Detection and Analysis of Potential Planetary Cave Entrances in Remote-sensing Data Using Machine and Deep Learning

Thesis submitted to the University of Kent for the degree of PhD in the subject of Physics

by Daniel Peter Patrick Le Corre

5th April 2025

Detection and Analysis of Potential Planetary Cave Entrances in Remote-sensing Data Using Machine and Deep Learning

by Daniel Peter Patrick Le Corre[1,2]

Supervisors:

Nigel Mason[1] Jeronimo Bernard-Salas[2] David Mary[3] Nick Cox[2]

[1] Centre for Astrophysics and Planetary Science (CAPS), **University of Kent**

[2] Centre d'Etudes et de Recherche de Grasse (CERGA), **ACRI-ST**

[3] Lagrange UMR 7293, Observatoire de la Côte d'Azur (OCA), **Université Côte d'Azur**

Acknowledgements

This PhD project was a collaboration between the University of Kent and ACRI-ST, co-funded in equal parts. I would like to first thank my primary supervisor from the University of Kent, Prof. Nigel Mason. His expertise, generosity of time, and encouragement to pursue various opportunities during this project were vital to its success.

ACRI-ST is an SME of the space sector with headquarters in Sophia-Antipolis, France. ACRI-ST, founded in 1989, provides services in Earth and Space Observation and data management. I am very grateful for the hospitality that I experienced from those at ACRI-ST during my work visits to the Centre d'Etudes et de Recherche de Grasse (CERGA)—an ACRI-ST establishment in Grasse, France. My supervisors from ACRI-ST, Dr Jeronimo Bernard-Salas and Dr Nick Cox, as well as my co-supervisor from the Observatoire de la Côte d'Azur in Nice, Prof. David Mary, have all been immensely helpful for their insights and approachability in spite of the physical distance between us for much of the project.

This project is also part of the Europlanet 2024 Research Infrastructure (RI) which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 871149. This PhD project was also closely linked to the EXPLORE project (https://explore-platform.eu)—a consortium of institutions aiming to promote the exploitation of space data—which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101004214. Much of the work in this thesis, such as model training and batch image processing, has been enabled by the Specialist and High Performance Computing systems provided by Information Services at the University of Kent.

To my incredible girlfriend, thank you for always believing in me and trying to make me do the same, even at my most stubborn. Lastly, I would like to pay tribute to my parents, as it was their dedication to providing me and my brother the most joyful childhood imaginable that has encouraged us in the pursuit of our dream careers.

Abbreviations and Acronyms

Below is an alphabetised summary of the abbreviations and acronyms commonly used throughout this thesis. The names of the operating space agency(ies) have been given in brackets alongside the abbreviations for any of their space missions or services. The same is true for the host spacecraft when one of its instruments has been abbreviated, except when the spacecraft's name is included within the abbreviation itself.

Abbreviation	Description
APC	Atypical pit crater
BAP	Bel'kovich A Pit
CaSSIS	Colour and Stereo Surface Imaging System (ExoMars TGO)
CDA	Crater detection algorithm
CDR	Calibrated Data Record
CE-1	Chang'e-1 (CNSA)
CE-2	Chang'e-2 (CNSA)
CNN	Convolutional neural network
CNSA	China National Space Administration
CTX	Context Camera (MRO)
DCNN	Deep convolutional neural network
DL	Deep learning
DEM	Digital elevation model
DTM	Digital terrain model
EDR	Experimental Data Record
EO	Earth Observation
ESA	European Space Agency
ExoMars TGO	ExoMars Trace Gas Orbiter (ESA)
FN	False negative
FP	False positive
FPN	Feature Pyramid Network
GDAL	Geospatial Data Abstraction Library
GRAIL	Gravity Recovery and Interior Laboratory (NASA)
HiRISE	High Resolution Imaging Science Experiment (MRO)
HRSC	High-Resolution Stereo Camera (MEX)
IMF	Impact melt fracture
IoU	Intersection over Union
IS	Instance segmentation
ISIS	Integrated Software for Imagers and Spectrometers
JAXA	Japan Aerospace Exploration Agency

JMARS	Java Mission-planning and Analysis for Remote Sensing
LOLA	Lunar Orbiter Laser Altimeter (LRO)
LRO	Lunar Reconnaissance Orbiter (NASA)
LROC NAC	Lunar Reconnaissance Orbiter Narrow Angle Camera
LROC WAC	Lunar Reconnaissance Orbiter Wide Angle Camera
LPA	Lunar Pit Atlas
MDIS-NAC	Mercury Dual Imaging System Narrow Angle Camera (MESSENGER)
MDIS-WAC	Mercury Dual Imaging System Wide Angle Camera (MESSENGER)
MGC^3	Mars Global Cave Candidate Catalog
MGS	Mars Global Surveyor (NASA)
ML	Machine learning
MESSENGER	Mercury Surface, Space Environment, Geochemistry, and Ranging
MEX	Mars Express (ESA)
MOC	Mars Orbiter Camera (MGS)
MOLA	Mars Orbiter Laser Altimeter (MGS)
MPO	Mercury Planetary Orbiter (BepiColombo)
MRO	Mars Reconnaissance Orbiter (NASA)
MTP	Mare Tranquillitatis Pit
OD	Object detection
PITS	Pit Topography from Shadows
PDS	Planetary Data System (NASA)
PSA	Planetary Science Archive (ESA)
PSR	Permanently shadowed region
QGIS	Quantum Geographic Information System
R-CNN	Region-based convolutional neural network
RDR	Reduced Data Record
RoI	Region of Interest
SC	Scene classification
SELENE	Selenological and Engineering Explorer, now known as 'Kaguya' (JAXA)
SIMBIO-SYS	Spectrometers and Imagers for MPO Integrated Observatory System
SMHP	South Marius Hills Pit
SPICE	Spacecraft, Planet, Instrument, C-matrix, and Events
SS	Semantic segmentation
TC	Terrain Camera (SELENE/Kaguya)
THEMIS	Thermal Emission Imaging System (Mars Odyssey)
TN	True negative
TP	True positive

Symbols

Below is a list of the symbols/abbreviations used throughout this thesis, along with a brief description and the equation, figure, or table in which they are defined.

Symbol	Description	Reference
h	Apparent depth	Eq. (4.18)
ω	Azimuth angle (spacecraft)	Fig. 4.7
ϕ	Azimuth angle (solar)	Fig. 4.7
D	Diameter	Fig. 4.10
ε	Emission angle	Fig. 4.9
F_1	F_1-score	Table 2.1
d_g	Ground distance	Fig. 4.9
γ	Ground phase angle	Fig. 4.9
α	Incidence angle	Fig. 4.10
IoU	Intersection over Union	Table 2.1
H	Maximum depth	Fig. 4.10
$\varepsilon_{\parallel}$	Parallel obliquity to Sun's line of sight	Eq. (4.12)
$\varepsilon_{\perp}$	Perpendicular obliquity to Sun's line of sight	Eq. (4.13)
P	Precision	Table 2.1
R	Recall	Table 2.1
d_h	Satellite altitude	Eq. (4.9)
S_{obs}	Shadow width (observed)	Fig. 4.7
S_{true}	Shadow width (true)	Eqs. (4.14) and (4.16)
S_{vis}	Shadow width (visible)	Eq. (4.15)
$\bar{s}_i$	Silhouette coefficient (average) of cluster i	Eq. (4.7)
d_s	Slant distance	Eq. (4.9)

Abstract

Machine and deep learning (ML/DL) techniques have been successfully applied to imagery across numerous fields, including Earth observation. However, prior to this thesis, the usage of ML/DL in the domain of remote-sensing data taken of the Moon and Mars was dominated by impact crater detection algorithms. This is in spite of the numerous other interesting and relevant features on these surfaces, which hadn't received the same treatment. This is addressed within this thesis through three separate projects that employ ML or DL techniques upon planetary surface features. The features in question are pits, skylights, and impact melt fractures, which are significant for astrobiological investigation and space exploration for their potential to be entrances to sub-surface cavities.

In the first project, the Pit Topography from Shadows (PITS) tool is described, which automatically derives apparent depth profiles of Lunar and Martian pits and skylights using only a single visual-band satellite image. PITS does this by utilising k-means clustering and silhouette analysis to detect shadow pixels (achieving precision and recall rates of 99.6 and 94.8% upon shadow-labelled HiRISE imagery) under a range of illumination conditions. The known image resolution and positions of the Sun and satellite are used to produce apparent depth measurements (corrected for non-nadir observations) along the entire length of the shadow. Thanks to the PITS tool, 10 Mars Global Cave Candidate Catalog features were discovered to exhibit signs of possible cave entrances in their depth profiles.

The second project trains Mask R-CNN (Region-based convolutional neural network) instance segmentation DL models to detect pits and skylights on the Moon, with the objective of bolstering the existing Lunar Pit Atlas. During testing, the best model (which was trained on Lunar, Martian, and synthetic Lunar data with a ResNet50 backbone) achieved F_1-scores (F_1) of 82.4 and 93.7% for the bounding boxes and predicted masks, respectively. Despite having only been applied to ≈1.9% of the total Lunar maria, this model—named ESSA (Entrances to Sub-Surface Areas)—has found two previously uncatalogued skylights on the Moon, with one situated ≈58.7° in latitude and the other being found in the previously mapped Marius Hills. ESSA has also been applied to the January 2025 release of HiRISE RDRV11 images.

The third project presents the results of training a DeepLabV3 semantic segmentation DL model to detect cooling fractures found within the melt deposits of impact craters on the Moon. In testing upon impact melt fractures (IMFs) found within the Copernicus and Virtanen F craters, the highest F_1 of 55.9% was achieved by the DeepLabV3 trained with a ResNet50 backbone and a batch size of 32—the combination of which is named IMFMapper. IMFMapper has been deployed to map the IMFs within the previously surveyed Crookes crater, where it found new candidate melt deposits within the crater's western and southern walls. Moreover, IMFMapper has produced the first mapping of any kind for Schomberger A crater, where IMFs may act as permanently shadowed regions due to its proximity to the Lunar south pole. This project also represents an avenue for future work due to the prospect of BepiColombo imaging IMFs on Mercury with greater detail.

Contents

List of Figures

List of Tables

Chapter 1. Thesis Scope and Objectives

The principal aim of the work reported in this thesis is to utilise machine and deep learning (ML/DL) to detect and analyse various surface features within planetary remote-sensing data, particularly of the Moon and Mars. This portfolio of research takes the form of three individual, self-contained projects with a common link that involves the automated detection or analysis of features which may allow access to the sub-surface environments of these two bodies.

The first project develops a tool called Pit Topography from Shadows (PITS), which employs a classical ML approach (k-means clustering) to automatically derive depth profiles and volume estimates of Lunar and Martian pits by measuring the shadows cast within them. By operating on as little as one image per pit, PITS aims to produce size estimates in scenarios where high-resolution elevation data is unavailable—within time frames unachievable without automation. The second project trains DL models, in the form of Mask region-based convolutional neural networks (R-CNNs), to detect Lunar pits and skylights in the context of searching for possible cave entrances to intact lava tubes. Knowledge of the locations of such features will be crucial for future space exploration, and PITS can give an indication of the accessibility of such features. Lastly, the use of DeepLabV3 semantic segmentation DL models for mapping cooling fractures within the melt deposits of impact craters on the Moon is presented. Impact melt fractures could provide another possible entrance to the Moon's sub-surface, where reserves of water ice may be sheltered from the Sun's illumination.

The aspiration of these three studies is that they improve upon existing planetary surface feature detection literature, as well as increase the number of planetary surface features which have received treatment from automated methods. In addition, the thesis should serve as an enabler and an incentive for future researchers to use automated techniques for their planetary

science research and vice versa for computer scientists searching for interesting ML/DL case studies.

The research projects within this thesis will be supported by a comprehensive introduction into the separate fields, such that readers with limited knowledge of planetary surfaces or experience with ML or DL techniques will be able to engage with the concepts discussed. Firstly, an overview of planetary surfaces will present the variety of features in separate sections according to their common genesis. Features are mostly grouped according to whether they are formed via interactions with other bodies, by historic or extant volcanism/tectonics, through the action of liquid on the surface, or due to meteorological activity within the host body's atmosphere. The formation mechanism(s), morphology, and relevance of each surface feature are described, with part of the reason being that they were factors which helped with deciding which ones to consider for this project. A second element of the introduction will outline the ML/DL principles, challenges and methods that are relevant for a project such as this, which involves the classification of imagery. This will start from simple concepts such as the need for separate, high-quality training, validation and test datasets and discuss more technical considerations such as current state-of-the-art model architectures.

After considering a number of factors, it was decided that the project would prioritise the use of existing architectures instead of designing entirely new algorithms from scratch. These well-established methods, possibly in similar fields such as Earth Observation (EO), can then be employed to help solve the challenge of detecting surface features on bodies much further afield. This strategy also offers the benefit that such models, which have proven effective in an EO context, will observe significantly less variation in the features present when faced with planetary surfaces. The aspiration is that this will lead to the high transferability of these existing models to this new domain. Furthermore, whilst there is certainly sufficient planetary data for automated techniques to be trained, evaluated, tested and inferred upon, there exists currently no 'benchmarking' datasets for models tasked with planetary surface feature detection. Benchmarking is the process of testing the performance of novel models upon a sufficiently large common dataset, allowing for direct comparison between their abilities

to recognise features and generalise across new data. For example, one such example of this is the Microsoft Common Objects in Context (COCO) dataset (Lin et al., 2014a), which contains object detection and semantic segmentation labels, as well as image captions, for nearly 200 target classes within over 200,000 images. No such dataset exists which is both of a comparable size and contains a similar variety of planetary surface features, and the time and consideration required to produce one could constitute an entire PhD thesis on its own.

Nonetheless, knowing which surface features to investigate with these existing models is another choice to be made. As will become clearer after Chapter 3, this project also aims to prioritise those which have not yet been investigated automatically—whether in detection or analysis. Thanks in part to their prevalence in the Solar System, but also the availability of large catalogues, the majority of the prior research was in the development of impact crater detection algorithms (CDAs, see Section 3.1). The current CDA research leaves only marginal room for improvement in terms of accuracy, with developments largely being in the minimum crater diameters that they can detect. Moreover, there is sufficient satellite data and several catalogues for non-impact-related surface features, thus leaving the door open for the creation of labelled training datasets for use with ML/DL. Despite the role of impact craters in dating geological surfaces, there are surface features, such as caves, which may play a greater role in future space exploration and be of more astrobiological significance.

Ultimately, the result of this work will be a set of tools and frameworks for automatically processing, detecting and analysing planetary surface features within satellite data. These tools should be able to reach reasonable, near-human level accuracies in order to justify the use of automated methods. However, current algorithms will not necessarily be able to learn the same level of context that an experienced planetary scientist will be able to employ. For example, a planetary scientist may use knowledge of the location on the host body or the position relative to other features of interest to inform their classification of the type of given surface feature. Whereas, awareness of these factors is not built into current models trained to make the same classifications. Therefore, the models should be able to achieve results on time-scales that are much shorter than can be completed by any manual method. Any

detections made by these tools should also be used for further analysis or investigation, where possible. This is as opposed to simply making detections for the sake of proving that the model is effective. Lastly, these tools will be made publicly available, simple to understand and easy to access, ensuring that future researchers do not have to repeat the same steps or fall into the same pitfalls.

These objectives can be summarised below:

1. To utilise ML/DL methods for the automatic detection or analysis of planetary surface features within satellite data.

2. To introduce researchers from the planetary or computer science communities to the other field and consider it for future projects or collaboration.

3. To prove the efficacy of existing architectures in a planetary remote-sensing context.

4. To investigate planetary surface features which have not received as much prior treatment from automated methods.

5. To make the tools, frameworks and datasets publicly available to foster future work and avoid repeating mistakes.

Chapter 2. Introduction

2.1 Overview of Planetary Surfaces

Understanding the types, distributions and formation of planetary surfaces can help us uncover the regional and global histories of their host planets, geologically- and meteorologically-speaking. On any given rocky surface in the Solar System, one can expect to see a range of geological landforms and surface expressions. The particular assortment of these features which are observed will depend on which geological processes were active at the time they formed. The propensity for these processes to leave their mark on the surface may also be driven, controlled, or even hampered by other external factors, such as the strength of the host body's gravitational field, whether it possesses a significant atmosphere, or even the wider space environment that it finds itself within. The morphologies of these features as they appear today will also be shaped by any such activity since their formation.

In this particular section of Chapter 2, the most significant surface features and landforms present throughout the Solar System will be outlined—along with their sizes, locations, morphologies, and formation mechanics. The relevance of these features in terms of subsequent geological investigations, future robotic/crewed space exploration, or astrobiological significance will also be discussed. Only a selection of the total population of surface features has been chosen for discussion in this work, since there are already several reviews and books which are able to give far more extensive descriptions of a more exhaustive list of features than would be within the scope of this thesis. For example, one such resource that has been referred to repeatedly during this project is the Encyclopedia of Planetary Landforms (Hargitai and Kereszturi, 2015). It gives a compilation of expert reviews of the overwhelming majority

of surface features and landforms that are found on the terrestrial (including Earth) and icy bodies within the Solar System.

2.1.1 Impact Structures

Perhaps the most instantly recognisable features found on the surfaces of rocky planets are impact structures, such as basins and craters. For example, if you were to ask an individual with little to no experience with planetary geology to draw a planet, they would almost certainly draw a sphere—either surrounded by rings or littered with craters. This is likely due in part to their prominence across the Solar System, even on bodies with thick atmospheres such as Earth. They are also identifiable by their circular-to-elliptical appearance, which is also a possible explanation for why the term 'crater' is very often used to describe any round depression, regardless of whether it was formed through impact. As such, this thesis will avoid using 'crater' as a broad term and will instead specify the origin when referring to circular depressions (such as impact craters or pit craters).

Impact structures are created when a planetary projectile, such as an asteroid or cometary nucleus, impacts the surface at hyper-velocity speeds (Hargitai et al., 2021). At the moment the impactor and target meet, a shock wave is generated that compresses both objects simultaneously. Once the compression wave has travelled through the entirety of the impactor, a rarefaction wave travels behind that of the compression to release the target material from its compacted state. The extreme pressures of impact events can lead to the vaporisation of the target material and the impactor itself. As the kinetic energy of the impact dissipates with distance from the event, the shocked material is excavated and projected radially up and away. The resulting pattern of ejected material (known as an ejecta ray or blanket) can be used to identify young impact events which have not yet undergone significant degradation. Some of the ejected material may also fall back in towards the maximum excavated depression known as the transient crater/cavity.

There are a number of factors, such as the impactor's relative velocity and size, that will dictate the diameter and morphology of the resulting impact structure. The impactor

velocity will range somewhere between the escape velocity of the target body and the sum of the target body's orbital velocity plus the escape velocity of the Solar System at the target body's distance from the Sun (r). By assuming the target body's orbit is circular and that $M_{sat} << M_p << M_\odot$, using the formulas for the escape velocity $\left(\sqrt{2GM/r}\right)$ and the orbital velocity $\left(\sqrt{GM/r}\right)$ it can be shown that an impactor could be travelling at approximately 11–72 km/s for the Earth-Moon system, 5–59 km/s for Mars, and 4–114 km/s for Mercury. In terms of impactor size, there is no minimum for bodies with no significant atmosphere, such as the Moon and Mercury, since all sizes of impactors will reach the surface. However, in the case of Mars, its present-day atmospheric pressure of 6.36 mbar means that the smallest impactor to reach the surface would be an iron meteoroid around 0.01 m in width, which leads to a minimum impact structure diameter of 0.3 m (Popova et al., 2003).

The result of smaller impacts is often a bowl-shaped depression which maintains the approximate shape of the transient crater—these are known as simple impact craters and are shown in Figure 2.1. It should be noted that all maps shown within this thesis are assumed to have north as directly up unless otherwise stated. Simple impact craters will have diameters of a few kilometres, while their depth-to-diameter ratios will be roughly 0.3, with these values varying according to the strength of the material being impacted and the gravitational attraction of the host body (McSween, Jr et al., 2019). As the diameters of impact craters increase, one will begin to observe characteristics that are indicative of complex impact craters. As can be seen in Figure 2.1, these are impact craters that have flat floors (often filled with impact-melted material); a central uplift in the form of a central peak, several peaks enclosing a central depression, or peak rings; and terraced crater rims (McSween, Jr et al., 2019). Figure 11.4 of McSween, Jr et al. (2019) plots the diameter range at which one can expect to see a transition from simple to complex impact craters on the Moon, Mercury, Mars and Earth with their different gravities.

While there is currently no fixed definition for what constitutes an impact 'basin', the term was proposed by Hartmann and Kuiper (1962) as is often used to describe impact structures with sufficiently large diameters ($\gtrsim$300 km on the Moon) which exhibit one or more

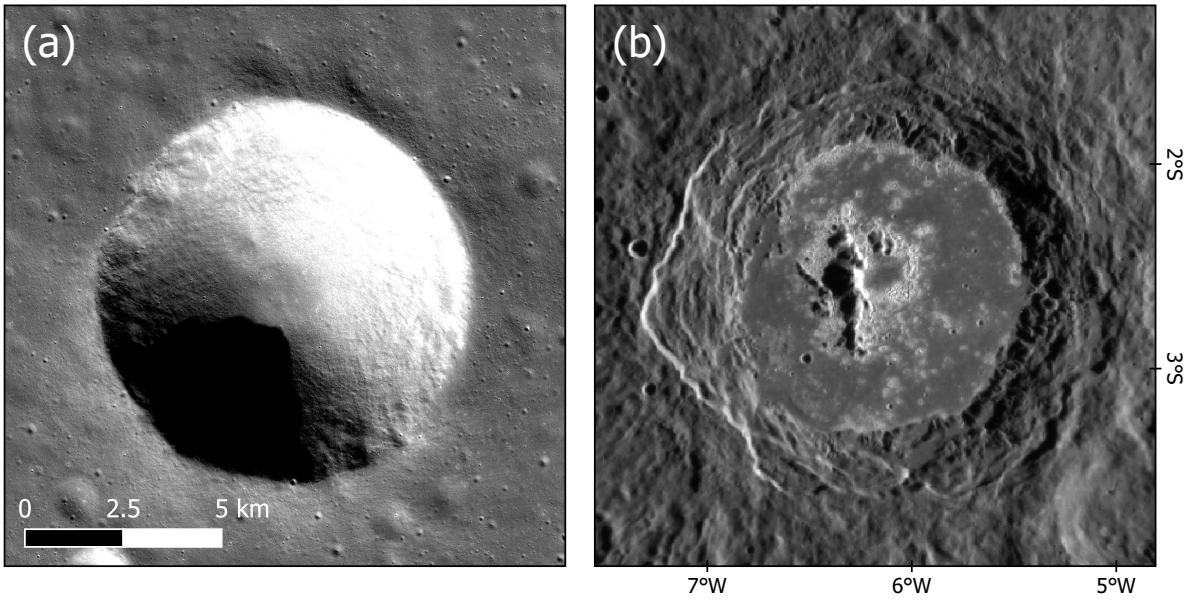

Figure 2.1: Examples of simple and complex impact craters on rocky planetary surfaces. (a) An unnamed simple impact crater on the Moon (at 59.236°N, 25.329°E) in the Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) feature mosaic NAC_ROI_MareFrigLoB_E583N0258 of Mare Frigoris. (b) The complex impact crater 'Warhol' on Mercury (at 2.55°S, 6.27°W) as seen in the MESSENGER (Mercury Surface, Space Environment, Geochemistry, and Ranging) Mercury Dual Imaging System (MDIS; Hawkins et al. 2007) 256 pixels per degree (ppd) global base-map.

inner ring features. Figure 2.2 displays examples of impact basins found on the Moon (Mare Orientale), Mars (Argyre basin), and Mercury (Rembrandt basin). Due to the increased role of the gravitational attraction of the host body at the scales of basin-forming events, it is uncertain what occurs after the transient crater has been formed in order to form these ring structures. Melosh (1989) suggest that they are formed by the fracturing of a thin and/or weak lithosphere, whereas Hydro-code simulations by Collins et al. (2002) show that ring features could be produced by the collapse of the transient crater (inward from the basin rim and outward from the central uplift).

There are several other subtypes of impact crater which, as with simple and complex, relate to the impact event itself, as well as to weathering from other surface processes. Firstly, if an impactor is incident upon the target at an oblique angle, it can lead to an elliptical impact crater (i.e. a noticeable difference between the crater's major and minor axes). According to surveys and models, elliptical impact craters constitute ≈5% of the population on Mars, with

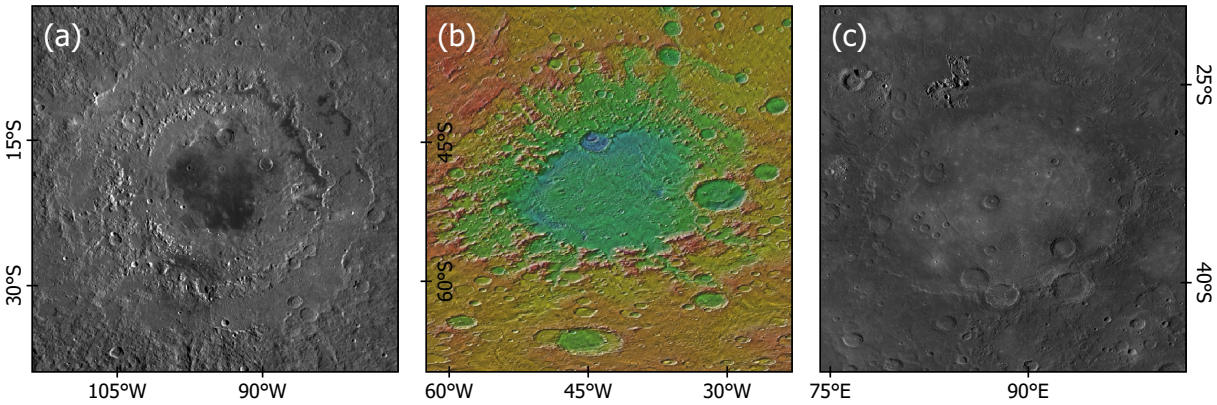

Figure 2.2: Examples of impact basins on rocky planetary surfaces. (a) Mare Orientale basin on the Moon (at 19.9°S, 94.7°W) in the Lunar Reconnaissance Orbiter Wide Angle Camera (LROC WAC; Robinson et al. 2010) 100 m/px global mosaic (Speyerer et al., 2011). (b) Argyre Planitia on Mars (at 49.8°S, 43.3°W) as seen in the Mars Odyssey Thermal Emission Imaging System (THEMIS; Christensen et al. 2004) Daytime IR mosaic colourised with Mars Global Surveyor (MGS) Mars Orbiter Laser Altimeter (MOLA; Smith et al. 2001) elevation (Hill et al., 2014). (c) Rembrandt basin on Mercury (at 32.8°S, 87.5°E) as seen in the MESSENGER MDIS 256 ppd global base-map (Hawkins et al., 2007).

similar proportions for the Moon and Venus (Bottke et al., 2000). Bottke et al. (2000) also derived a threshold for an impactor's incidence angle (θ_{ET}) above the surface of $\theta_{ET} \leq 12°$ for an impact crater to have an ellipticity above 1.1.

In situations where a sufficiently large meteoroid fragments into two or more pieces before impact, a linear chain of impact craters can be formed. Such a chain is often called a catena and is usually the result of the impactor being separated by tidal forces on approach (Richardson and Hagen, 2014). Impact crater chains can also be made up of secondary impacts, which is where debris from a primary impact event has fallen back to the surface with enough kinetic energy to cause another impact event.

Impact craters which have undergone severe degradation to the point of being only recognisable by their rim, even if incomplete, are known as ghost craters (De Hon, 2021). Impact craters can become 'ghosts' when they themselves have been hit by other impacts or the resulting ejecta has blanketed the pre-existing crater. Burial under lava flows or sediment from flowing water can also lead to the creation of ghost craters.

Aside from processes on the surface, impact craters can also be modified by ongoing mechanisms underground. More than 300 Lunar impact craters exhibit annular, radial, and

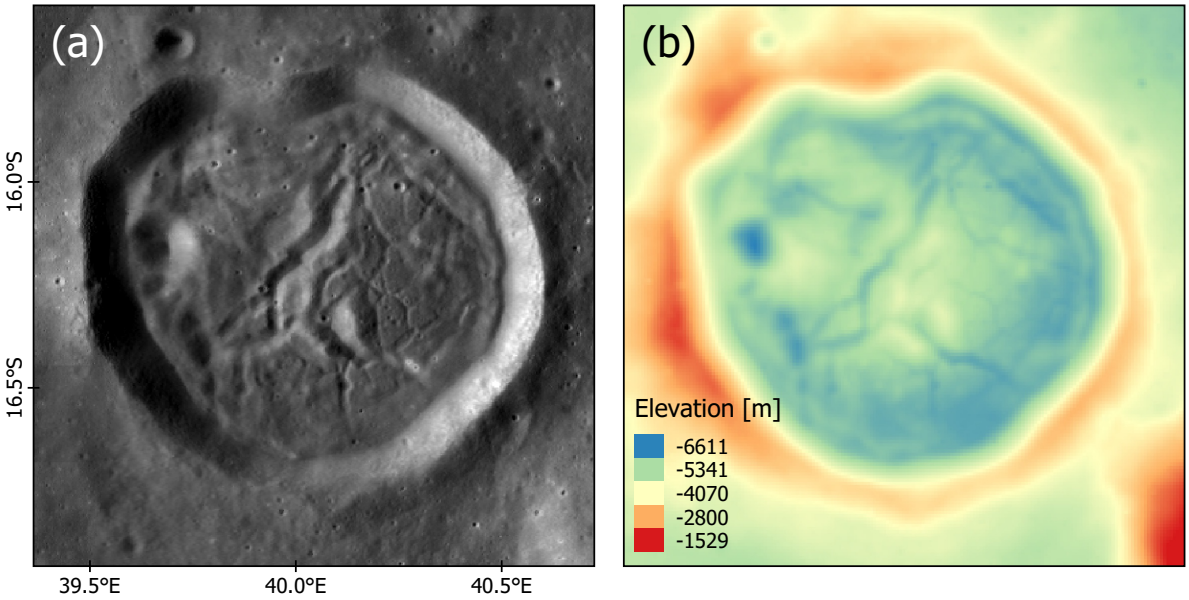

Figure 2.3: An example of a floor-fractured crater (FFC) on the Moon. (a) The Lunar crater 'Bohnenberger' (at 16.24°S, 40.06°E) within the LROC WAC global mosaic (Speyerer et al., 2011). (b) Bohnenberger crater shown in the LOLA/Kaguya DEM (Barker et al., 2016). This crater is also shown in Figure 4 of Jozwiak et al. (2012).

polygonal patterns of linear fractures caused when the crater floor uplifts due to the intrusion of lava from nearby magma chambers in the Lunar mare (Korteniemi et al., 2010; Korteniemi, 2021). Figure 2.3 presents an example of a floor-fractured crater (FFC) on the Moon. Similar FFCs are observed on Mars, which may have collapsed due to the existence of a volatile-rich crust at the time when the impacts were formed (Korteniemi, 2021).

On Mars, some impact craters between 0.5–5 km in diameter are situated within approximately circular plateaus, which are raised ≈20–100 m above the surrounding surface (Hargitai, 2021b). These pedestal impact craters are formed when the more erosion-resistant ejecta lands around the cavity and shields the finer-grained surrounding surface. The parts of the surface not covered by ejecta are then subjected to sublimation or aeolian erosion, which causes it to sink relative to the pedestal that the crater sits upon (Hargitai, 2021b).

Impact craters are one of the most significant surface features for a number of reasons. Firstly, impacts between planetesimals have been critical for the formation of our Solar System and the bodies within. An excellent example of this is the Moon, which is believed to have been formed via a colossal impact with Earth (Cameron and Ward, 1976; Halliday, 2000).

Furthermore, this is one aspect in which impact events have had an effect on the development of life on Earth, with the Moon's tidal forces causing transitional areas between purely ocean and land habitats. However, impacts have also been detrimental to organisms on Earth. The prime example is the Cretaceous-Palaeogene (K-Pg) extinction event, which saw ≈70% of all terrestrial life on Earth lost after the Chicxulub impactor struck what is now the Yucatán peninsula in Mexico (Alvarez et al., 1980).

There is also a present-day relevance to impact craters for planetary geologists, thanks to 'crater counting'. Crater counting is the practice of identifying impact craters within remote-sensing data to determine their distribution across a planetary surface. For long periods, this has been a solely manual task in which labellers would produce databases full of the positions and diameters of impact craters. Such crater databases have been produced for the Moon (Head et al., 2010; Povilaitis et al., 2018; Robbins, 2019), Mars (Robbins and Hynek, 2012a,b; Francis et al., 2020), Mercury (Fassett et al., 2011; Herrick et al., 2018), Venus (Herrick et al., 1997), and many of the other smaller bodies (see the 'Databases' section of Hargitai 2021b). However, automated crater detection algorithms have significantly reduced manual labelling times and led to detections of more small-diameter craters, as will be discussed in Section 3.1. The main purpose of crater counting is that knowledge of the distribution of craters of different sizes (called the crater size-frequency distribution, CSFD) can be used to derive surface ages of geological units. Surface ages have been estimated using crater counting on several bodies, including the Moon and Mars (Hartmann and Neukum, 2001; Neukum et al., 2001), by using known ages of rock samples that have reached Earth as constraints.

2.1.2 Consequences of Volcanism or Tectonics

The most obvious sign of past or present volcanic activity is the existence of volcanoes, which are not solely found on Earth. Venus has the most volcanoes of any Solar System body with ≈85,000 edifices (Hahn and Byrne, 2023), whilst Mars is host to the largest volcano in the Solar System, called Olympus Mons (shown in Figure 2.4), with a width and height of ≈550 and ≈21 km, respectively (McSween, Jr et al., 2019). However, the only body apart from

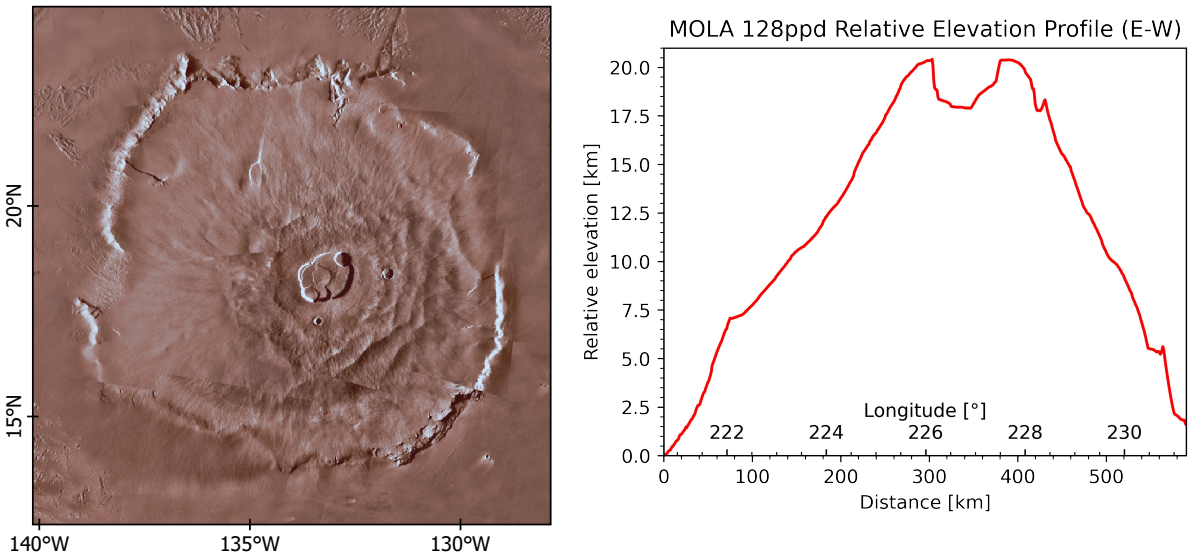

Figure 2.4: (Left) Shield volcano 'Olympus Mons' on Mars (at 18.65°N, 133.80°W), as seen in the Mars Viking Orbiter 232 m global colour mosaic[1]. (Right) Elevation profile taken longitudinally from east to west through the caldera at its centre using MOLA data at 128 ppd (Som et al., 2008). The distance and elevation axes are not to scale since Olympus Mons' diameter is approximately 20 to 30 times greater than its peak height.

Earth on which current volcanic eruptions have been observed is Jupiter's innermost moon, Io (Williams et al., 2011).

Several other signs of recent or historic volcanism can be observed in the vicinity of volcanoes, but also in the volcanic plains of bodies without volcanoes (Jagert, 2021), such as on the Moon (Head, 1976). The Lunar maria are low albedo regions found primarily on the Moon's nearside, which have been formed by multiple low-viscosity basaltic lava flows. Within these maria, one can find sinuous rilles, such as Hadley Rille, which was visited by Apollo 15. Rilles are thought to be ancient lava channels or even the remnants of the total collapse of an underground lava tube and are mostly found at the boundaries of Lunar maria (Hurwitz et al., 2013). Similar channel features are found on Mars, albeit it is difficult to deduce by remote observations alone whether they are a result of flowing lava or liquid water (Hargitai, 2021a). For more details on rilles and lava tubes on the Moon and Mars, see Chapter 5.

As mentioned previously, some surface features etched by volcanic or tectonic activity are easily misconstrued as impact craters. For example, calderas are approximately circular depres-

[1] The Mars Viking colourised global mosaic can be downloaded from the USGS Astropedia service at https://astrogeology.usgs.gov/search/map/mars_viking_colorized_global_mosaic_232m.

sions which can be formed during large volcanic explosions or via gravitational collapse into an evacuated near-surface magma chamber (Galgana, 2021). On Earth, volcanic depressions are known as calderas when they are greater than 1 km in diameter, whereas the threshold for other planetary bodies is roughly ten times that (Mouginis-Mark, 2001), partially due to lower gravity and more voluminous volcanoes. If the diameter of the depression does not meet this condition, then it is known as a pit (Mouginis-Mark, 2001). Pits are circular-to-elliptical depressions which are often found aligned into pit chains, leading to further collapses to form linear and sinuous chains or troughs (van der Bogert et al., 2014). Pits and pit chains have been observed on Earth (Okubo and Martel, 1998), Mars (Wyrick et al., 2004), Moon (Wagner and Robinson, 2014); Mercury (Gillis-Davis et al., 2009); and Venus (Davey et al., 2013). The proposed formation mechanisms for pits are either the gravitational collapse of a sub-surface cavity, subsidence due to sublimation, volcanic explosions, or tectonic faulting (van der Bogert et al., 2014). Pits which exhibit overhanging walls, which may act as an entrance to a sub-surface cavity, are also called skylights. Planetary caves and the features which allow access to them will be relevant for future space exploration as sources of natural shelter (Boston et al., 2004) and potentially water ice on Mars (Williams et al., 2010). Figure 2.5 gives examples of some of the pits, skylights, and pit chains found across the Solar System. Pits and skylights will be discussed in more depth in Chapters 4 and 5.

