Automated Lunar Crater Detection and Locating Using a Convolutional Neural Network and Digital Elevation Models

Jamie Johnston
jjohn42@ilstu.edu
Illinois State University
Department of Information Technology

Matt Caplan
mecapl1@ilstu.edu
Illinois State University
Department of Physics

1 INTRODUCTION

Traditionally, lunar crater counting has been done by visual inspection of images of the moon's surface. This method is time consuming and has poor inter-rater reliability for smaller craters. Automating this process using a Convolutional Neural Network (CNN) greatly improves the speed and reliability at which surface features can be detected and classified.

Using available high resolution Digital Elevation Model (DEM) images from the Lunar Reconnaissance Orbiter (LRO), we utilize a CNN to identify craters and locate them by translating their pixel coordinates to geospatial coordinates. Presently the population of small impactors is not well understood but improved detection of the smallest craters can constrain the size distribution of asteroids in the solar system.

In future work, we intend to search the data set for small craters with novel features that are inconsistent with traditional asteroid impacts to potentially constrain the moon's interaction history with MACHO dark matter from the Galactic halo.

2 MOTIVATION

As much as 10^{18}g of dark matter (DM) may have transited the moon in the history of the solar system, motivating us to consider ways to use the moon as a DM detector. Craters on the lunar surface hold a record of billions of years of impacts, some of which may be due to interactions with MACHO dark matter such as primordial black holes (PBHs) or strangelets.

While traditional asteroids create a point explosion at the surface which destroys the impactor on contact, relativistic compact DM fully penetrates due to its high relative density. Accretion onto a PBH during this transit drives an outward shock producing a 'line explosion' with markedly different cratering dynamics (Caplan & Yalinewich 2020) [1]. Ejecta from a collision with a PBH results in more vertical ejecta and less radial ejecta due to the explosion dynamics, producing a steeper ejecta blanket which may be detectable in high resolution lunar surface maps now available. We seek to identify candidate craters with anomalous ejecta profiles in Digital Elevation Models taken from the Lunar Reconnaissance Orbiter (LRO).

3 PRIOR WORKS

Prior work conducted using CNNs to identify surface features on the Moon have provided promising results, including those from A. Silburt & C Zhu in the DeepMoon [2] project and Wang et al [3] at East China Jiaotong University (ECJU) both achieving high levels of detection and prediction accuracy. These works analyze larger formations on the lunar surface than is the objective of this project.

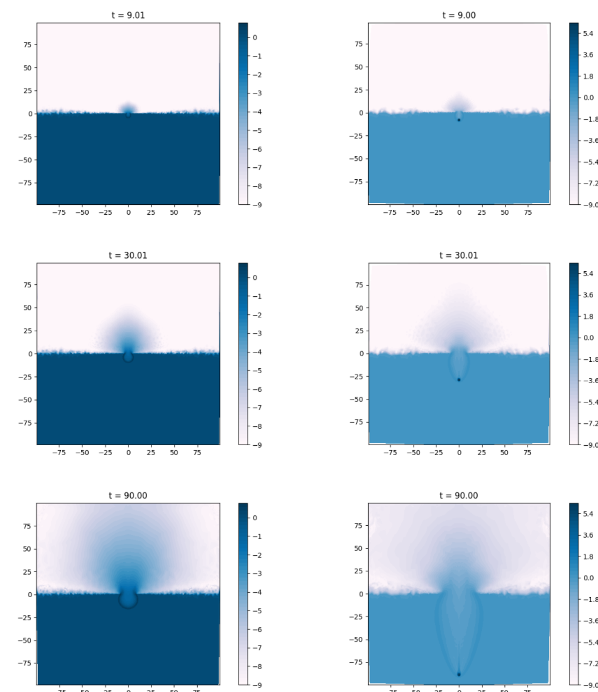

Figure 1: Log density snapshots from numerical simulations of a regular, head on impact (left) and an impact of a primordial black hole (right), both without gravity [1]

4 PIPELINE

Figure 2 illustrates the data pipeline utilized in this work. Digital Elevation Models are retrieved from the NASA's Planetary Geologygydessy Data Archive (PGDA) [4], split into much smaller tiles in preparation for the CNN. The Convolutional Neural Network will produce x,y image pixel coordinates of detections which are then translated into geospatial coordinates.

4.1 Sourcing Data

A major challenge for this project was sourcing the data needed. Finding high density, high resolution surface scans of the entire Lunar surface was challenging. There are multiple university and government websites providing data, however finding catalogued coverage of the moon in a suitable format was a significant hurdle to overcome.

