Thermophysical Change Detection on the Moon with the Lunar Reconnaissance Orbiter Diviner sensor

Silvia Bucci[*]
Polytechnic of Turin, Italy
silvia.bucci@polito.it

Jose Ignacio Delgado Centeno[*]
University of Luxembourg
jose.delgado@uni.lu

Ben Gaffinet[*]
RSS-Hydro, Luxembourg
bgaffinet@rss-hydro.lu

Ziyi Liang[*]
University of Southern California
ziyilian@usc.edu

Valentin Bickel
ETH Zürich VAW
bickel@vaw.baug.ethz.ch

Ben Moseley
ETH Zürich AI Center
benjamin.moseley@ai.ethz.ch

Miguel Olivares-Mendez
University of Luxembourg
miguel.olivaresmendez@uni.lu

Abstract

The Moon is an archive of the history of the Solar System, as it has recorded and preserved physical events that have occurred over billions of years. NASA's Lunar Reconnaissance Orbiter (LRO) has been studying the lunar surface for more than 13 years, and its datasets contain valuable information about the evolution of the Moon. However, the vast amount and heterogeneous nature of data collected by LRO make the extraction of scientific insights very challenging - in the past most analyses relied on human review. Here, we present NEPHTHYS, an automated solution for discovering thermophysical changes on the surface using one of LRO's largest datasets: the thermal data collected by its Diviner instrument. Specifically, NEPHTHYS is able to perform systematic, efficient, and large-scale change detection of present-day impact craters on the surface. Further work could enable more comprehensive studies of lunar surface impact flux rates and surface evolution rates, providing critical new information for future missions.

1 Introduction

The nature, magnitude, and frequency of present-day changes of the lunar surface are of key importance to understand and reconstruct the Moon's and the Solar System's history, to analyze the surface evolution of airless planetary bodies and to identify potential environmental hazards for future robotic and crewed For the Moon, the most extensive datasets that allow for the study of the surface are provided by NASA's Lunar Reconnaissance Orbiter (LRO) mission, which has been operating since 2009. Two of LRO's most important instruments are its NAC (Narrow Angle Camera, Robinson et al. (2010)), which acquires high-resolution ($\sim$0.5 m footprint) optical images, and Diviner (Paige et al. (2010)), which acquires intermediate-resolution ($\sim$200 m footprint) point measurements of the surface temperature.

[*]Equal contribution

Machine Learning and the Physical Sciences workshop, NeurIPS 2022.

Despite the abundance of these global datasets with long temporal baselines, our knowledge about present-day changes is limited. Past work has been able to identify a small number ($\lesssim 1000$) of such changes - predominantly fresh, small impact craters (Speyerer et al. (2016), Watters et al. (2022), Williams et al. (2018)), but has faced three major challenges: 1) the massive size of the existing datasets (hundreds of terabytes) which makes human mapping impossible, 2) highly variable illumination conditions, and 3) complex noise in the dataset which makes faint and small change signals hard to detect. For example, the most extensive change detection study so far (Speyerer et al. (2016)) relied on the matching of multi-temporal optical NAC images acquired at similar local times, which limited their spatial coverage to only 6.6% of the lunar surface.

To overcome such limitations, we turn to machine learning to provide more automated, scalable and efficient change detection, which accurately detects low signal-to-noise events. In particular, we focus on detecting thermophysical changes within the Diviner dataset; such changes are usually caused by physical events such as small impacts but can have significant spatial thermal expressions and remain visible for hundreds of thousands of years (Williams et al. (2018)). Specifically, we focus on detecting fresh impact craters and their surrounding ejecta blankets, which have different night-time cooling properties than the regular lunar surface, producing so-called cold spots (Williams et al. (2018)) (Fig. 1). Fresh impact craters are believed to represent the largest source of lunar surface change.

We introduce NEPHTHYS (New Event Perceiving High-Trust High-Yield System), which searches for cold spots by using two different neural networks (a CNN-based binary classifier and a normalizing flow based anomaly detector) to analyse preprocessed thermal images constructed from the Diviner thermal point measurements. The output of the networks includes a detection confidence level to increase their reliability. Our workflow is validated by scanning areas with known fresh impacts and then verifying candidates by hand using NAC optical images (Williams et al. (2018)).