2.1.3 Results of Meteorology

On planetary bodies which harbour an atmosphere, meteorological processes can cause the formation of new surface features, as opposed to simply degrading existing ones. For example, Martian dust devils can leave tracks more than 2 km in length, which appear as thin linear or swirling albedo features, as they move across the surface, kicking up dust (Balme et al., 2003).

Large regions of a planet's surface can show unique indications of geological processes, not just isolated features. For example, several planetary bodies such as Mars (McCauley, 1973), Venus (Greeley et al., 1995) and Titan (Lorenz et al., 2006) are observed to have dune

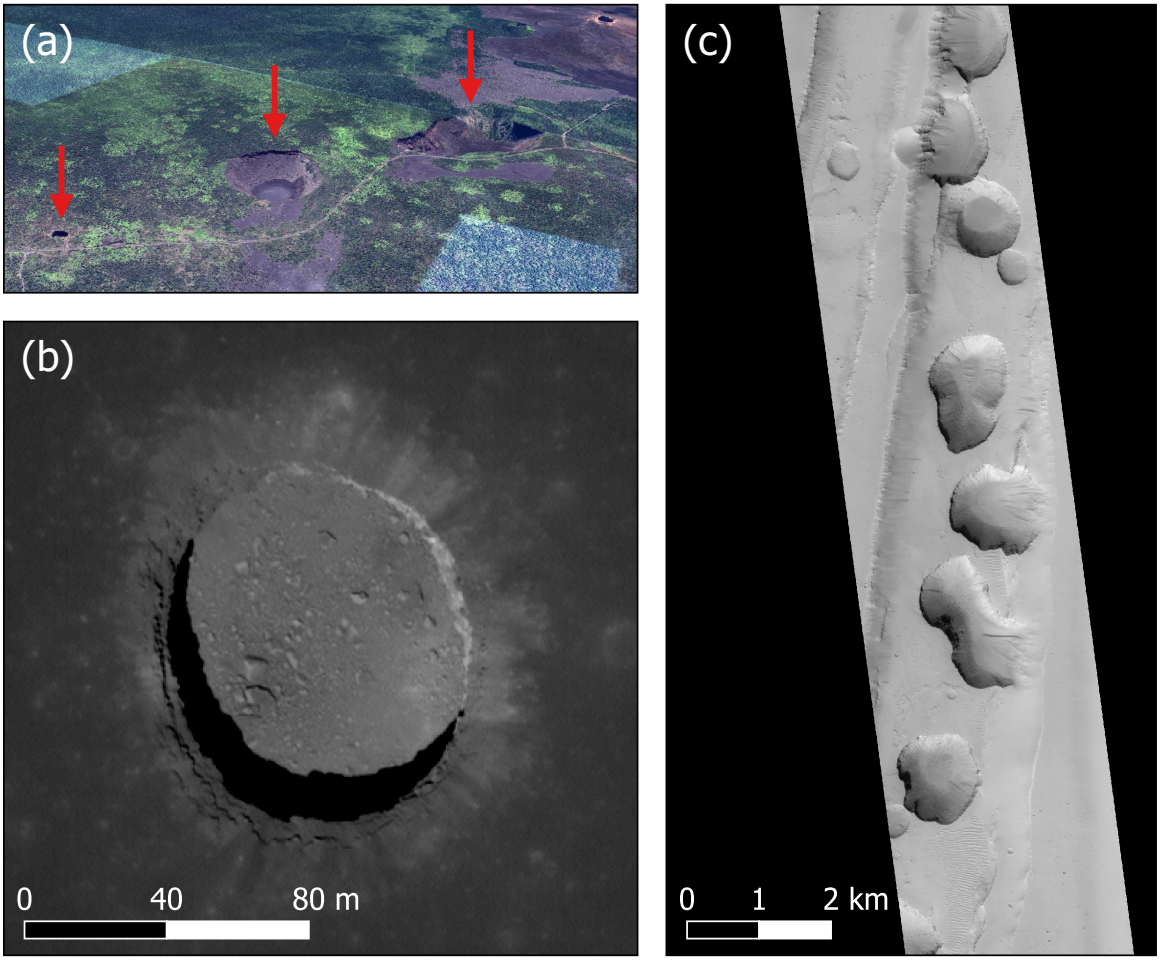

Figure 2.5: Examples of pits and pit chains across the Solar System. (a) Oblique view in Google Earth of the 'Hiiaka', 'Pauahi' and 'Devil's Throat' pits adjacent to the Kilauea volcano on the island of Hawai'i (labelled with red arrows). (b) LROC NAC image M155016845R of the famous Lunar skylight in Mare Tranquillitatis. (c) MGS Mars Orbiter Camera (MOC; Malin et al. 1998) narrow-angle image R1300250 of a Martian pit chain situated within a graben in the Tractus Fossae region and in line with several normal faults, suggesting tectonic origins (Wyrick et al., 2004).

systems (otherwise known as aeolian bedforms). Dune systems consist of patterns of linear ridges which have been formed by the delivery of loose, eroded material by aeolian forces. Dunes grow in size when the wind that was previously strong enough to carry the material periodically dies down, thus depositing the sediment in place (Cardinale et al., 2021).

2.1.4 Signs of Current or Previous Hydrology

Mars has been synonymous with the presence of previously flowing liquid ever since its dried-up channels and oceans were first mapped at the beginning of the twentieth century. A large ocean may have once occupied the lowlands of Mars' northern hemisphere (Baker et al., 1991). The elevations and geologic origin of the contacts along the perimeter of the smooth northern plains, which could be construed as a shoreline, do not support the existence of an ancient ocean (Carr and Head, 2003). Nonetheless, it is evident that smaller bodies of standing water, such as in-filled impact craters, were present on the surface of ancient Mars (Michalski et al., 2022).

As Mars' once-thicker atmosphere weakened over time, so did its ability to protect bodies of surface water from solar radiation. Consequently, the clearest evidence for previous fluvial activity comes in the form of valley networks (Alemanno et al., 2018) and deltas (De Toffoli et al., 2021). These features can often be difficult to separate from other linear yet non-hydrological features, such as grabens or lava channels. However, one identifying characteristic would be that they sometimes contain features that have been streamlined by hydrological action or deposition (Erkeling and Carling, 2021).

Whilst there are no large bodies of water on Mars today, there is still the possibility of some liquid on the surface in the form of gullies. Gullies are linear, 'v-shaped' depressions in the sides of steep slopes, which are possibly formed through erosion by flowing liquid or by seasonal frost sublimation (Heldmann et al., 2014).

2.2 Introduction to Machine and Deep Learning

2.2.1 Comparison Between Machine and Deep Learning

At its simplest, machine learning (ML) is the act of training a computer algorithm to learn patterns in examples of data such that it can detect the same patterns in new data that it has never 'seen' before. One of the main benefits of these ML algorithms (also called models) is

that they can detect patterns in large amounts of data that would take humans inconceivably long time-scales to process. This is, of course, assuming that the data on which the model was trained to learn these patterns (called the training data or dataset) is of sufficient quality and abundance. This is a drawback of ML in that it takes far more data for an algorithm to learn such patterns than it would for a human.

Training data must also be representative of the 'unseen' data that the model will be applied to after it has been trained. For example, an ML model trained on data of one specific breed of dog may not be able to detect all dogs in future instances. This process of applying a trained ML model to new data that it was not trained upon is often called 'prediction' (or 'classifying' in the case of classification algorithms; see Section 2.2.2).

ML and deep learning (DL) have become increasingly prevalent in analysing remote-sensing data in Earth observation (Persello et al., 2022), as well as other planetary systems. DL is a subset of ML whereby there is no need to provide information to a model on what patterns make up the target feature(s), unlike standard ML models. DL models are based on neural networks, which learn by transforming some input data through layers consisting of weights. These weights are updated as the network learns the patterns in the input data. They are called 'neural' networks because they mimic the way in which neurons in animal brains activate according to certain stimuli.

For example, a DL model tasked with detecting craters in satellite images should learn on its own what patterns (such as a circular shape and raised rims) distinguish them from other surface features. As mentioned at the start of Chapter 2, there is equal promise in applying ML and DL techniques in planetary science (Azari et al., 2021). Therefore, Sections 2.2.2 to 2.2.8 will explain some of the foundational terminology before the planetary surface feature detection literature can be reviewed in Chapter 3.

2.2.2 Classification Versus Regression

ML/DL algorithms are typically tasked with performing one of two tasks: classification or regression. A model that predicts what target classes some data belongs to is performing

classification. Whereas, one that is predicting a numerical value is performing regression. Confusion between the two can be common since classification models will often predict a probability value that a data point belongs to a certain class. In addition, logistic regression, which predicts the probability of a binary event occurring, is a popular regression method whose output can have a threshold applied to it in order to produce a classification.

This review will focus mainly on classification scenarios and methods since planetary surface feature detection operates by satellite images being classified based on what surface features are present in them.

2.2.3 Supervised Learning Methods

Before a model can begin to learn, the dataset must be split into three individual sets: train, validation, and test. The way in which ML/DL models learn is by repeatedly classifying the model on the data in the training set. The same is also done for the validation set in order to determine the model's performance on previously unseen data. The model is then iteratively updated so as to minimise a 'loss function' which measures how incorrect the model's predictions were during training. A fully trained model can then be applied to the data in the test set to visualise or further assess the detections. It is vital that the data in these three sets are all independent.

The learning processes can largely be broken down into three main categories: supervised, unsupervised and semi/weakly supervised learning. For a model to learn in a supervised manner, it needs to be given information on what the target features that require detection are. These sources of information are commonly called labels, or in the case of remote-sensing data, ground truth. The loss function is often a direct comparison between the label and the model's training output under this scenario. Therefore, once it has finished training, a supervised model will output a probability that the data belongs to each of the target classes.

Popular non-neural network supervised ML algorithms are logistic regression (as explained in Section 2.2.2), Support Vector Machines (SVMs) and decision trees. An SVM is a deterministic algorithm which takes data points that are labelled as belonging to one of two classes

and transforms them to a higher-dimensional space. SVMs attempt to maximise the gap between the two classes in this higher-dimensional space such that new data can be mapped there and classified into one of the two classes (Cortes and Vapnik, 1995). Whereas, decision trees are a non-parametric predictive model for splitting data into separate classes based on certain conditions or rules (Lindholm et al., 2022). They are trained by recursively splitting the training dataset until each division has the same value as the labels or until splitting yields no improvement. Multiple decision trees can be combined to form a 'forest'. The random forest classifier operates by randomly selecting which nodes in the trees should be split and then classifying the data by the class which was predicted by the most trees. This reduces the chance that every tree splits in the same fashion, which leads to the same suboptimal result in the predictions of the decision trees (Lindholm et al., 2022).

The most relevant forms of neural networks for feature recognition in imagery are the multi-layer perceptron (MLP) and deep convolutional neural networks (DCNNs). Both neural networks contain an input layer (where the training data is fed to), one or more hidden layers (where the learning takes place) and output layers that contain the model's predictions (Lindholm et al., 2022). DCNNs will have several of these hidden layers, hence the name. However, where MLPs and DCNNs differ in the context of image segmentation is that MLPs take an image as a flattened vector as an input, whereas the input for a DCNN can be an n-dimensional matrix instead (Lindholm et al., 2022). As a result, DCNNs can learn the spatial distribution of the training data. Convolutional neural networks (CNNs) get their name from the use of convolutional layers to extract the feature maps containing responses based on what patterns are present, such as edges, curves, and corners.

2.2.4 Unsupervised Learning Methods

Unsupervised learning is where these training labels are not provided to the model, and instead, it is trained to predict the probability that some data belongs to N number of classes (where N can be known or unknown depending on the method). Since there are no labels, the loss function must be defined by other means, typically by clustering or reconstructing the original

inputs.

Clustering algorithms work by separating data points into groups of data points called clusters. The clusters that are found by these algorithms will be those that have converged to a potentially local, but hopefully global, minimum of some loss function or those that were found after the maximum number of iterations. Clustering algorithms do not need labelled data since the loss functions they use are typically descriptions of the distribution of the data. For example, the k-means clustering algorithm seeks to find the k number of clusters (where k is known) that minimises the squared Euclidean distances of each data point from the centre of their assigned cluster (Lloyd, 1982). k-means will play a central role in the automated shadow extraction element in the Pit Topography from Shadows (PITS) algorithm presented in Chapter 4, where it will also be explained to a greater degree. Other popular clustering algorithms include Gaussian Mixture Model, DBSCAN, Agglomerative Hierarchy and Mean-Shift.

Auto-encoders are a type of neural network which learn a compressed version of the training data without the need for it to be labelled (Hinton and Zemel, 1993). These networks are divided into two main elements, the encoder and decoder. The former takes an input image and compresses it over a series of hidden layers into a lower-dimensional space. The decoder is then used to upscale this encoded data in the latent space to the original dimensions of the input data. Auto-encoders train by considering the difference between the input data and the reconstructed data after encoding and decoding.

Generative adversarial networks (GANs) use a decoder as a generator network and an encoder as a discriminator network for producing new, but artificial data (Goodfellow et al., 2014). The generator aims to produce a convincing fake example of what is drawn from the same distribution as the unlabelled data. Whereas the discriminator must be able to detect if the sample is fake or not. While this is a generative network and not a classifier, a fully trained GAN can store important information about the training dataset in the latent space. Simple vector analysis can be performed on these latent variables in order to produce new samples with certain features added/removed (Radford et al., 2015).

2.2.5 Other Forms of Learning

Semi-supervised learning is, in some ways, a combination of supervised and unsupervised learning, in that it can provide training labels to a model but does not do so constantly throughout the training process. This has a significant advantage in cases where the size of the training data outweighs the number of labels. Weakly-supervised learning is a learning process that provides a fully labelled training dataset, but it is one which is very low in detail. Due to the nature of these learning approaches falling somewhere between supervised and unsupervised, many of the algorithms described in Sections 2.2.3 and 2.2.4 can be used for semi- and weakly supervised learning.

Active learning is an approach which can help in scenarios where manual labelling is particularly laborious. Take two datasets where one is fully labelled, and one is not. Active learning is where an algorithm can suggest to the user to label new data from the unlabelled dataset while it is being trained on the labelled one (Settles, 2012). A specific piece of data is suggested for labelling if the algorithm estimates that it would improve its training/validation metrics. Therefore, active learning helps to ensure that only the data that would improve the model's performance is labelled, resulting in a more efficient training dataset generation. However, the method for deciding which data is selected for labelling can ultimately lead to different results since some data may be highlighted under one strategy but not another.

Despite the name, transfer learning does not describe how labels are fed to an ML/DL model during training. Instead, it relates to the act of taking an already-trained model, potentially on some independent related data, and re-training it from its current state on a new training dataset (Lindholm et al., 2022). The benefit here is that less training data (labelled or unlabelled) needs to be provided to a model in order to achieve the same performance as if it had been trained from scratch. In situations where the new training dataset is significantly smaller than the original, this process is often also called fine-tuning. If the data supplied to the model in the different trainings are related, such as two datasets of impact craters on different bodies, then this could help the model to better generalise on previously unseen data.

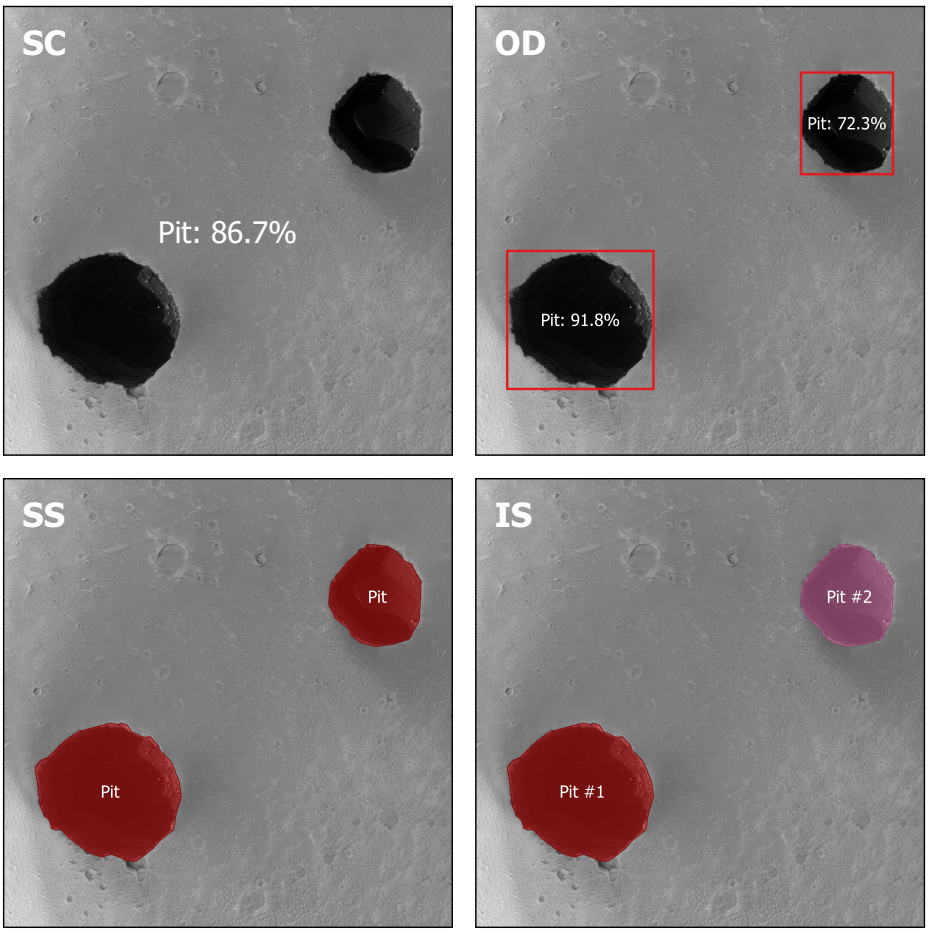

Figure 2.6: A hypothesised illustration of the various classification methods upon the Mars Reconnaissance Orbiter (MRO) High Resolution Imaging Science Experiment (HiRISE; McEwen et al. 2007) image ESP_019997_1975_RED containing two Martian pits from the Mars Global Cave Candidate Catalog (MGC^3; Cushing et al. 2015). For reference, SC: scene classification; OD: object detection; SS: semantic segmentation; IS: instance segmentation.

2.2.6 Classification Methods

Focusing on feature recognition in imagery, there are three different paths which define how ML and DL models are trained and how they perform detections. These concepts are known as scene classification, object detection and semantic/instance segmentation, and have been displayed visually in Figure 2.6 for an example of some Martian pits. Sections 2.2.6.1 to 2.2.6.3 will explain each approach to image classification, along with some of the common architectures which have been used in planetary surface feature detection literature.

2.2.6.1 Scene Classification

Scene classification is where an ML or DL model is trained upon imagery which has been labelled according to what target feature(s) or scenario(s) are present. However, the training labels for each image will not say where these targets are found within the image, only that they are present. Therefore, when the trained model is applied to new images it has never 'seen' before, it will give a prediction of what target class is present, but no information on its location. This is shown in (a) of Figure 2.6, whereby the hypothetical SC model only predicts that one or more pits are present within the image with a high confidence score.

Aside from the SVMs and decision trees discussed in Section 2.2.3, there is a plurality of CNNs which can perform SC upon imagery. AlexNet represents one of the earliest uses of CNNs for SC upon imagery (Krizhevsky et al., 2012), having been developed as part of the 'ImageNet Large Scale Visual Recognition Challenge'. As such, AlexNet contains only five convolutional layers.

VGG (Virtual Geometry Group) develops upon the AlexNet architecture by training more convolutional layers to improve performance, whilst also leading to increased training times (Simonyan and Zisserman, 2014). VGG is available in two sizes, VGG16 and VGG19, which contain 8 and 11 more convolutional layers, respectively. VGG models also reduce the kernel size (i.e. the window) of convolution operations compared to AlexNet in order to limit this increased model complexity.

ResNet (residual neural network) models are known for their use of residual blocks and skip connections (He et al., 2016). The output of a residual block is the sum of a convolutional layer's output and its original input, which allows the model to learn a series of residual functions and eases the training of deeper models. This development was also made to try to rectify the vanishing gradient problem. This is where a large number of successive layers will lead to the gradient of the loss function becoming increasingly small, meaning that there is very little scope to update the model's weights and improve performance. A 'skip connection' is the name given to the addition of the input to the residual block to its output—hence 'skipping' the convolutional layers. As is commonplace in many SC CNNs, ResNet models are

available with different quantities of layers (i.e. depths).

Huang et al. (2016) take the residual blocks popularised by ResNet and develop them to create 'dense blocks'. Dense blocks operate on a similar principle to residual blocks in that they allow output from previous layers to skip current ones. However, where they differ is that each layer within a dense block is the concatenation of the feature maps learnt from all previous layers. This improves the flow of information and, hence, the gradient of the loss function through the model since the feature maps from previous layers are not lost in the summation operations in ResNet models.

The various versions of Inception CNN models attempt to solve the issue of target features having different scales within a given image by feeding the input through several parallel convolutional layers with a range of kernel sizes (Szegedy et al., 2014). Since this addition of more layers significantly increases the number of parameters which require learning, later versions of Inception apply dimension-reducing convolution operations before these branches with different scales of convolution operations.

MobileNet is another CNN commonly used for image recognition (Sandler et al., 2018), particularly on devices with limited computational resources (hence the 'Mobile'). This is thanks to the depth- and point-wise convolution used within its architecture. This is where convolutional layers are applied to each channel of an input (i.e. each image band) individually and then recombined in order to reduce the number of model parameters.

Since the CNNs mentioned thus far all produce feature maps for the target classes found within an image, they are commonly used as 'backbones' within other DCNNs for performing other tasks, such as object detection or semantic segmentation. A model's backbone can often be an ambiguous description, but it is generally the element which is responsible for feature extraction.

In order to solve the problem of detecting objects with varying sizes, model backbones often implement a feature pyramid network (FPN). FPNs leverage the multi-scale nature of a CNN (wherein image resolution decreases but the semantic value increases as successive convolution operations are applied) to output feature maps for all layers (Lin et al., 2016).

2.2.6.2 Object Detection

The task of performing object detection (OD) is very different to SC. Firstly, the training labels that are required must include the location of the target feature in the form of a 'bounding box'. A bounding box is essentially the vertical and horizontal extents (or the minimum and maximum latitude-longitude coordinates) of the feature. This means that when an OD model is being applied to new data, it will output the class of the feature, the confidence score of the detection, and the bounding box of its predicted location. This is displayed in (b) of Figure 2.6.

A popular series of DCNNs that can perform OD within imagery is region-based convolutional neural networks (R-CNNs). R-CNNs operate by proposing several regions of interest that are likely to contain target features and then feeding them through a neural network, which aims to detect any such features should they be present (Girshick et al., 2013). Where previous R-CNN iterations used selective search to generate these region proposals (Girshick et al., 2013; Girshick, 2015), the Faster R-CNN implements a region proposal network (RPN), which is a fully convolutional network (FCN) that is trained to generate region bounds and their likelihood to contain target features Ren et al. (2015). Another iteration of the R-CNN series is Mask R-CNN (He et al., 2017). However, this particular DCNN will be explained further in Chapter 5.

Where R-CNNs operate in a two-stage OD framework (stage 1: region proposal, stage 2: detection), You Only Look Once (YOLO) OD DCNNs work without the need for proposals (Redmon et al., 2015). YOLO works by training a traditional DCNN architecture and splitting the images into a grid in order to perform SC on each cell. A fully connected layer then predicts the class and extent for several bounding boxes for the detected object. The YOLO network then learns by finding which of those box proposals achieved the highest Intersection over Union (IoU, see Table 2.1) with the corresponding training label. Non-maximum suppression (NMS) is also an important element of the YOLO architecture, which is responsible for the removal of low-scoring overlapping detections. NMS is explained in more detail in Chapter 5.

2.2.6.3 Semantic and Instance Segmentation

Models that practice semantic segmentation (SS) require each pixel in the input image to be labelled with a particular class, including whether it is a background pixel. Then, when an SS model is used to detect features in new images, it will output a probability map which contains the likelihood of each pixel belonging to each target class.

Regularly used methods for performing SS on imagery are clustering algorithms, as described in Section 2.2.4, but also the U-Net and DeepLab DCNNs. The former is named as such for the 'U' shape typically seen in algorithm diagrams due to its use of an encoder-decoder network (Ronneberger et al., 2015). The encoder network of the U-Net architecture reduces the spatial dimension and increases the number of feature maps, while the decoder performs the reverse upon the spatially reduced, high semantic value state of the data. Bridge connections between the encoder and decoder networks allow for the flow of gradients between the two, as the feature maps from the former are concatenated to the latter. DeepLab architectures are another form of DCNN used for SS, which, instead of using an encoder-decoder system, implements Spatial Pyramid Pooling to classify regions of pixels at a range of scales (Chen et al., 2016). Chapter 6 will discuss DeepLab in more detail.

In addition to SS, instance segmentation (IS) is where each pixel is classified according to which target class it belongs to, but they are also given an instance classification based on how many of these features are present in an image. Incidentally, one of the most common IS models is the Mask R-CNN mentioned in Section 2.2.6.2 and discussed in Chapter 5.

2.2.7 Hyper-Parameters

When developing DCNNs, as Chapters 5 and 6 later will, there are several parameters which require some user decision or input prior to beginning the training process. These parameters, often referred to as hyper-parameters, will influence how the chosen model will train.

The batch size of a given training dataset is defined as the number of training samples, or images in the case of this thesis, which have been packaged together to be fed through the model at one time. The purpose of training on batches is to limit the effect of a single

training sample on how the model is learning. Therefore, a greater batch size reduces the risk of a model over-fitting to the training data. 'Over-fitting' is the term often used to describe a model which is performing significantly better upon the training data than on data it has not seen before. When this occurs, there is often said to be a 'generalisation gap'.

The number of epochs that a model will be trained over is defined as the number of times that all batches have repeatedly been fed through the model. Therefore, training over a larger number of epochs will take longer and, depending on the size of the training data, could also be a cause of over-fitting.

The optimiser is the algorithm which decides how the weights of the model should be updated after each batch is passed through in order to minimise the given loss function. Popular loss functions and optimisers include Cross-Entropy Loss (Mao et al., 2023) and Adam (Kingma and Ba, 2017), or derivatives of Adam, such as AdamW (Loshchilov and Hutter, 2017).

As suggested by the name, the learning rate regulates the degree to which the optimiser should be allowed to change the model's weights (i.e. how much it is allowed to learn). A well-chosen learning rate will mean that the weights are not updated too drastically or too little. The former may mean that the model's performance during training will vary significantly, whereas the latter may mean that more epochs are required for the model to converge. Both instances could mean that the optimiser gets stuck within a local minimum for the loss function.

2.2.8 Performance Metrics

Numerous metrics have been devised by the ML/DL community for assessing a given model's performance during the training, validation and testing phases of its development. These performance scores are essential for comparing the results of one algorithm to another when applied to the same (or at least relevant) data. Observing how a particular performance metric changes from model training through to testing can also be highly informative. For example, if the metrics calculated during a model's training suggest a high degree of accuracy but return significantly poorer values when that same model is applied to unseen data during testing,

then the model may have 'over-fit' or become 'specialised' to the training data. This means that the model has effectively 'memorised' the training data instead of learning the patterns that constitute the target classes.

Table 2.1 provides a list of performance metrics that are commonly used in training ML/DL models for planetary surface feature detection. These metrics are derived by comparing the model's prediction with the labelled ground truth. The majority of the metrics in Table 2.1 depend upon the numbers of true positives (TPs), false positives (FPs), true negatives (TN) and false negatives (FNs), which are often collectively known as the 'confusion matrix'. A confusion matrix details where the model has correctly predicted (TP), incorrectly predicted (FP), correctly omitted (TN) or incorrectly omitted (FN) a target class.

Metric Name	Definition	Formula
Precision (P)	The proportion of predictions that were correct	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}$
False Discovery Rate (FDR)	The proportion of predictions that were incorrect	$\frac{\mathrm{FP}}{\mathrm{TP}+\mathrm{FP}}$
Recall (R)	The proportion of relevant instances that were predicted	$\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}$
False Negative Rate (FDR)	The proportion of relevant instances that were omitted	$\frac{\mathrm{FN}}{\mathrm{TP}+\mathrm{FN}}$
True Negative Rate (TNR)	The proportion of irrelevant instances that were omitted	$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}$
False Positive Rate (FPR)	The proportion of irrelevant instances that were predicted	$\frac{\mathrm{FP}}{\mathrm{TN}+\mathrm{FP}}$
Neg. Predictive Value (NPV)	The proportion of omissions that were correct	$\frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FN}}$
False Omission Rate (FOR)	The proportion of omissions that were incorrect	$\frac{\mathrm{FN}}{\mathrm{TN}+\mathrm{FN}}$
Mean Average Precision (mAP)	The unweighted average of the Average Precision (AP) over all target classes	$\frac{1}{N}\sum_{i=1}^{i=N}\mathrm{AP}_i$
Accuracy	The ratio between correct and total predictions/omissions	$\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}$
F_1-score	The harmonic mean of the precision and recall	$2\times\frac{\mathrm{P}\times\mathrm{R}}{\mathrm{P}+\mathrm{R}}$
Intersection over Union (IoU)	The overlap ratio between the areas of a prediction and the target	$\frac{\mathrm{Area}(X\cap Y)}{\mathrm{Area}(X\cup Y)}$

Table 2.1: List of performance metrics commonly used to assess ML models in planetary surface feature detection literature. The majority of these metrics are derived from confusion matrices, which consist of the true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN).

Chapter 3. Planetary Surface Feature Detection with Machine and Deep Learning - A State-of-the-Art

In this chapter, the existing literature that uses machine and/or deep learning (ML/DL) for the purpose of automatically identifying planetary surface features and landforms will be reviewed. Whilst the use of ML and DL is prevalent in other areas of planetary science (see Nagle-McNaughton et al. 2022 for ML/DL on Mars), the focus will be on the domain of planetary remote-sensing data. This can include traditional visual imagery, such as that taken by the Mars Reconnaissance Orbiter (MRO) High Resolution Imaging Science Experiment (HiRISE; McEwen et al. 2007); digital elevation/terrain models (DEMs/DTMs), like that acquired by the Lunar Reconnaissance Orbiter (LRO) Lunar Orbiter Laser Altimeter (LOLA; Smith et al. 2010); as well as infrared imagery, with the daytime infrared images retrieved by the Mars Odyssey Thermal Emission Imaging System (THEMIS; Christensen et al. 2004) being an example. There has also been work on using DL to enhance the resolution of planetary remote-sensing data (Handmer, 2021; Sun and Chen, 2022) and rover imagery (Wang et al., 2021) through the process of super-sampling. The principle of change detection has also been applied through the means of DL to determine how certain planetary surface features change in a temporal dimension (Kerner et al., 2019). These are uniquely interesting uses of DL in planetary science, yet such works are not of sufficient relevance and will therefore not be discussed in further detail.

This review will predominantly include research articles and in-proceedings that have been

published before or shortly after the commencement of the post-graduate study leading towards this thesis (September 2021). However, there has been a body of research with a direct relevance to the thesis topic (i.e. surface features that may act as entrances to sub-surface environments) that has been published since then. Newer research relating to sub-surface entrances will be described in the context of the work in this thesis, either in this state-of-the-art or in the relevant project chapter.

This chapter will be broken down into three sections according to the surface features with which the literature is concerned. The sections are titled: 'Crater Detection Algorithms (CDAs)' (Section 3.1), as the majority of the literature involves automatic crater counting; 'Other Feature Detection' (Section 3.2), since there are comparatively few efforts to detect other surface features; and lastly 'Automated Geologic Mapping' (Section 3.3), which discusses the research which aims to segment entire regions or surfaces according to their geologic context or geomorphological appearance. At the end of this chapter in Section 3.5, there are three tables which summarise the research efforts that have been reviewed in each of the respective sections. Tables 3.1 to 3.3 list the data sources used, the ML/DL method applied, the performance during testing, and any other notable remarks. Precision (P) and recall (R) rates are quoted throughout this chapter in order to better compare different studies where an inconsistent range of performance metrics has been utilised. As such, P and R have been derived from other metrics where they have not been explicitly given, yet perhaps an analogue metric or a confusion matrix of true and false positives/negatives (TP, FP, TN, FN) were made available.

3.1 Crater Detection Algorithms (CDAs)

As the spatial resolutions and coverage of satellite imagery and elevation data have improved, they have revealed a plurality of impact events. With global datasets that resolve craters on kilometre scales, entire planetary surfaces are now open to crater cataloguing. Meanwhile, the abundance of impact craters in the solar system scales exponentially with decreasing

diameters, meaning that identifying smaller craters becomes increasingly cumbersome to do manually. Therefore, several automated crater detection algorithms (CDAs) which use ML and DL techniques have been devised to reduce the burden of these two factors while still aiming to achieve human-level accuracy.

The study by DeLatte et al. (2019b) reviews research efforts implementing convolutional neural networks (CNNs) for the sole purpose of automated crater detection and counting—such is the heavy dominance of CDAs in the field of planetary surface feature detection. While previous reviews had discussed the use of simple ML methods in crater counting (Stepinski et al., 2012; Patil and Kini, 2015), this was the first to consider the efficacy of CNNs. DeLatte et al. (2019b) included research papers, in-proceedings and conference abstracts which were written in English, with the majority having been published between 2015 and 2019. DeLatte et al. (2019b) also focused on the most significant advances, particularly in the context of using pre-trained networks and reducing crater candidate post-processing times. To avoid repeating the same work, the following is a summary of the research considered for review by DeLatte et al. (2019b), which only applies to the Moon and Mars.

A specific improvement noted by DeLatte et al. (2019b) in the development of CDAs had been in the use of pre-trained CNNs. For example, Norman et al. (2018) trained a GoogLeNet-OverFeat object detection (OD) CNN, which had been pre-trained on the ImageNet dataset (Deng et al., 2009), to detect craters in visible-band THEMIS images. Meanwhile, Emami et al. (2018) compares the ability of pre-trained (also on ImageNet) VGG, GoogLeNet and ResNet models to make classifications of the presence of impact craters in LRO imagery. semantic segmentation (SS) has also been utilised for crater detection with DeLatte et al. (2019a) and Silburt et al. (2019) training U-Net-based CNNs (Ronneberger et al., 2015) for this purpose. The former was trained using THEMIS DIR images between $\pm 30^{\circ}$ latitude on Mars and annotations of craters between 2–32 km in diameter, resulting in a chosen model (denoted E-4) with a P and R of 87.42 and 66.72%, respectively (DeLatte et al., 2019a). Whereas, the latter used a customised U-Net architecture (called DeepMoon) and trained it on the LOLA/Kaguya co-registered DEM (Barker et al., 2016), achieving a P and R of 56

and 92%, respectively. This same trained model was then inferred on the global DEM of Mercury taken by the MESSENGER (Mercury Surface, Space Environment, Geochemistry, and Ranging; Solomon et al. 2001) spacecraft (Becker et al., 2016), showing promising visual results (Silburt et al., 2019).

Since the review by DeLatte et al. (2019b), advances have been made in the quality of CDAs. The following sections will attempt to separate the research discussing recent CDA advances into subsections depending on whether they focus on the surfaces of the Moon or Mars. However, as several CDAs have been trained on or inferred on data taken of both bodies, a judgment was made as to which section they should primarily fall in. There have also been several CDAs published in the time since this literature review was completed in early 2022, and when the general projects in this thesis had already been decided. For this reason, they have not been discussed in Sections 3.1.1 and 3.1.2, but the following may be of interest as further information for the reader. Moon: Tewari et al. (2022); La Grassa et al. (2023) and see Chaini and Jha (2024); Mars: Zhao and Ye (2024); Charon: Ali-Dib (2022).

3.1.1 Mars Crater Detection Algorithms

Following on from research reviewed by DeLatte et al. (2019b), (Lee, 2019; Lee and Hogan, 2021) proposed a CDA which consisted of three stages: separate SS detection of Martian craters; confidence score calculation by classification with an encoder network; and finally, removal of false positives by a Gradient Boosting classifier. This approach was designed to operate on the THEMIS DIR mosaic (Edwards et al., 2011b,a) and a DEM blended between Mars Global Surveyor (MGS) Mars Orbiter Laser Altimeter (MOLA; Smith et al. 2001) and Mars Express (MEX) High/Super Resolution Stereo Colour imager (HRSC; Jaumann et al. 2007) data (Fergason et al., 2018). A Res-U-Net architecture was used in the initial SS detection step, which is a form of U-Net with the implementation of skip connections to improve training and classification times (Zhang et al., 2018). Compared to a Martian crater catalogue (Salamunićcar et al., 2012), this CDA returned a P, R and F_1-score (F_1) of 87, 80, 84%, respectively (see Section 2.2.8 for a definition of F_1). In addition, every one of these

metrics was higher for the THEMIS DIR detections compared to those of the MOLA-HRSC blended DEM. Since one of the defining characteristics of impact craters is their depth profiles, it is interesting that the highest performing CDA in this scenario was the one trained upon infrared imagery rather than elevation data. This could be a result of the different resolutions of the datasets: 100 and 200 m/px for the THEMIS DIR images and MOLA/HRSC DEM, respectively (Edwards et al., 2011b; Fergason et al., 2018).

Automatic crater counts have also been used to calculate geological surface ages, with the first occasion being the work by Benedix et al. (2020). In their study, a You Only Look Once (YOLOv3; Redmon et al. 2015) OD CNN was trained to detect craters within images taken from the THEMIS DIR mosaic (Edwards et al., 2011b) between $\pm 35^{\circ}$ latitude. During validation on a subset of the training data, the CDA achieved a high performance with a P and R of 90% and 94%, respectively. Moreover, when the model was applied to the THEMIS mosaic between $\pm 65^{\circ}$, it detected 83% of all craters (of all diameters) in the Mars Crater Database (Robbins and Hynek, 2012a). The same collaboration from Benedix et al. (2020) then fine-tuned their model upon MRO HiRISE imagery (McEwen et al., 2007) to extract sub-kilometre Martian craters (Lagain et al., 2021b). The model was then applied to imagery of 10 named Amazonian or Hesperian-era impact craters (with diameters over 20 km) to find smaller craters in their ejecta blankets. By performing a tailored clustering analysis, they were able to remove clusters of secondary craters from the automatic crater counts since they would lead to overestimates in surface ages. This led to more accurate surface ages of the 10 named craters, which agreed with literature estimates, than if secondary craters were not removed. This same analysis was later used to predict that the Tooting and 09-00015 craters from the Tharsis region were the most likely sources of depleted shergottites, which reached Earth after ejection from Mars 1.1 million years ago (Lagain et al., 2021a), representing a significant step in the power of DL in planetary science.

A multi-layer perceptron (MLP) has been trained upon calculated mean and Gaussian curvatures of Martian craters from MOLA DEMs when they are represented as a surface mesh (Christoff et al., 2020). This is a breakaway from the typical procedure of training on just

imagery or elevation data. This change in domain returned a P between 81.1 and 92.1%, and an R ranging from 76.3–96.1% across three different test sites when compared to the Mars Crater Database by Robbins and Hynek (2012a). P and R were not used by Christoff et al. (2020), yet the confusion matrix that they provided meant they could be calculated. These three sites were used by previous work, which used an AdaBoost classifier on a series of feature descriptors (Di et al., 2014). Christoff et al. (2020) observed that their model had improved in terms of R across all regions, yet their other metrics were showing a higher degree of FPs being retrieved.

