and Talukdar 2014; Monti et al. 2017; Ma et al. 2019; Jiang et al. 2019). The G-SSL optimization objective is a convex classification function $Z = (1-\omega)(I - \omega D^{-\sigma} W D^{\sigma-1})^{-1} Y$, where W_{ij} is a similarity matrix of image pixels (in our case obtained using radial basis functions on 2D pixel coordinates), $D_{ii} = \sum_{j=1}^{N} W_{ij}$ is a diagonal matrix, and ω is a regularization parameter. Let K be a number of classes, then define a $N \times K$-label matrix Y such that $Y_{ij} = 1$ if vectorized image $X_i \in \mathbb{R}^{N^2 \times 1}$ is labeled as class j, and 0 otherwise. We then incorporate G-SSL into GCN as follows. First, we consider Generalized Sigma-based (GS) filter $\mathbf{f}_{GS}(\omega, \sigma) = (1-\omega)(I - \omega L_\sigma)^{-1}$, where $L_\sigma = D^{-\sigma} W D^{\sigma-1}$, $0 < \omega, \sigma \leq 1$. Second, we use Taylor approximation of $\mathbf{f}_{GS}(\omega, \sigma)$ with the order of $\lceil 4\omega \rceil$, yielding the best performance on the training dataset. We now obtain the general classification function $\mathcal{X} = \mathbf{f}_{GS}(\omega, \sigma) X = (1-\omega)(X^{'})_i$, where $(X^{'})_i = X + \omega L_\sigma (X^{'})_{i-1}$, $(X^{'})_0 = X$, $i \in \mathbb{Z}_{i>0}$. Finally, we provide an implementation function of GS filter as a convolutional layer

$$H^{(t+1)} = \phi\left(\mathbf{f}_{GS}(\omega, \sigma) H^{(t)} \Theta^{(t)}\right),$$

where $H^{(t+1)}$ is the hidden layer output, $H^{(0)} = X$, $\phi(\cdot)$ is the adopted activation function, and $\Theta^{(t)}$ is the trainable weight in the t-th layer. This G-SSL GCN architecture allows to take any image data as an input. To better account for highly sophisticated terrain patterns of planetary surfaces in a robust to uncertainties manner, we also propose to integrate G-SSL GCN with a topological layer.

Topological Layer to GDL. Pressure sensor images from the tactile wheel prototype demonstrate a complex local terrain structure. To better explore such local distinct properties manifested both from an airborne instrument and from under a wheel, we also complement the GCN model with a new topological layer. The topological layer allows us to integrate information on the terrain shape structure, that is, terrain features which are invariant under continuous transformations such as bending, stretching, and compressing. We collect information on the terrain shape using tools of persistent homology, i.e., by counting certain patterns such as loops, holes, and cavities, within terrain, evaluated through multiple user-defined similarity scales. To systematic define what patterns to include and how to count them efficiently, we equip the pressure sensor images with a combinatorial object, namely a cubical complex, and use grayscale thresholds as varying similarity thresholds. Formally, a cubical complex $\mathcal{K}$ in $\mathbb{R}^n$ is a union of elementary cubes, where each elementary cube is a finite product of elementary intervals $\mathcal{I}_1 \times \mathcal{I}_2 \times \mathcal{I}_n \subseteq \mathbb{R}^n$ and an elementary interval is a subset $\mathcal{I} \in \mathcal{R}$ such that either $\mathcal{I} = [l, l+1]$ or $\mathcal{I} = [l, l]$, $l \in \mathbb{Z}_{>0}$. Now let f be a real-valued function (here the grayscale value) that maps every simplex to the maximum function value of its vertices. Let $\mathcal{K}_r = f^{-1}(-\infty, r]$, $r \in \mathbb{R}$ be a cubical complex formed by pixels of grayscale value of not higher than r. Varying thresholds r results in a nested sequence of cubical complexes $\mathcal{K}_{r_1} \subset \mathcal{K}_{r_2} \subset \ldots \mathcal{K}_{r_m}$, called a *lower-star filtration* of cubical complex (Edelsbrunner and Morozov 2012). We then track patterns (i.e., loops, cavities etc) appearing and disappearing in this lower-star filtration. To input the extracted topological terrain information into GCN, we use its finite-dimensional vector representation as persistence image (PI) which is derived by the weighted kernel density estimation (Adams et al. 2017). PI is a finite-dimensional vector representation of a persistence diagram (PD) and can be computed through the following steps:

- *Step 1:* we map the PD to an integrable function $\rho : \mathbb{R}^2 \to \mathbb{R}^2$, which is called a persistence surface. The persistence surface ρ is given by sums of weighted Gaussian functions that are centered at each point in PD.
- *Step 2:* we take a discretization and linearisation of a subdomain of persistence surface in a grid.
- *Step 3:* the PI, i.e., a matrix of pixel values, can be obtained by subsequent integration over each grid box.

Formally, the value of each pixel z within a PI is defined as

$$PI(z) = \iint_z \sum_{[b_i, e_i] \in I} g\left(b_i, e_i\right) \times \frac{1}{2\pi\sigma_x\sigma_y} exp - \frac{1}{2}\left(\frac{(x-b_i)}{\sigma_x^2} + \frac{(y-e_i)}{\sigma_y^2}\right) dydx,$$

where $g(b_i, e_i)$ is a weighting function, which depends on the distance from the diagonal, σ_x and σ_y are the standard deviations of the Gaussians in x and y direction. For the G-SSL GCN, given persistence images PI_X, the corresponding initial topological layer is given by $H^{(1)} = \phi\left(\mathbf{f}_{GS}(\omega, \sigma) \mathrm{PI}_X \Theta^{(0)}\right)$. We emphasize that the G-SSL GCN architecture can be integrated with any image type, i.e., with raw images, PIs or both. When applied to both image types simultaneously, i.e., raw pressure sensor images and their PI counterparts, we call the approach TOPO-GCN. Figure 5 depicts the flowchart for the proposed TOPO-GCN architecture: the TOPO-GCN input contains original rock images and corresponding PIs generated by the lower-star filtration of cubical complex; the upper part shows that we pass the PIs to the G-SSL GCN architecture which has multiple convolutional layers with the GS filter; the lower part shows the deep neural network (DNN) architecture which takes the original rock images as an input; we then aggregate outputs of G-SSL GCN and DNN. To classify terrain, the aggregated features are then processed via a CNN framework.