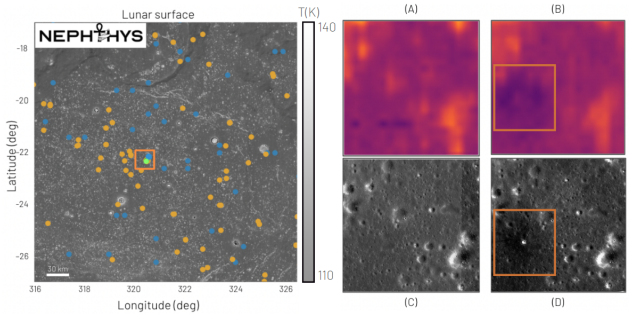

Figure 1: An example of impact detection with NEPHTHYS. Left: thermal Diviner mosaic (showing average night temperature) of a scanned region (280x280 km): orange and blue dots indicate potential candidates for fresh craters found by the binary classifier and anomaly detector, respectively. Green dot inside the bounding box indicates a known recent impact, the target of the search. Validation of the detection: (A),(B) our preprocessed thermal image *before* and *after* the impact, (C),(D) NAC optical image *before* and *after* the impact.

2 Methodology

Our high-level detection workflow consists of two main steps. First, we carry out preprocessing of the raw Diviner point measurements (See Appendix A for a detailed description of the Diviner dataset). Specifically, given an area of interest on the surface, we bin and aggregate the point measurements into two images, representing a *before* and *after* image of the surface temperature spanning the entire operational window of the instrument (Fig. 2). Secondly, we feed these images into two different detection networks, which independently predict whether an impact has occurred by searching for an associated cold spot in the images. Finally, these networks can be scanned over the entire lunar surface to provide an efficient and large-scale detection method.

2.1 Data Preprocessing Workflow

Following Moseley et al. (2020), we first sort and store all of the available Diviner point measurements into three-dimensional "data cubes", indexed by location (in 0.5° latitude and 0.5° longitude intervals) and local time (6-hour intervals), for computational retrieval efficiency. We then produce one *before* and one *after* image for each region of interest: *before* images are populated using the oldest available Diviner measurements over the respective location; *after* images are populated using the most recent Diviner observations over the respective location. Importantly, we currently only use night time point measurements acquired between 22:00 and 06:00 local time to avoid the complex changes in temperature associated with daytime illumination. All data over 2009-2022 is used. Populating the *before* and *after* images consists of two steps: a) *Data normalization* and b) *Binning and aggregation*. *Gaussian smoothing* is also required as third step before creating a *difference* image, used as the input of one of our ML solutions (Section 2.3).

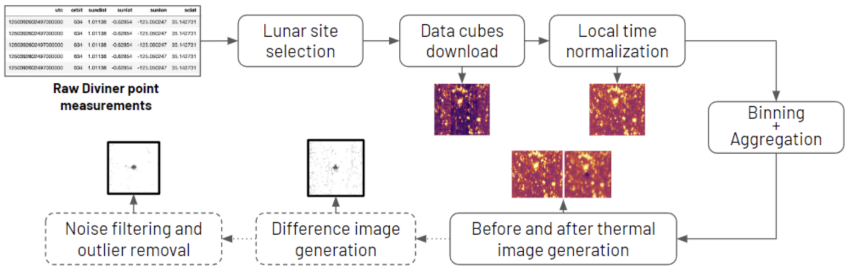

Figure 2: Data preprocessing workflow used by NEPHTHYS.

In Fig. 2 we show a raw temperature image (after Data cubes download). We note that heterogeneous stripes of different temperatures can be observed. This is because some point measurements are taken at an earlier local time (within the 22:00 - 06:00 window used) when the surface has not cooled down completely (brighter color). The darker regions instead correspond to measurements taken at a later local time. To fix this issue, we normalize the measurements to a reference time (precisely, 00:00 local lunar time) computed using Hayne's model (Hayne et al. (2017)), which simulates the background temperature variation with local time for a given latitude on the Moon. After this step, we obtain a more uniform image (Fig. 2, after Local time normalization).