A scale-aware CDA has been devised that feeds information on the size of the target crater to the algorithm in a bid to improve the adaptability of CDAs across different crater diameters on Mars (Hsu et al., 2021). This is done by extracting the crater sizes from THEMIS images using the Circular Hough transform—an image processing technique for extracting circular features. The result of this is then fed to a feature pyramid network (FPN; Lin et al. 2016), which is running parallel to another FPN that is trained on the original input image. These feature maps learnt by the FPNs are then used to train the region proposal network (RPN) within a region-based CNN (R-CNN; Girshick et al. 2013). With a detection confidence threshold of 50%, this proposed three-step scale-aware CDA yielded a higher P (83.50%), R (86.62%) and mean Average Precision (mAP) compared to using a regular Faster R-CNN (Ren et al., 2015).

3.1.2 Moon Crater Detection Algorithms

In a similar fashion to DeLatte et al. (2019b), Chaini and Jha (2024) have conducted a review of CDA research which utilises DL for the automated detection of impact craters, this time just for the Moon. This review was not published at the time when CDA literature was being surveyed in the scope of this project. Therefore, there will be some overlap between the papers reviewed in this section and those in the review paper by Chaini and Jha (2024). Although Chaini and Jha (2024) serves as crucial further reading on the state-of-the-art CDA research on the Moon.

Hashimoto and Mori (2019) presented an early example of DL being applied in a CDA for the Moon. They use a YOLOv3 network structure to detect impact craters in Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) imagery of the Moon's south pole. However, they segment images into grids to be able to deal with the issue of having multiple craters per input image. This returned an R of 80.5% and a P between 70.2 and 77.3% depending on the varying confidence of the human-made labels.

Unsupervised methods have also been applied for automatically generating Lunar crater candidates for supervised classification (Emami et al., 2019). This process, dubbed hypothesis generation (HG), was conducted using a range of image-processing techniques to negate the need for classifying entire satellite images. The techniques used for HG were convex grouping, Hough transform, highlight-shadow region detection, and interest points. The classification of these crater candidates, called hypothesis variation (HV), was performed using three different CNN architectures with varying input image sizes and the number of fully connected layers. This two-step HG-HV algorithm was trained separately using two different labelled image sets: randomly selected image patches and HG-suggested image patches from LROC NAC data. These two trained algorithms were then applied to two test datasets of the same format as the training ones. Applying the four HG and three HV methods found that using interest points generally returned the preferred combination of P (92% or higher) and R (over 81%) across the two test sets. Meanwhile, the preferred CNN for HV was based on the LeNet architecture but used three fully connected layers and an input sample size of 32×32 pixels.

Active learning has also been used for detecting craters in satellite imagery and elevation data of the Lunar and Martian surfaces (Wang and Wu, 2019). An AdaBoost classifier is trained on Harr-like features extracted from a small subset of positive samples of Lunar and Martian craters. An AdaBoost (Adaptive Boosting) algorithm is made up of several weak classifiers that have each been trained to correct the errors of the one previous, in order to form one robust classifier (Freund and Schapire, 1996). Several of these classifiers are stacked in series in a process called 'cascading' (Viola and Jones, 2001) to reduce the processing cost of obviously cratered or non-cratered images. Where active learning comes in is that these

classifiers are initially trained on a small subset, and then they are classified on the full set of unlabelled positive and negative crater samples. The positive classifications are then verified by a Support Vector Machine (SVM) based on the crater profiles in the corresponding DEM. Correct detections, according to the SVM, are labelled as positive crater samples and added to the training set, and similarly, wrong detections are added to the negative crater training set. This cascading AdaBoost classifier was repeatedly re-trained until the performance metrics did not improve by more than 1%. This approach, after 6 rounds of active learning input, yielded a True Detection Rate (TDR), which is equivalent to R, of 94.01 and 92.80% for the Moon and Mars, respectively. Whereas the False Detection Rate (FDR), which is equivalent to $1 - \mathrm{P}$, was 11.09% for the Moon and 4.72% for Mars.

A certain degree of weakly supervised learning has also been employed for Lunar crater detection by automatically generating pixel-level crater labels from latitude-longitude coordinates and diameters from a known crater catalogue (Ali-Dib et al., 2020). This dataset generation follows a similar process as that for the DeepMoon model (Silburt et al., 2019), except that the labels produced are binary masks made of disks rather than rings. This is partly due to the requirement of pixel-level labels for training a Mask R-CNN model (He et al., 2017) to detect Lunar craters. The predicted masks outputted by this form of R-CNN allowed the proposed CDA not only to detect the location of Lunar craters but also to extract their shapes and ellipticity. The P, R, and F_1 of this trained Mask R-CNN were all noticeably lower than DeepMoon (after removing repeat observations of craters). However, the lower P could be due to the identification of even more new craters that were missed in the crater catalogue (Head et al., 2010; Povilaitis et al., 2018) used for testing.

DL has also been used for automatic age estimation, as well as crater identification on the Moon (Yang et al., 2020). In this work, a pre-trained (on ImageNet) region-based fully convolutional network (R-FCN; Dai et al. 2016) with a ResNet101 backbone was re-trained to detect impact craters within data taken by the Chang'e-1 (CE-1; Zuo et al. 2021) satellite. Digital ortho-mosaics (DOMs) were fused with slope and curvature information derived from DEM data to provide morphological and topographical context to the model. In addition to

this, a set of 12 deep CNNs (including ResNet101) was trained on the DOM, while a simple feed-forward network (FFN) was trained on crater attribute information such as diameter and depth. The traditional CDA element, when tested on CE-1 data, achieved an R of 94.71%. The trained CDA was then directly inferred on fused data from the Chang'e-2 (CE-2; Zuo et al. 2021) mission, reaching a comparable value of 93.35% in terms of R. After merging the CE-1/2 detections and removing duplicates, it was found that the False Positive Rate (FPR) when manually inspecting a subsample was 95.51% for craters with $1 < D < 100$ km and 95.33% for those with $D > 100$ km. This resulted in a new catalogue of 117,240 Lunar impact craters with diameters ranging between 0.9 and 532 km. Meanwhile, the automatic age estimation from the 12 CNNs was tested on ≈140 CE-1 and 502 CE-2 images of dated craters, which found that 88.97 (CE-1) and 89.04% (CE-2) of craters were estimated to belong to the correct epoch.

Efforts have also been made to increase the efficiency of CDA architectures such that they could be deployed onto mobile devices with limited graphical capacity (Wu et al., 2021). This has been accomplished by compressing a U-Net CNN by dropping model weights via a gradual pruning method, resulting in a network called PRU-Net. PRU-Net was trained on the same Lunar DEM data as previous U-Net-based CDAs in terms of data source and training sample size. One such U-Net-based CDA was the ERU-Net, named after its effective use of residual blocks in the U-Net model. ERU-Net was also trained to detect Lunar craters in LOLA elevation data, achieving an R of 83.59% and a P of 84.80% on 3,000 test images. Performing this model pruning and then testing on the same unseen DEMs yielded a F_1 only 0.55% lower than ERU-Net (Wang et al., 2020), yet with a model which was only 5.12% of the size in terms of model weights.

Meanwhile, the topology of Lunar craters has also been derived post-detection, whereby 3D morphological parameters were derived from crater predictions in Lunar elevation data (Wang et al., 2021). An SVM classifier was trained on Histogram-Oriented Gradient (HOG) feature descriptions of DEMs (Barker et al., 2016), which were labelled as either containing or not containing craters. This automated analysis found that several highland regions on the Moon

had reached cratering saturation (for crater diameters 2.5–5 km) and that smaller craters (1–5 km) in the Lunar mare were deeper on average than in the highlands. Less variation in R and FDR was seen compared to a Martian CDA also trained using DEMs (Christoff et al., 2020). However, it is difficult to compare results between the two bodies, in addition to the two CDAs being tested on different numbers of sites.

3.2 Non-Impact-Related Feature Detection

Whilst the field of planetary surface feature detection is populated mostly by CDAs, impact craters are not the only feature to have received ML/DL treatment. As has been explained in Section 2.1, other surface features apart from impact craters can benefit from automated detection due to their own scientific and/or engineering potentials. This has been reflected by several studies, which this section will detail, in detecting non-crater surface features with ML/DL. The reader will notice that the majority of these efforts focus on the surface of Mars, possibly due to the combination of high-quality datasets and a more varied surface morphology.

For example, one of the earliest uses of ML in detecting a planetary surface feature was in the case of volcanic and hydrological plumes (Lin et al., 2014b). Thanks to observations made by passing space missions, Io and Enceladus are both known to have active volcanic (Morabito et al., 1979) and water vapour plumes (Hansen et al., 2006), respectively. Due to the relative size of the plumes to the target body, these transient occurrences have appeared in Voyager, Galileo, New Horizons (Io) and Cassini (Enceladus) images. These data sources were used to train a k-nearest neighbour (k-NN, not to be confused with k-means clustering or neural networks) algorithm to classify plumes. k-NN in a classification scenario operates by predicting the class of a given data point (X) by finding the k closest training data points. Of these k nearest neighbours, the modal class is then chosen and applied to X. Using Scale Invariant Feature Transform (SIFT) feature descriptors, the k-NN method could achieve a recall of 95% for Cassini images of Enceladus and an average recall of 82.3% across Voyager, Galileo and New Horizons images of Io.

One study has used an AdaBoost ML algorithm to detect dark slope streaks (DSSs) on Mars (Wang et al., 2017b). A DSS is a thin, low-albedo feature on the inclined surfaces that occasionally appears to fan outwards as the elevation decreases (Schörghofer, 2015). They are of particular interest since they are one of the few observable active surface features on Mars. Training an AdaBoost classifier on Local Binary Patterns (LBPs) derived from red-band HiRISE images returned a True Positive Rate (TPR) of 79.1% and a False Positive Rate (FPR) of 3.7%. This classifier was also compared to an SVM and (the equivalent of) an MLP trained on the same data. The SVM had the worst FPR and TPR overall, whereas the MLP achieved the highest TPR. However, the AdaBoost still had the highest accuracy of the methods attempted. Wang et al. (2017b) suggest that the algorithm would be best suited to eliminating data which does not contain DSSs due to the low FPR.

DL has been employed to detect rockfalls on the surface of the Moon and Mars (Bickel et al., 2019, 2020). Rockfalls are a good visual representation of activity, as mechanical processes above or below the ground can stimulate their fall. The initial work saw a RetinaNet OD CNN model trained on a dataset of 2,932 LROC NAC images of Lunar rockfalls, which reached a P and R of 79 and 65%, respectively, upon a testing dataset with confidence and Intersection over Union (IoU) thresholds of 50% (Bickel et al., 2019). The latter involved a total of six RetinaNet models being trained, with two on a combination of HiRISE images of Martian rockfalls and LROC NAC images from Bickel et al. (2019) and the other four trained solely on the former. The models themselves range in depth, as well as the complexity exhibited by the training imagery. Of these CNNs, the ones using a training dataset of only the clearer Martian rockfall examples exhibit the highest P at the expense of a significant decrease in R. As a result, the best F_1-scores achieved across these four CNNs when a confidence threshold of 50% was applied only reached 36% (for the CNN trained with more complex data and a ResNet101 backbone). However, the average time taken for these CNNs to process a single HiRISE image was $\approx$45 s, which is an order of magnitude faster than a manual labeller.

ML algorithms have also been applied in Valles Marineris on Mars to detect extra-terrestrial landslides for morphometric analysis (Rajaneesh et al., 2022). 682 landslide events in this

region were identified and labelled with polygons in MRO Context Camera (CTX; Malin et al. 2007), Mars Orbiter Mission (MOM; Negi et al. 2019) Mars Colour Camera (MCC) and Google Mars imagery. These events were separated into 3 subclasses: debris flows, rock avalanches and slumps. Morphometric parameters, including length, width and relative relief, were derived for newly classified landslides by analysing corresponding HRSC-derived DEMs and THEMIS thermal inertia data. Five different ML methods were tested for producing pixel-level landslide classifications: simple logistic, logistic regression, a meta-classifier, MLP and Sequential Minimal Optimisation (SMO). The meta-classifier involves a classifier learning from the predictions produced by another classifier before it, with both classifiers being logistic regression. SMO is a method of solving the quadratic problem faced when training SVMs, leading to less complex and intensive calculations (Platt, 1998). The technique which proved most accurate was simple logistic regression, which returned the highest F_1 of 81.1%, as well as a TDR and FDR of 81.1 and 11.3%, respectively. It should be noted that these metrics were weighted averages across the three subclasses, but the exact weights are unknown. It was also found that in terms of recall, debris flows were the most difficult subclass of landslide to detect.

A recent study has proposed a method of automatically detecting Martian dust storms with Mask R-CNNs (Alshehhi and Gebhardt, 2022). Again, while not strictly surface features, dust storms do have a role in creating them or changing the morphology of already existing features. Previous work has attempted to detect dust storms on Mars, but they have either not provided the locations of their detections (Maeda et al., 2015) or have only considered regions with high frequencies of these events (Gichu and Ogohara, 2019). After training the Mask R-CNN on Mars Daily Global Maps, which are not time-series data products but global weather maps that cover all longitudes between $\pm 60^\circ$ in latitude, it is said to have detected the majority of instances where images contain dust storms. However, with a mAP between 67.8–68.1% (depending on training/validation data content and split) at an IoU threshold of 25%, the model struggled with areas with missing data and with misclassifying other atmospheric events as dust storms.

Preliminary work has also been presented for automatically detecting boulders on the comet 67P/Churyumov–Gerasimenko, with the approach being to use a general-purpose circle detector based on a Mask R-CNN called the Astrophysical Circle Detector (ACID; Ali-Dib et al. 2022). However, after the original completion of this review, a different group of researchers have more recently trained a Mask R-CNN model to detect boulders in satellite imagery of Earth, the Moon, and Mars (Prieur et al., 2023). With detections overlapping their corresponding label by less than 50% removed, Prieur et al. (2023) achieved a P and R of 72 and 64%, respectively—although this drops to 40 and 35% with an IoU threshold of 75%.

In a similar fashion to boulders, at the time this literature review was being conducted, only preliminary results had been reported for the automatic detection of pits on Mars (Nodjoumi et al., 2021). A modified YOLOv5 OD algorithm was trained on 130 red-band HiRISE images and inferred on a mixture of labelled MRO CTX and HiRISE data. This data was sourced by finding corresponding imagery which contained features, found mainly in the Tharsis region, from the Mars Global Cave Candidate Catalog (MGC^3; Cushing et al. 2015). While the initial effort was to bolster this database of Martian pits, early results have shown that some manual verification is needed to remove false positives. This is possibly due to the very small training dataset as well as the complexity of the features themselves in terms of their similarity to impact craters and different degradation states. The same team published the DeepLandforms framework for training a range of DL models on planetary remote-sensing data (Nodjoumi et al., 2023). The aim of DeepLandforms is to be able to perform image pre-processing, labelling, training and inference while maintaining the geolocations of the new detections. In the evaluation of the framework, pits/skylights on Mars were used as a case study whereby a Mask R-CNN was trained on HiRISE imagery of MGC^3 features. Attempting to sub-categorise Martian pits into 4 classes according to their morphology and proximity to other pits proved to be too complicated for the Mask R-CNN, as it could only achieve an R of 16.2%. Whereas, when the model was trained to detect all pits/skylights as a single class, along with impact craters as a second, the P and R both increased significantly to 95.2 and 70.6%, respectively. This work will be discussed further in Chapter 5.

There has also been work into automatically detecting larger geological landforms (Palafox et al., 2017), bridging the gap between surface feature detection and automated geological mapping. The study proposes an algorithm called MarsNet, which consists of a series of five CNN architectures running in parallel. Each CNN has been trained to detect target landforms of different sizes, such as volcanic rootless cones (VRCs) and transverse aeolian ridges (TARs), for training and comparison to SVMs. Unlike most work in geological mapping, which produces a single map with labels for each separate geological class, MarsNet produces multiple probability maps for each target landform. MarsNet outperformed SVMs trained using HOG feature descriptors in terms of the accuracy of detections (see Section 2.2.8) in the detection of VRCs in both HiRISE and CTX imagery. This accuracy reached 93.6 and 91.9% for MarsNet on HiRISE and CTX imagery, respectively. Accuracies are also stated to be consistent when MarsNet was trained to detect TARs in HiRISE data, although training MarsNet to detect multiple classes at once is suggested as a way to improve the robustness of the algorithm across different landforms.

3.3 Automated Geological Mapping

Similar to crater counting, another area of planetary science that has long been a manual task is geological mapping. Understanding the terrain of a planetary surface is an essential step before any space mission, robotic or human, can be sent there. Identifying the types of landforms present on a surface is also one of the methods of deducing what geological processes either were or are currently active. Therefore, as with other fields, the advent of increasing spatial resolution and more abundant quality remote-sensing data has been taken advantage of by multiple studies into the automated geological mapping of landforms.

In cases where the target classes are unknown, unsupervised learning approaches may be a solution. This was true for one such study, which employed ISO clustering of raster images made up of DEM-derived parameters to geologically map three test areas on the Moon (Wang et al., 2017a). These parameters were elevation, filled elevation, slope, filled slope, relief

amplitude, and filled relief amplitude. Therefore, ISO clustering was tasked with segmenting images made up of 6 raster bands (one for each parameter) into 20 clusters (Wang et al., 2017a). Comparing the results of clustering on three test areas (USGS quadrangles LQ-8, 11, and 20) to a reference map found that the overall accuracies fell between 70.12 and 83.34% (Wang et al., 2017a). However, the applicability of the method to other regions or planetary bodies is uncertain since the clustering in a brightness/colour dimension is not considered, and the threshold for merging similar clusters would likely change.

Weakly supervised learning has also been employed by one study in order to geomorphologically map the surface of Mars (Wilhelm et al., 2020). 16,150 cropped tiles made from 163 larger CTX images are given scene classification labels based on which of 15 total geomorphological classes (such as crater fields or aeolian terrain) are present. The ImageNet dataset was used to pre-train a range of off-the-shelf SC CNNs. To output consistent geomorphological maps, the class with the highest detection score is then classified for the entire tile, and Markov Random Fields (MRFs) are used to smooth the boundaries between classes. The F_1 reached between 93.4 and 93.6% according to whether the true/false positives/negatives were calculated for each class individually or across all classes at once. This method was then applied to imagery of Jezero crater and Oxia Planum, which are landing sites for the Perseverance and ExoMars rovers, in order to produce geomorphological maps of both regions.

Where a CNN has been trained for scene classifications to solve the issue of automated geological mapping, another has also been developed for semantic segmentation of the Martian surface (Barrett et al., 2022). As part of the Novelty or Anomaly Hunter (NOAH) project, DL was used to be able to classify each pixel in an area of interest as one of 14 pre-defined ontological classes. However, this required pixel-level labels, which were produced for 1,504 128x128 m cropped HiRISE imagery of the Mawrth Vallis and Oxia Planum candidate ExoMars landing ellipses. The NOAH Dataset Annotation Tool was used only by the authors to make these labels (Wallace et al., 2017), as opposed to labelling being conducted through citizen science. Google DeepLab was the model of choice for training and evaluating with this data since its convolutions use more context for making classifications without increasing the layer

size. Across three different training runs, where the labelled data was increased and/or refined, the mean IoU was found not to significantly improve since the proportion of labelled data for each class affected their detection accuracies. A best mean IoU across the full list of classes of 74.2% was found in the final training run, whereas it reached 92.5% in the initial run when the 14 classes were further categorised into 5 common groups. This showed that more training data was required for each class, but also a proportional spread of data between classes is needed, where accuracies of some classes lags behind others.

3.4 Current Limitations

The following section will detail some of the limitations of the literature reviewed within this chapter. Some of these challenges have been highlighted in previous reviews (DeLatte et al., 2019b; Chaini and Jha, 2024). However, many of these still apply and are also relevant in the case of planetary surface feature detection as a whole, as opposed to just one feature. Therefore, they will be addressed in this new context.

Perhaps the greatest challenge for planetary surface feature detection is to overcome the dominance of research on impact crater detection. In fact, Figure 3.1 shows that CDA research makes up more of the literature than automated geological mapping and all other non-impact-related literature combined. Figure 3.1 also presents a breakdown of the literature in each application by the target body, proving that there is currently a larger focus towards Martian surface feature detection between non-CDA and CDA research. This is likely due to Mars' surface having a wider variety of surface features, particularly in terms of genetic processes. Nonetheless, as portrayed by Section 2.1, planetary surfaces are not just covered in craters, but also a variety of scientifically interesting and important features and landforms. Therefore, many more surface features deserve the same level of ML treatment as has been offered to craters in order to understand them more deeply.

Concerns have also been raised previously about the inconsistency in CDA research and the lack of benchmark training datasets (DeLatte et al., 2019b). This is a particular concern for

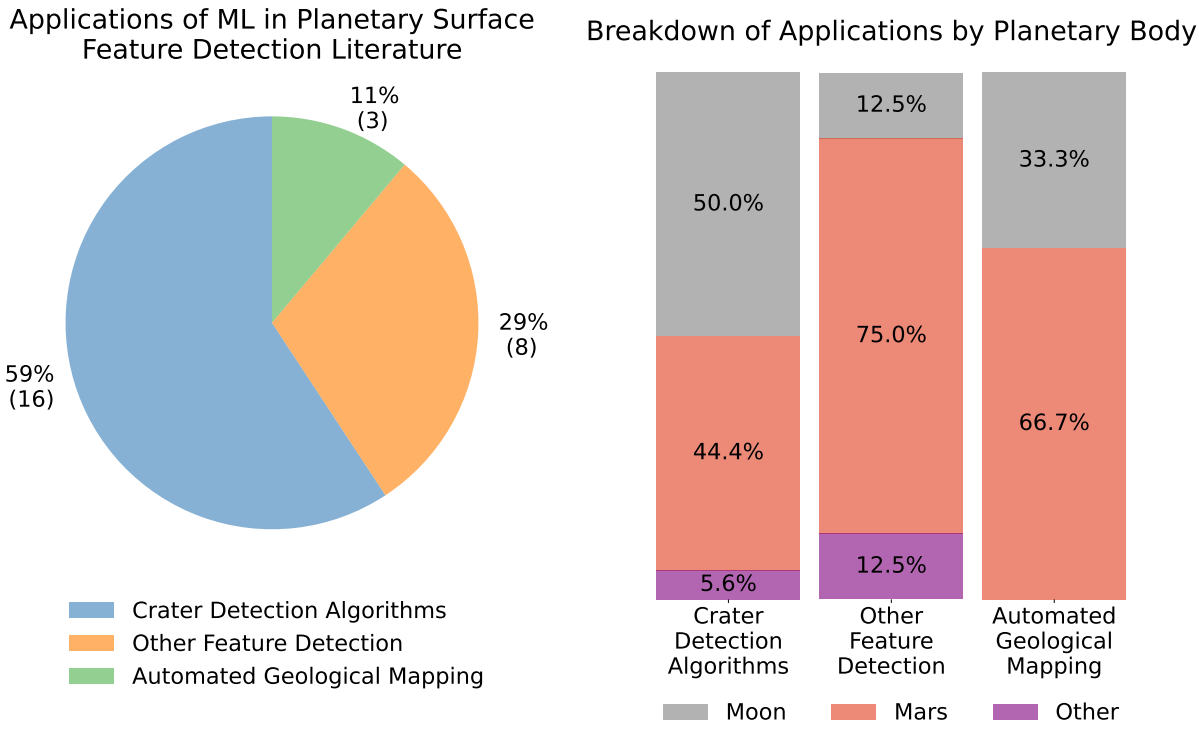

Figure 3.1: Breakdown of the most significant literature cited in this review by application and planetary body of interest. It should be noted that some papers, such as Silburt et al. (2019), have used data from multiple bodies.

CDA research since the majority of planetary surface feature detection literature is attempting to detect the same feature (as is shown by Figure 3.1). Therefore, without a dataset (of remote-sensing data and a feature catalogue) that the community can agree to test their algorithms against, it is difficult to compare the methods to determine if the field is, in fact, advancing. As more research into detecting other surface features is performed, this will become an issue in this application as well.

Inconsistency is also an issue for all applications of planetary surface feature detection. As can be seen in this chapter, many papers report a different set of training, validation and testing metrics. In addition to this, the distinction between these three sets of metrics is not always clear. It is crucial that the final training accuracies and precisions are stated alongside the metrics when tested on unseen data in order to determine how applicable the model is in new scenarios. Even in situations where the same metrics are available, different confidence or IoU thresholds are sometimes chosen for defining the sufficient output probability or overlap with a label for a prediction to count as a detection. For example, a broad and inconsistent

range of thresholds is being reported in CDA literature (from 20% in Benedix et al. 2020 to 95% in Wang et al. 2018 in terms of confidence, and 30% in Emami et al. (2019) to 50% in Ali-Dib et al. 2020; Benedix et al. 2020; Yang et al. 2020; Wang et al. 2018 in terms of IoU). What this means is that it is trickier to recognise if a model is genuinely more precise than another or whether it is just using a more conservative threshold.

Aside from any planetary geology or ML/DL challenges, there is a problem with the discoverability of planetary surface feature detection research. For example, the title of an article can have a big impact on how easy it is to find papers which have used a similar algorithm or target feature. However, the term 'automatic detection' may not be exclusive to ML or DL and can occasionally lead to confusion with research which uses simple image processing techniques. The literature is also spread between computer science and typical planetary/astrophysical science journals since there is no obvious, unified location to submit this type of research to. There are also occasions (Silburt et al., 2019; Yang et al., 2020) where specific terminology is incorrectly quoted in the title or text of research, such as the use of transfer learning being wrongly used to describe inferring a trained model on data from a different domain (i.e. from imagery to elevation data or from Moon data to that of Mercury).

3.5 Summary Tables

As explained at the beginning of this chapter, Tables 3.1 to 3.3 summarise the various planetary surface feature detection literature reviewed in Sections 3.1 to 3.3, respectively. Table 3.1 concerns only CDAs using ML and DL techniques for automated crater counting. Table 3.2 lists the comparatively fewer uses of such methods for the detection of non-impact-related features. Table 3.3 gives the three studies which performed automated geologic mapping.

Table 3.1: Summary of the crater detection algorithms (CDAs) discussed within Chapter 3. Where they have been provided, and unless otherwise specified, the precision (P) and recall (R) scores stated were calculated on testing data (i.e. not used for training). Literature occasionally used synonyms or other metrics from which P and R could be derived. Where available, the confidence score and IoU overlap thresholds are also quoted. Abbreviations are defined in Chapter 3. L = Moon, Ma = Mars and Me = Mercury.

Reference	Data Source (Target)	Architecture	P [%]	R [%]	Note(s)
DeLatte et al. (2019a)	THEMIS DIR images (Ma)	U-Net	87.42	66.72	Metrics calculated for 'E-4' model with a 40% confidence threshold.
Emami et al. (2019)	LROC NAC images (L)	3 CNNs (2 LeNet-based)	95.18	90.80	Image processing to suggest crater candidates, then CNNs for crater verification. Metrics calculated with a 30% IoU threshold.
Hashimoto and Mori (2019)	LROC NAC images (L)	YOLOv3	70.2–77.3	80.5	Attempted successive OD & SS approach.
Lee (2019)	MOLA/HRSC DEM (Ma)	U-Net	78	67	Silburt et al. (2019) model achieved an F_1 of 73% on the same test set.
Silburt et al. (2019)	LOLA/Kaguya (L) & MESSENGER (Me) DEMs	U-Net	56	92	Inferred trained model on Me data.
Wang and Wu (2019)	Kaguya TC images & LOLA/Kaguya DEM (L), HRSC images & DEMs (Ma)	AdaBoost & SVM on DEM profiles	88.91 95.28	94.01 92.80	Metrics (given for the L & Ma, respectively) calculated after 6 rounds of active learning.
Ali-Dib et al. (2020)	LOLA/Kaguya DEM (L)	Mask R-CNN	40.2	85.1	R after post-processing was best for craters with $10 < D < 20$ km (90.1%), but Mask R-CNN generally underperformed compared to DeepMoon.
Benedix et al. (2020)	THEMIS DIR images (Ma)	YOLOv3	90	94	Metrics calculated with IoU & confidence thresholds of 50 & 20%. Infers their model on CTX images & achieves accurate crater ages.

Table continues on following page

Table 3.1: (Continued)

Reference	Data Source (Target)	Architecture	P [%]	R [%]	Note(s)
Christoff et al. (2020)	Mean & Gaussian curvatures of MOLA DEM (Ma)	MLP & CHT for rim extraction	92.2 85.0 81.1	76.3 81.5 96.1	Metrics calculated from confusion matrices of 3 different test sites compared to the catalogue by Robbins and Hynek (2012a).
Wang et al. (2020)	LOLA/Kaguya DEM (L)	U-Net	85.3	83.3	ERU-Net reduces the no. of input channels in the layers before residual blocks to speed up training.
Yang et al. (2020)	CE-1/2 images & DEMs (L)	R-FCN (ResNet101)	n/a	94.71 93.35	R given separately for CE-1 & 2 test data. Infers their model trained on CE-1 data to that of CE-2.
Hsu et al. (2021)	THEMIS DIR images (Ma)	Faster R-CNN	84.50	86.62	Best results were with the inclusion of an FPN, CHT & scale-aware classifier.
Lagain et al. (2021b,a)	CTX & HiRISE images (Ma)	YOLOv3	n/a		Retrains model from Benedix et al. (2020) upon imagery of sub-1 km-diameter craters. Uses clustering analysis to remove secondaries from crater counts, further improving derived ages.
Lee and Hogan (2021)	MOLA/HRSC DEM & THEMIS DIR images (Ma)	Res-U-Net	87	80	Metrics are compared to the Salamunićcar et al. (2012) catalogue.
Wang et al. (2021)	LOLA & LOLA/Kaguya DEMs (L)	SVM classifier on HOG	91.05 88.73	94.03 93.91	Produced a global crater catalogue with morphometric information. Metrics are for two different test sites.
Wu et al. (2021)	LOLA/Kaguya DEM (L)	Res-U-Net	83.03	81.95	Used pruning to reduce model size & complexity.

Table 3.2: Summary of the research efforts discussed within this literature review that utilise machine and deep learning (ML/DL) methods to automatically detect non-impact process-related planetary surface features. The format of this table is identical to that of Table 3.1.

Reference	Data Source (Target)	Architecture	P [%]	R [%]	Note(s)
Lin et al. (2014b)	Voyager, Galileo, New Horizons & Cassini images (plumes, Io & Enceladus)	k-NN classifier on SIFT	n/a	84.39	No. of FPs are given but not the total targets (i.e. TP + FN).
Wang et al. (2017b)	HiRISE images (DSS, Ma)	AdaBoost classifier on LBP	n/a	79.1	Only FPR available (3.7%).
Palafox et al. (2017)	CTX & HiRISE images (VRCs & TARs, Ma)	5 CNNs with differing scales	n/a	94.86 85.98	Metrics given for HiRISE & CTX test data, respectively. Only TNR available (92.43 & 85.02%).
Bickel et al. (2019)	LROC NAC images (rockfalls, L)	RetinaNet	79	65	Metrics calculated with confidence & IoU thresholds at 50%.
Bickel et al. (2020)	HiRISE images (rockfalls, Ma)	RetinaNet	73	24	Metrics given for model with best F_1. Same thresholds as Bickel et al. (2019).
Rubanenko et al. (2021)	CTX images (barchan dunes, Ma)	Mask R-CNN	n/a		Only AP available (77%).
Nodjoumi et al. (2023)	HiRISE images (pits/skylights, Ma)	Mask R-CNN	95.2	70.6	Pits were a case study for a framework for developing DL models on planetary remote-sensing data
Watson and Baldini (2024)	CTX mosaic screenshots (potential cave entrances, Ma)	Simple CNN with 2 convolutional layers	84.71	82.76	P calculated from confusion matrix in supplementary material with 50% confidence threshold. 0.78% of all detections made during the survey were PCEs.

Table 3.3: Summary of the research efforts discussed within this literature review that utilise machine and deep learning (ML/DL) methods to automatically segment entire regions or surfaces according to their geologic context or geomorphological appearance. The format of this table is identical to that of Table 3.1.

Reference	Data Source (Target)	Architecture	P [%]	R [%]	Note(s)
Wang et al. (2017a)	LOLA DEM (20 landforms, L)	ISO Clustering	n/a		Only average accuracy across 5 classes available (between 70.12–83.34%).
Wilhelm et al. (2020)	CTX images (15 landforms, Ma)	5 CNNs (AlexNet, ResNet, VGG, DenseNet)	n/a		F_1 given (min: 76.53%, max: 98.62%, average: 93.44%), but confusion matrix available for DenseNet161.
Barrett et al. (2022)	HiRISE images (14 surface ontologies, Ma)	Google DeepLab	n/a		P & R available for only some classes. IoU available as an average over all 14 classes (74.2%) or when merged into 5 common groups (92.5%).

Chapter 4. Automatic Depth Calculation of Lunar and Martian Pits with the Pit Topography from Shadows (PITS) Tool

4.1 Introduction

As was briefly discussed in Section 2.1.2, pits are circular-to-elliptical depressions on rocky planetary surfaces which are generally formed by the gravitational collapse of surface material (Ferrill, 2015). Pits can often be misconstrued as impact craters due to their circular morphology and negative elevation relative to the surrounding surface. Nonetheless, pits are most distinguishable from impact craters by the clear lack of a raised rim or any ejecta rays—formed by the impact ejecta settling on the surface. Pits can also range from being bowl-shaped (similar to the interior of impact craters) to being either more regularly conical or cylindrical in shape (Ferrill, 2015). Pits exhibiting steep-sided walls and overhanging rims (potentially allowing lateral entrances to a sub-surface cavity) are commonly called skylights (van der Bogert and Ashley, 2015). For the purpose of conciseness, pits and skylights are both referred to as pits within this chapter. However, Chapter 5 will make this distinction between the two categories. For several examples of terrestrial, Martian and Lunar pits, see Figure 2.5.

There are several proposed mechanisms for how pits are formed on planetary surfaces. 'Pit crater' is a commonly used term for defining a pit that has been created through volcanic,

tectonic or faulting-related processes (van der Bogert et al., 2014). Pits can also be formed due to the loss of sub-surface volatiles through chemical dissolution, impact melting or sublimation (Spencer and Fanale, 1990; Ferrill, 2015). The different formation processes can result in several pits occurring in linear or sinuous series, often called 'pit chains' (Hagen, 2015). For example, a survey of Mars Global Surveyor (MGS) Mars Orbiter Camera (MOC; Malin et al. 1998) wide-angle images (with resolutions ≈250 m/px) of Martian pit chains with diameters (D) greater than 930 m suggested that dilational normal faulting was likely to be their most prominent formation pathway (Wyrick et al., 2004). This is where normal faults are refracted as they pass through unconsolidated layers, causing dilation under the surface (Ferrill and Morris, 2003) through which surface material may drain. By comparison with terrestrial examples, the partial collapse of lava tube roofs has been proposed as the cause of several pit chains on the Moon and Mars (Sauro et al., 2020). Lava tubes are underground conduits formed by the current or previous flow of lava (Gadányi and van der Bogert, 2015) and will be discussed in more depth in Chapter 5. Diagrams of proposed pit chain formation processes, such as collapsing lava tubes and dilational normal faulting, are given in Figure 4.1.

Pits are not just present on Earth (Okubo and Martel, 1998), the Moon (Greeley and Gault, 1979; Wagner and Robinson, 2022) and Mars (Wyrick et al., 2004; Cushing et al., 2015); but have also been observed on Mercury (Gillis-Davis et al., 2009), Venus (Davey et al., 2013) and Phobos (Horstman and Melosh, 1989). The work in this chapter will mostly focus on Martian pits due to their prominence in satellite imagery taken by the Mars Reconnaissance Orbiter (MRO) High Resolution Imaging Science Experiment (HiRISE; McEwen et al. 2007) sensor, as well as their role as potential entrances to underground caves. Martian pits have been proposed as potential cave entrances due to the thermal properties of their interiors resembling those which would be expected of sub-surface material (Cushing et al., 2007). Using Mars Odyssey Thermal Emission Imaging System (THEMIS; Christensen et al. 2004) infrared imagery, the interiors of seven candidates were found to emit these wavelengths more strongly than the exposed surrounding surface during the night (Cushing et al., 2007). Figure 4.2 gives an example of this phenomenon for the candidate nicknamed 'Annie' in Cushing et al. (2007).

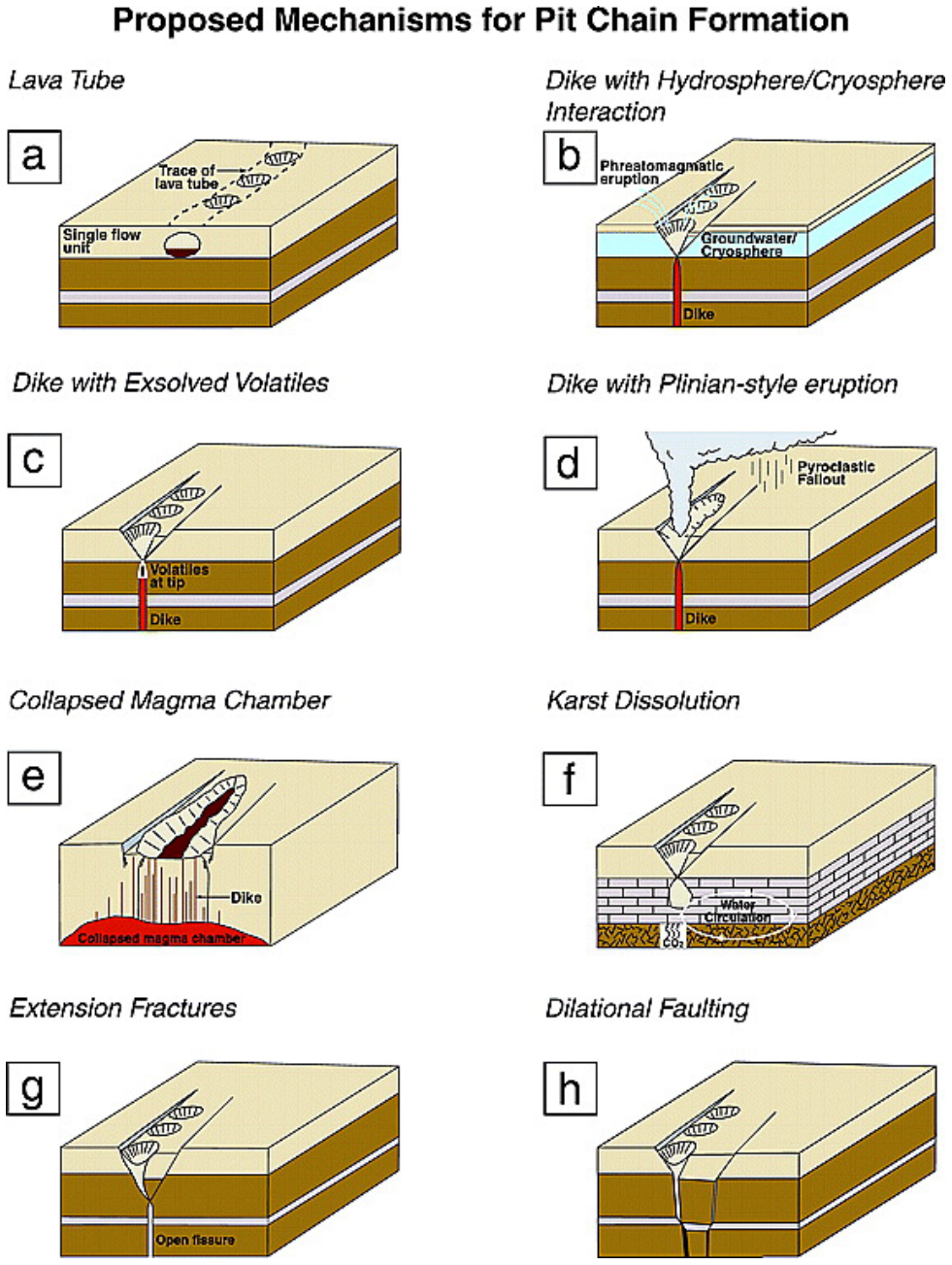

Figure 4.1: Schematics of the various proposed formation mechanisms for pit chains. This is Figure 2 of Wyrick et al. (2004), and has been used with permission.