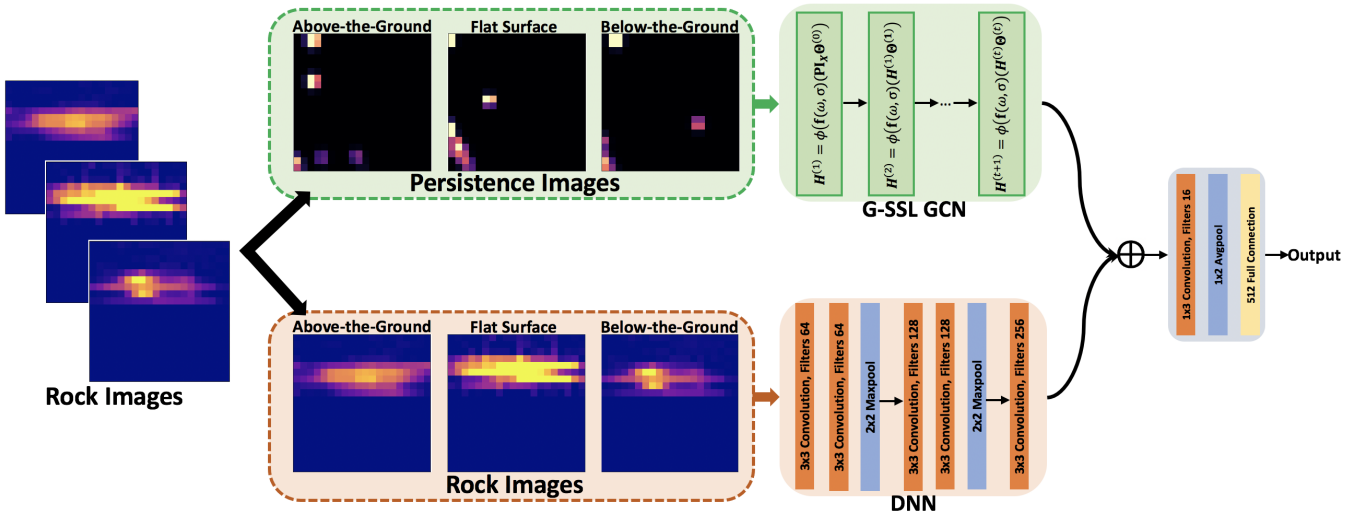


Figure 5: TOPO-GCN architecture for the rock image classification model. The upper part is the model architecture using proposed G-SSL GCN for PIs and the lower part is the model architecture using DNN for rock images.

Results using Pressure Grid "Unwrap" Images

We now discuss the results for (i) rock binary classification, and (ii) rock depth analysis with various DL methods. The dataset is comprised of 163 rock experiments across different materials, amounting to 92,177 data points, i.e. pressure thumbnails obtained from unwrap images by stacking. For each data point, a binary label is assigned, i.e., $y_t \in \{0, 1\}$, and the corresponding rock depth is measured. To estimate the test error, we use group K-fold cross-validation and set the number of folds to 10. As performance metric for rock binary classification, we use misclassification error, based on evaluating the differences between true (*ground truth*) and predicted label vectors, i.e., $\text{ME} = 1/T \sum_t I\left(y_t \neq \hat{y}_t\right)$, where T is the number of data points in the test set.

We compare TOPO-GCN to several DL benchmarks, including GCN with convolutional ARMA filters (ARMA-GCN) of (Bianchi et al. 2019), VGGNet (VGG-16) of (Simonyan and Zisserman 2015), and DeepNet with topological signatures (PD-CNN) of (Hofer, Kwitt, and Niethammer 2019), as well as with the currently employed SGB model (see discussion in Marchetti et al. 2020) with 300 hand-engineered features. For all considered models, we report misclassification errors (see Table 1) and relative gains in respect to SGB.

Performance comparison of rock binary classification. As Table 1 and Fig. 6 indicate, TOPO-GCN delivers most competitive overall accuracy over all considered models and is followed by another GDL method, i.e. ARMA-GCN. Furthermore, the correct classification of rock is a more important task as well as substantially more challenging than classification of flat surfaces. Hence, we also investigate performance of all models under the following three scenarios, i.e., (i) flat surface, (ii) above-the-surface rock, and (iii) below-the-surface rock. The below-the-surface rock scenario (i.e., the corresponding to rock depth higher than 0) is the most challenging classification task.

Method	Overall	Flat surface	Above rock	Below rock
TOPO-GCN	**97.1 (0.07)**	97.7 (0.09)	**91.8 (0.04)**	**31.0 (0.01)**
ARMA-GCN	96.0 (1.07)	98.4 (1.20)	88.4 (0.02)	29.0 (0.03)
VGG-16	94.0 (0.10)	96.8 (0.15)	84.4 (0.05)	28.5 (0.02)
PD-CNN	80.0 (0.09)	98.0 (0.10)	28.9 (0.01)	12.0 (0.07)
SGB	95.3 (0.01)	**99.3 (0.01)**	84.7 (0.03)	6.7 (0.04)

Table 1: Average accuracy (in %) of image classification approaches for rock images; () is standard deviations.