In order to produce *before* and *after* images from the normalized point measurements, we need to go through the second preprocessing step: *Binning and aggregation*. For each image, we bin all the available point measurements onto 100×100 m grids. To create a *before* image we take the measurements from the earliest year available in each bin and for the *after* image, we take the measurements from the most recent year. Lastly we average the selected measurements in each bin to get the final image pairs. To create a *difference* image, we subtract the *after* image from the *before* image. We note that Diviner's pointing accuracy includes slight errors, occasionally resulting in a spatial misalignment of measurements, resulting in spatial noise in the *difference* images. Gaussian filtering (Chung (2020)) is employed to smooth out this noise after the binning step in both images. The last two pairs of images in Fig 2 show the effect of the Gaussian filtering.

2.2 Binary Classification

A straightforward solution to detect the presence of a fresh cold spot by analyzing the *before* and *after* images separately is to check for the presence of a cold spot in both - if the *before* image does not contain a cold spot but the *after* image does, then there was a recent change. We note there is a difference between *old* cold spots and *fresh* cold spots: the former refers to impacts that occurred before 2009 (the year the LRO diviner mission launched); the latter refers to impacts that happened after 2009 - those are actual present-day changes.

We use a Deep Neural Network (DNN) (He et al. (2016)) as a binary classifier to determine whether an image contains a cold spot or not. Our training data is a catalog of $\sim$ 2000 known, *old* cold spots as positive samples (taken from Williams et al. (2018)) and $\sim$ 2000 random points as negative samples taken over the entire lunar surface. We can safely use random points as negative samples given that the probability of picking a point with an old or fresh cold spot is negligibly small (Williams et al.

(2018)). The trained classifier can be deployed on the entire lunar surface: for each point we produce the *before* and *after* images, use the model to classify both, and evaluate for changes.

2.3 Anomaly Detection

Ideally, if there were no changes, the *before* and *after* images of the same location should be identical. However, minor measurement localization inaccuracies add noise (*background noise*). This is the reason why applying simple thresholding techniques on the *difference* images is often challenging for change detection. As an alternative approach, we propose to formulate the problem as an *anomaly detection* task (Hendrycks et al. (2018), Yang et al. (2021)), shown in Fig. 3. Given a *difference* image, we want to assess whether the image lies within the distribution of difference images where no fresh cold spots are present. Our solution is based on *DifferNet* Rudolph et al. (2021), a recent approach in which a Normalizing Flow (NF) Rezende and Mohamed (2015) is applied to the features extracted from a DNN to compute such an *anomaly score*.

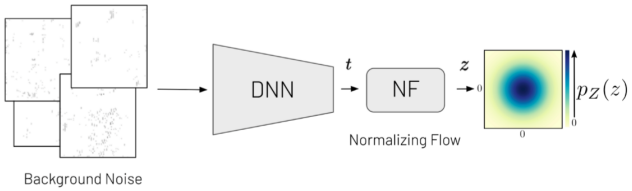

Figure 3: Overview of the architecture used for *anomaly detection* Rudolph et al. (2021)

Let X_T be the set of training data composed of the *difference* images of locations without fresh cold spots (the 2000 random locations used for the BC). For each $x_i \in X_T$, let t_i be the features extracted from the DNN in Fig. 3, and let z_i be the features extracted from the NF. Following Rudolph et al. (2021), we maximize the likelihood of the extracted features t which are quantifiable in the latent space Z. After a change-of-variables and the use of the negative log-likelihood (Rudolph et al. (2021)) obtain the following loss function to train the NF:

$$L(t_i) = \frac{\|z_i\|^2}{2} - \log \left| \det \frac{\partial z}{\partial t} \right| \tag{1}$$

Intuitively, we want that NF maps t as close as possible to $z = 0$ for all of the training examples. During evaluation we can compute the *anomaly score* through the negative log-likelihoods for each test sample $x_j \in X_{eval}$:

$$\tau(x_j) = E[-\log p_Z(z_j)] \tag{2}$$

An image is classified as anomalous if the anomaly score is above some threshold value, Θ.