Since then, MRO Context Camera (CTX; Malin et al. 2007) and HiRISE images have been surveyed to produce a database of possible cave entrances on Mars called the Mars Global Cave Candidate Catalog (MGC^3; Cushing 2015). This catalogue contains the locations of 1,062 candidate cave entrances and subdivides them into a number of categories, including atypical pit craters (APCs), lava tube skylights, and non-pit-related species. APCs are good candidates for cave entrances due to their vertical or overhanging walls (Cushing et al., 2015),

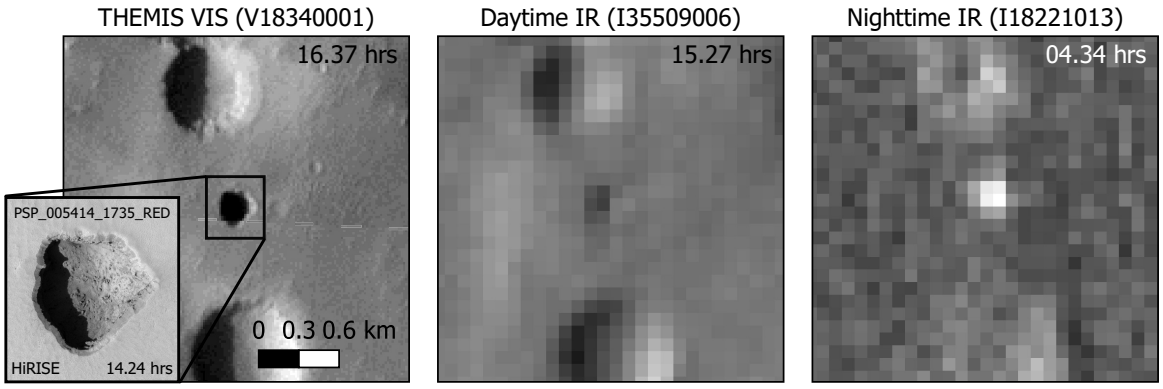

Figure 4.2: Mars Odyssey THEMIS VIS (left) and HiRISE observations (left inset) of a Martian pit nicknamed 'Annie' by Cushing et al. (2007) and code-named APC091 in MGC^3. THEMIS daytime (middle) and nighttime (right) IR images highlight the temperature difference between the surface and the pit interior during the night. The local solar times at the time of observation are also stated. This is the same pit that is shown in Figure 4 of Cushing et al. (2007).

as opposed to the more conical/bowl-shaped examples observed by Wyrick et al. (2004).

A catalogue similar to the MGC^3 also exists for pits on the Moon in the form of the Lunar Pit Atlas (LPA; Wagner and Robinson 2021). Although Chapter 5 will discuss the LPA in greater depth, within the context of intact Lunar lava tubes, one particular LPA feature—the Mare Tranquillitatis Pit (MTP, see Figure 2.5)—will be a focus of this work. The MTP is a near-circular Lunar pit (88–100 m in diameter) found at 8.3355°N and 33.2220°E, with vertical walls and overhanging rims on its north, east and western walls (Wagner and Robinson, 2021). The floor of the MTP sits approximately 105 m below the edge of its funnel, although it has been seen to slope downward in the direction of its eastern side (Wagner and Robinson, 2021). Along with the Mare Ingenii Pit (MIP, see Figure 5.3), the MTP was first discovered within SELENE (Selenological and Engineering Explorer, now commonly known by its nickname 'Kaguya'; Kato et al. 2008) Terrain Camera (TC) imagery (Haruyama et al., 2010). The MTP was one of the five LPA features which were the subject of 3D reconstruction using mono- and stereoscopic analysis of several Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) observations at varying satellite emission angles (Wagner and Robinson, 2022). Zhou et al. (2024) have further improved upon the density of the point cloud produced by Wagner and Robinson (2022) by using photogrammetry techniques. The

MTP is well-known for its interest and potential for future exploration, with it having been the target of the unsuccessful mission proposal 'Moon Diver' (Nesnas et al., 2023) and suggested to provide access to a sub-surface cavity due to radar observations (Carrer et al., 2024). Therefore, the MTP seemed a suitable case study for this work, as will be explained in further sections.

Caves have long been a target for human space exploration and scientific investigation (Boston et al., 2004), particularly on Mars. Without sufficient shielding, prospective astronauts on the Martian surface will be limited by how long they can spend there due to exposure to cosmic radiation. Over time, this dosage will lead to an increased chance of cancer development, nervous system deterioration and acute radiation syndrome (Atri et al., 2022). In spite of this need, it will not be feasible (in terms of fuel and cost) to transport the amount of shielding necessary for protecting human life over long durations. Therefore, underground caves have been identified as one of the "best solutions" for a Martian habitat for humans, in terms of eliminating the effect of radiation exposure (Atri et al., 2022). The depths and volumes of such features will indicate the amount of shelter that is provided, as well as the stability of their cavities. Numerical simulations have also suggested that caves are able to harbour stable ice deposits (i.e. no net loss of ice after one Martian year) across much of the Martian surface (Williams et al., 2010). Within the simulation, a variable that increased the propensity for ice stability was the cave's ceiling thickness (Williams et al., 2010). It can be assumed that for a cave to have a thicker ceiling, it must be sufficiently deep and voluminous to cause the roof to become unstable and collapse. In the case of Lunar caves, they have been simulated to maintain near-constant temperatures of around 290 K all year round (Horvath et al., 2022). Reserves of water ice and consistent temperatures would be particularly relevant for longer-duration crewed missions to the Moon and Mars, but also for astrobiological investigations on the latter.

Consequently, knowledge of the depths of planetary pits will be relevant for knowing which features to target for follow-up remote observation with higher resolution sensors, as well as which ones to explore for scientific or habitation purposes. Typically, the relative depths of

planetary surface features are found in elevation data produced via radar/laser altimetry or from observing shading changes in stereo image pairs. It is the latter method that generally produces digital elevation/terrain models (DEMs/DTMs) with the highest spatial resolutions. For example, Mars Global Surveyor (MGS) Mars Orbiter Laser Altimeter (MOLA; Smith et al. 2001) DEMs have high surface coverage, but only a best spatial resolution of ≈460 m/px (Som et al., 2008). Whereas DTMs produced from HiRISE stereo image pairs can have spatial resolutions as good as 1 m/px (Kirk et al., 2008). In order to produce this form of elevation data, more than one image is required to have been taken of the same region, and it also requires significant computational effort to estimate the elevation. Holes will also appear within DEMs where the stereo observations all had shadows in the same places. Therefore, datasets of such elevation data typically have low coverage over the entire surface. For example, only eight MGC^3 features were contained within HiRISE DTMs at the time of writing.

A method for estimating the depths of pits that does not require such elevation data is to observe the size of the pit's shadow as it appears in visual satellite imagery. The principle is that a deeper pit will cast a wider shadow than a shallower one, assuming that the Sun is in the same relative position for both features. Wyrick et al. (2004) propose a method for measuring the width of a pit's shadow (S) as a means of calculating its apparent depth (h)—the relative depth at the edge of the shadow—and estimating their volumes by assuming all pits are conical. According to the MGC^3 documentation (Cushing, 2015), a similar approach, which is outlined in Cushing et al. (2015), is also "generally" used to provide h estimates for all MGC^3 APCs within HiRISE imagery. Both Wyrick et al. (2004) and Cushing et al. (2015) manually measure S using GIS software. However, Cushing et al. (2015) apply an additional correction to S for all images with an emission angle (ε) greater than 5° by assuming that the satellite (MRO) and the Sun are pointing either due east or west at all times. An explanation of ε and a description of the method by which h can be derived from a measurement of S is given in Section 4.3.2.

Machine and deep learning (ML/DL) are currently being used to automatically detect

planetary surface features in remote-sensing imagery, as discussed in Chapter 3, including work on detecting Martian pits (Nodjoumi et al., 2021, 2023; Watson and Baldini, 2024). These techniques will serve to bolster existing catalogues in terms of their quantities and spatial completeness. Therefore, a method of automatically estimating pit depths from visual satellite imagery will be critical since it has been shown that having corresponding high-resolution DEMs is unlikely and costly to produce.

In this work, a tool is presented that can automatically calculate the apparent depths of Martian and Lunar pits by extracting their shadows and measuring the widths. This application, named Pit Topography from Shadows (PITS), is a Python-based framework that only requires a single cropped remote-sensing image to operate. PITS can also work with single- or multi-band imagery. Furthermore, the tool can automatically derive profiles of the apparent depths for entire datasets of pits without the need for any manual calculation or elevation data. The main objectives for developing PITS were to i) devise an appropriately simple and robust method for automatically measuring shadow widths, ii) prove that the tool achieves near-human level accuracy, and iii) ensure that the entire method is quicker to run than it would take a human to complete.

The remainder of this chapter is divided into the following sections. Section 4.2 is a description and justification of the datasets used for the development and testing of PITS. Section 4.3 outlines the methodology used to automatically produce apparent depth profiles. Section 4.4 details and discusses the testing performance of PITS' automated shadow extraction, as well as the results of applying the tool to MGC^3 APCs on Mars and the MTP on the Moon. The development, testing, and application of PITS to MGC^3 APCs was published in the *'Royal Astronomical Society's Techniques and Instruments (RASTI)'* journal in August 2023 (Le Corre et al., 2023a).

4.2 Dataset

The PITS tool has been specifically developed for use on remote-sensing visual imagery due to its availability at sufficient resolution and coverage, as well as for how shadows appear within it. Shadows are expected to be more distinct in imagery taken in visible wavelengths rather than in the infrared range, since some infrared will still be radiated from the surface underneath the shadow. The amount of infrared radiation observed will also vary depending on a number of factors, such as latitude, time of day/Martian year and meteorology. In this work, the focus has also been on ensuring the tool can be used with geo-referenced remote-sensing data. This is because it is essential to know where the analysed pit is present on the surface, especially for the search for potential cave entrances.

During development and testing, PITS has been applied to MRO HiRISE Reduced Data Record Version 1.1 (RDRV11) imagery of Mars (McEwen, 2007). This dataset was chosen for its high spatial resolution (0.25–0.5 m/px) and the fact that its products are already calibrated and map-projected (McEwen et al., 2007). The products in this dataset were also already known to contain sufficient MGC^3 features for PITS to be applied to. HiRISE RDRV11 products are available in single-band images taken in red wavelengths (570–830 nm) and multi-band colour images comprised of red, blue-green (below 580 nm) and near-infrared (above 790 nm) filtered passes (McEwen et al., 2007).

19 red-band images (shown in Figure 4.3) were selected from the HiRISE RDRV11 dataset for determining PITS' accuracy when automatically extracting shadows. These images were acquired from NASA's Planetary Data System (PDS; see McEwen 2007) by finding the MGC^3 APCs that had corresponding HiRISE imagery and selecting the ones which exhibited good variation in size and morphology. APCs were chosen since they were one of the most abundant MGC^3 classes and appeared to be the most visually synonymous with potential cave entrances. Of these 19 red-band images, 12 had corresponding colour versions, allowing for comparison between the different data types.

These testing images were manually labelled by the author using polygons in QGIS (Christensen et al., 2009) to produce geo-referenced ESRI shapefiles. The polygons in these 'vali-

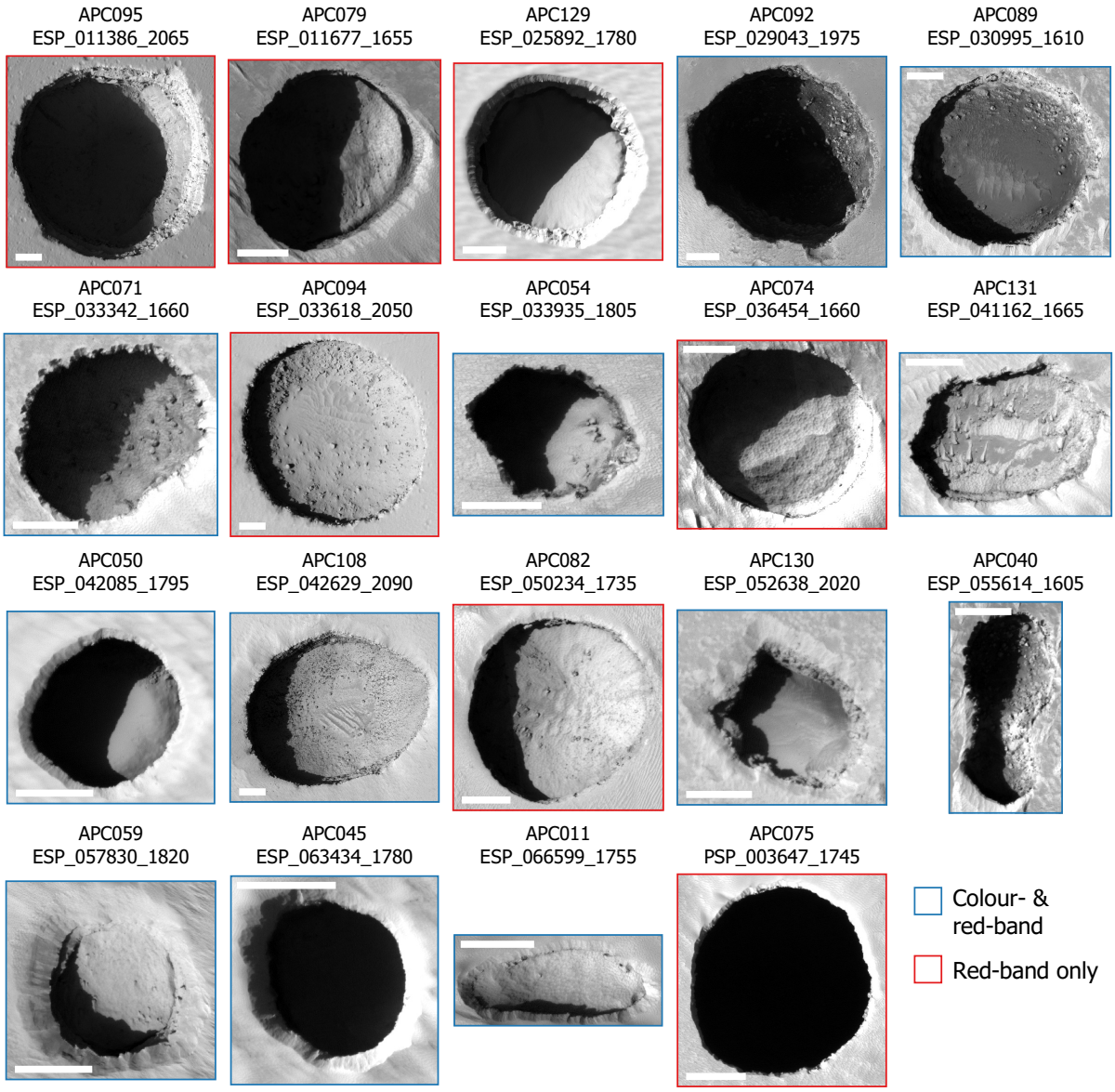

Figure 4.3: 19 cropped red-band Mars Reconnaissance Orbiter HiRISE Reduced Data Record Version 1.1 (RDRV11) images of Mars Global Cave Candidate Catalog (MGC^3) atypical pit craters (APCs). The shadows in these images have been manually labelled in order to assess the performance of the different automated shadow extraction methods detailed in Section 4.3.1. 12 of these 19 red-band images also have corresponding HiRISE colour versions. The scale bars in the top or bottom left of each image are all equal to 50 m.

dation shapefiles' were assigned a value according to what that region of pixels represented. A value of 1 was assigned to the largest continuous shadow in the image, which had clearly been cast by the pit's rim. A value of 2 was given to any bright features which were wholly contained within the shadow polygon. The reasons for labelling bright features are explained in Section 4.3.1.4. All remaining regions, which constituted the background, were given a

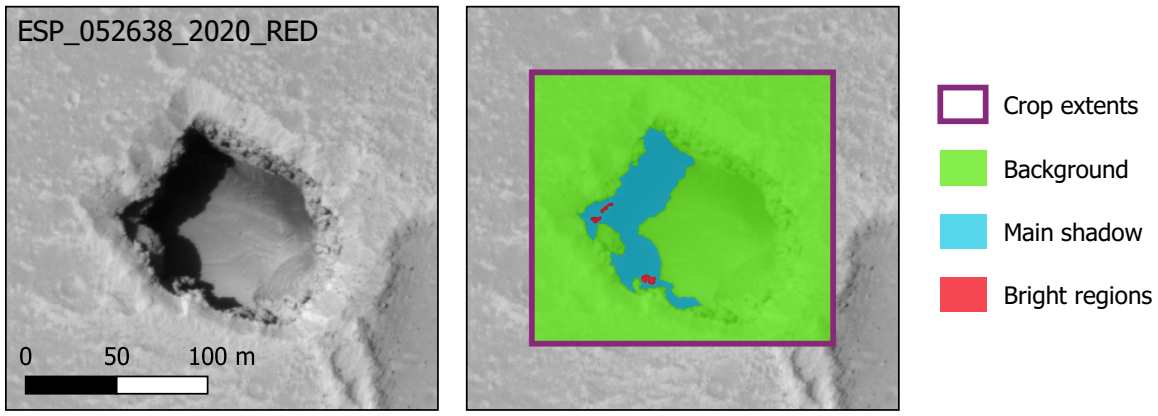

Figure 4.4: Example of the polygons produced when cropping and labelling the testing image ESP_052638_2020_RED of MGC^3 feature APC130 for its main shadow and any bright regions contained within it.

value of 0. These polygons were then rasterised such that each pixel in the input image could be compared to see if a correct shadow detection had been made. Figure 4.4 gives an example of this labelling procedure for the testing image ESP_052638_2020_RED containing MGC^3 feature APC140.

Since PITS extracts shadows from images cropped to the extents of a pit, a functionality exists to automatically crop larger image products using user-inputted rectangular ESRI shapefile labels. Using these 'location shapefiles', as opposed to simply indexing the pixel coordinates, ensures that the results of PITS can be geo-referenced to the pit's position on the surface. The location labels must contain the entire pit while making sure that no other features exhibiting shadows of comparable sizes (such as impact craters) are included. The location shapefiles should also ideally minimise the proportion of the surrounding surface, but different crop sizes were shown not to significantly affect the resulting shadow detection, as long as the above condition was met.

All HiRISE RDRV11 images which contained MGC^3 APCs were acquired from PDS in order to compare the depths calculated by PITS with literature estimates. Due to the testing results discussed in Section 4.4.1, as well as the limited availability of colour versions, only red-band versions were acquired for this depth comparison. From this selection, APCs were removed if they did not exhibit a significant shadow or if they partially fell outside the images' extents.

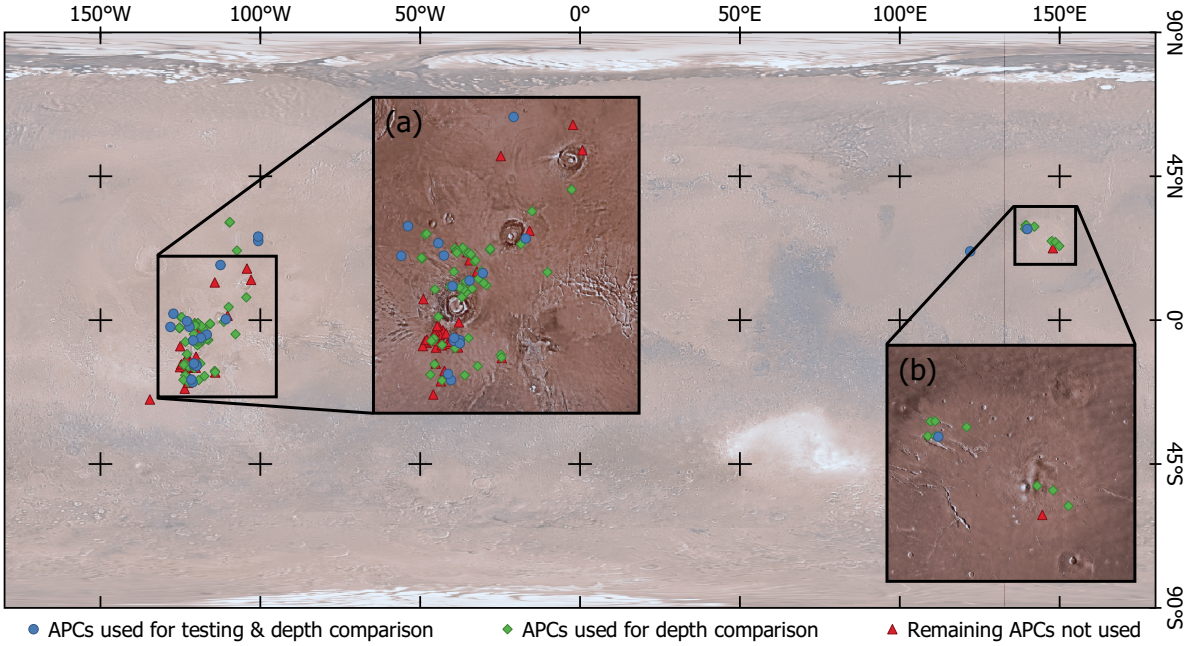

Figure 4.5: Global map of Mars showing the 88 MGC^3 APCs that were used in the development of the PITS tool. The 19 APCs that were used both for testing the shadow extraction performance (Section 4.4.1) and for comparing with depths from the MGC^3 (Section 4.4.2) are given as blue circles. The 69 APCs that were only used for depth comparison are shown as green diamonds, whilst the locations of the remaining 44 APCs not used are given as red triangles. The inset maps focus on (a) Tharsis Montes, since the majority of the APCs used in this work are located here, and (b) Elysium Mons. The imagery used here is the Mars Viking colourised global mosaic at ≈232 m/px at the equator.

This resulted in 123 HiRISE observations of 88 MGC^3 APCs, from which PITS extracted the shadows using k-means clustering when maximising $\bar{s_0}$. A map of the Martian surface overlaid with the locations of the MGC^3 APC features used for both shadow extraction testing and depth comparison can be found in Figure 4.5. As well as this depth comparison, a survey was conducted upon these 88 APCs to determine which, if any, were possible cave entrances given the shapes of the h profiles produced by PITS.

For the investigation of applying PITS to the MTP, LROC NAC images were identified that overlapped it with solar incidence and satellite emission angles of less than 90 and 20°, respectively. This was to ensure that both the Moon's surface was illuminated and the floor was visible to the pit, even if it was completely blanketed in shadow. Experimental Data Record (EDR) versions of these images were acquired from PDS and self-processed into calibrated and map-projected GeoTiff images using the ISIS (Integrated Software for Imagers and Spec-

trometers; Laura et al. 2023) and GDAL (Geospatial Data Abstraction Library; GDAL/OGR contributors 2025) software suites. This process is described in further detail in Section 5.2 of Chapter 5. The processed images were then surveyed to filter out any images in which the MTP was completely covered in shadow (which generally occurred for solar incidence angles greater than 35°), or the only shadow visible was cast by the opposite rim from the Sun's illumination. This resulted in 10 LROC NAC images of the MTP, ranging between 0.47–1.11 m/px in resolution, which could be cropped to its extents and fed through the PITS tool.

4.3 Methodology

The following section will detail the various elements of the PITS algorithm that, when combined, can derive a profile of h along the length of a pit's shadow from a single cropped satellite image. Figure 4.6 shows a flowchart of the whole algorithm applied to the same HiRISE image and pit as in Figure 4.4.

4.3.1 Automated Shadow Extraction

As discussed in Section 4.1, Wyrick et al. (2004) provide a method for estimating the depth of a pit by manually measuring the width of the shadow along the Sun's line of sight in GIS software. In this work, image segmentation is chosen to automate this process, whereby a cropped satellite image of a pit is separated into pixels that contain shadow and those that are

Figure 4.6: (Following page) A flow diagram of the Pit Topography from Shadows (PITS) algorithm. Blue and red arrows represent iterative processes and the points where input data is first used. All parameters of the PITS algorithm are stated, including the range of k values used for silhouette analysis, the maximum number of iterations of the k-means algorithm, the number of random initialisations of cluster centroids, the factor used to scale down images and cluster labels when calculating $\bar{s}_0$, and the tolerance (in squared Euclidean distances) for declaring convergence of the k-means algorithm. The diagram also shows how, at the end of the algorithm, the user is returned the h profile for that image (as a CSV table and plotted) and the geo-referenced shadow shapefile.

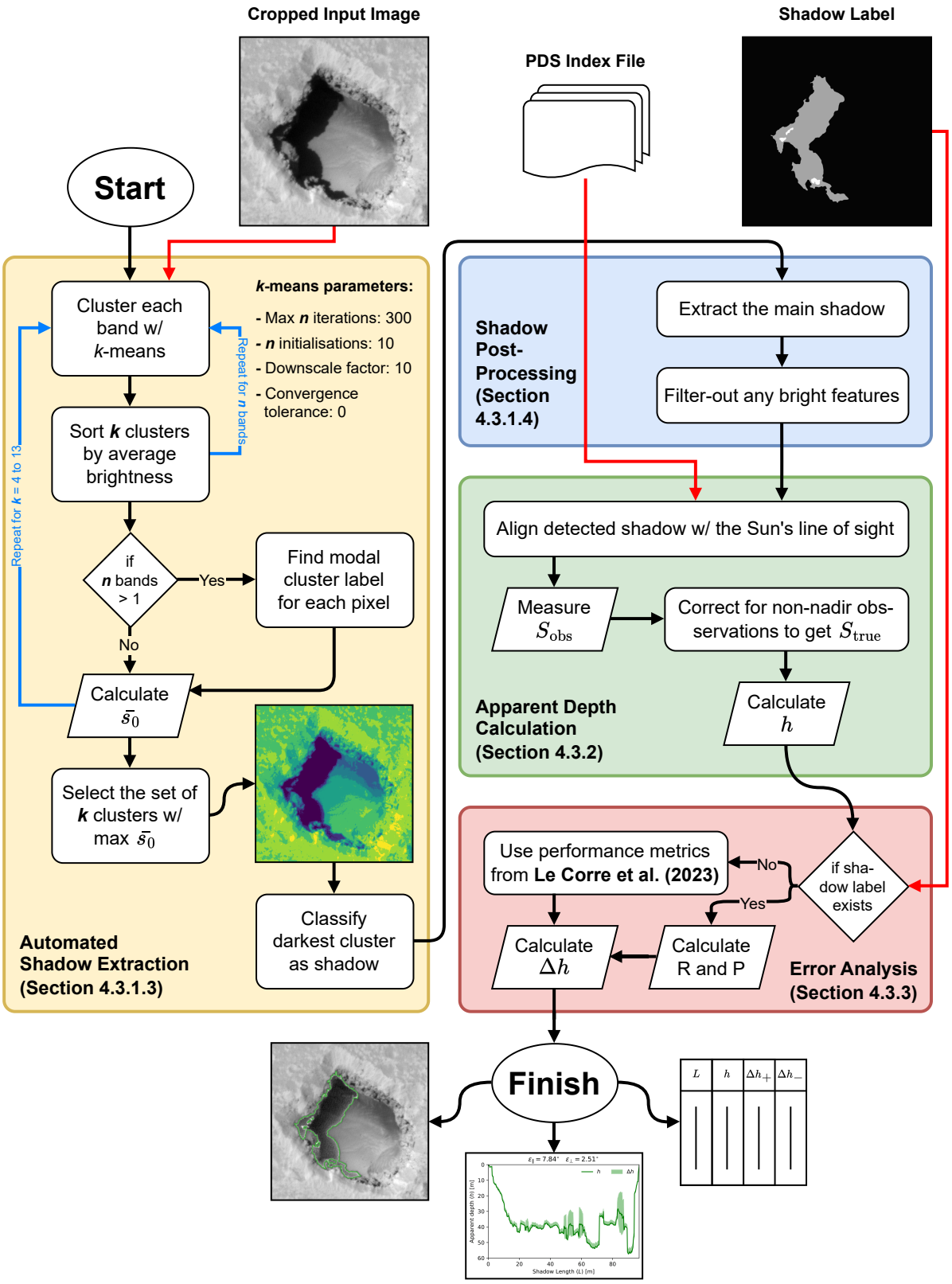

Figure 4.6: (Caption on previous page)

non-shadow background. This would result in a binary shadow mask consisting of detected shadow and background pixels (assigned a 1 and 0, respectively).

Image segmentation methods can range in complexity from applying simple thresholds up to trained semantic/instance segmentation ML or DL models. As mentioned previously, the aim when developing PITS was to devise a method that reflected the simplicity for a human to delineate a shadow. This section details the approaches to image segmentation that were considered for automated shadow extraction: Otsu's method and k-means clustering with two different types of silhouette analysis. It also details the necessary post-processing steps, as well as how these methods can be compared to each other on the shadow-labelled images described in Section 4.2.

4.3.1.1 Applying an Image Threshold with Otsu's Method

Otsu's method is a non-parametric, unsupervised method of image segmentation that determines a threshold for separating a grey-scale image into foreground and background pixels (Otsu, 1979). The threshold found by Otsu's method is the one which maximises the inter-class variance in intensity between the foreground and background pixels. As shadows should be visibly distinct from the surrounding surface, Otsu's method was considered for use in extracting them.

The full mathematical procedure is as follows: calculate a histogram of the pixel values with L bins, where L is the number of grey levels. The number of pixels in the bin i is given by n_i whilst the total number of pixels is $N = n_1 + n_2 + ... + n_L$. Following Otsu's method, a probability can be assigned to each bin i as $p_i = n_i/N$, where $p_i \geq 0$ and $\sum_{i=1}^{L} p_i = 1$. A given threshold value (t) will separate the pixels into two classes (C_0 and C_1), whose probability distributions are given by Equations (4.1) and (4.2), respectively.

$$w_1 = \sum_{i=1}^{t} p_i = 1 - w_2 \tag{4.1}$$

$$w_2 = \sum_{i=t+1}^{L} p_i = 1 - w_1 \tag{4.2}$$

The mean of classes C_0 and C_1 (μ_1 and μ_2, respectively) are given by Equations (4.3)

and (4.4), where the total mean of all pixels is $\mu_T = w_1\mu_1 + w_2\mu_2$.

(4.3) $$\mu_1 = \sum_{i=1}^{t} ip_i/w_1$$

(4.4) $$\mu_2 = \sum_{i=t+1}^{L} ip_i/w_2$$

The threshold selected by Otsu's method is the one that maximises the inter-class variance (σ_B^2) of pixels between the two classes, which is given by Equation (4.5).

(4.5) $$\sigma_B^2 = w_1(\mu_1 - \mu_T)^2 + w_2(\mu_2 - \mu_T)^2$$

Therefore, any pixels which have an intensity below this threshold are classified as a shadow pixel. It should also be noted that Otsu's method requires a grey-scale image to operate. Therefore, colour images have been averaged across all bands and the averages rounded to produce a single grey-scale image of integer pixel values.

4.3.1.2 k-means Clustering

k-means is an unsupervised clustering algorithm whose objective is to separate N data points into the k clusters (defined as $\mathbf{C} = \{C_0, C_1, ..., C_{k-1}\}$) which locally minimise the intra-cluster sum of squared Euclidean distances. Therefore, k-means clustering finds a local minimum of the objective function $J(k)$, given in Equation (4.6), where $\boldsymbol{\mu}_i$ is the mean (also called the 'centroids') of the data points $\mathbf{x}$ in cluster C_i.

(4.6) $$J(k) = \sum_{i=0}^{k-1} \sum_{\mathbf{x} \in C_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2$$

k-means finds only the local minimum of $J(k)$ since the centroids need to be initialised—often randomly. Once the centroids have been defined for the first time, the N data points are then each assigned to the cluster to which their centroid is the closest. These centroids are then iteratively updated by calculating the new mean of the data points $\mathbf{x}$ for each cluster and then repeating this assignment step. This is repeated until the improvement in $J(k)$ no longer meets a given tolerance or the chosen number of iterations is exhausted. k must also

be known a priori for k-means clustering to operate.

k-means clustering was chosen above other clustering methods for its simple-to-interpret procedure and since this algorithm was found to have a suitable run-time when dealing with the high numbers (as well as dimensionality) of data points found in high-resolution remote-sensing imagery. In the case of PITS, the data points that k-means must segment are pixel intensities, meaning that the algorithm is tasked with clustering in three dimensions: the longitude, latitude, and pixel intensity axes.

With the implementation of the k-means algorithm into PITS, the result is an array of the same size as the input image with each pixel being assigned an integer value according to which cluster it belongs to (0, ..., $k-1$). However, due to how the cluster centroids have to be randomly initialised, these integer values are not assigned consistently across multiple runs, even with identical parameters. In order for the tool to extract specific clusters, clusters are reassigned integer values according to their average brightness in the input image (from 0 being the darkest, to $k-1$ the brightest). Only the darkest cluster is then selected, and all constituent pixels are classified as shadow pixels.

In the case of colour images, a single shadow mask had to be derived from the resulting clusters from applying k-means to each of the bands. This was performed by clustering each band individually and then taking the modal cluster label for each pixel across the three bands. For example, if a pixel was assigned to clusters C_0, C_1 and C_1 across three bands, then this pixel would be reassigned to the modal cluster of C_1. Where the modal count across all bands for a cluster assigned to a given pixel is 1 or equal to that of another cluster, pixels are preferentially assigned to darker clusters.

4.3.1.3 k-means Clustering with Silhouette Analysis

As was discovered in preliminary testing, using a constant value of k for all images resulted in highly variable results. Therefore, a method was necessary for automatically deciding which value of k to use for a given image. This would allow PITS to adapt the k-means clustering algorithm to images which may be best described by more or fewer clusters than others.

Silhouette analysis is an approach for automatically suggesting a k value whereby it finds the value of k which returns the clusters whose data points are most consistent and distinct from those in other clusters (Rouseeuw, 1987). This can be said to be the k that yields the most appropriate clusters. The degree to which a given data point in cluster C_i is appropriate is defined by the silhouette coefficient (s_i) and is given by Equation (4.7), where a_i is the mean distance of a point in cluster C_i to all other points in C_i. b_i is the shortest mean distance of a point in C_i to all of the points in another cluster which it does not belong to.

$$s_i = \frac{b_i - a_i}{\max\{a_i, b_i\}} \tag{4.7}$$

The value of s_i can range between -1 and 1, with the goal being to maximise s_i since a larger value corresponds to a more consistent and distinct cluster. This means that instead of applying one value of k, PITS iterates its k-means clustering over a range of k values (4 to 13 in integer intervals). This range was selected as testing found that using a smaller or larger k never produced the most accurate results. The tool then calculates s_i for each pixel within down-scaled (by a factor of 10) versions of the input image and the array of assigned cluster values. This factor was chosen since it significantly improves the run-time during this silhouette analysis while not changing the suggested k values for each image during testing.

s_i is then averaged over all the data points within the same clusters to produce k average silhouette coefficients ($\bar{s_i}$). $\bar{s_i}$ describes, on average, how consistent all pixels in cluster C_i are with each other and how distinct they are from those in other clusters. The mean of $\bar{s_i}$ can then be taken again, this time to produce an overall silhouette score, which describes how appropriate all clusters are for this value of k. This silhouette analysis has been tested by selecting the value of k that maximised the overall silhouette score and also that which maximised the average silhouette coefficient for the pixels in the darkest cluster ($\bar{s_0}$).

4.3.1.4 Post-Processing of Shadow Masks

In order to ensure correct shadow width measurement, the extracted shadow masks (produced by Otsu's method and k-means clustering with silhouette analysis) undergo post-processing. Once the binary shadow masks are produced, the largest continuous feature is chosen while all other smaller shadow detections are removed. The reason for this is to remove any shadows from other objects in the image, such as boulders, which could interfere with the shadow width measurement. The 'main' shadow, upon which the width is measured, is the largest shadow, provided the feature is not highly irregular or degraded such that multiple similarly sized shadows are present. This condition is manually verified before using PITS to extract the main shadow in a given image.

Once the main shadow has been found, the next step is to identify any bright features that are wholly contained within it. These are likely to be either noisy pixels or detritus protruding above the shadow. Whilst any holes in the shadow mask due to the former should be filled, filling in the latter would lead to overestimation in the shadow width since they will be casting their own shadows. As a result, PITS performs morphological closing on all holes with an area less than 10 px, such that any remaining larger holes are likely to be protruding bright features that are casting their own shadows.

4.3.1.5 Comparing Automated Shadow Extraction Methods

In order to determine which of the above approaches to automated shadow extraction are most appropriate, a consistent way of comparing them is necessary. All three of i) Otsu's method, ii) k-means clustering when maximising the overall silhouette score, and iii) k-means clustering when maximising $\bar{s_0}$ have been applied to the 19 shadow-labelled HiRISE RDRV11 images described in Section 4.2. Their results have been compared to the rasterised validation shapefiles, whereby a confusion matrix could then be calculated for each image. It should be noted that the results are compared after the shadow has undergone the post-processing described in Section 4.3.1.4. From this matrix of true positive (TP), false positive (FP), true negative (TN) and false negative (FN) shadow pixel classifications, the precision (P), which is

the ratio of correctly detected shadow pixels to all detected shadow pixels, can be calculated. The recall (R) is also calculated, which is defined as the ratio of correctly detected shadow pixels to all true shadow pixels. For reference, the formulas for P and R have been given previously in Table 2.1.

The reason for calculating P and R is that F_1-score, which is the harmonic mean between the two and a very common performance metric for classification algorithms, is used in this work to assess the accuracy of the various shadow extraction methods in this work. As such, the method that achieves the highest F_1-score in detecting shadow pixels is the one chosen for inclusion in the PITS algorithm. F_1-score, also commonly shortened to F_1, is also given in Table 2.1.

As explained in Section 4.3.1.3, k-means clustering is performed for a range of k values in order for silhouette analysis to be possible. Therefore, when testing k-means clustering with silhouette analysis, the k values which produced the maximum F_1 in each testing image (not necessarily the one suggested via silhouette analysis) could also be recorded. The F_1 generated when using these 'target' k values represents the highest achievable performance when using k-means clustering for shadow extraction. This helps assess if silhouette analysis is an appropriate method of automatically suggesting k values. However, they will, of course, not be known in future instances where the shadows have not been manually labelled.