Overall, GDL models such as TOPO-GCN tend to be the most competitive approaches. For example, for above-the-surface classification (see Fig. 6c), TOPO-GCN yields relative improvement of 7.4% and relative gain of TOPO-GCN for the below-the-surface classification is 77.4% (see Fig. 6d). In turn, conventional CNN (i.e., VGG-16) yields no improvement in classification of the above-the-surface objects and delivers a similar performance as ARMA-GCN for the below-the-ground classification. Finally, topological CNN (PD-CNN) of Hofer, Kwitt, and Niethammer (2019) tends to be outperformed by all other models, including SGB, in all but the below-the-surface classification where PD-CNN results in 41.7% of relative gain. All classification models perform similarly for flat surface classification.

The more competitive performance of GDL (i.e., TOPO-GCN and ARMA-GCN) could be attributed to the fact that both GDL better accounts for both local and global structure of the terrain data, while TOPO-GCN has an edge due to coupling the graph-convolutional approach with topological and geometric layers. That is, TOPO-GCN integrates the complementary information on terrain persistent shapes (i.e., analysis of persistent homology on images).

Rock depth analysis. We further separate the misclassified instances from each model into 3 groups of various rock depth, S_1 (rock depth $\in [0, 1)$), S_2 (rock depth $\in [1, 2)$), and S_3 (rock depth $\in [2, 6)$), as shown in Fig. 7. We find that

TOPO-GCN tends to outperform other classifiers when the rock depth is within the range of $[0, 2)$, which makes sense since the pressure sensor cannot detect the rock well when it is positioned deeper below the ground. In turn, the fewer misclassified cases in the $[2, 6)$ category could indicate that a model picks up spurious signals or overfitting.

The hyperparameter optimization for the TOPO-GCN model is performed by grid search. We use the *Adam* optimizer with a learning rate of 1×10^{-3} and the fully connected layer is followed by a dropout layer with a dropout rate of 0.2 and a batch-normalization layer. Parameter setting of the TOPO-GCN model will be described as follows:

G-SSL GCN. Regularization weight $L_2 = 5 \times 10^{-4}$. Regularization parameters in GS convolutional layer $\omega = 0.1$ and $\sigma = 0.1$. Two GS convolutional layers, where the number of hidden units are 128 and 5 in the first and second GS convolutional layers respectively.

DNN. The number of hidden units in three convolutional layers are 64, 128, 256. The convolution kernel size is same for all layers and is set to 3 with stride set to 1. We set the pool size to 2 and stride to 2 in Max-Pooling layers.

Finally, we record and report results for the average training time per epoch for all models used in the experiment. The quantitative comparison in terms of computational cost on GPU p3.8xlarge (64 GB GPU memory) is following: (i) TOPO-GCN - 21.0s, (ii) ARMA - 12.0s, (iii) VGG-16 - 27.0s, and (iv) PD-CNN - 14.0s.

Lessons Learned

The tactile wheel is a new experimental concept that has been tested in only a controlled environment and is not hardened for space or for rigorous use in harsh conditions. However, it has a great potential to be deployed on rovers in the near and far future, especially as more exploration missions are introduced, such as those to Europa and other icy ocean worlds. Our data-taking experiments and the development of a new TOPO-GCN model, has presented us with several lessons that will be considered for our future work as well as for other onboard tactile sensing frameworks:

Topological and geometric properties of the tactile data. Geometric DL models (e.g., TOPO-GCN), which are designed to account for non-Euclidean objects such as manifolds and graphs, tend to deliver the highest accuracy, especially for more sophisticated terrains such as the above-the-ground and below-the-ground rock scenarios. In general, the more complicated the terrain structure, the more classification performance benefits from integrating topological and geometric layers within the DL model, such as TOPO-GCN. There appears to be no gain in using any DL model for simpler flat surface terrain, and more conventional ML tools are preferred. Hence, the choice of classification tools for terrain learning and prediction might incorporate prior knowledge on the expected terrain type, the desired accuracy, acceptable uncertainty levels, and the ability of the model to capture either local or global terrain properties.