NASA's Planetary Geology, Geophysics and Geochemistry Laboratory [5] host an archive of their data and readings from the

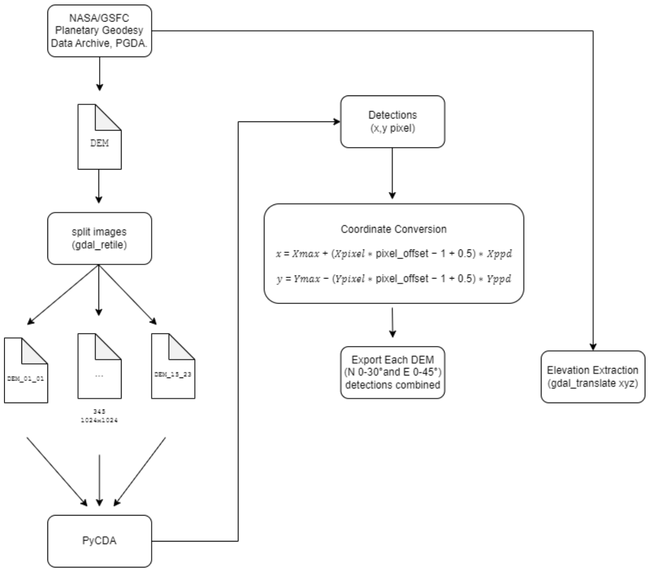

Figure 2: Data Flow Diagram

LRO/LOLA which proved to be exactly what was needed. The merged LOLA Kaguya Lunar Topography (Azimuth) catalog was selected for this project.

4.2 Splitting the Digital Elevation Models

The Geospatial Data Abstraction Library (GDAL) [7] library provides a suite of utilities for manipulating and analyzing raster and vector geospatial data formats, which was used to split the DEM files into a size that the CNN could handle. Each DEM was split into 345 .tif files, each 1024x1024 resulting in 16,650 individual images to be processed. Using gdal_retile was an essential step in the process as it retains the elevation data contained within the file, while most other solutions lose this metadata.

One concern that is easily addressed in the tiling process is edge losses, potentially missing detections that fall on the bordering pixels of two images. In order to combat this issue, we remove 512 pixels from the top, and from one side of each DEM and run this new set of images through the pipeline. The result of this process gives us a set of tiles where the center of each tile falls on the border of the 4 adjacent tiles from the previous set. Duplicate detections can then be eliminated easily, resulting in total coverage without any edge losses.

4.3 PyCDA

PyCDA [8] is a Crater Detection Algorithm (CDA) utilizing a trained Convolutional Neural Network (CNN) to both identify and classify craters on celestial bodies. Inspired by research in applying Convolutional Neural Networks to crater detection (Benedix et al.) [6] and crater candidate classification (Cohen et al.) [9], PyCDA is aimed at making CDA research modular and usable.

This model was trained to detect craters with a diameter of 80 pixels or less, and the data set used for analysis was published by NASA/GSFC Planetary Geodesy Data Archive, (PGDA) [4], obtained by the Lunar Orbiter Laser Altimeter (LOLA) an instrument onboard the LRO. These surface scans collected in 2015 each cover 30 degrees of latitude, and 45 degrees of longitude, at 512 pixels per degree (59.2 m/pixel) meaning the trained model can identify sub kilometer craters with significantly higher accuracy than can be achieved by human inspection.

Table 1: Important Values

Name	Value	Comments
mpd	59.225294	Meters Per Degree
X_{max}	-5458203.07634699	(Top) Left Most Coordinate
Y_{max}	1819401.02544900	(Top) Left Most Coordinate
X_{mpd}	59.2252939999999	Meters Per Degree on X axis
Y_{mpd}	59.2252939999999	Meters Per Degree on Y axis

Analyzing these images requires us to divide each image into smaller tiles, of 1024x1024 pixels and processing those images individually. A "U-Net" Convolutional Neural Network is used to produce these detections; this type of model was selected because it gives highly location-specific predictions.

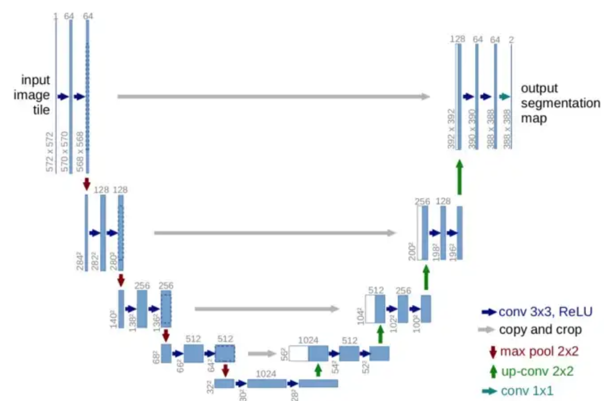

Figure 3: U-Net Architecture Diagram [10]

The reason it is able to localize and distinguish borders is by binarizing the input, converting to black and white pixels only, and performing classification on every pixel, so the input and output share the same size. Once the detection stage is complete, craters can be correlated with Digital Elevation Model (DEM) images to study crater profiles, extracting elevation data from these images allows for the analysis of the ejecta blanket for PBH collision candidates.