4.3.2 Apparent Depth Calculation

4.3.2.1 Automatically Measuring Shadow Widths

The task of manually measuring a shadow's width along the Sun's line of sight is relatively straightforward thanks to map-projected satellite data and GIS software tools. On the other hand, making the same measurement automatically upon the binary shadow masks (extracted via the methodologies in Section 4.3.1) is not as simple. This is due to the fact that the Sun's line of sight can come from a range of angles.

In order to account for this, PITS utilises the solar azimuth angle (ϕ), which is regularly given in the metadata for planetary remote-sensing imagery. ϕ is the clockwise angle between

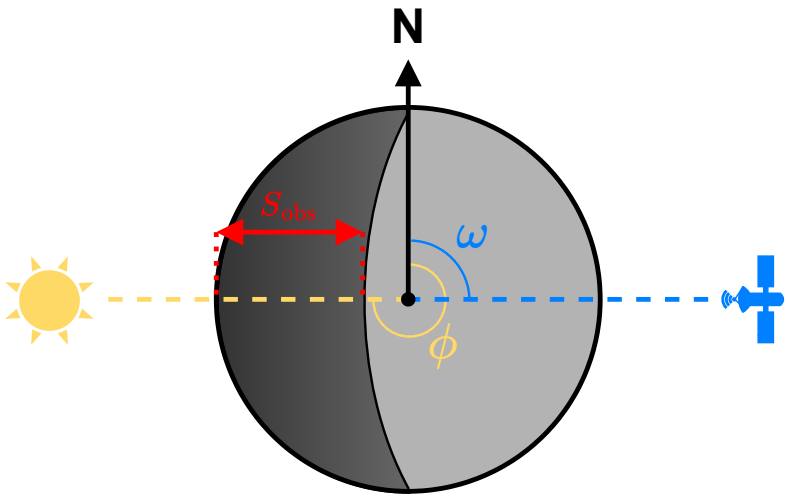

Figure 4.7: Top-down schematic of an idealised, circular pit, displaying where the observed shadow width (S_{obs}) of a pit along the Sun's line of sight corresponds to when being measured by the PITS tool. The solar azimuth angle (ϕ) is the clockwise angle between due north (in the case of PITS) and the sub-solar point. Similarly, the satellite azimuth angle (ω) is the clockwise angle between the same reference point and the sub-satellite point.

some reference point (due north in PITS' case) and the sub-solar point—the location directly below the Sun on the surface. Figure 4.7 shows how the observed shadow width (S_{obs}), as seen by the HiRISE camera, can be measured by knowing ϕ.

PITS can retrieve ϕ for each image from the index files hosted on PDS for both the Moon and Mars. These comma-separated files provide information on the locations of the Sun and the satellite at the time of sensing for all images within the HiRISE RDRV11 and LROC NAC EDR datasets. Since there were over 2.6 million LROC NAC images on PDS at the time, the cumulative index files required filtering down to just those being fed to PITS. Technical Note TN-1 in Appendix A presents some BASH code for filtering a HiRISE RDRV11 or LROC NAC EDR cumulative index file to just those within a given directory. Therefore, PITS can rotate the binary shadow masks anti-clockwise about their centre by ϕ to horizontally align the shadow such that the Sun's line of sight now passes directly through the mask from bottom to top. Nearest-neighbour interpolation is used when aligning the shadow masks so as to return the pixels to a gridded array while retaining their binary values, which would not be the case if a higher order of interpolation were used.

Aligning shadows in this manner means that PITS can now measure the distance between

the first and last shadow pixels along the Sun's line of sight (i.e. the shadow pixels adjacent to the pit's rim and shadow's edge, respectively). This shadow width (in pixels) is then multiplied by the image's resolution to acquire S_{obs} (in metres). This measurement of S_{obs} is taken at every pixel along the shadow's length (i.e. the axis perpendicular to the Sun's line of sight) in order to produce a profile of the shadow width.

However, if any bright features are found, as per the method in Section 4.3.1.4, then the aligned binary shadow mask requires further treatment before measuring S_{obs}. For those bright features that are found to be closer to the pit's rim than the shadow's edge, all shadow pixels between them and the rim along the Sun's line of sight are removed. This effectively means that the calculated apparent depth is now the relative depth from the bright feature to the shadow's edge.

Alternatively, if the feature is closer to the shadow's edge than the pit's rim, then all shadow pixels beyond it are removed, meaning the apparent depth is now the relative depth from the rim to the bright feature. Due to the risk that detected bright features are only artefacts of the image or caused by the reflection of a sunlit object elsewhere in the pit, S_{obs} is also measured upon a shadow mask where all holes are filled. This provides a range of possible values to display within the apparent depth profiles.

4.3.2.2 Correcting Shadow Widths for Non-Nadir Observations

As mentioned in Section 4.1, Cushing et al. (2015) apply a correction factor to the manual shadow width measurements made within all HiRISE images with an emission angle (ε) greater than 5°. ε is the angle between the satellite and a normal drawn from the surface. The satellite will also have its own azimuth angle (ω), which is the clockwise angle between the same reference point as for ϕ and the point directly below the satellite (see Figure 4.7). ε plays an important role in correctly calculating h since the camera will observe a distorted shadow within any image where ε is non-zero.

The motivation for applying a 'ε-correction' is that when the HiRISE camera is pointing in the same direction as the Sun's illumination, it will observe a foreshortened shadow due to the

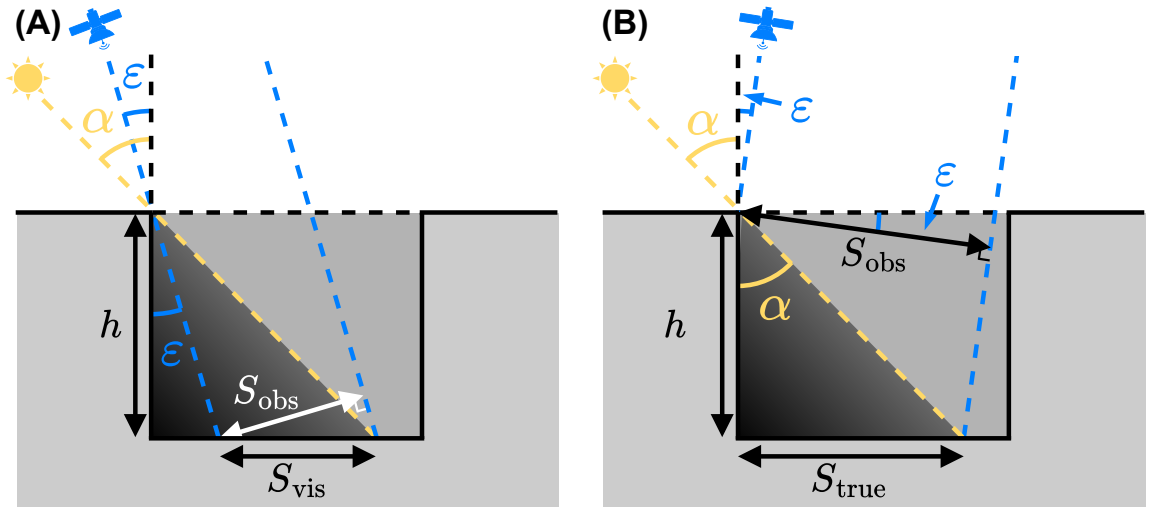

Figure 4.8: Diagrams showing the ideal scenarios where the Sun and the satellite are pointing in precisely the same (A) or opposite (B) directions. This has been adapted from Figure 9 of Cushing et al. (2015).

pit's rim partially obscuring its view. Whereas the shadow will appear enlarged when HiRISE and the Sun are looking in opposite directions, as some of the shadowed pit wall will now be visible. The scale of these distortion effects will depend on the degree of ε.

In their correction, which is portrayed within the diagrams in Figure 4.8, Cushing et al. (2015) assume that the Sun and the HiRISE camera are always pointing either due east or west (i.e. ϕ and ω must equal 90 or 270°). This may be suitable for HiRISE imagery at non-polar latitudes, but this may not be the case for all spacecraft in orbit around Mars or other planetary bodies. Therefore, a generalisation of this correction method is presented in this work, such that it can be employed for all values of ϕ, ω and ε.

The aim is to retrieve the true shadow width (S_{true}) that would be observed if ε was zero. This is achieved by calculating the obliquity of the satellite parallel and perpendicular to the Sun's line of sight ($\varepsilon_{\shortparallel}$ and $\varepsilon_{\perp}$, respectively, and shown in Figure 4.9). The purpose of calculating these separate variables is that a non-zero $\varepsilon_{\shortparallel}$ will lead to a distorted shadow width measurement and, thus, an incorrect h. Whereas a non-zero $\varepsilon_{\perp}$ will cause an incorrect shadow length measurement. In order to get S_{true}, the ε-correction presented in this section adapts the equations presented in Cushing et al. (2015) by replacing ε for $\varepsilon_{\shortparallel}$.

The first step in deriving $\varepsilon_{\shortparallel}$ and $\varepsilon_{\perp}$ is to calculate the ground distance (in metres) between

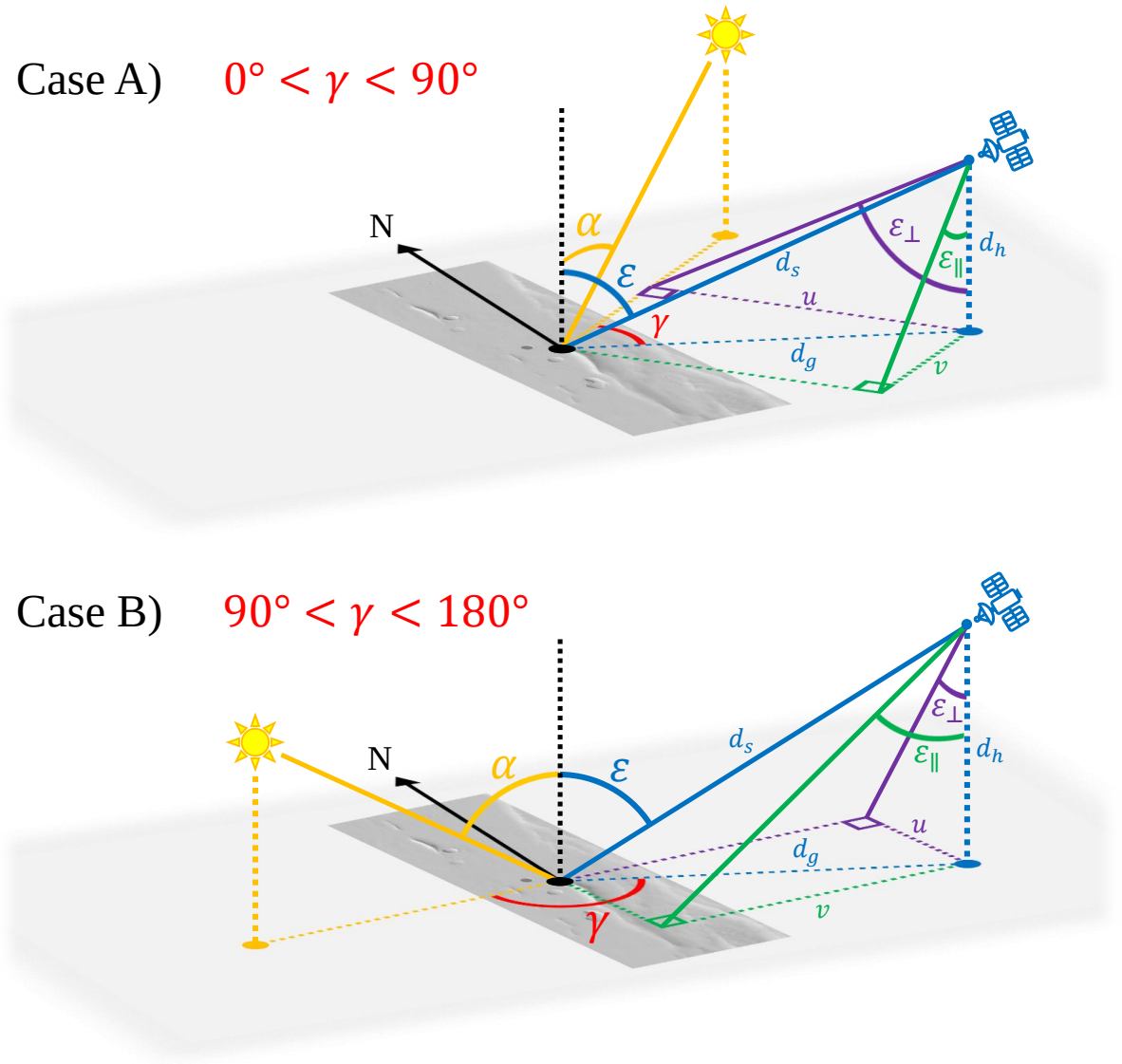

Figure 4.9: Schematic (overlaid on HiRISE image ESP_011531_2065_RED) of the sensing parameters in two examples of the cases A) and B) where the satellite and Sun are pointing in approximately the same or opposite directions, respectively. These symbols are defined in Section 4.3.2.2. A) and B) dictate how the shadow width observed by the satellite (S_{obs}) is corrected to obtain the true shadow width (S_{true}) as if the emission angle (ε) were zero. The black, yellow and blue dots represent the image centre and the sub-solar and sub-satellite points, respectively. This is Figure 6 of Le Corre et al. (2023a).

the centre of the image and the sub-satellite point (d_g). This can be done using the law of haversines to find the great circle distance for any two points upon a spherical surface, as shown in Equation (4.8), where r is the target body's radius, and (φ_1, λ_1) and (φ_2, λ_2) are the latitude-longitude coordinates of the image's centre and the sub-satellite point (found in

the PDS index files). For reference, the haversine function is defined as $\mathrm{hav}\,(\theta) = \sin^2(\theta/2)$.

$$d_g = 2r \arcsin\left(\sqrt{\mathrm{hav}\,(\varphi_2 - \varphi_1) + \cos\varphi_1 \cos\varphi_2 \mathrm{hav}\,(\lambda_2 - \lambda_1)}\right) \tag{4.8}$$

Another necessary parameter is the satellite's altitude above the sub-satellite point at the time of sensing (d_h). This information is provided within the PDS index files for the LROC NAC EDR dataset. However, only the areodetic altitude relative to the ellipsoid of Mars is known for HiRISE RDRV11 images, meaning that d_h must be derived using Equation (4.9), where d_s is the slant distance between the satellite and the centre of the image. d_s is given in the PDS index files for both datasets.

$$d_h = d_s \cos\varepsilon \tag{4.9}$$

Figure 4.9 shows how the parallel ground distance from the sub-satellite point to the centre of the image (v) and the perpendicular ground distance from the sub-satellite point to the Sun's line of sight (u) can be calculated using Equations (4.10) and (4.11). γ is the ground phase angle between the Sun and satellite, which is defined as the smallest absolute difference between ϕ and ω, meaning that γ can range between 0 and 180°.

$$v = d_g \mid \cos\gamma \mid \tag{4.10}$$

$$u = d_g \sin\gamma \tag{4.11}$$

Therefore, Figure 4.9 shows how $\varepsilon_{\shortparallel}$ and $\varepsilon_{\perp}$ can now be found using Equations (4.12) and (4.13), respectively.

$$\varepsilon_{\shortparallel} = \arctan\,(v/d_h) \tag{4.12}$$

$$\varepsilon_{\perp} = \arctan\,(u/d_h) \tag{4.13}$$

In case (A) where $0 \leq \gamma < 90°$ (i.e. the Sun and the satellite are pointing roughly in the same directions), S_{obs} can be corrected to find S_{true} using Equation (4.14), which has been

adapted from Equations (3) to (5) of Cushing et al. (2015).

$$S_{\text{true}} = S_{\text{vis}} + \frac{S_{\text{vis}} \tan \varepsilon_{\shortparallel}}{\tan \alpha - \tan \varepsilon_{\shortparallel}} \tag{4.14}$$

The solar incidence angle (α) is defined as the angle between the Sun and a normal drawn from the surface, and S_{vis} is the true width of the visible portion of the shadow not blocked from view of the satellite by the pit's rim. S_{vis} is given by Equation (4.15).

$$S_{\text{vis}} = \frac{S_{\text{obs}}}{\cos \varepsilon_{\shortparallel}} \tag{4.15}$$

Meanwhile, in case (B) where $90 < \gamma \leq 180°$ (i.e. the Sun and satellite are pointing in approximately opposite directions to one another), S_{true} is found using Equation (4.16), which has been adapted from Equations (7) to (9) of Cushing et al. (2015).

$$S_{\text{true}} = \frac{S_{\text{obs}} \sin \alpha}{\cos\left(90° - \alpha - \varepsilon_{\shortparallel}\right)} \tag{4.16}$$

In the unlikely event that γ equals exactly 90°, $\varepsilon_{\shortparallel}$ will be zero, meaning that there will be no distortion in the observed shadow width (i.e. $S_{\text{true}} = S_{\text{obs}}$). While γ does not equal 0 or 180°, a correction can also be applied to the observed length of the detected shadow (L_{obs}) using Equation (4.17) to retrieve the true shadow length (L_{true}). L_{obs} is simply the length of the detected shadow (in px) multiplied by the image's resolution (in m/px). Correcting L_{obs} does not affect the h profile directly, but rather serves as a reference for where the profile is taken along the length of the pit.

$$L_{\text{true}} = \frac{L_{\text{obs}}}{\cos \varepsilon_{\perp}} \tag{4.17}$$

When this ε-correction was applied to all 123 HiRISE images used in this work, the difference between the h values (the derivation of which is detailed in Section 4.3.2.3) before and after ε-correction was found to be ≈14.5% on average and ≈231.4% at its maximum (with a

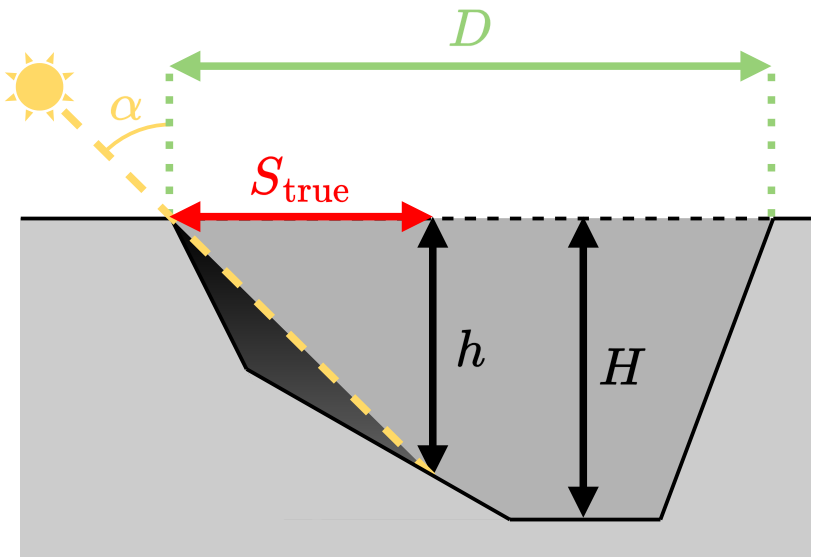

Figure 4.10: The cross-section of a simplified pit with an uneven floor as taken along the Sun's line of sight. This figure shows how the apparent depth (h) can be found using the solar incidence angle (α) and the true shadow width (S_{true}). The diameter (D) is not measured by PITS, nor does it attempt to find the maximum depth (H), like in Wyrick et al. (2004), since this requires the assumption that all pits are conical.

ε_{II} of 28.9°).

4.3.2.3 Calculating Apparent Depth Profiles

Since S_{obs} has been automatically measured and corrected for non-nadir observations to obtain S_{true}, the apparent depth (h) can now be found. Figure 4.10 takes a cross-section of a simplified pit along the Sun's line of sight, showing how a value for S_{true} can be combined with knowledge of α (also retrieved from the PDS index file) to calculate h using Equation (4.18).

$$h = S_{\text{true}} \tan \alpha \tag{4.18}$$

Unlike Wyrick et al. (2004), PITS does not assume that all pits are conical, since this was visibly untrue for the majority of the APC features observed in this work. Moreover, no suitably robust and similarly computationally cheap method of automatically measuring the diameter (D) could be found, which is necessary for deriving the maximum depth (H). Instead, the tool focuses on extracting the full shadow using image segmentation such that

S_{true} can be measured at every pixel along the entire length of the shadow—rather than just at one location. This means that a profile of h is produced, whereas previously, h has been provided as a single value (Wyrick et al., 2004; Cushing et al., 2015). This h profile is built up of the relative depths between the pit's rim and the shadow's edge along the Sun's line of sight, taken at each pixel in the shadow's length.

It should be noted that regions of PITS' h profiles that correspond to where the shadow's edge is adjacent to the opposite wall or rim of the pit should be treated as a minimum depth since the full width could not be measured. This is particularly relevant for deep features whose floors are completely covered in shadow.

4.3.3 Error Analysis

As can be seen from the many terms within Equations (4.14), (4.16) and (4.18), the uncertainty in h (Δh) ultimately depends on a variety of factors.

The first step towards finding Δh is determining the uncertainty in S_{true} (ΔS_{true}), which itself depends on the uncertainties in S_{obs}, ε and α (ΔS_{obs}, $\Delta\varepsilon$ and $\Delta\alpha$). The values for ε and α, which are given within the PDS index files, are taken at the centre of the HiRISE image. Therefore, there will be an uncertainty in these values when using them to calculate h, since the pit could fall anywhere within the minimum and maximum latitude/longitude coordinates of the image. However, due to the significant distance between the Sun and Mars, $\Delta\alpha$ was found to be negligible. Even for the HiRISE image with the largest footprint of those described in Section 4.2, $\Delta\alpha$ represented $\approx 1\times10^{-5}\%$ of the value for α.

Whilst $\Delta\varepsilon$ may be negligible, the altitudes of planetary remote-sensing cameras render $\Delta\varepsilon$ too significant to ignore. Thanks to the PDS index files, both the range of latitude/longitude coordinates of the image and the distance between the camera and the image centre are known for every HiRISE RDRV11 product. PITS is then able to use these values to calculate the maximum and minimum possible values for ε for each image to use as $\Delta\varepsilon$. A maximum value for $\Delta\varepsilon$ of 0.063° was calculated across all 123 HiRISE images. As a result, $\Delta\varepsilon$ can become relatively high compared to ε when the satellite is very close to looking through nadir.

Since PITS measures a shadow's width via automated means, this will contribute an uncertainty to S_{obs} which needs to be accounted for. Factors contributing to this uncertainty are the expected proportion of true shadow pixels that are missed, as well as the expected rate at which non-shadow pixels are incorrectly classified as shadow. This expected uncertainty has been estimated by taking the P and R calculated in Section 4.4.1 and averaging them across all 19 testing images. The Point Spread Function (PSF) of the HiRISE camera, which is typically only 2 px wide at Full Width at Half Maximum (FWHM), is expected to be negligible compared to this uncertainty (McEwen et al., 2007).

As a result of potentially different average precision and recall rates, the upper and lower limits of ΔS_{obs} are not necessarily equal. The upper limit ($\Delta S_{\text{obs},+}$) is a function of the average ratio of missed shadow pixels to all true shadow pixels and is equal to $1-\text{R}$. Therefore, $\Delta S_{\text{obs},+}$ can be found by using Equation (4.19).

$$\Delta S_{\text{obs},+} = S_{\text{obs}}(1 - \text{R}) \tag{4.19}$$

The lower bound $\Delta S_{\text{obs},-}$ is a function of the ratio of falsely detected shadow pixels to all shadow detections and is equal to $1 - \text{P}$. $\Delta S_{\text{obs},-}$ is found using Equation (4.20).

$$\Delta S_{\text{obs},-} = S_{\text{obs}}(1 - \text{P}) \tag{4.20}$$

Now that ΔS_{obs} and $\Delta\varepsilon$ are known, ΔS_{true} can be calculated using Equation (4.21). The full derivation for $\Delta\varepsilon_{\shortparallel}$, as well as the partial derivatives of S_{true} with respect to S_{obs} and $\varepsilon_{\shortparallel}$ depending on whether case (A) or (B) applies, can be found in Appendix B.1.

$$\Delta S_{\text{true}} = \sqrt{{\Delta S_{\text{obs}}}^2 \left(\frac{\partial S_{\text{true}}}{\partial S_{\text{obs}}}\right)^2 + {\Delta\varepsilon_{\shortparallel}}^2 \left(\frac{\partial S_{\text{true}}}{\partial \varepsilon_{\shortparallel}}\right)^2} \tag{4.21}$$

$$\Delta\varepsilon_{\shortparallel} = \Delta\varepsilon \left(\frac{|\cos\gamma|\sec^2\varepsilon}{\cos^2\gamma\tan^2\varepsilon + 1}\right) \tag{4.22}$$

Method	P [%]	R [%]	F_1 [%]
Otsu's method	76.8 (22.1)	**100.0** (**0.0**)	84.6 (18.4)
Silhouette score	**97.7** (**2.2**)	93.4 (10.0)	95.1 (5.4)
Silhouette coefficient ($\bar{s_0}$)	94.8 (4.2)	99.6 (1.1)	**97.1** (**2.6**)

Table 4.1: Average precision (P), recall (R) and F_1-scores (F_1) of shadow pixel detections with Otsu's method and k-means clustering with silhouette analysis (separately for maximising the silhouette score or the silhouette coefficient of the darkest cluster, $\bar{s_0}$) across 19 shadow-labelled HiRISE red-band images of MGC^3 APCs. The maximum value achieved for each score is highlighted in bold. The standard deviations of each metric across all images are given in brackets.

Since $\Delta\alpha$ is negligible, as mentioned previously, the uncertainty in h (Δh) is simply a function of ΔS_{true} and is given by Equation (4.23).

$$\Delta h = \Delta S_{\text{true}} \tan\alpha \tag{4.23}$$

4.4 Results and Discussion

4.4.1 Testing Shadow Extraction on HiRISE Imagery

4.4.1.1 Shadow Extraction Performance on Red-Band Imagery

Table 4.1 presents the average P, R and F_1 of shadow pixel detections when applying Otsu's method and k-means clustering with silhouette analysis to automatically extract shadows from 19 shadow-labelled HiRISE red-band images of MGC^3 APCs. Results are given separately according to whether the k-means algorithm has been applied using the k values which maximised the overall silhouette score or $\bar{s_0}$.

As Table 4.1 shows, extracting shadows by applying a threshold via Otsu's method resulted in an F_1 of 84.6%. While this may suggest high performance, this means that the vast majority of true shadow pixels will be detected, but roughly a quarter of all shadow detections will be false. Figure 4.11 displays the shadows that were extracted by each method presented in Section 4.3.1 for three out of the 19 HiRISE red-band testing images. For example, the threshold generated by Otsu's method for ESP_030995_1610_RED is evidently too weak

due to the large number of FP detections. The poor P and high variability of Otsu's method suggest that the number of visibly distinct regions in satellite images of pits can vary. Different illumination conditions, surface materials and morphologies mean that two classes are often insufficient to appropriately extract a pit's shadow.

According to Table 4.1, using k-means clustering with the k values which maximised $\bar{s_0}$ achieved the highest F_1 and R across all 19 red-band testing images (97.1 and 99.6%, respectively). The F_1-scores achieved by this approach also possessed the lowest standard deviation (SD). However, the highest P (97.7%) was achieved by k-means clustering when the overall silhouette score was maximised. Using k-means clustering with the target k values (as discussed in Section 4.4.2.2) upon the 19 shadow-labelled HiRISE red-band images produced an F_1 of 98.1% with a SD of 2.2%, which is not significantly higher than the F_1 in Table 4.1 achieved by both approaches to k-means clustering with silhouette analysis. As a result, silhouette analysis can be said to be a suitable method for k suggestion when using k-means clustering for automated shadow extraction.

It was also expected that maximising $\bar{s_0}$ would achieve the highest F_1 across the shadow-labelled red-band and colour HiRISE imagery. This is a result of PITS classifying the darkest cluster as the shadow, meaning that it was critical for this cluster to be as appropriate as possible. Moreover, the objective of silhouette analysis when maximising $\bar{s_0}$ is synonymous with the definition of a shadow (i.e. uniformly dark but also distinct from its surroundings).

4.4.1.2 Shadow Extraction Performance on Colour Imagery

PITS' automated shadow extraction has also been tested upon HiRISE colour imagery of Martian pits to determine if more image bands will lead to more accurate shadow pixel detections. Table 4.2 shows the P, R and F_1 when applying Otsu's method and k-means clustering with silhouette analysis across 12 shadow-labelled HiRISE colour images of MGC^3 APCs. This confirms that applying k-means clustering with the k which maximised $\bar{s_0}$ achieved the highest and least variable F_1 across both red-band and colour images.

In order to be able to directly compare the performance on red- and multi-band data, the

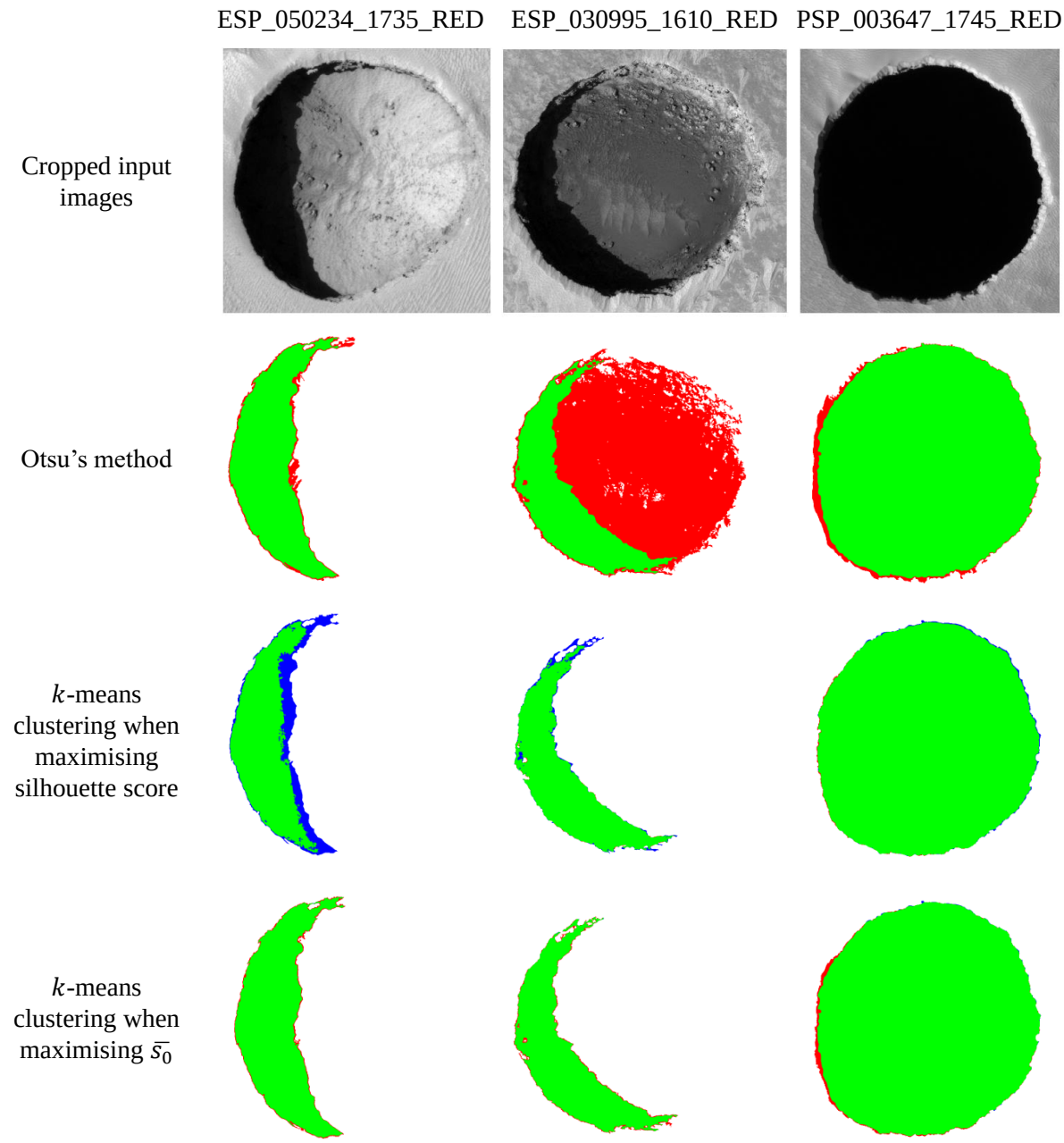

Figure 4.11: Examples of the shadow pixel detections using each of the methods described in Section 4.3.1. The detections and omissions of each extracted shadow have been colour-coded: true positives are green, false positives are red, false negatives are blue and true negatives have been blended into the white background. This is Figure 9 of Le Corre et al. (2023a).

average performance metrics were calculated for only the 12 HiRISE red-band images that had corresponding colour versions. Otsu's method achieved an F_1 of 79.4% (with a SD of 21.1%) upon these 12 red-band images. Whereas k-means clustering with the k suggested by maximising the silhouette score or $\bar{s}_0$ achieved F_1 values of 93.2 and 94.9% (with SD of

Method	P [%]	R [%]	F_1 [%]
Otsu's method	69.4 (24.6)	**100.0** (**0.0**)	78.9 (21.1)
Silhouette score	92.6 (11.0)	98.9 (3.0)	95.2 (7.0)
Silhouette coefficient ($\bar{s_0}$)	**94.1** (**3.8**)	99.4 (1.4)	**96.6** (**2.5**)

Table 4.2: Average performance scores of shadow pixel detections with Otsu's method and k-means clustering with silhouette analysis across 12 shadow-labelled HiRISE colour images of MGC^3 APCs. This table follows the same format and notation as in Table 4.1.

7.2 and 4.1%), respectively. Comparing these scores with those in Table 4.2 shows that using HiRISE colour imagery only yielded a marginal improvement in shadow extraction performance when using k-means clustering, but not when using Otsu's method. This suggests that the appearance of a pit's shadow is consistent across multiple HiRISE colour bands, but not all three. This is likely because Otsu's method operates on an image that is averaged across all colour bands, whilst k-means clustering assigns pixels to a cluster based on the modal cluster label across all bands.

The average run-time when applying PITS to colour imagery was found to be approximately three times longer than when using red-band versions. This is a result of HiRISE colour images having three times the number of bands (and, thus, pixels) as a red version. Therefore, due to the already high performance and improved run-time, the recommendation is to apply PITS to red-band versions when dealing with Martian pits imaged by HiRISE.

4.4.2 Applying PITS to HiRISE Imagery of Mars Global Cave Candidate Catalog (MGC^3) Atypical Pit Craters (APCs)

As previously mentioned in Section 4.2, all HiRISE RDRV11 red-band images which fully contained MGC^3 APCs that exhibited shadows were acquired. PITS was then applied to these 123 HiRISE observations of 88 APCs to automatically calculate their h profiles. Due to it achieving the highest F_1 during testing, as detailed in Section 4.4.1, k-means clustering when maximising $\bar{s_0}$ was the chosen method of automated shadow extraction. A CSV file containing the results for all 123 HiRISE images can be found in the online supporting material of Le Corre et al. (2023a), along with the corresponding MGC^3 code names and sensing information.

In order to reduce the h profiles to singular values, the h at the shadow's centre (h_c) and the maximum h (h_m) were both recorded.

This investigation has also allowed the recording of the run-time in order to suggest to a future user the approximate speed at which the tool can analyse their dataset. It was found that across the full dataset of 123 HiRISE images, the tool exhibited a minimum and maximum of $\approx$5 and 500 s for images with approximately 0.02 and 4.80 Mpx, respectively. This was achieved using a laptop with an 11$^{\text{th}}$ generation Intel Core i5 processor and 8 GB of RAM.

Appendix B.2 gives a list of the h_c and h_m values (and their respective uncertainties) calculated for all 123 HiRISE images of the 88 MGC^3 APCs. This table is also available as a CSV file in the supporting information of Le Corre et al. (2023a).

4.4.2.1 Investigating Apparent Depth Variation

Of these 88 APCs, 25 APCs were contained in one or more HiRISE RDRV11 images. This allowed for an investigation into the effect of non-zero ε values on the h values that are calculated. This was made possible by replacing S_{obs} for S_{true} in Equation (4.18) to get the h that would be observed before applying any ε-correction (h_{obs}). As before, h_c and h_m were also recorded from the h_{obs} profile in order to compare the variation in h profiles derived upon HiRISE images containing the same MGC^3 features before and after ε-correction. The expectation is that APCs with multiple HiRISE observations having high variation in their ε, but little variation in the Sun's position (i.e. α and ϕ), will also experience high degrees of variation in the h values that are calculated.

Figure 4.12 compares the relative SD in the h_c values that are calculated for the 25 APCs with multiple HiRISE observations (σ_h/h) before and after ε-correction. σ_h/h is the SD in the h_c values divided by their average, which avoids only larger APCs appearing to have high variation in h_c. Figure 4.12 shows that before correction, σ_h/h reached as high as $\approx$37% for some MGC^3 APCs. This is partly due to the fact that a variation in the illumination conditions (α and ϕ) will lead to the shadow's edge being cast to a different place within the pit, which can cause a different h that is calculated for features with uneven floors/rims. High variation

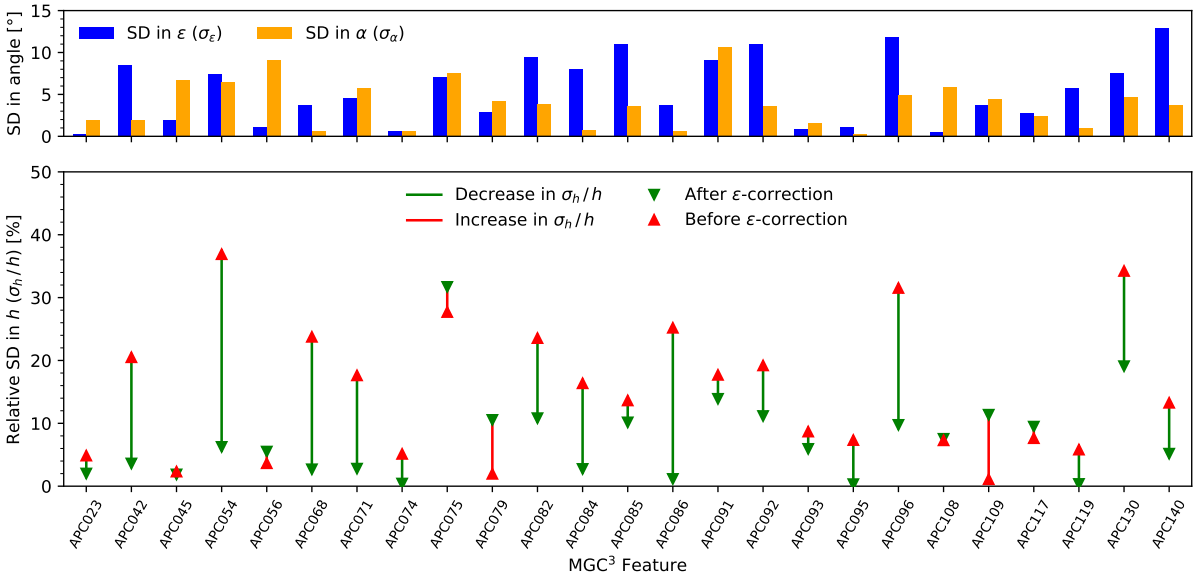

Figure 4.12: Comparison of the relative standard deviations (SD) in the apparent depth values (σ_h/h) calculated by PITS upon multiple images taken of the same MGC^3 APC before and after correcting for non-zero ε values. The σ_h/h before and after correction are represented by red and green triangles, respectively. Instances where σ_h/h has increased or decreased after correction are denoted by a red or green line, respectively. The SD in the images' emission (σ_ε) and incidence (σ_α) angles are also given for context about possible causes of high σ_h/h.

in h is also caused by these multiple images of the same APC having a larger range of ε. This is supported by the fact that the σ_h/h was significantly reduced after ε-correction for APCs whose images had a low σ_α. In essence, the Sun's position has not significantly changed, so the corrected h values are now consistent.