3 Results and validation

To evaluate the quantitative performance of our binary classifier, we consider a validation set made of 400 images (half positive, half negative): it is able to achieve about 95% of accuracy on them. For the anomaly detector we consider the known fresh cold spots (enhanced with augmented versions of them) and random points: it assigns a score 4-times higher on average for the difference images that contain a cold spot on them. A downstream task for the validation of our approach can be seen in Fig. 1. Here we chose an area of 10x10 degrees in latitude and longitude that contains one known fresh impact that occurred in 2013 (Williams et al. (2018)). The two detectors identify a number of fresh cold spot candidates (orange dots: binary classifier, blue dots: anomaly detector). Importantly, both approaches were able to find the target of the experiment, the known recent impact (green dot). We repeated the same experiment for 4 additional sites with a-priori known fresh impacts: NEPHTHYS was able to consistently find every known fresh impact. Unfortunately, the locations of most existing known fresh impacts have not been made publicly available by LROC, and only these 5 were available from the Diviner team (Williams et al. (2018)) and recognizable in Diviner data.

4 Conclusion, limitations and future steps

We have implemented NEPHTHYS: an ML-based approach to detect present-day thermophysical changes on the lunar surface. A pre-processing pipeline was implemented to convert point measurements from the Diviner instrument into images suitable for ML-driven change detection. Two independent ML solutions were developed, which were able to efficiently scan a region of the Moon and produce a list of potential surface changes. Verifying the detected sites was one of the biggest limitations of our approach due to the lack of known impacts and significant human effort required to check the candidates with NAC imagery. In future work, we would like to submit target acquisition requests to the LROC team to further verify NEPHTHYS' candidates. We would also like to assess whether NEPHTHYS can detect other thermophysical changes, such as moonquakes and landslides. Newly discovered surface changes will provide key insights into the history of the Moon and the Solar System and contribute essential information to the planning of upcoming lunar missions.

Acknowledgement

This work has been enabled by the Frontier Development Lab Program (FDL USA). FDL is a collaboration between SETI Institute and Trillium Technologies Inc., in partnership with NASA, USGS, Google Cloud, NVIDIA, PLANET and the Luxembourg Space Agency (LSA).

5 Broader impact

The main goal of our work is enabling planetary science researchers to process and analyze NASA's Lunar Reconnaissance Orbiter (LRO) data in an automated manner. We believe that our work could have a great impact in the society. Helping in having successful lunar missions, our work contributes to the exploration of the universe whose goal is also to find other planets suitable for life. Considering the climate change and all the other problems that our planet is facing, a backup plan becomes urgent.

References

Chung, M. K. (2020). Gaussian kernel smoothing.

Hayne, P. O., Bandfield, J. L., Siegler, M. A., Vasavada, A. R., Ghent, R. R., Williams, J.-P., Greenhagen, B. T., Aharonson, O., Elder, C. M., Lucey, P. G., and Paige, D. A. (2017). Global regolith thermophysical properties of the moon from the diviner lunar radiometer experiment. *Journal of Geophysical Research: Planets*, 122(12):2371–2400.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Hendrycks, D., Mazeika, M., and Dietterich, T. (2018). Deep anomaly detection with outlier exposure. *arXiv preprint arXiv:1812.04606*.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90.

Moseley, B., Bickel, V., Burelbach, J., and Relatores, N. (2020). Unsupervised learning for thermophysical analysis on the lunar surface. *The Planetary Science Journal*, 1(2):32.

Paige, D., Foote, M., Greenhagen, B., Schofield, J., Calcutt, S., Vasavada, A., Preston, D., Taylor, F., Allen, C., Snook, K., et al. (2010). The lunar reconnaissance orbiter diviner lunar radiometer experiment. *Space Science Reviews*, 150(1):125–160.

Rezende, D. and Mohamed, S. (2015). Variational inference with normalizing flows. In *International conference on machine learning*, pages 1530–1538. PMLR.