Complexity of the DL model and image sizes. As is well known in statistical sciences and ML, model performance, model complexity and the amount of available data are closely interrelated. Our results indicate that not all DL methods integrating potentially valuable topological information on images may be an appropriate classification choice for a given size of observed images. For instance, topological CNN (PD-CNN) of Hofer, Kwitt, and Niethammer (2019) tends to yield suboptimal performance, which can be attributed to the fact that 9×20 pressure images do not allow us to estimate homology above order one as in the examples presented in Hofer, Kwitt, and Niethammer (2019).

Multiple object identification. The pressure sensor image dataset included a small number of rock experiments with dense rock distribution, i.e. rocks were laid out sequentially close to each other, so that the wheel might be touching more than one rock at a time, however, the rocks were not positioned side to side. Such experimental set up is due to the fact that the classical ML model developed for the tactile wheel would not be able to handle multiple objects explicitly, and more features would need to be developed and extracted to account for that. As a result, this dataset does not capture such scenarios that could further benefit from GDL. This GDL research thrust, however, opens the possibility of more complex terrain experiments with the tactile wheel that would be handled automatically by GCN models (Chen et al. 2019). In real-life exploration scenarios, the terrain will not be homogeneous. Rocks of different sizes and shapes, combined with other landforms, will be present.

Non-image sensor data. In case of tactile and in-situ sensing, the data from instruments other than a 2D pressure sensor array will be essential for such systems to be effective. For example, the EIS sensor on the tactile wheel takes highly valuable measurements of soil hydration, which consist of amplitudes and phases across various frequencies. The ability to tackle non-image data, such as spectra, will benefit any model framework if to be used for in-situ sensing. Performance of GCN models for image classification recently has been proven to be enhanced by introducing additional data sources as in natural language processing (Garcia and Bruna 2018; Yan, Xiong, and Lin 2018). The TOPO-GCN, ARMA-GCN and other GDL approaches are potentially more suited for combining data sources.

Path to Deployment

While we show that DL, and especially TOPO-GCN, can achieve substantial gains in rock classification accuracy, these methods retain the challenge for future onboard systems. Currently, machine learning applications for space robotics are limited due to the strict hardware, computation and energy resource constraints. There are, however, successful applications and deployments that are based on machine learning principles (Chien et al. 2003; Estlin et al. 2014; Thompson et al. 2014; Doran, Thompson, and Estlin 2016). The state of the art are simple models or algorithms developed on the ground, rigorously validated and tested, and integrated into flight software with strict requirements. Model re-training is not currently feasible onboard of instruments or spacecraft and is limited to feature extraction and prediction. Pre-trained models is a possible intermediate solution to this problem which requires further analysis

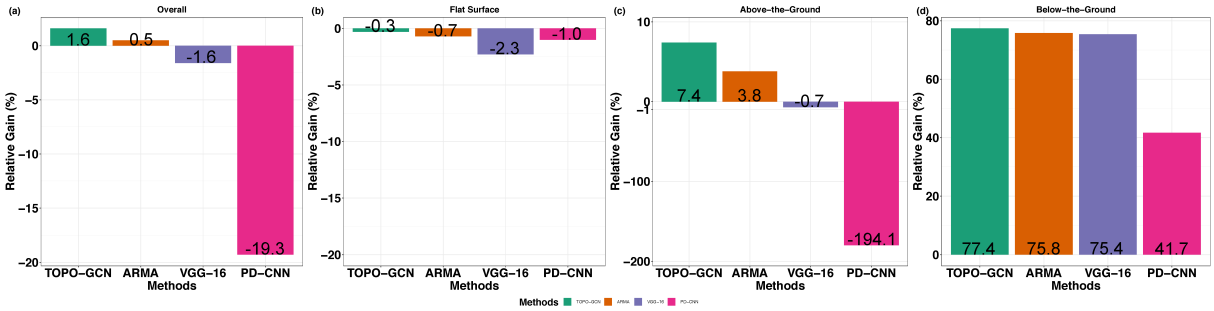


Figure 6: Relative gains (%) compared with SGB as baseline on TOPO-GCN, ARMA-GCN, VGG-16, and PD-CNN. Otherwise, the baseline result is reported in Table 1.