Despite σ_h/h decreasing after ε-correction in the majority of cases, Figure 4.12 shows how σ_h/h increases for six APCs: APC056, APC075, APC079 and APC108, APC109 and APC117. An increase in σ_h/h for these six APCs was also observed when plotting the relative SD in the h_m values before and after correction, as opposed to h_c. APC075 is a sufficiently deep pit that, despite a difference in α of nearly 20° between the three HiRISE images, its floor was completely covered in shadow in every image (i.e. S_{true} did not change). As a result, the change in α led to a high σ_h/h, which was then exacerbated when correcting for ε. APC056, APC079 and APC108 were all contained within two HiRISE images each, with one of the two images having higher α and ε values than the other, and vice versa. The images taken of these features also had γ values close to 0 or 180°. This meant that while the larger α was casting a

wider shadow in one image, the larger ε was making this width appear thinner. This resulted in consistent S_{obs} measurements despite the clear variation in α placing the shadow's edge at a different depth. Correcting for the change in ε then revealed the correct S_{true}, ultimately causing σ_h/h to increase. For the features APC109 and APC117, there was one image for which γ was close to 180° that also had the largest ε, meaning that $S_{\text{obs}} > S_{\text{true}}$. Whereas the other image containing the two features had a smaller ε, or a γ close to 90°, meaning that the difference between S_{obs} and S_{true} was not as great in this particular image. Therefore, the σ_h/h increased after applying ε-correction.

4.4.2.2 Comparing PITS and MGC^3 Apparent Depths

As previously discussed in Section 4.1, it is not possible to validate the h profiles derived by PITS upon all 123 HiRISE images of 88 MGC^3 APCs due to insufficient coverage of high-resolution elevation data. Even in the few instances where such data is available, there is the additional problem that the shadow's edge within a HiRISE image may not correspond to the same location within an ortho-rectified DEM. However, the h profiles derived by PITS can be compared with the manually produced h measurements made for each APC in the MGC^3 database (Cushing, 2015).

Figure 4.13 compares (on a logarithmic scale) the ε-corrected h_m values calculated automatically by PITS with the h values provided by MGC^3 for the HiRISE RDRV11 images that were the only such observation of that MGC^3 APC. This is to ensure that the same shadow is being extracted, and consequently, the depth is taken at the same location by PITS as was measured manually by Cushing et al. (2015). Cushing et al. (2015) state that they "typically" chose the HiRISE image with the smallest α to calculate h, but since this is not an absolute statement, only unique HiRISE observations of the MGC^3 APCs have been plotted.

Figure 4.13 shows that PITS and MGC^3 agree on a number of occasions. However, the h_m values calculated by PITS are less than the MGC^3 h for the majority of images, despite being the maximum h found within the PITS profile. In fact, there are four APCs (APC001, APC029, APC048 and APC135) for which there is more than a 200% difference between the

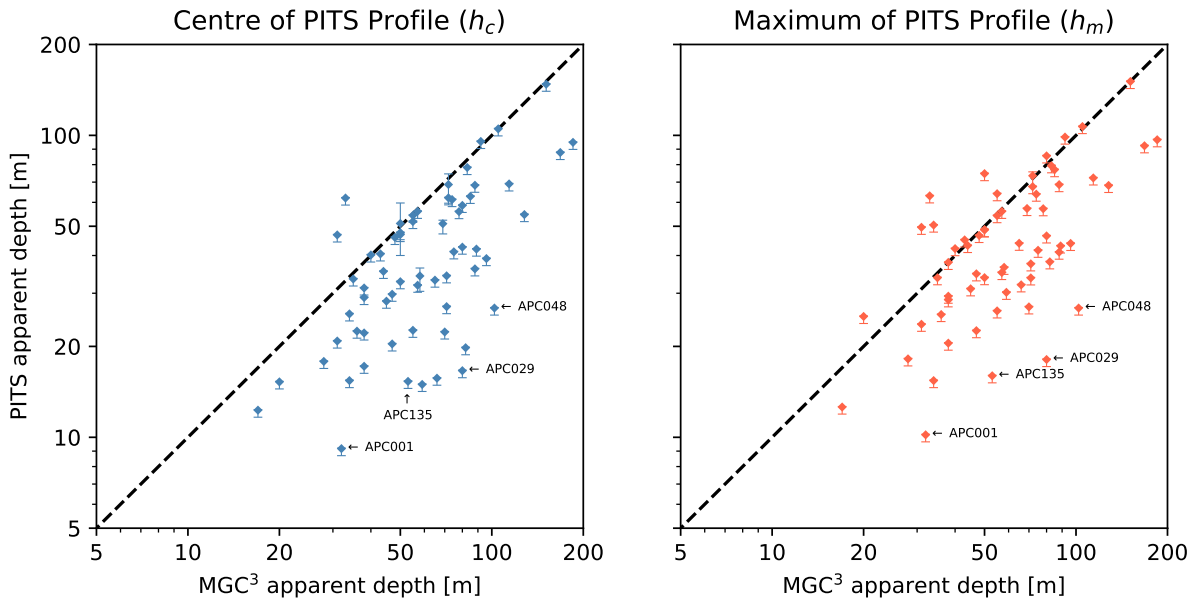

Figure 4.13: Apparent depths calculated by PITS compared to those in the MGC^3 catalogue for the HiRISE RDRV11 images that were the only such observation of that APC. The PITS h values have been corrected for all values of ε, whereas Cushing et al. (2015) apply a correction to all images with $\varepsilon > 5°$.

values predicted by MGC^3 and PITS. The shadows extracted by PITS for these APCs were individually inspected and exhibited no visible errors that would cause such deviation. Manual measurements of S_{obs}, followed by the depth calculation and ε-correction procedure in this work, yielded similar results. This was also the case when the correction procedure utilised by Cushing et al. (2015) was followed for only those images with $\varepsilon > 5°$. There are several potential explanations for this discrepancy. Firstly, since Cushing et al. (2015) measured S_{obs} manually within JMARS GIS software (Christensen et al., 2009), it may not have been taken correctly parallel to the Sun's illumination direction. This may have led to a wider width being measured and, ultimately, a larger h being predicted compared to PITS. Furthermore, the depths provided in MGC^3 may also be estimates of the maximum depth of the pit (H) as opposed to h—perhaps similar to the procedure used by Wyrick et al. (2004). This is supported by the statement made in the aforementioned MGC^3 documentation, which says that apparent depths were only "generally" derived (Cushing, 2015). It may also be that h was derived by Cushing et al. (2015) within self-processed HiRISE Experimental Data Record (EDR) images, which were not available within the RDRV11 dataset at the time of this work.

This would place the shadow at a different location within the pit and lead to a different value for h. However, given the difference in time between the two studies, this is less likely.

4.4.2.3 Possible Cave Entrances Suggested by PITS

In applying PITS to the 123 HiRISE RDRV11 images containing 88 MGC^3 APCs, the results of 11 images suggest the possible presence of a cave entrance. Since two of the 11 images contained the same feature, this resulted in 10 APCs that can be considered as potential primary targets for the exploration of cave entrances.

These features were identified according to three factors: the suggested presence of overhanging rims, the topography exhibited in their h profiles, and the wider context provided by the input image. Figure 4.14 displays the resulting h profiles for all 11 images of these 10 MGC^3 APCs. The MGC^3 codenames of these 10 features, from top to bottom in Figure 4.14, are: APC042, APC129, APC054, APC074, APC050, APC074 (again), APC032, APC022, APC006, APC015 and APC025.

Figure 4.14 also shows the original input image, along with an aligned version such that the Sun's line of sight passes from bottom to top. Overlaid upon this aligned image is the shadow edge and pit rim detected by PITS, which serves as the reference points between which S_{obs} has been measured. Elements of the h profiles in Figure 4.14 have been manually plotted in red to highlight regions where the shadow has not reached the pit floor (due to it being cast against the opposite wall) or is being cast by the opposite wall itself, and should be ignored.

Firstly, the geo-referenced shapefile containing the main shadow detected by PITS was used to crop the input HiRISE image to the shadow's perimeter. This allowed for the contrast of the pixels within the shadow to be enhanced in order to look for deeper shading indicating steep or overhanging rims, all the while retaining the context of the surrounding image. Figure 4.15 gives an example of this for the overhanging rim found within the HiRISE RDRV11 image ESP_066942_1735_RED taken of MGC^3 feature APC015. Further justification for an overhanging rim is if no pit wall is visible in the image and there is a notable difference in albedo and morphology between the pit's floor and the surrounding surface.

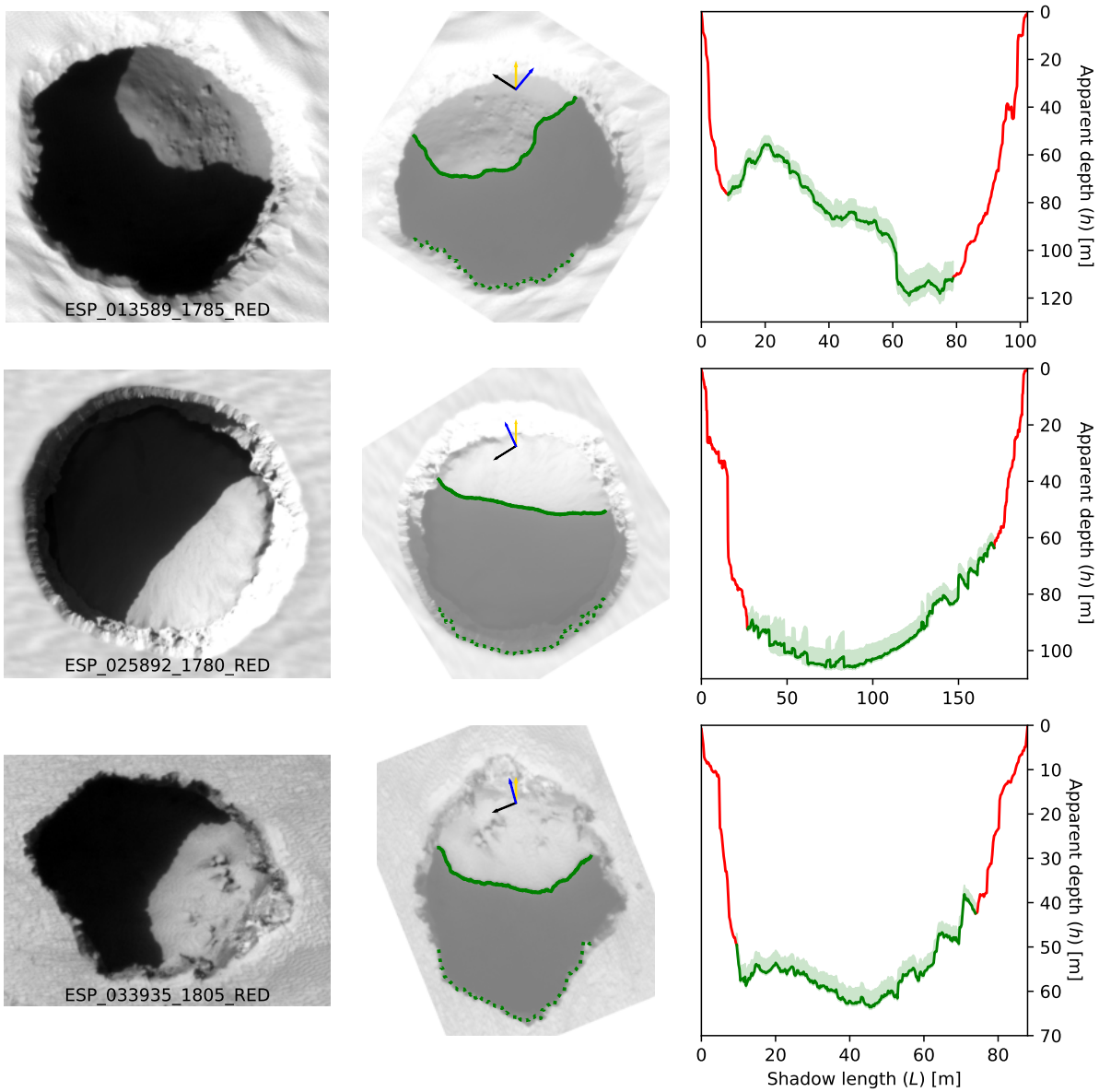

Figure 4.14: (Continuing onto next two pages) Apparent depth (h) profiles generated by PITS upon 11 cropped HiRISE RDRV11 images, which suggest that 10 MGC^3 APCs could host possible cave entrances. (Left) Input images and their product IDs. (Middle) Input images aligned with the Sun's line of sight such that it passes from bottom to top (also shown by a yellow arrow). Overlaying the rotated images are the detected shadow edges and pit rims (solid and dotted green lines, respectively), as well as the directions of due-north and the satellite's azimuth angle (as black and blue arrows, respectively). (Right) h profiles plotted in different colours to highlight the regions of the profile which should (in green) or should not (in red) be considered.

If the contrast-enhanced shadow and/or input image suggest the presence of an overhanging rim, the APC remains in consideration if its h profile showed a gradual or sudden decline towards the suggested location of the overhanging rim. The input image is examined again to

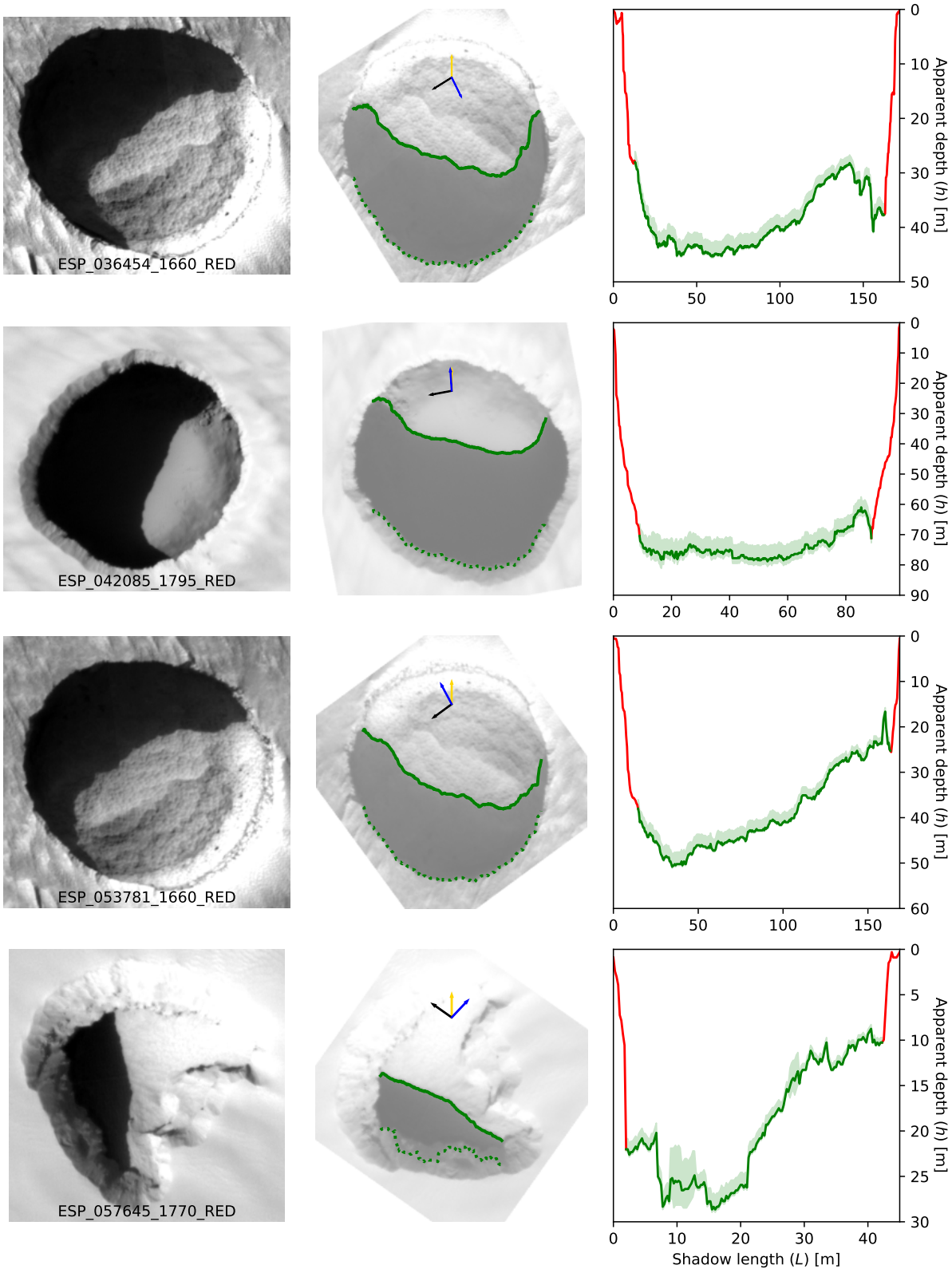

Figure 4.14: (Continued)

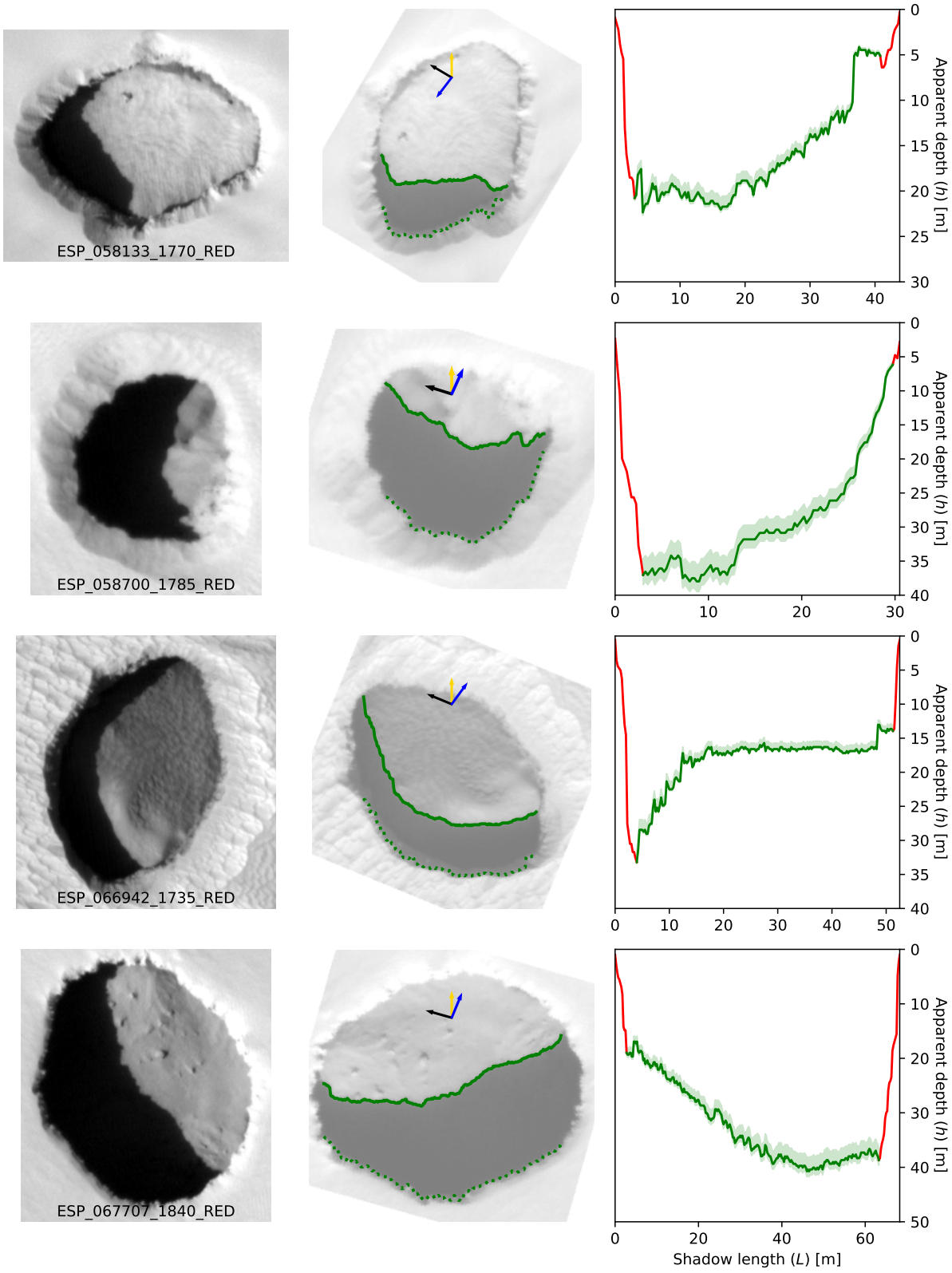

Figure 4.14: (Continued)

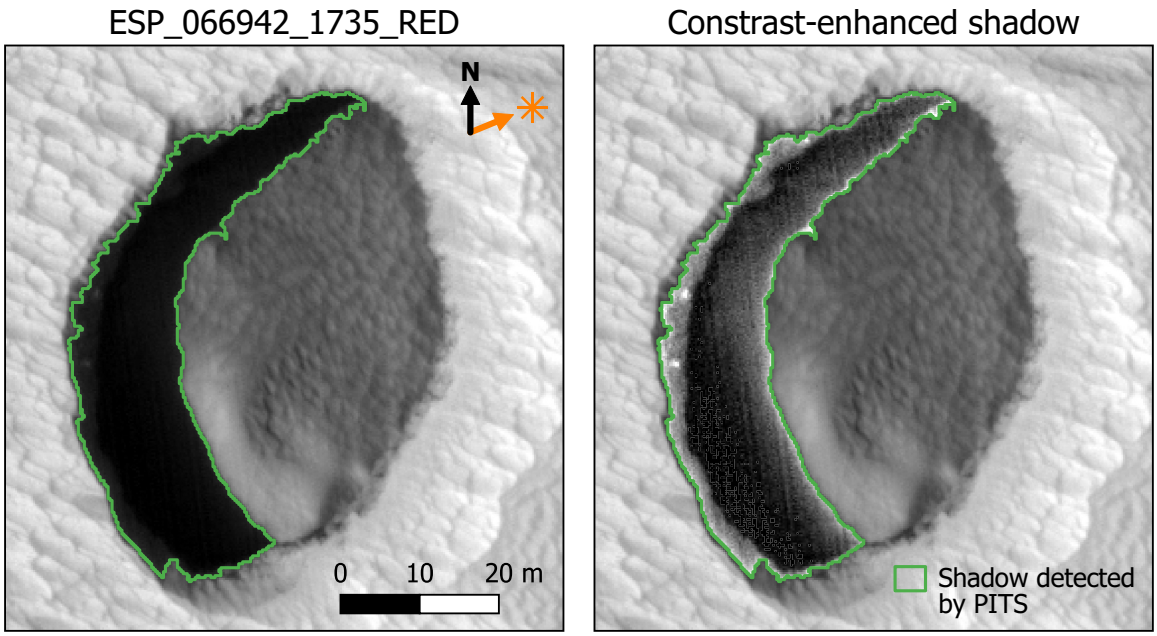

Figure 4.15: The overhanging rim of MGC[3] feature APC015 revealed by enhancing the contrast of shadow pixels in QGIS using the geo-referenced shadow shapefile automatically detected and outputted by PITS.

ensure that this feature of the h profile is not simply a product of an uneven pit rim casting a wider or thinner shadow.

The h profile derived from ESP_066942_1735_RED for APC015, shown in the penultimate row of Figure 4.14, exhibits a sudden decline towards the overhanging rim revealed in Figure 4.15. The near-zero $\varepsilon_\perp$ (0.02°) and the decline to the left of the h profile of approximately 15 m suggest a potential cave entrance. If $\varepsilon_\perp$ was high, then the satellite may be peering below an overhanging rim, which is casting its own shadow directly beneath.

4.4.3 Applying PITS to LROC NAC Imagery of the Mare Tranquillitatis Pit (MTP)

Figure 4.16 displays the results of applying PITS to 10 LROC NAC images containing the Mare Tranquillitatis Pit (MTP). For all but 2 out of the 10 observations (M137332905R and M155016845R), the direction of the Sun's illumination is pointing westward. This means that there are only two instances where the shadow has been cast onto the eastern side of the

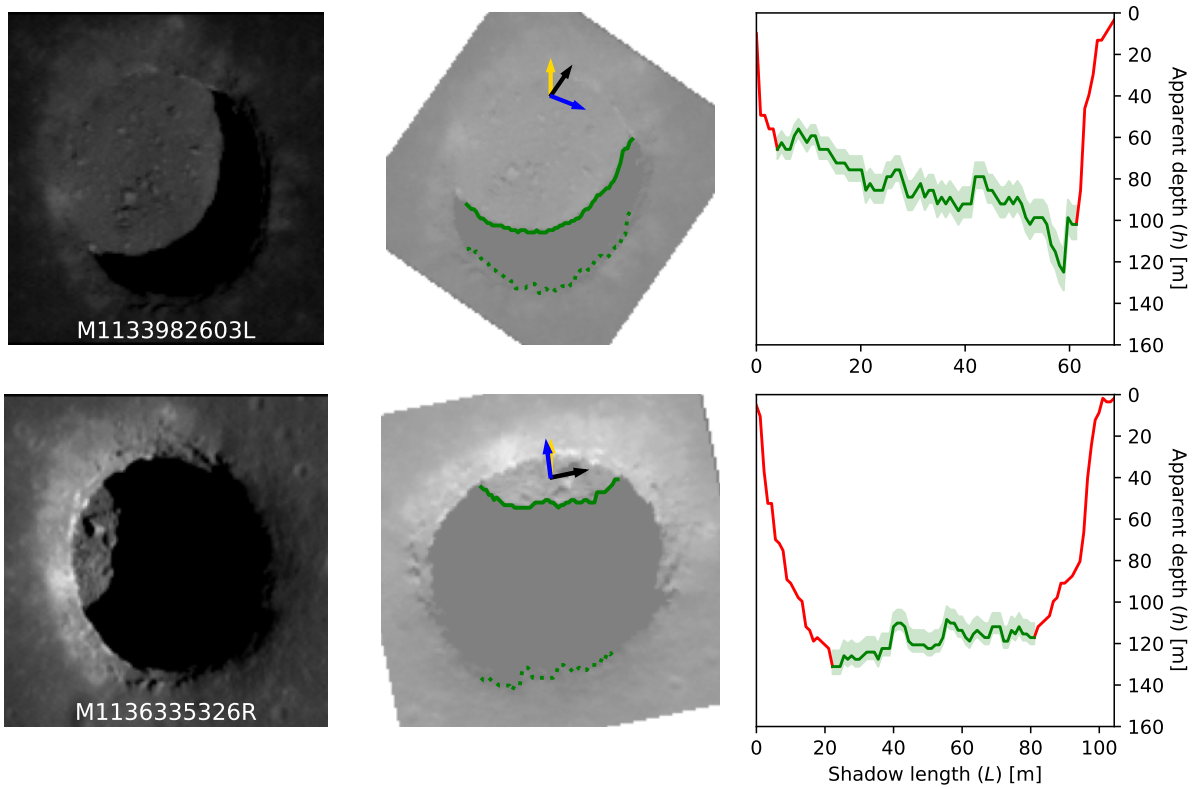

Figure 4.16: (Continuing onto next two pages) Apparent depth (h) profiles automatically generated by the PITS tool from cropped 10 LROC NAC observations of the Mare Tranquillitatis Pit (MTP). This figure follows the same formatting and symbology as Figure 4.14.

MTP's floor, whereby PITS could corroborate the sloping observed by Wagner and Robinson (2021).

From the detected shadow edges and pit rims shown across the 10 examples in Figure 4.16, it is clear to see that the PITS' automated shadow extraction methodology of unsupervised k-means clustering and silhouette analysis has transferred directly to LROC NAC imagery of the MTP—despite having initially been developed for Martian pits. This would have been aided by the fact that the shadows being cast are well-defined and prominent across all 10 images. However, the contrast between the surrounding surface and the interior of the MTP does differ. For example, within M1256379501L, the floor is much brighter due to the small values for ε and α. Whereas, the combination of a higher α while the camera is still pointing relatively nadir means that M126710873R shows a floor which is more consistent with the surrounding surface, as well as a lack of back-reflected light from the MTP's funnel as seen in

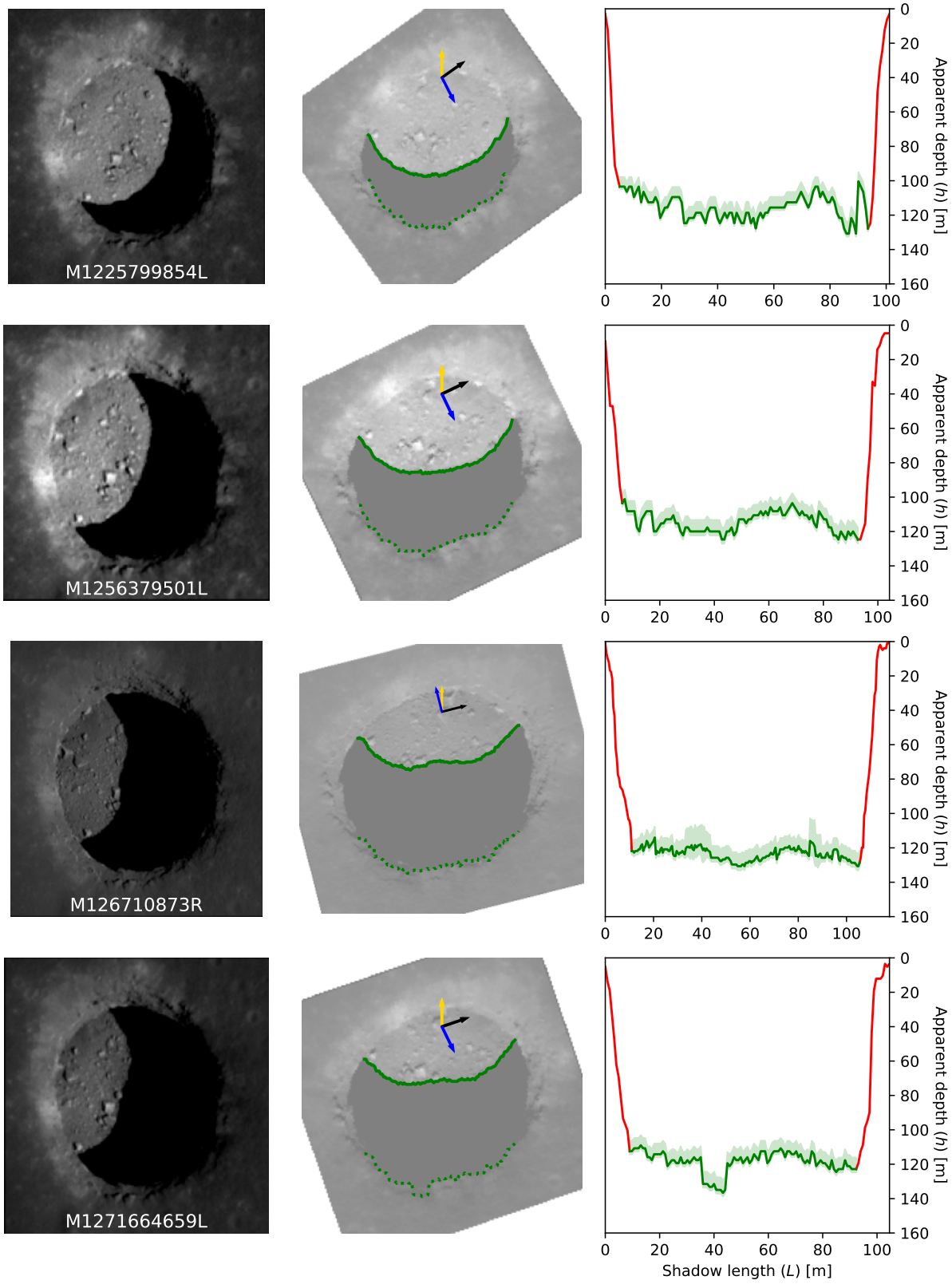

Figure 4.16: (Continued)

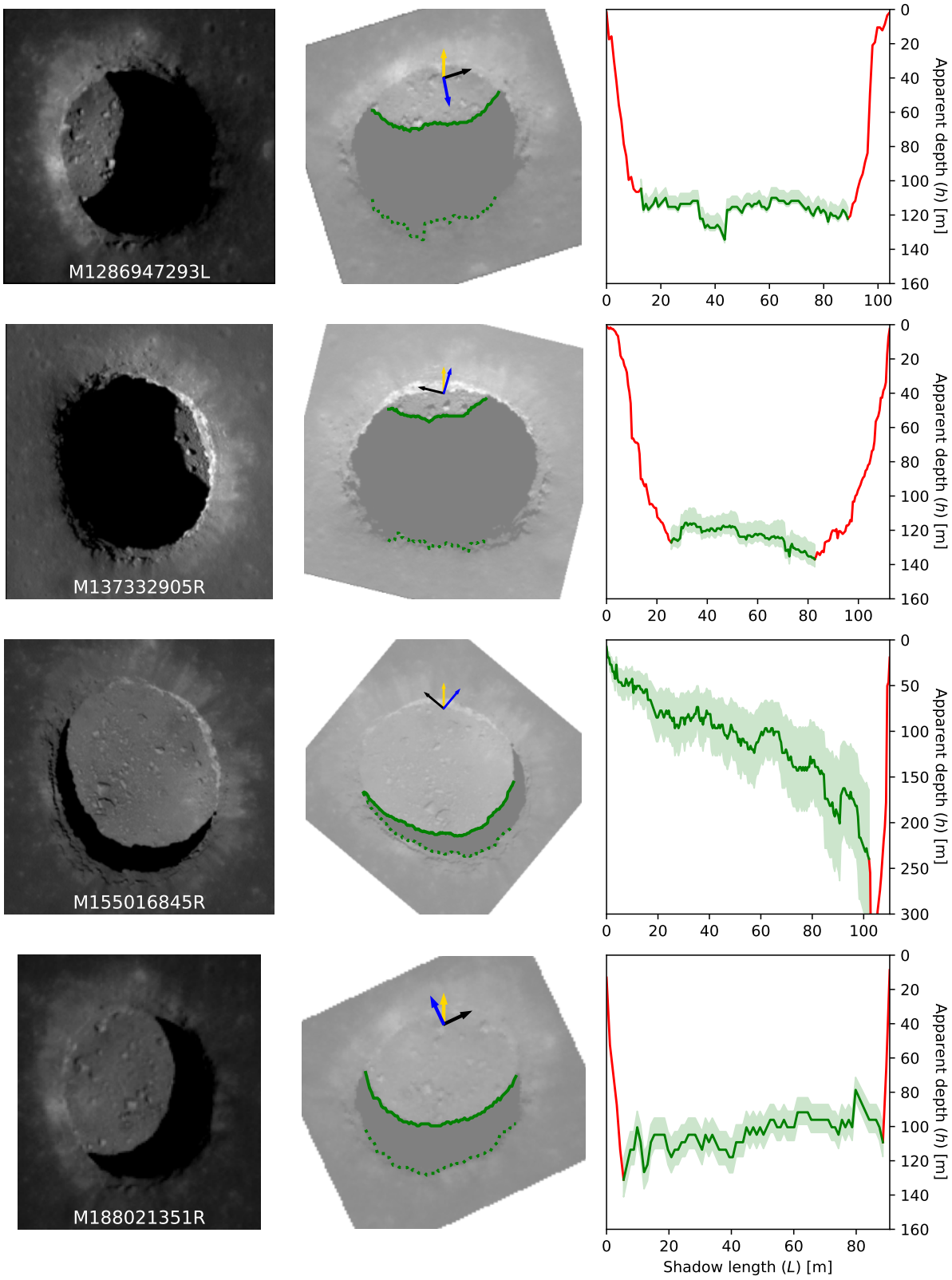

Figure 4.16: (Continued)

other images. PITS has been able to perform this domain change thanks to the iterative use of silhouette analysis since it can adapt to the number of distinct regions within a given image. The same quantitative testing procedure as in Section 4.3.1.5 will be required for further LPA pits to be able to say that PITS' automated shadow extraction performs as well, if not better, on Lunar pits within LROC NAC imagery.

In terms of the h profiles calculated by PITS, the images in which the shadow width is approximately equal to half of the MTP's D all predict a flat floor with depths ranging between 100–140 m. This depth range is partly due to shadows being cast from different stages of the funnel, meaning that the relative apparent depth is being taken from an initially higher elevation. Accounting for Δh, this range agrees with singular literature values of approximate funnel and inner pit depths of 20 and 105 m, respectively (Wagner and Robinson, 2014), and a total depth of 133 m (Wagner and Robinson, 2022). This is in spite of M1136335326R having an ε of 4.36°, suggesting that the ε-correction proposed in this work has also reduced the h variation of a Lunar pit.

In the case of the two images (M137332905R and M155016845R) exhibiting an eastwards illumination angle, with ω and ε angles ranging between 267.38–270.01 and 4.83–8.37°, respectively, LROC NAC is able to marginally peer underneath the eastern overhanging rim of the MTP. In doing so, the h profile from M137332905R can observe a gradual sloping towards the eastern rim with a maximum elevation change across the entire profile of ≈20.5 m. Again, considering Δh, this would fall within the range of 10–20 m quoted in Wagner and Robinson (2022). Interestingly, the left-most region of the h profile for M137332905R corresponds to where the shadow is cast into the north-eastern area of the floor, which, according to Wagner and Robinson (2022), is a topographic high-point at 115 m in depth. At this location, the minimum h predicted by PITS was $115.6^{+3.7}_{-7.1}$ m, thus agreeing with their value (see Section 4.3.3 for a description and derivation of these uncertainties).

Despite having a lower ε than M137332905R, PITS predicts a much greater h of close to 250 m within the image M155016845R. However, Δh is much more significant in this image due to the larger γ, of which $\Delta\varepsilon_{\shortparallel}$ (and consequently Δh) is a factor. In fact, the maximum

h predicted by PITS is a value of $239.6^{+74.6}_{-75.6}$ m. Therefore, given this uncertainty and the fact that this is the only near-nadir observation to date where the shadow edge falls against the south-eastern rim of the MTP, it is not possible to know from this profile alone whether the scales of these depths are genuine. Although, were they to be, they may be an indication of the sub-surface cavity entrance predicted by the point clouds of Wagner and Robinson (2022); Zhou et al. (2024) and the radar simulations of Carrer et al. (2024).