Robinson, M., Brylow, S., Tschimmel, M., Humm, D., Lawrence, S., Thomas, P., Denevi, B., Bowman-Cisneros, E., Zerr, J., Ravine, M., et al. (2010). Lunar reconnaissance orbiter camera (lroc) instrument overview. *Space science reviews*, 150(1):81–124.

Rudolph, M., Wandt, B., and Rosenhahn, B. (2021). Same same but differnet: Semi-supervised defect detection with normalizing flows. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1907–1916.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., et al. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3):211–252.

Speyerer, E. J., Povilaitis, R. Z., Robinson, M. S., Thomas, P. C., and Wagner, R. V. (2016). Quantifying crater production and regolith overturn on the moon with temporal imaging. *Nature*, 538(7624):215–218.

Watters, T., Speyerer, E., and Robinson, M. (2022). Recent landslides and their relation to young thrust fault scarps on the moon. *LPI Contributions*, 2678:1626.

Williams, J.-P., Bandfield, J., Paige, D., Powell, T., Greenhagen, B., Taylor, S., Hayne, P., Speyerer, E., Ghent, R., and Costello, E. (2018). Lunar cold spots and crater production on the moon. *Journal of Geophysical Research: Planets*, 123(9):2380–2392.

Yang, J., Wang, H., Feng, L., Yan, X., Zheng, H., Zhang, W., and Liu, Z. (2021). Semantically coherent out-of-distribution detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8301–8309.

Appendix

A LRO Diviner dataset

The LRO Diviner instrument (Paige et al. (2010)) is a passive radiometer on board NASA's LRO satellite which has been gathering point measurements of the lunar surface temperature along the ground track of the satellite over the past 13 years. Measurements are collected across nine different wavelength channels from optical to thermal wavelengths. Over multiple orbits, LRO Diviner has covered the entire surface of the Moon, providing multiple measurements at each surface location taken over the full range of local times (noon over midnight to noon). All collected data points are publicly available on NASA's PDS [2]. We use a processed version of the raw Diviner data that was pointing-corrected and calibrated (Level 1), specifically including channels 7, 8, and 9 (covering 12.5 to 200 μm), provided by the Diviner team (Paige et al. (2010)).

B Implementation details

Binary Classification. We build our approach over a standard ResNet-18 (He et al. (2016)) pretrained on ImageNet (Russakovsky et al. (2015)). We train our model for 2000 epochs using batch size 32, learning rate 0.001 and Adam optimizer. We use the standard cross-entropy loss to finetune only the last block of the network keeping the rest frozen with the ImageNet weights. As data preprocessing we first resize the images to 256, then we crop every image into 50x50 crops covering with 121 windows the entire image. Lastly, we resize every crop to 224 and feed the images into the network.

Anomaly Detection. We used as feature extractor AlexNet (Krizhevsky et al. (2017)) followed by the Normalizing Flow block as proposed in DifferNet (Rudolph et al. (2021)). We update the weights only for the Normalizing Flow block keeping the rest frozen with the ImageNet pretrained weights. We first resize the images to 448 and then crop the image to four 224 sub images with no overlap, then resize each sub image to 448. Considering the performances on the validation set, we set the threshold value Θ to 0.8.

For the training of both of the architectures described in the previous section, a NVIDIA Tesla V100 was used. The lightweight nature of the networks used made it possible to run multiple iterations of the training for optimization purposes in a short amount of time. For inference, a similar hardware setup will make it possible to scan specified lunar surface regions with the change detection algorithms.

[2] `https://pds-geosciences.wustl.edu/missions/lro/diviner.htm`

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes]
   (c) Did you discuss any potential negative societal impacts of your work? [Yes]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
   (e) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
   (f) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
   (g) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes]
   (h) If your work uses existing assets, did you cite the creators? [Yes]
   (i) Did you mention the license of the assets? [N/A]
   (j) Did you include any new assets either in the supplemental material or as a URL? [N/A] No new assets
   (k) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] The NASA dataset is publicly available
   (l) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]