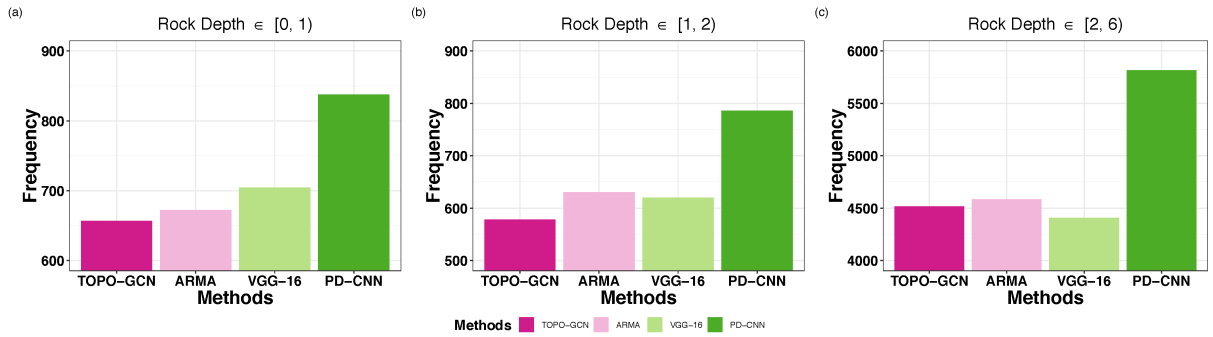


Figure 7: Distribution of misclassified instances as a function of rock depth. Lower values are preferred.

on transferability. In turn, topological methods in GDL have a potential for such interpretability and transferability properties (Gabrielsson et al. 2020).

High Performance Space Computing (HPSC) processor (Doyle et al. 2013; Powell 2018; Schwaller, Holtzman, and George 2019) is a new promising technology in development that could potentially overcome the limitations of the current flight processors. RAD750 (Lentaris et al. 2018) is the most advanced space computer, designed to operate in high-radiation environments and is currently installed on the Curiosity and Perseverance Rovers and the Juno mission. It operates at 132 MHz with 128 megabytes of DRAM local memory. HPSC, however, would be a high-performance, multi-core processor, a future "laptop in space", and would potentially increase computing power by two orders of magnitude. It could allow prediction based on advanced and deep machine learning models, which would greatly benefit space and spacecraft autonomy applications (McGovern and Wagstaff 2011), including terrain recognition for rover missions. As the need for more advanced methodologies in space exploration and instrument autonomy grows due more far-away and complex missions (Castano et al. 2007; Wagstaff et al. 2014, 2019), HPSC framework would potentially make it possible to take TOPO-GCN into space for future planetary image processing and classification needs.

Conclusions

We have explored utility and limitations of various DL tools for the analysis of terrain properties in a novel tactile sensing dataset. Our findings have indicated that classification of more sophisticated terrain types tend to benefit more from the DL methodology which systematically integrates both topological and geometric properties of the underlying data. In the future, we plan to explore slip pressure images and more complex surface patterns, e.g. dunes, gullies and other landforms along with their combinations. We will also seek to integrate non-image sensor data, such as spectra, into GDL methods for terrain as well as soil property learning. We believe that these methodologies have the potential to enable smart, autonomous sensors on future planetary exploration vehicles.

Acknowledgments

Gel has been supported in part by NSF DMS 1925346, NSF IIS 1633331 and NASA 20-RRNES20-0021. Chen has been supported by 2020 NSF INTERN under NSF ECCS 1824716. The research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

References

Adams, H.; Emerson, T.; Kirby, M.; Neville, R.; Peterson, C.; Shipman, P.; Chepushtanova, S.; Hanson, E.; Motta, F.; and Ziegelmeier, L. 2017. Persistence images: A stable vector representation of persistent homology. *JMLR* 18(1): 218–252.