4.5 Key-Point Summary

- Pit Topography from Shadows (PITS) is an automated tool for calculating apparent depth (h) profiles of Martian and Lunar pits/skylights using just a single satellite image. It does so by employing k-means clustering and silhouette analysis to detect shadow pixels that are then measured along the shadow's width and corrected for non-nadir observations.
- Thanks to its automated shadow extraction methodology, PITS calculates h profiles along the entire length of the shadow, whereas previously only single values for h have been provided (Wyrick et al., 2004; Cushing et al., 2015; Wagner and Robinson, 2021).
- PITS also generates geo-referenced shapefiles of its detections, which can be used to increase the contrast of shadow pixels in the search for overhanging rims and possible cave entrances.
- Across 19 shadow-labelled red-band HiRISE images of atypical pit craters (APCs) from the Mars Global Cave Candidate Catalogue (MGC^3), PITS' automated shadow extraction achieved a recall (R) of 99.6% and a precision (P) of 94.8%, equating to an F_1-score of 97.1%.
- PITS has been applied to 123 red-band HiRISE images of 88 APCs, which revealed an improvement in the h variation due to emission angle correction, and also found 10 good candidate cave entrances on Mars due to their h profiles. This also highlighted the issue with comparing different h values—in that they may have been derived from width measurements where the shadow was cast to a different location within a pit's uneven floor.
- PITS has also successfully analysed the Mare Tranquillitatis Pit (MTP), proving that its shadow extraction methodology has transferred to a Lunar example. Across 10 LROC NAC images, PITS agrees with several depth estimates and topographic observations available from the various literature.

Chapter 5. Searching for New Candidate Cave Entrances on the Moon with Deep Learning

5.1 Introduction

Pits and skylights are particularly timely for their potential to signify the presence of and provide access to underground cavities such as lava tubes. Lava tubes are sub-surface conduits found sub-parallel to terrestrial surfaces that have been formed by the current or historic flowing of lava (Gadányi and van der Bogert, 2015). Intact lava tube systems can reveal themselves to the surface when their roofs become unstable and collapse to form circular, coalesced or elongated pits that can appear along the path of the tube. As a result, lava tubes have been suggested as possible locations for future Lunar bases (Coombs and Hawke, 1992).

Lava tubes are most commonly formed in smooth pahoehoe basaltic (smoother, lower viscosity, slower velocity) lava flows, but can also be found in aa (rougher, higher viscosity, higher velocity) flows (Léveillé and Datta, 2010). The six main mechanisms for the formation of lava tubes are: i) over-crusting of lava channels by coagulation of floating cooled slabs; ii) lateral accretion and eventual closure of lava channel shelves; iii) shallow inflation between pahoehoe sheets and downward erosion; iv) inflation and draining of an aa flow nucleus; v) deep inflation between older lava flow layers enlarged by thermal erosion and breakdown; vi) thermal erosion along a sealed fracture (see Sauro et al. 2020, Figure 2). Processes i) and ii) are also not mutually exclusive and can occur in conjunction with each other (Sauro et al.,

2020).

Speleologists have explored numerous lava tubes on Earth in locations known for volcanic activity, such as the Canary Islands (Sauro et al., 2019), Korea (Kim et al., 2022) and Sicily (Calvari and Pinkerton, 1999). The Ronald Greeley Centre for Planetary Studies at Arizona State University hosts a digital database of 1,084 terrestrial lava tubes from 34 countries across six continents (RGCPS, 2025). On the Moon, the total collapse of lava tubes has long been proposed as an explanation for the formation of sinuous trough-like structures called 'rilles' (Greeley, 1971). The same may also be true for the 'canali' seen on the Venusian surface (Gregg and Greeley, 1992). Martian analogues are also believed to be emanating from the various shield volcanoes (Léveillé and Datta, 2010).

In scenarios where physical exploration is not readily possible, such as on the Moon or Mars, the identification of the various collapse morphologies mentioned above relies upon remote-sensing data. Currently, there are two main catalogues for pits and skylights on the Moon and Mars. The Mars Global Cave Candidate Catalog (MGC^3) has been discussed in detail within Chapter 4. Whereas Chapter 4 briefly mentioned the existence of the Lunar Pit Atlas (LPA; Wagner and Robinson 2021). The Lunar Pit Atlas (LPA) gives the positions, morphometric parameters, and descriptions of 281 Lunar pits (Wagner and Robinson, 2021). Despite the majority of LPA features being found within impact melt ponds and thus unrelated to lava tubes, there remain 21 pits found within the highlands or maria.

In terms of larger collapses, Hurwitz et al. (2013) catalogued 193 sinuous rilles in mare regions of the Moon that could be observed in Lunar Reconnaissance Orbiter Wide Angle Camera (LROC WAC; Robinson et al. 2010) and Kaguya Terrain Camera (TC; Kato et al. 2008) visual imagery (100 and 10 m/px, respectively). 6 of these sinuous rilles also appear to be related to partial lava tube collapses. In a similar study for Mars, Crown et al. (2022) mapped the western flank of Alba Mons for potential lava tube collapses using the Mars Reconnaissance Orbiter (MRO) Context Camera (CTX; Malin et al. 2007) global mosaic (Robbins et al., 2023), resulting in 331 examples found to have a mean length of 36.2 km. Sauro et al. (2020) performed a comprehensive morphometric comparison between known lava

tubes on Earth and proposed analogues on the Moon and Mars. Thanks to the availability of high-resolution satellite imagery and digital elevation models (DEMs), Sauro et al. (2020) found that Martian and Lunar lava tubes could be 1 to 3 times more voluminous than their terrestrial counterparts. The presence of intact Lunar lava tubes deep underneath the surface has also been inferred through mass-deficit (Chappaz et al., 2017) and radar data (Kaku et al., 2017).

With regard to the Moon, potential caves have already been the subject of a proposed space mission in Moon Diver (Nesnas et al., 2023). An unsuccessful submission to the 2019 NASA Discovery program, Moon Diver is a mission concept in which a rotor-craft tethered to a lander would land near the Mare Tranquillitatis Pit (MTP), abseil down its wall and determine the composition and mineralogy of the Lunar mare bedrock. Three context cameras would also provide information on whether the void beneath the MTP is, in fact, an evacuated intact lava tube that could become a future Lunar base. As such, the structural stability of lava tubes (Blair et al., 2017; Theinat et al., 2020) and the possibility of pressurising a small example with breathable air (Martin and Benaroya, 2023) have both been modelled for a range of roof thicknesses and materials.

Perhaps due to the abundance of high-quality satellite data, there have been multiple studies employing automated techniques to detect and analyse planetary surface features related to the collapse of underground cavities. Firstly, the aforementioned LPA was derived in part thanks to the automated PitScan tool (Wagner and Robinson, 2014). PitScan operates by applying a threshold to Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) Calibrated Data Record (CDR) images to extract shadow regions on the Moon and then considers how the brightness of pixels changes along the Sun's line of sight across them to determine if they are positive (e.g. boulders) or negative elevation features. However, in order to restrict the types of elevations that can cast significant shadows, only images with a solar incidence angle (α) of less than 50° could be fed through PitScan. This meant that PitScan could not be used far outside the ±50° latitude range. Figure 5.1 plots an estimate of the coverage of LROC NAC images with $\alpha < 50°$ prior to the publication of

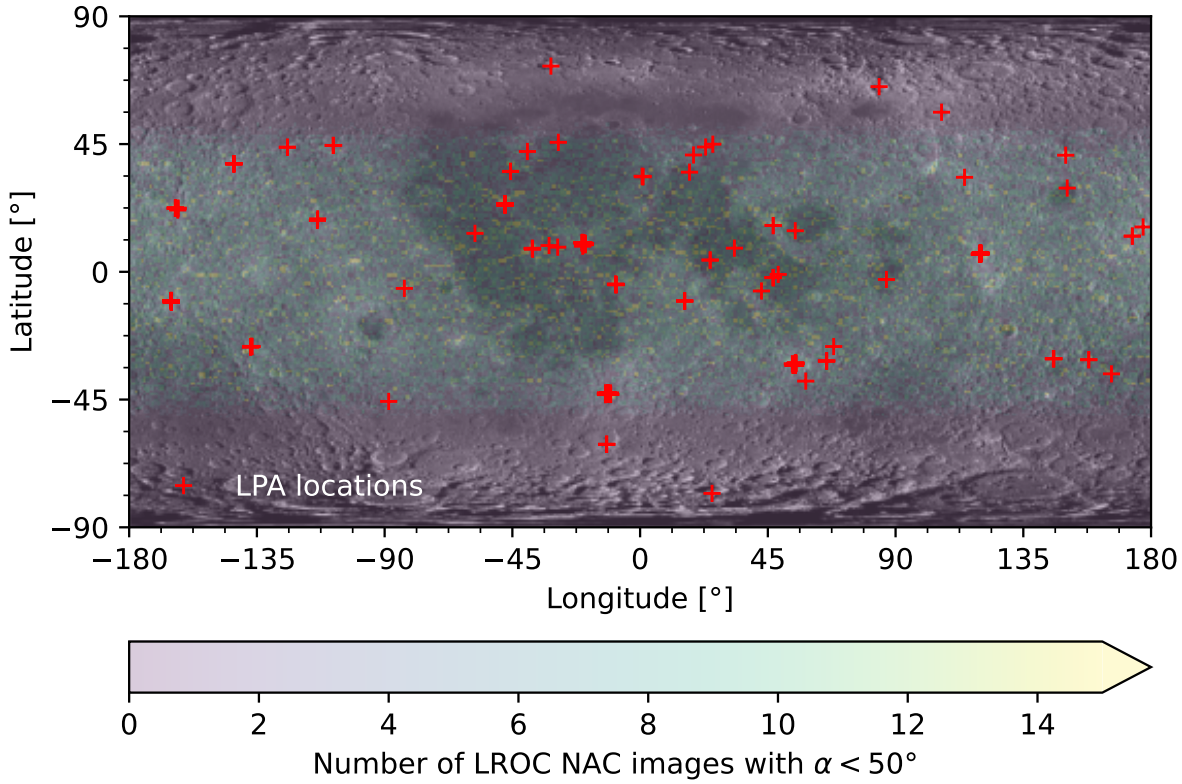

Figure 5.1: Estimate of the global coverage of LROC NAC images with solar incidence angles (α) of less than 50°. The locations of all Lunar Pit Atlas (LPA) features are given as red crosses. Base-map is the LROC WAC global mosaic (Speyerer et al., 2011).

PitScan and the LPA (Wagner and Robinson, 2014).

Nodjoumi et al. (2021, 2023) have developed the DeepLandforms platform for training deep learning (DL) models to detect planetary landforms in satellite imagery. DeepLandforms provides a workflow to go from data preparation and labelling to model training, validation and inference. The first use case for DeepLandforms was training a Mask R-CNN model to detect Martian pits by labelling MGC^3 features in MRO High Resolution Imaging Science Experiment (HiRISE; McEwen et al. 2007) imagery. Nodjoumi et al. (2023) specifically use the `mask_rcnn_R_50_FPN_3x` model from Facebook AI's Detectron2 model zoo (Wu et al., 2019), which has a ResNet50 backbone with a feature pyramid network that has been pre-trained on the ImageNet dataset. While the model struggled in testing when it had been trained on several morphologically similar classes, training on only skylights and impact craters achieved a testing F_1-score of 81.1%.

Alternatively, Watson and Baldini (2024) trained a simple convolutional neural network (CNN) on screenshots of MGC^3 features within the MRO CTX global mosaic (Robbins et al., 2023) in JMARS (Christensen et al., 2009) in order to detect further potential cave entrances (PCEs) on Mars. Whilst training/validation accuracies appeared high, 61 new and 24 already known (out of a total of 341 previously catalogued) PCEs were found out of a total of 10,834 detections when applied to five volcanic regions—resulting in precision and recall rates of just 0.78% and 7.03%, respectively. This represents a decent method for identifying regions to look for PCEs, but lacks the completeness for absolute regional mapping.

Whilst there has been no work yet on automatic detection of Lunar pits or skylights, Zhang et al. (2024) have used a combination of elevation (LOLA/Kaguya co-registered DEM at 59 m/px; Barker et al. 2016), slope (derived from DEM) and visual imagery (LROC WAC upscaled to DEM resolution) data to train a DL model based upon the DeepLabv3+ encoder-decoder architecture to detect Lunar sinuous rilles. This multi-modal approach achieved an F_1-score of 93.67% during testing, with significant improvement compared to if the model was trained on a single channel of the above sources.

Insomuch as there has been great effort at manually or automatically mapping pits, skylights and other lava tube-related collapses, there is yet to be a study using DL to search for pits and skylights on the Moon that may play a role as cave entrances. Given the global availability of high-resolution LROC NAC images and the aforementioned mapping exercises, there is high-quality data for creating a training dataset. As such, the aim of this work was to train DL models upon a combination of Lunar and Martian imagery in order to detect individual Lunar pits and skylights. In order to maximise the chance of detecting previously uncatalogued features which could provide access to sub-surface cavities, these models were applied to select mare regions which satisfy one or more of the following conditions: i) the region had not been completely mapped with PitScan—either due to high-latitudes or the unavailability of data at the time; ii) it appears in the global LROC WAC mosaic to exhibit sinuous or linear depressions; iii) they are known to have large sub-surface mass-deficit according to Chappaz et al. (2017). The PITS tool, as described in Chapter 4 and Le Corre et al. (2023a), is also

Body	Source	Feature Class	Images	Labels
Moon	Lunar Pit Atlas (LPA)	Impact melt	42	68
		Mare	47	53
		Highland	15	15
	Hurwitz et al. (2013)	Lava tube-related rilles	11	37
	Klem et al. (2014)	Apollo 11–16 sites	11	10
	Haase et al. (2019)	Apollo 17 site	1	0
	Sauro et al. (2020)	Lava tube collapses	27	25
Mars	Mars Global Cave Candidate Catalog (MGC^3)	Atypical pit crater (APC)	126	235
		Generic pit (pit)	4	37
		Skylight (sky)	20	66
		Small, rimless pit (srp)	15	23
	Watson and Baldini (2024)	Potential cave entrance	2	17

Table 5.1: Sources of training, validation and testing data of pits and skylights for the Moon and Mars, along with the numbers of acquired LROC NAC/HiRISE images and labelled features. The LPA mare quantities include the 5 LROC NAC images containing the Mare Tranquillitatis Pit (MTP), which were used for testing. The number of images given for the proposed lava tube collapses in Sauro et al. (2020) are the amount used to produce the mosaics used for testing data. All other data was used for training and validation.

used to derive apparent depth profiles where there are appropriate LROC NAC observations of any prospective detections. Finally, the models will be directly inferred upon a recent release of HiRISE RDRV11 images in order to determine their preliminary performance on the Martian surface. The following work presented in this chapter (prior to Section 5.2.3.2) was published in the journal *'Icarus'* in June 2025 (Le Corre et al., 2025a).

5.2 Dataset

5.2.1 Training and Validation Data Sources

A combination of several manually- and automatically-produced catalogues was used in this study as a source of 'ground truth' for pits and skylights found on the Moon and Mars. A full breakdown of the numbers of LROC NAC/HiRISE image products and labelled features amassed from each of the following sources of Lunar and Martian training, validation and testing data is given in Table 5.1.

5.2.1.1 Moon

The documentation accompanying the LPA gives the details of the 2 or 3 LROC NAC images which were used in cataloguing and measuring each impact melt, mare and highland feature. Thus, these images are of sufficient resolution and have appropriate sensing conditions to be useful for model training. The raw Experimental Data Record (EDR) versions of these LROC NAC images were obtained from NASA's Planetary Data System (PDS; see Robinson 2009). However, the small size of impact melt features meant that only the highest resolution image available for each feature was utilised, as is explained further in Section 5.3.1.

To train models on more elongated pits, LROC NAC EDR images were also acquired for all six lava tube-related sinuous rilles (#24, 37, 39, 75, 83 and 175) mapped by Hurwitz et al. (2013). Since there has been repeated LROC NAC coverage of these regions, LROC NAC EDR products were selected so that each feature had roughly one high-sun and one low-sun observation each.

All EDR images were then processed from raw data to high-level map-projected images using the Integrated Software for Imagers and Spectrometers (ISIS) and Geospatial Data Abstraction Library (GDAL) software packages. EDR products were first converted to ISIS cubes (using `lronac2isis`), then echo-corrected (`lronacecho`), radiometrically-corrected (`lronaccal`), map-projected (`cam2map`), and converted to GeoTiff format (`gdal_translate`). Technical Note TN-2 in Appendix A provides some BASH code for performing these processes for a batch of raw LROC NAC EDR image products.

Due to a user-defined shape model being used when map-projecting, there was a misalignment between the processed images and the LROC NAC shapefile footprints that are available through the PDS' Lunar Orbital Data Explorer (ODE). An accurate footprint of the image as it appears on the surface is necessary for ensuring that no-data regions are not seen during training. Therefore, new footprint shapefiles were created by taking the processed image and setting all pixels with intensities greater than 0 to a value of 1 (using `gdal_calc.py`) to produce a new binary image. This binary image was then polygonised (with `gdal_polygonize.py`) to convert from raster to vector data.

In addition to self-processed LROC NAC imagery, controlled LROC NAC mosaics of the Apollo program landing sites (Klem et al., 2014), as well as the Apollo 17 controlled 0.5 m/px ortho-mosaic (Haase et al., 2019), have been utilised as a source of unlabelled (i.e. no target features present) images of the Lunar maria—with the aim of training on such data being a reduced rate of false positives. However, the Hadley Rille seen in the Apollo 15 mosaics was labelled as an example of an elongated pit (see Section 5.3.1 for more detail).

5.2.1.2 Mars

Knowledge of the locations of Martian pits and skylights was also required to be able to deduce whether there is any improvement when training with imagery of a different planetary surface. The primary source of 'ground truth' for Mars was MGC^3 features with relevance to the collapse of sub-surface cavities (i.e. no cracks, fractures or karst pits). Therefore, MRO HiRISE Reduced Data Record Version 1.1 (RDRV11) images, which contained MGC^3 atypical pit craters (APCs), generic pits (pit), skylights (sky) and small rimless pits (srp) were acquired from PDS (see McEwen 2007). This was done by cross-referencing the catalogue within QGIS (QGIS, 2025) with the MRO HiRISE RDRV11 footprint shapefiles available in the Mars ODE.

The same cross-referencing was performed for the 61 PCEs identified by Watson and Baldini (2024), although just 3 were found to have corresponding HiRISE RDRV11 imagery. Furthermore, only 2 of the 3 features appeared clearly as either a pit or a skylight under the increased resolution of HiRISE. Unfortunately, no useful HiRISE RDRV11 images could be found for the lava tubes mapped by Crown et al. (2022). This is likely because only the start, end and mid-points of the lava tube path are available (Crown et al., 2022), but also due to many of the features being heavily modified by aeolian processes.

As the surveys above have focused mostly on the features which could provide access to caves, catalogued features were often found within chains of, or in proximity to, other uncatalogued pits. Therefore, the total number of labelled Martian pits and skylights is significantly larger than the number of corresponding HiRISE RDRV11 image products.

Whilst RDRV11 products are map-projected, there is an occasional error when viewing

and analysing them in GIS software, where an offset in the latitude direction is observed. This misplacement was corrected by replacing the values of the 'standard parallel' and 'centre latitude' parameters in the Proj4 string and reprojecting the image using the `gdal_translate` command line program. This meant the labelling could be carried out in QGIS, as is discussed further in Section 5.3.1, with the knowledge that the labels were in the correct location on the surface. Technical Note TN-3 in Appendix A outlines some BASH code for performing this correction for a batch of HiRISE RDRV11 images.

5.2.2 Testing Data Sources

In addition to separate training and validation datasets, it is also vital to have an additional dataset upon which the final, trained model can be tested as a means of determining its real-world performance. This testing dataset should be representative of the features that the model is tasked with detecting, which in this case are Lunar pits and skylights, which may be related to the collapse of lava tubes.

As a result, LROC NAC EDR images, which covered the sinuous collapse chains identified by Sauro et al. (2020) as potential lava tubes, were also collected. These include the Gruithuisen, Marius Hills A and Marius Hills B sinuous collapse chains. It was made certain that none of the lava-tube-related rilles that were seen during training were also included in this testing dataset. The EDR products containing these lava tube candidates were processed using the same procedure as those described in Section 5.2.1.1, except with the additional step of photometric correction (`lronacpho`). Photometric correction was performed to reduce the visibility of seams occurring when these individual processed images were stitched together with `automos` to form three separate uncontrolled mosaics. These self-produced uncontrolled mosaics are shown later in Figure 5.8.

5.2.3 Inference Data Sources

5.2.3.1 Moon

As mentioned in Section 5.1, eight Regions of Interest (RoIs) in the search for previously uncatalogued Lunar pits and skylights were selected from the various mare units on the Lunar surface as they satisfied a series of conditions. RoIs were chosen if they had not been completely mapped by PitScan due to being situated at latitudes beyond $\pm 50^{\circ}$ or insufficient coverage by LROC NAC at the time. These RoIs may have also exhibited linear collapses in the global LROC WAC mosaic or are regions with high degrees of sub-surface mass deficit according to Chappaz et al. (2017).

The chosen RoIs were the following: the Imbrian mare deposits inside the (a) Antoniadi, (b) Bel'kovich A, (c) Endymion, and (d) Lyot impact craters; (e) the Eratosthenian mare deposit inside Plato crater; (f) the Marius Hills region inside Oceanus Procellarum (between 10–15°N and 55–60°W), which hosts two known skylights (Haruyama et al., 2009) and high degrees of sub-surface mass deficit (Chappaz et al., 2017); (g) the high-latitude Mare Humboldtianum and (h) the Imbrian mare deposits within Poincaré Basin. In terms of area, these RoIs (labelled with the same letters in Figure 5.2) constitute ≈1.92% of the Lunar maria and just ≈0.23% of the entire surface.

The LROC NAC images (with resolutions of at least 1.5 m/px, $\alpha \leq 70^{\circ}$, and satellite emission angles, ε, of less than 5°) which overlapped each RoI were found by importing the geo-referenced shapefile of the corresponding mare unit(s) from the global geologic map by Fortezzo et al. (2020) into JMARS (Christensen et al., 2009). In using the 'Mosaic Select' tool within JMARS, the fewest number of LROC NAC images which would provide the maximum degree of coverage of the RoI could be selected. The raw products were then acquired through PDS and processed as in Section 5.2.1. In order to preserve disk space, however, the ISIS cubes were clipped to the wider latitude-longitude bounds of the corresponding RoI. Considering the relatively small size of these RoIs, a total of 2,157 raw LROC NAC images were downloaded and processed using the above procedure for these eight sites. Even with the specialist and

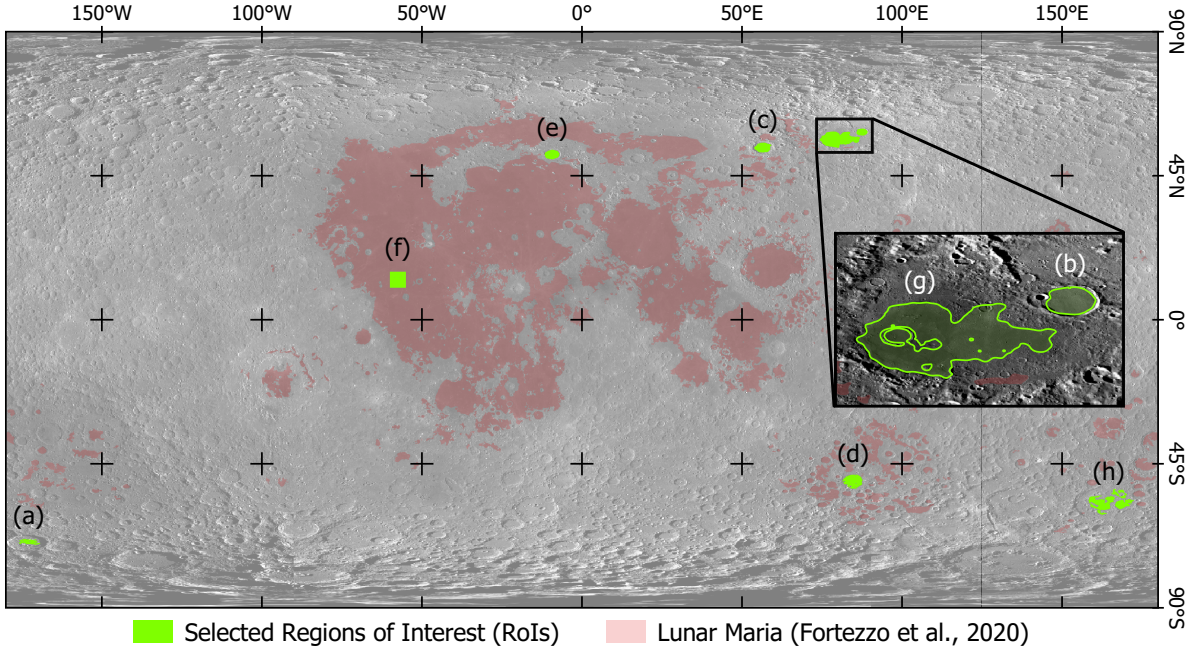

Figure 5.2: Global map of the Regions of Interest (RoIs) in the search for previously uncatalogued Lunar pits and skylights. (a, c, d, e) Antoniadi, Endymion, Lyot and Plato craters; (f) Marius Hills and (h) Poincaré Basin. The inset map highlights (b) Bel'kovich A crater adjacent to (g) Mare Humboldtianum. The Lunar maria from Fortezzo et al. (2020) yet to be surveyed are given in red. Base-map is the 100 m/px LROC WAC global mosaic (Speyerer et al., 2011).

high-performance computing systems provided by Information Services at the University of Kent, these products were processed at a rate of around 25 to 30 images per hour.

5.2.3.2 Mars

As mentioned in Section 5.1, the HiRISE RDRV11 red-band images disseminated within the January 2025 release have been acquired from PDS. This batch contains a total of 287 images with the same appropriate sensing conditions as above. Whilst this dataset still required downscaling to the common 1.5 m/px resolution, the recent creation of the images means that none suffered the same latitude misplacement in their embedded map-projection, as described in Section 5.2.1.2. However, large no-data regions appear in recent HiRISE images, since the middle RED4 channel failed in 2023 (McEwen et al., 2024)

5.3 Methodology

5.3.1 Training Dataset Creation

The first step in creating a labelled dataset of LROC NAC and HiRISE images was to address the difference in spatial resolutions between the two sensors. Between September 2009 and December 2011, while the Lunar Reconnaissance Orbiter (LRO) was still in its original 50 km circular polar orbit, the majority of LROC NAC images had a resolution of ≈0.5 m/px (Wagner and LROC Team, 2025). Since then, LRO has been in a fuel-conserving elliptical polar orbit with its periapsis near the Lunar south pole, meaning that images taken in the northern and southern hemispheres range in resolutions from 1.0–2.0 and 0.4–1.0 m/px, respectively (Wagner and LROC Team, 2025). Given that HiRISE has a best resolution of 0.25 m/px, there is a maximum of 8 times more detail in its images when compared to the worst available with LROC NAC. As such, all processed LROC NAC EDR images and projection-corrected MRO HiRISE RDRV11 images were downscaled to a common resolution of 1.5 m/px using cubic spline interpolation. This value was selected to maximise detail and minimise the number of instances where the LROC NAC resolution was any poorer. In the few cases where this was true, it meant that they were, in fact, upscaled, with the poorest resolution being ≈1.83 m/px. Downscaling to a consistent resolution was aimed at enabling the model to learn a sense of the scales of Lunar and Martian pits. However, this did mean that some of the catalogued features mentioned in Section 5.2 (particularly Lunar impact melt pits) could not be resolved at this scale, nor could they be confidently labelled.

Thanks to their map projection, these downscaled image products could then be loaded into QGIS and labelled by drawing polygons around the target feature so as to include the funnel, rim and floor regions of the feature. These polygons were then saved to a separate shapefile layer for each image, meaning that image products with multiple pits or skylights will have as many polygons in their shapefile. Each polygon was assigned an integer according to whether it appeared to be a 1) skylight (no raised rim; steep and/or overhanging walls; deep with a pronounced shadow) or a 2) pit (no raised rim; often larger, coalesced, or channel-

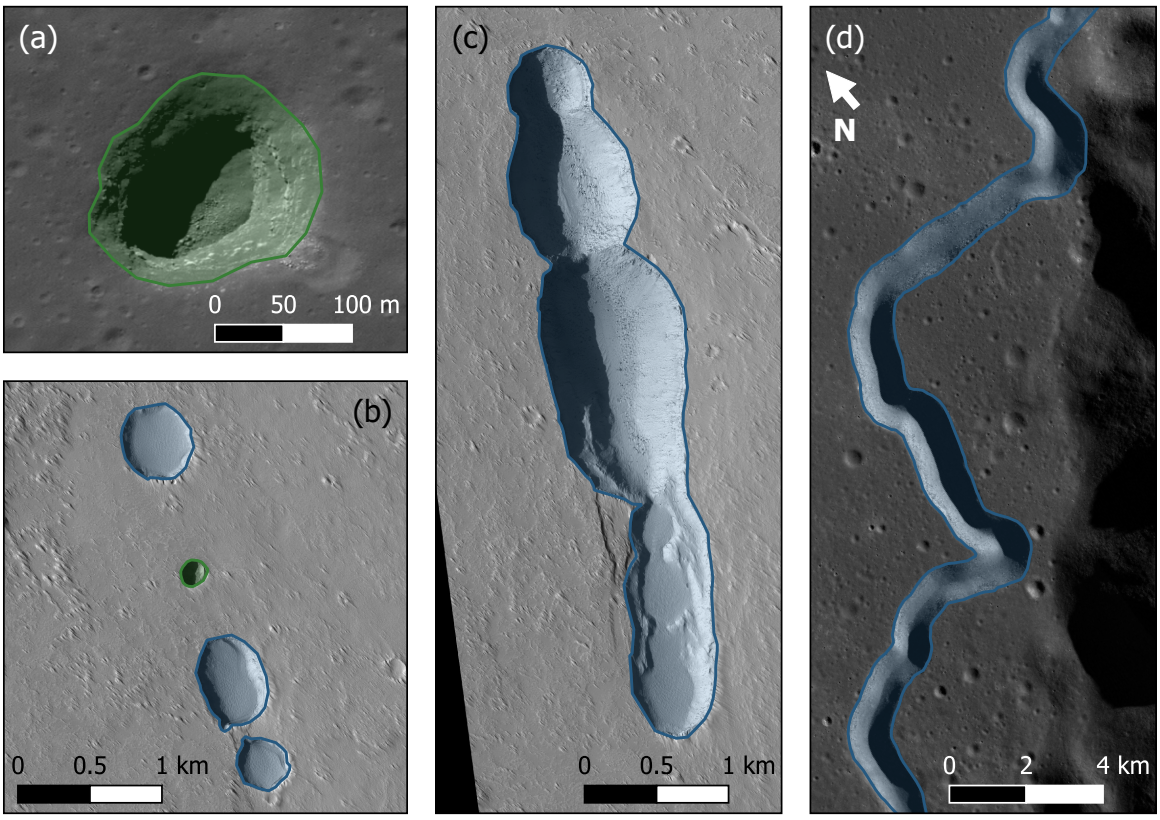

Figure 5.3: Examples of pits (blue) and skylights (green) on the Moon and Mars, along with their corresponding labels produced in this work. (a) LROC NAC image M138819477R of the skylight known as the Mare Ingenii Pit (MIP) in the Lunar Pit Atlas (LPA); (b) a Martian pit chain, as seen in HiRISE image ESP_011677_1655_RED, containing the Mars Global Cave Candidate Catalog (MGC^3) feature APC079; (c) a series of coalesced Martian pits in the same chain and image as (c); (d) A section of Hadley Rille close to the Apollo 15 landing site within the LROC NAC mosaic NAC_ROI_APOLLO15LOB_E259N0038.

like; all walls are visible; and may not have a pronounced shadow). Examples of the labelling procedure for a series of Lunar and Martian pits or skylights are given in Figure 5.3.

Since LROC NAC and HiRISE images are far too large to pass through a DL model (even after downscaling), they both require cropping into smaller tiles. This was done with a Python script that would loop through all features in the image and randomly generate coordinates for a 2,048×2,048 px tile, which aims to fully encapsulate it. For features such as Lunar rilles, whose extents may be larger than this given tile size, a coordinate grid is generated with 75% overlap between consecutive tiles in the longitudinal and latitudinal directions. A tile was then only included in the training dataset if one of the following two conditions was met for all

features found within it: i) 90% of a feature's area is contained within the tile, or ii) the area of intersection between the feature and the tile takes up more than 10% of the tile's area. This approach was necessary to ensure that as many features were appropriately contained within a tile while also maximising the number of features that could be seen during training. As mentioned in Section 5.2, when tiling the controlled LROC NAC mosaics of the Apollo landing sites, tiles which did not overlap any features whatsoever were allowed to be exported.

Tiling was also necessary for the imagery used for model testing and inference upon the eight RoIs. For the uncontrolled testing mosaics and five LROC NAC images containing the MTP, an overlap of 25% was utilised. An overlap between testing tiles allowed the Mask R-CNN model the maximum opportunity to detect some of the larger collapses, whilst not misrepresenting the amount that it would see when being applied to the images covering the RoIs. With regard to the RoIs, the processed images covering each site were tiled with no overlap whatsoever, since there was often a large degree of overlap between adjacent image products. Any tiles which did not improve the overall overage of a particular RoI were then removed from the final inference dataset. Tiling with no overlap and then removing tiles that did not provide additional coverage were both performed due to concerns regarding storage limits, since the total surface area of the RoIs is significantly larger than that of the testing regions. Therefore, as shown later in Section 5.2.3.1, some detections made within the Lunar RoIs have been made in multiple tiles originating from different LROC NAC image products. The same tiling with no overlap was also used for the HiRISE RDRV11 red-band images in the January 2025 PDS release.

As per the requirements for the DL model of choice (see Section 5.3.3), the polygon labels produced in QGIS were rasterised separately for each target class to create up to two images (also with a x-y size of 2,048 px) which contain pixel-level annotations of the target features found within the tile. In these annotation images, the background was set to 0, and the pixels corresponding to each instance of the given class were assigned a sequentially increasing integer value starting from 1.

Given the limited number of catalogued features, some data augmentation is necessary to

overcome the training dataset deficiency, but also to not risk over-fitting. As such, rotated versions of each image (clockwise by 90, 180 and 270°) were also included in training, thus increasing the number of images by a factor of 4. Images which were earmarked for training as opposed to validation also had the possibility of undergoing three further steps of data augmentation, each with a probability of 0.5 and no mutual exclusivity. Firstly, images could undergo Contrast-limited Adaptive Histogram Equalisation (CLAHE), which is a variant of Adaptive Histogram Equalisation which clips an image's histogram to a chosen value so as to minimise the amplification of noise in uniform regions (Pizer et al., 1987). Secondly, Gaussian noise could also be applied to the images with a σ randomly selected from a uniform distribution between 0.0075–0.0125. Lastly, the perspective of the image could also be transformed so as to simulate different viewing angles. This was done using the `perspective` function within PyTorch and a distortion scale randomly selected from a uniform distribution between 0.5–1.5.

After downscaling, tiling and data augmentation, the resulting LROC NAC and HiRISE tiles had to be split between those used for model training and those used for model validation. A split of 4:1 was chosen, with the proportion of annotated and unannotated tiles also being divided equally across the training and validation sets. Seeing as nearly twice as many unannotated tiles were exported, unannotated tiles were randomly removed from the final training/validation datasets such that the ratio of annotated to unannotated tiles was approximately equal. The resulting training and validation datasets are summarised in Table 5.2, with breakdowns of the number of image tiles which contained target features.

5.3.2 Creating Synthetic Moon Data

Up to this point, efforts have been made to make up for the relative lack of known pits and skylights on the Moon compared to Mars, including data augmentation and utilising images taken of the same features under different illumination conditions. However, the suspicion was that this would be insufficient. Therefore, the possibility of creating artificial or synthetic labelled training data for Lunar pits was explored to counteract this.

There is the possibility of generating a series of artificial DEMs containing Lunar pits by

Dataset	Purpose	Labelled	Empty
Moon + Syn. Moon	Training	1,152	1,149
	Validation	284	287
Moon + Mars	Training	1,156	1,152
	Validation	284	288
Moon + Mars + Syn. Moon	Training	1,808	1,802
	Validation	444	450

Table 5.2: Breakdown of the resulting training/validation datasets after downscaling, tiling, and data augmentation. The 'Moon + Syn. Moon' dataset consists of genuine LROC NAC data of Lunar pits and skylights, plus synthetic Moon samples generated by implanting augmented Martian features into empty LROC NAC images. 'Moon + Mars' contains only genuine LROC NAC and HiRISE data. 'Moon + Mars + Syn. Moon' contains all available training data created in this work.

drawing from real elevation data and illuminating it with randomly selected solar incidence and azimuth angles (α and ϕ). However, illuminating artificial elevation models would require significant computational effort to generate enough DEMs from stereo imagery and would likely require training an additional model (such as a generative adversarial network, see Section 2.2.4). Instead, synthetic Moon training samples were generated by implanting augmented Martian target features into LROC NAC tiles known not to include any.

This process involved extracting tiles which did not contain any target features (i.e. could be considered 'empty') from the larger LROC NAC products containing LPA (only mare and highland) features and lava-tube-related rilles. Empty tiles were not retrieved from LROC NAC images containing LPA impact melt pits since the surrounding surface was often too heavily fractured. A Python script was then written, which would loop through all empty LROC NAC tiles and randomly select a HiRISE tile from the training dataset. If the HiRISE tile contained multiple target features, then only one was randomly selected at a time for implantation.

The HiRISE tile corresponding to the chosen feature would then be cropped to the extent of the target feature using the training annotations described in Section 5.3.1. The cropped tile is then rotated by any angle drawn from a uniform distribution. The rotated, cropped tile is then resized in its x and y directions by separate factors, both drawn from a normal distribution with a mean and σ of 1 and 0.15, respectively. This allows for circular Martian

features to become more elliptical and vice versa. These same transformations are applied to the training annotation so that the pixel-level labels are placed correctly for the rotated and resized feature. Using the training annotation as a binary mask allows for the pixels solely corresponding to the target feature to be extracted. The transformed Martian feature is then implanted into the empty LROC NAC tile, thereby replacing the original pixel values at a random position such that the entirety of the feature falls within it. By using the result of passing the feature's binary mask through a Gaussian filter as weights, the implanted HiRISE pixels and the surrounding empty LROC NAC image are averaged to blend the two regions.