4.4 Converting Pixels to Spatial Coordinates

A Digital Elevation Model (DEM) is a representation of the bare ground topographic surface. Each DEM contains x,y,z spatial and elevation information, covers 30°x 45°, and is 15,360x30,720 pixels. A total of 471,859,200 pixels per image. We see sometimes over 150,000 detections per DEM where the detection is given as an x & y pixel coordinate.

We use the below functions to convert pixel coordinates to spatial coordinates as an offset from the Prime Meridian and Equator, values and descriptions can be found in Table 1. Providing surface coordinates is significantly more valuable than pixel coordinates, allowing researchers to locate detections on the Lunar surface for further research.

The center point of the Moon where the Equator intersects the Prime Meridian would be x = 0, y = 0. Therefore negative X coordinates cover West of the Prime Meridian, and positive Y values are North of the Equator, and vice versa.

$$\text{pixel\_offset} = \text{image\_offset} + (\text{dem\_offset} * \text{resolution}) \quad (1)$$

Calculate X spatial Coordinate:

$$x = X_{max} + (X_{pixel} * \text{pixel\_offset} - 1 + 0.5) * X_{mpd} \quad (2)$$

Calculate Y spatial Coordinate:

$$y = Y_{max} - (Y_{pixel} * \text{pixel\_offset} - 1 + 0.5) * Y_{mpd} \quad (3)$$

4.5 Extracting Elevation Data

Elevations were extracted from a N 60°and S 60°DEM Merge, originally computed by subtracting the lunar reference radius of 1,737,400m from the surface radius measurements. Thus elevation values are the distance above or below the reference sphere. This DEM is constructed from over 4.5 billion geodetically-accurate topographic heights from the LOLA, which were co-registered 43,200 stereo-derived DEMs (each 1°×1°) from the TC (1010 pixels total). [11].

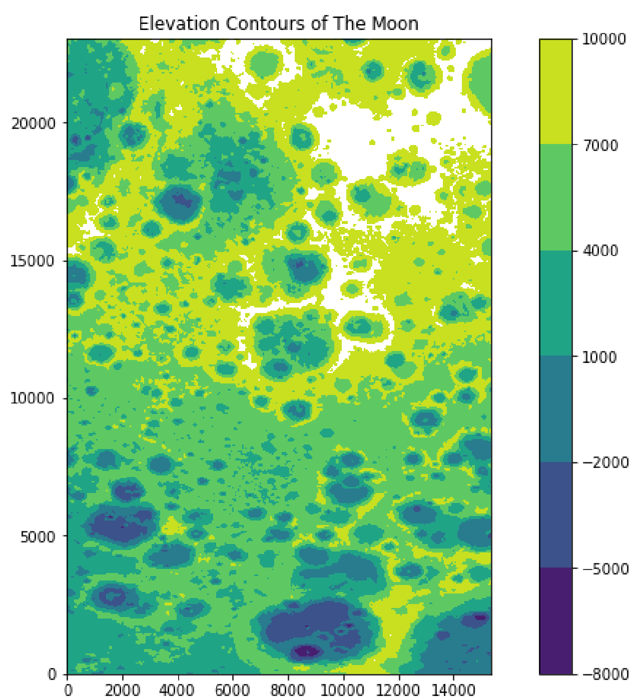

Figure 4: Elevation Contours of the Moon Fill, Lunar_LRO_DEMmerge_60N60S_512ppd

GDAL (gdal_translate) was used to extract XYZ values from the merged DEM, producing a 487 GB CSV file. A data set to correlate PyCDA detections to, allowing us to locate formations on the Lunar surface, and map the terrain.

GDAL Retile was also be used to split the merged DEM into identical proportions as the 48 individual DEMs, allowing the creation of 48 XYZ files, each containing a 1 to 1 comparison of the visible image tiles used for detection.

4.6 Parallelization

The pipeline to produce this work was serialized due to physical limitations of the computational resources available. The version of PIL used in PyCDA contains a bug where images are not unloaded from memory, limiting how many tiles can be processed before memory needs to be flushed. Based on observations during computation, it would take 64 - 128GB of RAM to process the entire data set without flushing memory. With more available RAM, such as on the Illinois State University High Performance Compute resource and Terabytes of RAM available, this process can be parallelized, significantly improving run time. Current run time was 12+ hours, on an HPC we would expect this to drop below an hour.

4.7 Output

Collectively, our results appear consistent with other similar research regarding crater detection and morphology analysis using Convolutional Neural Networks, showing fair agreement in the size distribution of crater detections.

Importantly, our results provide data that can be used for further research and analysis of crater formation morphology.