Arvidson, R. E.; Iagnemma, K. D.; Maimone, M.; Fraeman, A. A.; Zhou, F.; Heverly, M. C.; Bellutta, P.; Rubin, D.; Stein, N. T.;

Grotzinger, J. P.; and Vasavada, A. R. 2017. Mars science laboratory curiosity rover megaripple crossings up to sol 710 in gale crater. *J. of Field Robot.* 34(3): 495–518.

Bianchi, F. M.; Grattarola, D.; Alippi, C.; and Livi, L. 2019. Graph neural networks with convolutional ARMA filters. *arXiv:1901.01343* .

Bronstein, M. M.; Bruna, J.; LeCun, Y.; Szlam, A.; and Vandergheynst, P. 2017. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine* 34(4): 18–42.

Castano, R.; Wagstaff, K. L.; Chien, S.; Stough, T. M.; and Tang, B. 2007. On-board analysis of uncalibrated data for a spacecraft at Mars. In *ACM SIGKDD*, 922–930.

Chen, Z.-M.; Wei, X.-S.; Wang, P.; and Guo, Y. 2019. Multi-label image recognition with graph convolutional networks. In *IEEE CVPR*, 5177–5186.

Chien, S.; Sherwood, R.; Tran, D.; Castano, R.; Cichy, B.; Davies, A.; Rabideau, G.; Tang, N.; Burl, M.; Mandl, D.; et al. 2003. Autonomous science on the EO-1 mission .

Doran, G.; Thompson, D. R.; and Estlin, T. 2016. *Precision instrument targeting via image registration for the Mars 2020 rover*. JPL NASA.

Doyle, R.; Some, R.; Powell, W.; Mounce, G.; Goforth, M.; Horan, S.; and Lowry, M. 2013. High performance spaceflight computing (HPSC) next-generation space processor (NGSP): A joint investment of NASA and AFRL. In *FSW*.

Edelsbrunner, H.; and Morozov, D. 2012. Persistent homology: theory and practice. In *Eur. Congress of Math.*

Estlin, T.; Gaines, D.; Bornstein, B.; Schaffer, S.; Tompkins, V.; Thompson, D. R.; Altinok, A.; Anderson, R. C.; Burl, M.; Castaño, R.; et al. 2014. Automated targeting for the MSL rover ChemCam spectrometer. In *i-SAIRAS*, 17–19.

Gabrielsson, R. B.; Nelson, B. J.; Dwaraknath, A.; and Skraba, P. 2020. A topology layer for machine learning. In *AISTATS*, 1553–1563.

Garcia, V.; and Bruna, J. 2018. Few-shot learning with graph neural networks. In *ICLR 2018*.

Hofer, C. D.; Kwitt, R.; and Niethammer, M. 2019. Learning Representations of Persistence Barcodes. *JMLR* 20(126): 1–45.

Jiang, B.; Zhang, Z.; Lin, D.; Tang, J.; and Luo, B. 2019. Semi-supervised learning with graph learning-convolutional networks. In *IEEE CVPR*, 11313–11320.

Kerner, H. R.; Wellington, D. F.; Wagstaff, K. L.; Bell, J. F.; Kwan, C.; and Amor, H. B. 2019. Novelty detection for multispectral images with application to planetary exploration. In *AAAI*, 9484–9491.

Lentaris, G.; Maragos, K.; Stratakos, I.; Papadopoulos, L.; Papanikolaou, O.; Soudris, D.; Lourakis, M.; Zabulis, X.; Gonzalez-Arjona, D.; and Furano, G. 2018. High-performance embedded computing in space. *J. of Aerospace Inf. Systems* 15(4): 178–192.

Ma, J.; Tang, W.; Zhu, J.; and Mei, Q. 2019. A Flexible Generative Framework for Graph-based Semi-supervised Learning. In *NeuRIPS*, 3281–3290.

Marchetti, Y.; Lightholder, J.; Junkins, E.; Cross, M.; Fraeman, A.; and Mandrake, L. 2020. Barefoot Rover: a Sensor-Embedded Rover Wheel Demonstrating In-Situ Engineering and Science Extractions using Machine Learning. In *IEEE ICRA*, 6001–6006.

McGovern, A.; and Wagstaff, K. L. 2011. *Machine learning in space: extending our reach*. Springer.

Monti, F. 2020. Geometric Deep Learning. http://geometricdeeplearning.com. Accessed on December 28, 2020.

Monti, F.; Boscaini, D.; Masci, J.; Rodola, E.; Svoboda, J.; and Bronstein, M. M. 2017. Geometric deep learning on graphs and manifolds using mixture model CNNs. In *CVPR*, 5115–5124.

Nagatani, K.; Ikeda, A.; Sato, K.; and Yoshida, K. 2009. Accurate estimation of drawbar pull of wheeled mobile robots traversing sandy terrain using built-in force sensor array wheel. In *IROS*, 2373–2378.

Potts, C.; Yang, L.; Oyen, D.; and Wohlberg, B. 2019. A Topological Graph-Based Representation for Denoising Low Quality Binary Images. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 1788–1798.

Powell, W. A. 2018. *High-performance spaceflight computing (HPSC) project overview*. NASA.

Pun, C. S.; Xia, K.; and Lee, S. X. 2019. Persistent-Homology-based Machine Learning and its Applications–A Survey. *arXiv:1811.00252* .

Rothrock, B.; Kennedy, R.; Cunningham, C.; Papon, J.; Heverly, M.; and Ono, M. 2016. Spoc: Deep learning-based terrain classification for mars rover missions. In *AIAA SPACE 2016*, 5539.

Schwaller, B.; Holtzman, S.; and George, A. D. 2019. Emulation-Based Performance Studies on the HPSC Space Processor. In *IEEE AeroConf*, 1–11.

Senatore, C.; Stein, N.; Zhou, F.; Bennett, K.; Arvidson, R.; Trease, B.; Lindemann, R.; Bellutta, P.; Heverly, M.; and Iagnemma, K. 2014. Modeling and validation of mobility characteristics of the Mars Science Lab Curiosity Rover. In *i-SAIRAS*.

Shirai, T.; and Ishigami, G. 2015. Development of in-wheel sensor system for accurate measurement of wheel terrain interaction characteristics. *Journal of Terramechanics* 62: 51–61.

Simonyan, K.; and Zisserman, A. 2015. Very deep convolutional networks for large-scale image recognition. *ICLR 2015* .

Subramanya, A.; and Talukdar, P. P. 2014. Graph-based semi-supervised learning. *Synthesis Lec. on AI and ML* 8(4): 1–125.

Thompson, D. R.; Allwood, A.; Assad, C.; Flannery, D.; Hodyss, R.; Knowles, E.; and Wade, L. 2014. Adaptive sampling for rover x-ray lithochemistry. In *i-SAIRAS*.

Wagstaff, K. L.; Doran, G.; Davies, A.; Anwar, S.; Chakraborty, S.; Cameron, M.; Daubar, I.; and Phillips, C. 2019. Enabling Onboard Detection of Events of Scientific Interest for the Europa Clipper Spacecraft. In *ACM SIGKDD*, 2191–2201.

Wagstaff, K. L.; Lu, Y.; Stanboli, A.; Grimes, K.; Gowda, T.; and Padams, J. 2018. Deep Mars: CNN Classification of Mars Imagery for the PDS Imaging Atlas. In *AAAI*, 7867–7872.

Wagstaff, K. L.; Thompson, D. R.; Bue, B. D.; and Fuchs, T. J. 2014. Autonomous real-time detection of plumes and jets from moons and comets. *The Astrophysical Journal* 794(1): 43.

Wu, X. A.; Huh, T. M.; Mukherjee, R.; and Cutkosky, M. 2016. Integrated ground reaction force sensing and terrain classification for small legged robots. *IEEE Robot. Aut. Lett.* 1(2): 1125–1132.

Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; and Philip, S. Y. 2020. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems* .

Yan, S.; Xiong, Y.; and Lin, D. 2018. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *AAAI*, 7444–7452.