5 RESULTS

We report 35,346 crater radii detected in this work using PyCDA between N 0-30°and E 0-45°(approximately 1.8×10^6 km^2), and compare with the 19,337 craters with $R_{crater} \geq$ 5km reported in Wang et al. (2020). [3] The distribution of crater sizes is roughly a power-law as expected and shows fair agreement for crater sizes between 5 and 10 km.

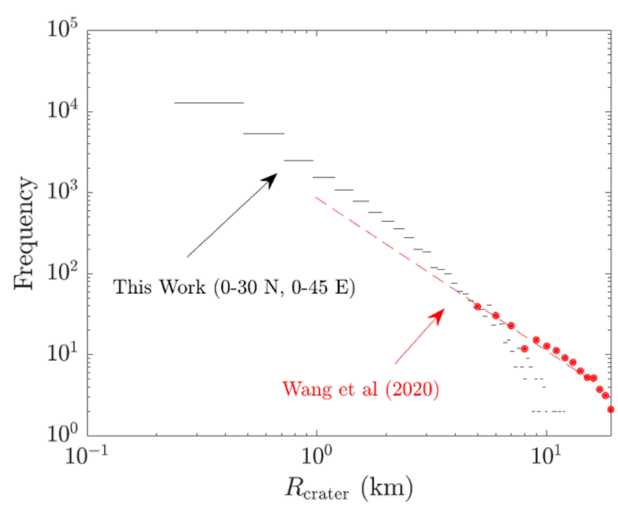

Figure 5: Detected Crater Distributions

While Wang et al. successfully identify craters of 5-20km in size, PyCDA was trained to detect, and performs best on small craters, with detections as small as 0.25 km radius (diameters of 8 pixels). While many of the smallest identified craters may be spurious, the use of DEMs to inspect ejecta blankets around the crater will aid in classifying the successful detections.

6 FUTURE WORK

In future work, developing a statistical model to find best fit slopes for the detected craters could prove valuable, improving the precision of the work output, validating sub kilometer craters and cataloging atypical formations as potential MACHO collisions using the elevation data presented here.

Applying different extractor methods, such as the watershed method could improve the pipeline, more accurately identifying intersecting craters as individual formations. PyCDA also allows the user to develop custom classifiers to be used by the model, allowing for further improvements.

It is possible, and quite likely to due to the size of a large number of the detections that some are missed as they sit on the border of a split tile. To improve the detection rate, the pipeline can be re-run with different sized, or offset tiles to capture the missing pixel values for analysis.

ACKNOWLEDGMENTS

Support for this work was provided by the School of Information Technology, and the Department of Physics at Illinois State University

REFERENCES

[1] Yalinewich, A., & Caplan, M.E. (2021, March 31). Crater morphology of Primordial Black Hole Impacts. Notices of the Royal Astronomical Society: Letters, Volume 505, Issue 1, July 2021, Pages L115–L119,

[2] Silburt, A., Ali-Dib, M., Zhu, C., Jackson, A., Valencia, D., Kissin, Y., Tamayo, D., & Menou, K. (2018, November 12). Lunar Crater identification via Deep Learning from https://arxiv.org/abs/1803.02192

[3] Wang, S., Fan, Z., Li, Z., Zhang, H., & Wei, C. (2020, August 20). An effective lunar crater recognition algorithm based on Convolutional Neural Network. MDPI from https://www.mdpi.com/2072-4292/12/17/2694

[4] NASA's Planetary Geologydessy Data Archive (PGDA), https://imbrium.mit.edu/

[5] NASA's Planetary Geology, Geophysics and Geochemistry Laboratory , https://astrogeology.usgs.gov/search/map/Moon/LRO/LOLA

[6] Benedix, G.K., Norman, C.J., Bland, P.A., Towner, M.C., Paxman, J., & Tan, T. (n.d.). Meeting Planning Services. 49th Lunar and Planetary Science Conference from https://www.hou.usra.edu/meetings/lpsc2018/

[7] 1998-2022 Frank Warmerdam, Even Rouault, et al, MIT License, https://gdal.org/

[8] Michael Klear, University of Colorado Boulder, MIT License, https://github.com/AlliedToasters/PyCDA

[9] Cohen, J.P, Lo, H.Z, Lu, T, & Ding, W. (2016, January 5). Crater detection via Convolutional Neural Networks from https://arxiv.org/abs/1601.00978v1

[10] Zhang, J. (2019, October 18). UNet line by line explanation. Medium from https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5

[11] Moon LRO Lola - Selene Kaguya TC Dem Merge 60n60s 59M V1, USGS Astrogeology Science Center. Moon LRO LOLA - SELENE Kaguya TC DEM Merge 60N60S 59m v1 | USGS Astrogeology Science Center. (n.d.) from https://astrogeology.usgs.gov/search/map/Moon/LRO/LOLA