Prior to the step where the HiRISE tile is cropped to the extent of the selected feature, histogram-matching was conducted in order to better assimilate with the brightness and contrast of the Lunar surface. Histogram-matching is the process of adjusting the pixel values of an input image (HiRISE tile) such that the Cumulative Distribution Function (CDF) of its histogram is equal to that of a reference image. The Python script then displayed five options according to how the histogram-matching was conducted to allow for a manual decision on the one, if any, that was most convincing. These options were: i) no histogram-matching; ii) histogram-matched to the randomly selected empty LROC NAC tile; iii) average between i) and ii); iv) histogram-matched to the average histogram of all annotated LROC NAC tiles in the training dataset; v) average between ii) and iv).

This endeavour produced an additional 812 training images, with each having one feature. Figure 5.4 displays some of the synthetic Moon tiles that were generated as part of this work and subsequently used for training a DL model to detect Lunar pits and skylights. Significantly more synthetic images were able to be created using Martian skylights (628) compared to Martian pits (184). This is a result of two factors. Only Martian pits or skylights that were fully contained within the extents of the source HiRISE tile were allowed to be implanted, and many of the coalesced, elongated pits were not. Of those that satisfied this condition, several pits exhibited aeolian features in their floors, which were too prominent to be convincing as a Lunar pit.

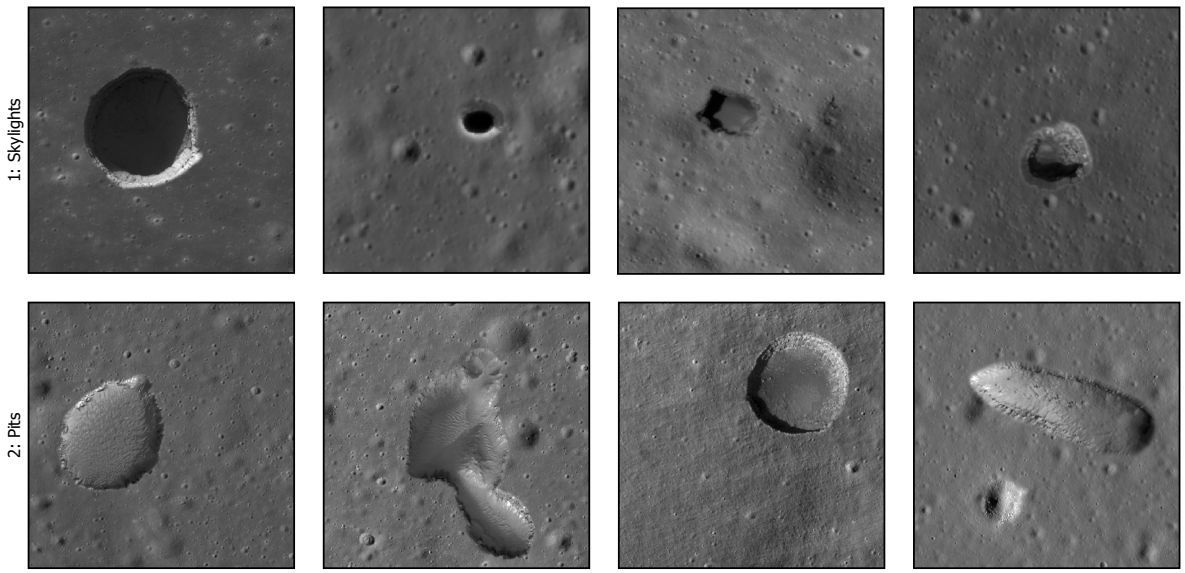

Figure 5.4: Examples of synthetic Moon training samples produced by implanting transformed and histogram-matched Martian pits/skylights into 'empty' LROC NAC tiles. The samples shown here have been cropped to show the implanted Mars feature in detail within the context of the Moon's surface and, as a result, are not to scale.

5.3.3 Model Architecture and Evaluation

The chosen model for this work is a Mask region-based convolutional neural network (Mask R-CNN; He et al. 2017). As discussed in Chapter 2, R-CNNs are a series of deep convolutional neural networks (DCNNs) which perform object detection (OD) by proposing several regions of interest that are likely to contain target features and feeding these through a neural network for detection and, in some cases, classification (Girshick et al., 2013). Mask R-CNN further develops upon Faster R-CNN (Ren et al., 2015) by including additional fully convolutional networks that are tasked with predicting a mask of the target feature simultaneously with its bounding box (He et al., 2017), thus performing instance segmentation (IS). A two-stage (region proposal and classification) OD model, such as Mask R-CNN, was chosen over other single-shot models, such as YOLO, as maximum performance in detecting pits and skylights was preferred over speed. The additional output of a predicted mask with Mask R-CNN also estimates the size of the footprint of a pit or skylight, as opposed to the latitude-longitude extent provided by a bounding box. Mask R-CNN has also seen extensive use in detecting surface features in planetary remote-sensing data (Ali-Dib et al., 2020; Alshehhi and Gebhardt,

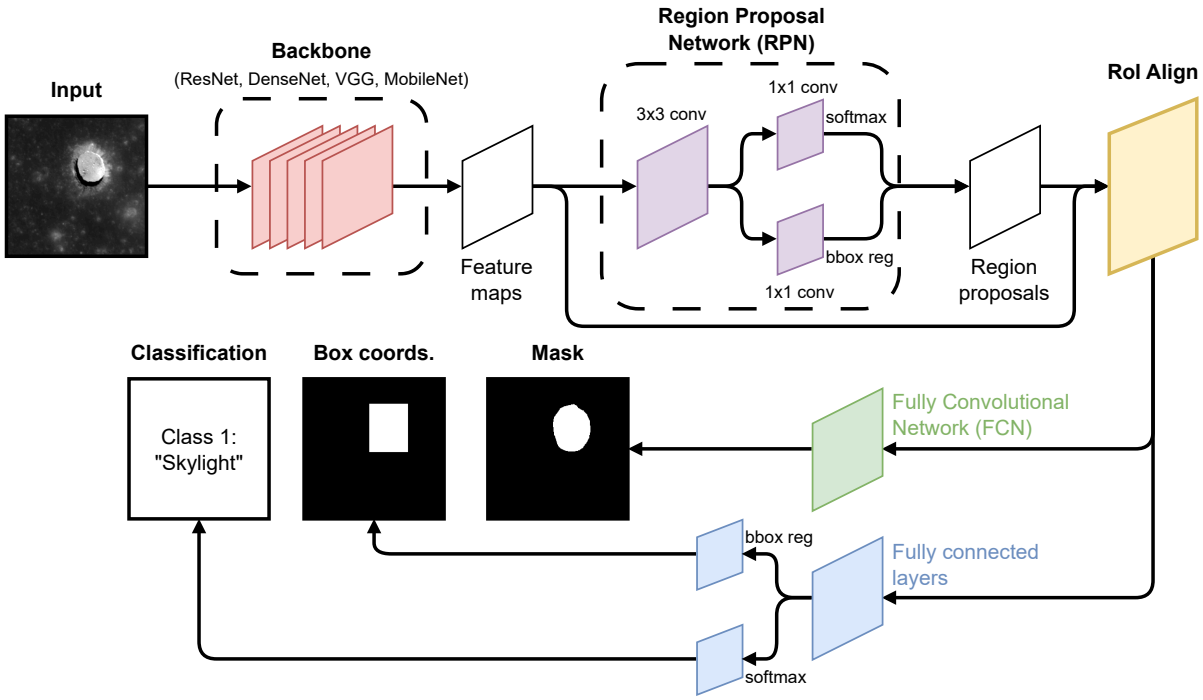

Figure 5.5: Diagram of the Mask R-CNN (region-based convolutional neural network) model in the context of detecting pits and skylights on the Moon (He et al., 2017). The hypothetical example shows the Mare Tranquillitatis Pit (MTP) being correctly detected as a 'skylight' with a bounding box and predicted mask.

2022; Prieur et al., 2023; Rubanenko et al., 2021; Nodjoumi et al., 2023). A diagram of the Mask R-CNN model, within the context of detecting pits and skylights within satellite imagery of the Moon and Mars, is given in Figure 5.5.

As the training framework has been built from scratch in Python using PyTorch[1], there is the ability to swap out the Mask R-CNN's default backbone model (ResNet50) for a series of other CNNs, which are all available with pre-trained model weights (on the ImageNet dataset; Deng et al. 2009). The backbone model is responsible for learning the feature maps that will be used by the rest of the model for region proposal, detection and classification. DenseNet121, 169 and 201 (Huang et al., 2016); VGG16 and 19 (Simonyan and Zisserman, 2014); as well as ResNet101 and 152 (He et al., 2016) were all chosen to be trained since they were compared by Dawson et al. (2023) for scene classifications of different carbonate rocks. Dawson et al. (2023) also assessed the InceptionV3 network, but persistent errors meant that this could not

[1] The Mask R-CNN model is provided by the PyTorch framework with a default, pre-trained ResNet50 backbone and a feature pyramid network (FPN). See `https://pytorch.org/vision/main/models/mask_rcnn.html` for more details.

be implemented as the backbone for Mask R-CNN in this work. The ResNet18, ResNet34 and MobileNetV2 (Sandler et al., 2018) models were also trained as they were quick to implement and train (due to having few convolutional layers). By default, the ResNet backbones had the ability to be trained with the implementation of feature pyramid networks (FPNs). However, similar errors to those for InceptionV3 meant that they could not be used for the non-ResNet backbones.

Before evaluating the model's performance, all detections on the validation samples undergo non-maximum suppression (NMS). NMS is a technique for iteratively removing repeated bounding box detections of the same objects made by OD models. The algorithm starts by sorting all detections by confidence score (from highest to lowest) to create set A. An empty set, B, within which the filtered detections will be placed, is also initialised. NMS then iterates through A and determines whether the bounding box A_i overlaps any of the lower-scoring boxes by a sufficient Intersection over Union (IoU, see Table 2.1) threshold. All of these overlapped lower-scoring boxes are removed from set A, and the box A_i is added to set B.

Performance metrics have been calculated over the course of the training process and exported to TensorBoard for simple visualisation. These metrics include the training losses for the various elements of the Mask R-CNN model (classifier, box regressor, mask, object-ness, RPN box regressor); the average recall (R), precision (P) and F_1-score (F_1) of the bounding boxes and predicted masks upon the training samples; and the average R, P and F_1 of the boxes and masks upon the validation samples. Recording P, R and F_1 separately in this way allows easier assessment of whether the model has over-fit to the training data, wherein the performance on the validation samples would be significantly worse than on the training samples.

Whilst P, R and F_1 are defined in Table 2.1, the numbers of true positives (TP), false positives (FP) and false negatives (FN) are derived differently for the bounding boxes and predicted masks. A bounding box is assessed to be a TP or FP depending on whether it was detected with the correct class, a confidence score over 50%, and the bounding box overlaps the extent of the corresponding label with a sufficiently strict IoU of 65% or more.

Any labelled target features with no detections meeting these criteria are denoted as FNs. Meanwhile, the TPs, FPs and FNs for the predicted masks are calculated in the same manner as in Section 4.3.1.5, whereby they are compared with the corresponding label to find where the pixel classifications agree or not. Of course, the confusion matrix for the predicted masks can only be calculated in cases where a mask has been predicted for a feature which has a label (i.e. a TP bounding box detection).

Having been introduced in Section 2.2.7, there are a series of hyper-parameters (such as the optimiser, learning rate, batch size and the number of epochs) which need to be defined before training any DCNNs. AdamW was chosen as the optimiser (Loshchilov and Hutter, 2017), with a learning rate of 0.0001. A batch size of 16 was used as it was the maximum value possible with the large tile size and computing resources available. All models were trained for 100 epochs (the number of times all batches have been passed through the model), but this was terminated early if the average between validation F_1-scores for the bounding box and predicted mask worsened after three consecutive epochs.

5.4 Results and Discussion

5.4.1 Training Performance

The extent to which the contribution of Mars and synthetic Moon data improved the Mask R-CNN's ability to detect Lunar pits and skylights required assessment before the models could be deployed upon the Lunar RoIs. For this, three Mask R-CNNs were trained (within the parameter space described in Section 5.3.3 and with the default ResNet50 backbone) upon different combinations of the various data sources. The first Mask R-CNN was trained on genuine and synthetic Moon data, the second on genuine Moon and Mars data, and the third on genuine and synthetic Moon plus Mars data. No model was trained on just genuine Moon data since there would be far too few training samples to lead to a successful model. The F_1-scores achieved by each of these models upon the validation samples over the course of 100 epochs are shown separately for the bounding boxes and predicted masks in Figure 5.6.

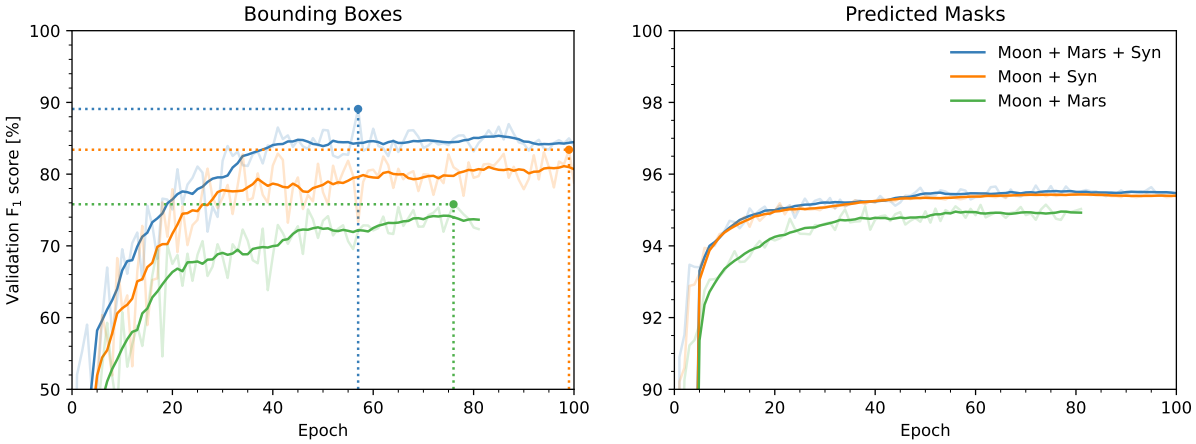

Figure 5.6: The F_1-scores (F_1) achieved upon validation samples across 100 epochs by three separate Mask R-CNN models (with a ResNet50 backbone). These models were trained on different combinations of Moon (LROC NAC), Mars (HiRISE) and Synthetic Moon (HiRISE pixels implanted into 'empty' LROC NAC images) data. (Blue) Moon, Mars and Synthetic Moon; (orange) Moon and Synthetic Moon; (green) Moon and Mars. The solid lines each represent the rolling average over 10 epochs of the fainter profile. The dotted lines annotate the F_1 that was achieved for the boxes and masks when the average between the two was at its maximum (i.e. the best version of each model). These metrics were calculated using confidence and IoU thresholds of 50 and 65%, respectively.

Figure 5.6 shows that the highest validation F_1 was achieved by the Mask R-CNN trained upon all available datasets, reaching 89.1 and 95.6% for the bounding boxes and predicted masks (at epoch 57), respectively. A high performance on the Moon, Mars and synthetic Moon datasets was to be expected since there is more data to train with in this instance. Ideally, there would be sufficient examples of Lunar and Martian pits that this comparison could be made independent of dataset sizes. As a result, it cannot be said whether the addition of Martian and synthetic data would improve the performance of these models in all cases, but it has in this scenario, where training data was limited. However, it should be noted that the entire training/validation dataset was randomly split irrespective of whether the data was of the Moon (genuine or synthetic) or Mars. What this meant is that a Mask R-CNN trained upon Moon and Mars data was also validated upon data of the Moon and Mars. Therefore, it may be that synthetic Moon or Martian pits and skylights were easier to detect than their genuine Lunar analogues, rather than their inclusion leading to improvements in the overall performance. Nevertheless, the deficiencies of the training datasets that do not include Martian or synthetic Moon data are still expected to be the predominant reason for

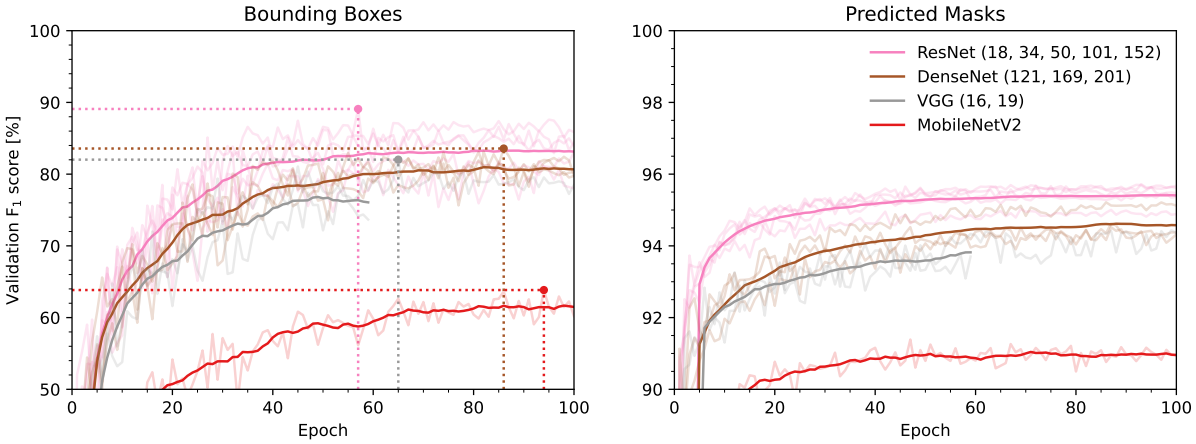

Figure 5.7: The F_1-scores (F_1) achieved upon validation samples across 100 epochs by several Mask R-CNN models using different backbones for feature map generation. The various model families are: (pink) ResNet, (brown) DenseNet, (grey) VGG, and (red) MobileNet. The line-style meaning and the confidence/IoU thresholds used are all identical to those in Figure 5.6. These models were all trained on Moon, Mars and Synthetic Moon data since this achieved the highest performance in Figure 5.6.

the better performance of the Mask R-CNN trained on all available data.

Following this investigation of the contribution of different data sources was a comparison of the various backbones described in Section 5.3.3. Figure 5.7 plots the validation F_1 of the bounding boxes and predicted masks for each of these backbones. Figure 5.7 shows that the ResNet family performed the best in this specific training scheme, with the ResNet50 model achieving the highest F_1-scores. Incidentally, this is also the same model which performed best in Figure 5.6, as shown by the blue profile.

The higher performance of the ResNet family could be attributed to the residual blocks within its architecture and the inclusion of an FPN, thus allowing for region proposals of pits and skylights at a range of scales, with greater knowledge of slight differences between other features, such as impact craters. The ResNet family performing best in this work contradicts what was observed by Dawson et al. (2023), which was that the ResNet models (ResNet50, 101, 152) performed the worst (in terms of accuracy averaged over all target classes) out of those trained upon the smallest dataset of 7,000 carbonate core images (i.e. the closest in size to this work). The poor performance of the ResNet models was consistent for both of the optimisers that they trained with: Stochastic Gradient Descent (SGD; Ruder 2016) and

Adam (Kingma and Ba, 2017). The best-performing CNNs in Dawson et al. (2023) were, in fact, VGG16 and 19. However, the training of VGG16 in this work was terminated early due to a decrease in performance in three successive epochs. This disparity between which model family performed best highlights how the difference in dataset sizes, optimisers (with AdamW being used in this work) and reported metrics can cause direct comparisons between the various CNNs used here as Mask R-CNN backbones and by Dawson et al. (2023) for image recognition to be tenuous.

5.4.2 Testing Performance

Since the Mask R-CNN model trained on Lunar, Martian and synthetic Moon data with a ResNet50 backbone has been shown in Section 5.4.1 to achieve the highest average F_1 upon the validation dataset, this model now requires testing on the data described in Section 5.2.2 to get a picture of its real-world accuracy before it is applied to the data of the eight Lunar RoIs. This particular model is referred to as ESSA (Entrances to Sub-Surface Areas), which is also the Cornish (*Kernewek*) name for the author's home-town of Saltash, Cornwall (*'Essa'*).

Figure 5.8 gives a visual representation of the results of applying ESSA on this data, which contains the Gruithuisen, Marius A and Marius B lava tube collapse chains proposed by Sauro et al. (2020), as well as five separate LROC NAC images of the MTP under different illumination conditions. It should be noted that some figures in this chapter display detection confidence scores of 100.0%. This is not an assertion that the probability of the detection being incorrect is zero, but rather that they were made with a score of 99.95% or more. The results (calculated with the same confidence and IoU thresholds as during training) equate to an F_1 of 82.4 and 93.7% for the bounding boxes and predicted masks, respectively. The

Figure 5.8: (Following page) Visual performance of ESSA on Lunar pits and skylights. The bounding boxes and predicted masks are overlaying self-produced mosaics of the lava tube candidates: Gruithuisen, Marius A and B; and five LROC NAC observations of the Mare Tranquillitatis Pit (MTP). True positives (TPs) are those with the correct classification, a confidence score over 50%, and an IoU with the label of 65% or more.

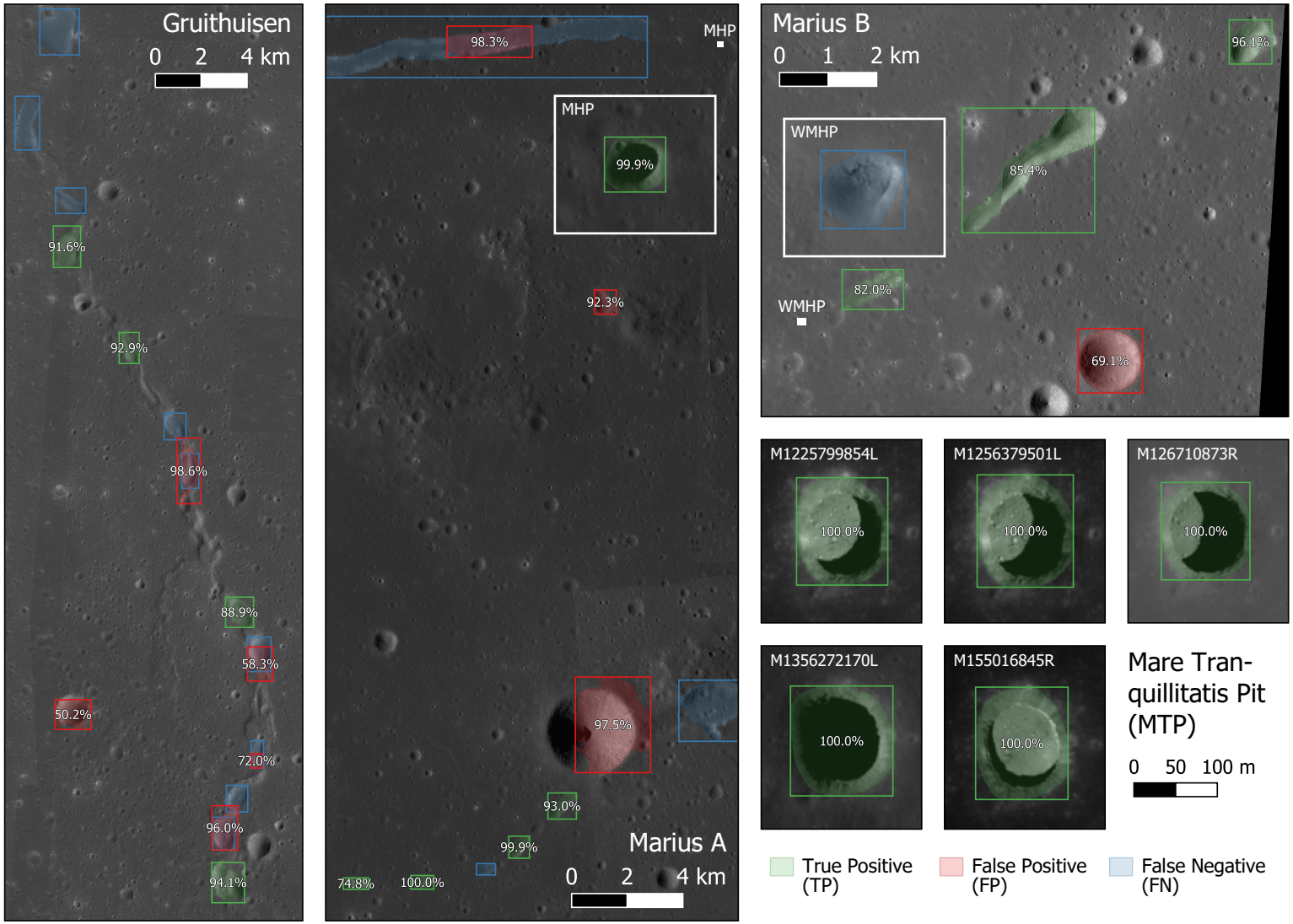

Figure 5.8: (Caption on previous page)

F_1-scores for the bounding boxes were calculated for each site individually and then averaged (to negate the inequality in the number of targets), whereas for the masks, they were averaged globally. Whilst the P and R of the predicted masks remained comparably high (over 91% for both), there was some difference between the P (86.3%) and R (78.8%) of the bounding boxes.

These discrepancies between the training and testing performance metrics were expected as the testing data contained several features which were likely to be more confusing for ESSA. For example, the skylight within the Marius B mosaic (known as the West Marius Hills Pit, WMHP, in the LPA; Wagner and Robinson 2021) exhibits no visible shadow, which, as one of the main characteristics of a skylight, partly explains why it was not detected. A lack of a significant shadow was true for several of the other missed pits in the Gruithuisen and Marius A chains. Moreover, impact craters on the same scales as the ≈3 km-diameter example detected as an FP within Marius A were rarely seen during training. Five out of the nine FPs Figure 5.8 also overlap an FN, but by an insufficient IoU.

Nonetheless, the MTP was detected in all five LROC NAC images with a score above 99.9%, with no FPs found elsewhere in any of the images. Meanwhile, the skylight in the Marius A mosaic (known as the Marius Hills Pit, MHP, in the LPA; Wagner and Robinson 2021) was also confidently detected, despite the fact that it was completely covered by shadow. Sufficient pits have also been detected in the proposed lava tube collapses such that a sinuous alignment can be seen, which is necessary to be able to search for any aligned chains in the RoIs.

This testing has also allowed for reassessment of the confidence score threshold. A higher score threshold limits the number of FPs and hence reduces the time taken to manually check whether the detections were genuine. Although the selection of this threshold should not be at the expense of missing too many true pits or skylights. Figure 5.9 plots the change in the P, R and F_1 of both target classes detected and shown in Figure 5.8 against an increasing score threshold. The decrease in R, as shown in Figure 5.9, is due to the increasing threshold causing TPs to require higher and higher confidence scores to remain as such. However, the decrease

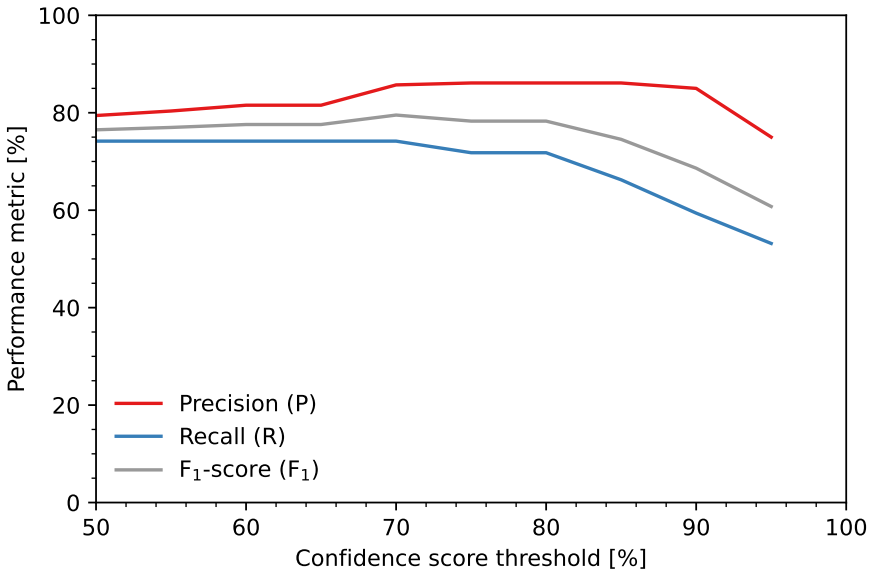

Figure 5.9: A plot of the change in the precision (P), recall (R) and F_1-score (F_1) of the testing results displayed in Figure 5.8 due to an increasing score threshold. The P, R and F_1 were calculated individually (with an IoU threshold of 65%) for each image, then averaged across all images and both target classes.

in P is a result of this reduction in the TPs, despite some high-scoring FPs still meeting the threshold. A score threshold of 80% appears to be the loosest approximate threshold before which R begins to decrease noticeably, leading to true pits or skylights being missed. Therefore, detections with a confidence score below 80% were removed when applying ESSA to the eight RoIs.

5.4.3 Inference on Eight Lunar Mare Regions of Interest (RoIs)

Across the eight Lunar RoIs, ESSA detected a total of 420 features: 116 as class 1 (skylights) and 304 as class 2 (pits). These detections required manual inspection to confirm whether they were genuine Lunar pits or skylights. Although it should be stated that many of these were repeat detections of the same features within multiple overlapping tiles or were separate sections of the same elongated pit or rille.

The same morphological descriptions used during labelling were also used during this manual inspection. Other factors, such as the Sun's position relative to the surface, were also

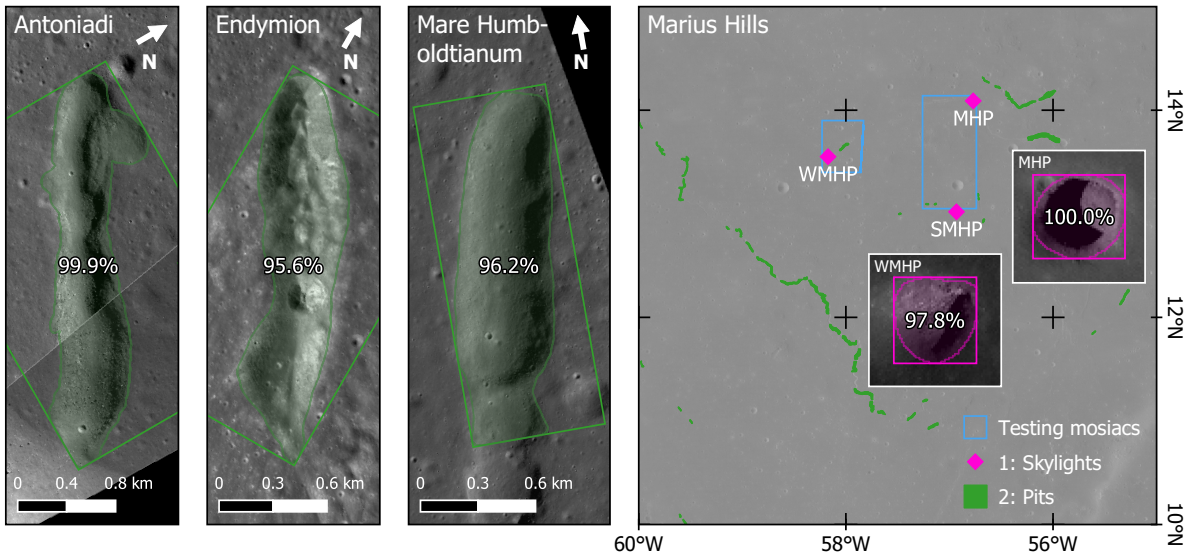

Figure 5.10: Examples of some of the elongated pits and rilles detected in four out of the eight RoIs. The map of Marius Hills plots the detected pits (in green) and skylights (in pink) upon the LROC WAC global mosaic (Speyerer et al., 2011). The inset maps zoom in on the detections of the Marius Hills and West Marius Hills skylights (MHP and WMHP). The footprints of the Marius A and B mosaics used in testing are also given in blue.

considered. The Sun's position was not known by ESSA when making detections, which makes sense of some of the FPs observed. Examples of some of the common false positives were small and/or debris-ridden impact craters, wrinkle ridges, and boulders, which were not frequently seen during training.

Figure 5.10 presents some of the elongated pits and Lunar rilles, which were correctly detected by ESSA across the eight RoIs, including a mapping of such features across the entire Marius Hills region. Regarding Marius Hills, applying ESSA to all 4,905 image tiles covering the region took only 4.35 mins, proving that it can analyse data at far quicker time-scales than could be reached through manual methods. Unfortunately, no new, clear instances of linearly or sinuously aligned pits, suggesting the presence of a partially collapsed lava tube, were detected by ESSA within the RoIs. However, given the relatively small sample size considered in this work when compared to the total Lunar maria, this is not to say that no more sinuous pit chains exist. Despite this, Figure 5.10 shows how ESSA was able to detect the MHP and WMHP within the vicinity of the Marius A and B collapse chains, despite the WMHP being missed in Section 5.4.2.

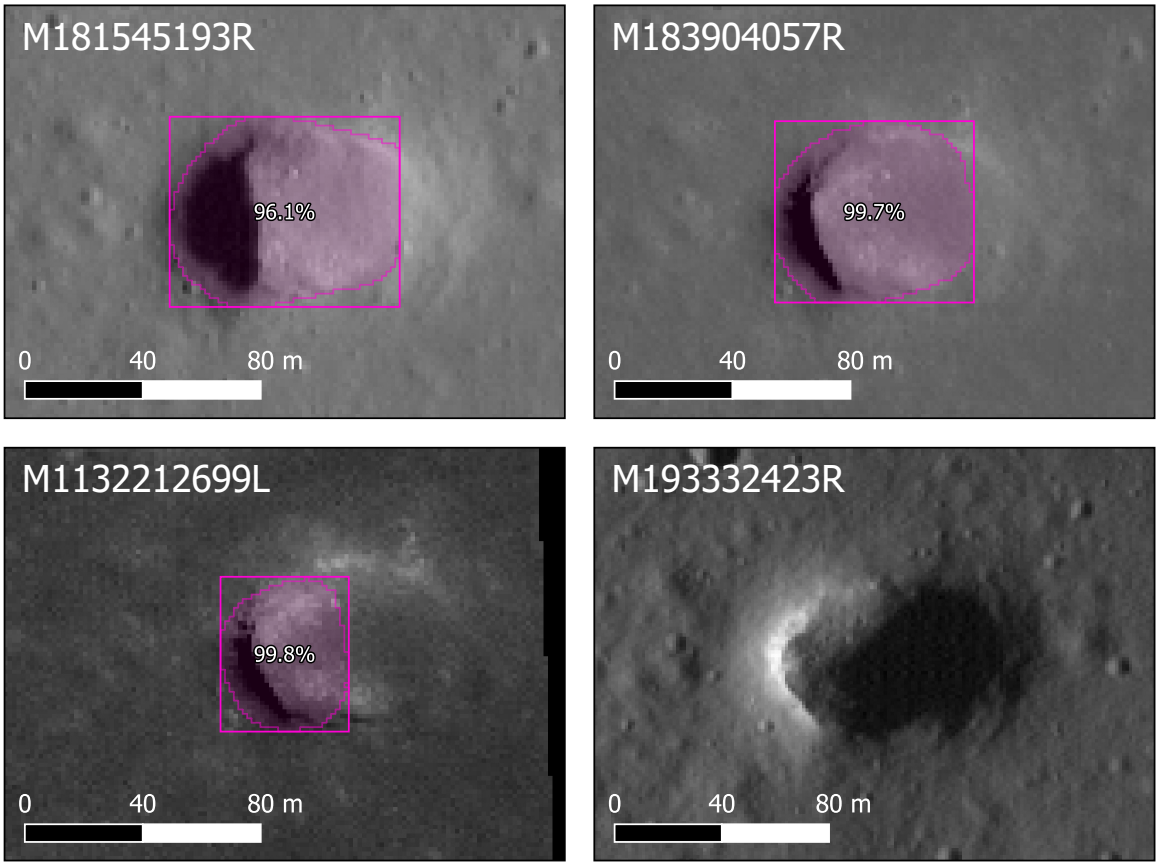

Figure 5.11: Four LROC NAC images of the skylight (nicknamed the 'South Marius Hills Pit', SMHP) detected by ESSA within the Marius Hills region of the Moon. (Top row, bottom left) The LROC NAC images in which the SMHP was detected. (Bottom right) An observation of the SMHP with the Sun's illumination coming from the opposite direction and revealing a possible overhanging rim.

Figure 5.10 also presents the location (at 13.0188°N and 56.9322°W) of a previously uncatalogued skylight detected by ESSA within the Marius Hills region. The presence of an uncatalogued skylight at such a low latitude is intriguing since there was LROC NAC imagery of the feature available (albeit with lesser coverage) at the time when PitScan was used in the creation of the LPA (Wagner and Robinson, 2021). It is also surprising that the skylight was not mentioned by Sauro et al. (2020) given its proximity (within ≈4.3 km) to the Marius A proposed lava tube collapse. Figure 5.11 displays this skylight, which is suggested to be named the South Marius Hills Pit (SMHP) in convention with the LPA, within the three LROC NAC images that it was detected within by ESSA—all with scores above 96.1%. SMHP is an elliptical skylight, with semi-major and minor axes of roughly 80 and 60 m, respectively.

A possible explanation for why the SMHP had been missed up until now is the lack of contrast between the feature's floor and the surrounding surface, as exhibited by the images in Figure 5.11. Therefore, the sensitivity of the PitScan tool may have been too weak to detect the shadow as being caused by a depression. One instance in which this is not true is M193332423R, shown in Figure 5.11, which was not likely to have been passed to PitScan due to its significant α (75.6°). In this near-nadir image ($\varepsilon = 1.2°$), the westward direction of the Sun's illumination reveals that its western wall is not visible, which may mean that it is overhanging or providing access to a sub-surface cavity. Figure 5.12 plots the apparent depth (h) profiles calculated by the PITS tool (see Chapter 4) upon M193332423R and another LROC NAC image of the SMHP having sufficient resolution. Both h profiles suggest a bowl-shaped floor with a maximum h of $14.0^{+0.1}_{-0.7}$ and $15.4^{+0.1}_{-0.8}$ m for M181545193R and M193332423R, respectively (see Section 4.3.3 for a derivation of these uncertainties). With the Sun's illumination coming from opposite directions, and the shadow being cast to a similar location within the floor, it can be said that the rims to the east and west of SMHP are approximately level. The northeastern rim of the SMHP may also provide the safest access to any robotic or crewed explorers with an average slope of ≈31° down towards the floor.

Aside from Marius Hills, another skylight was detected by ESSA within Bel'kovich A crater. At first glance, the skylight appeared as an impact crater itself, due to the wide (≈170 m in diameter) funnel around the detection and the fact that shallow depressions at such high latitudes can exhibit prominent shadows. However, upon closer inspection of the highest resolution LROC NAC image available (M128727218L), shown in Figure 5.13, it appears as though the western flank of the shadow is being cast against a steep or overhanging wall due to the high reflectance immediately opposite the shadow edge. Therefore, this feature is named the Bel'kovich A Pit (BAP), which is the closest to the Lunar north pole of any skylight found within a mare unit (at 58.6979°N, 87.5990°E).

As is also true for the SMHP, further observations at high resolutions, with different sensing conditions, are required to determine whether BAP potentially harbours a cave entrance. However, this is unlikely to be possible for BAP with current visual imagery datasets due to

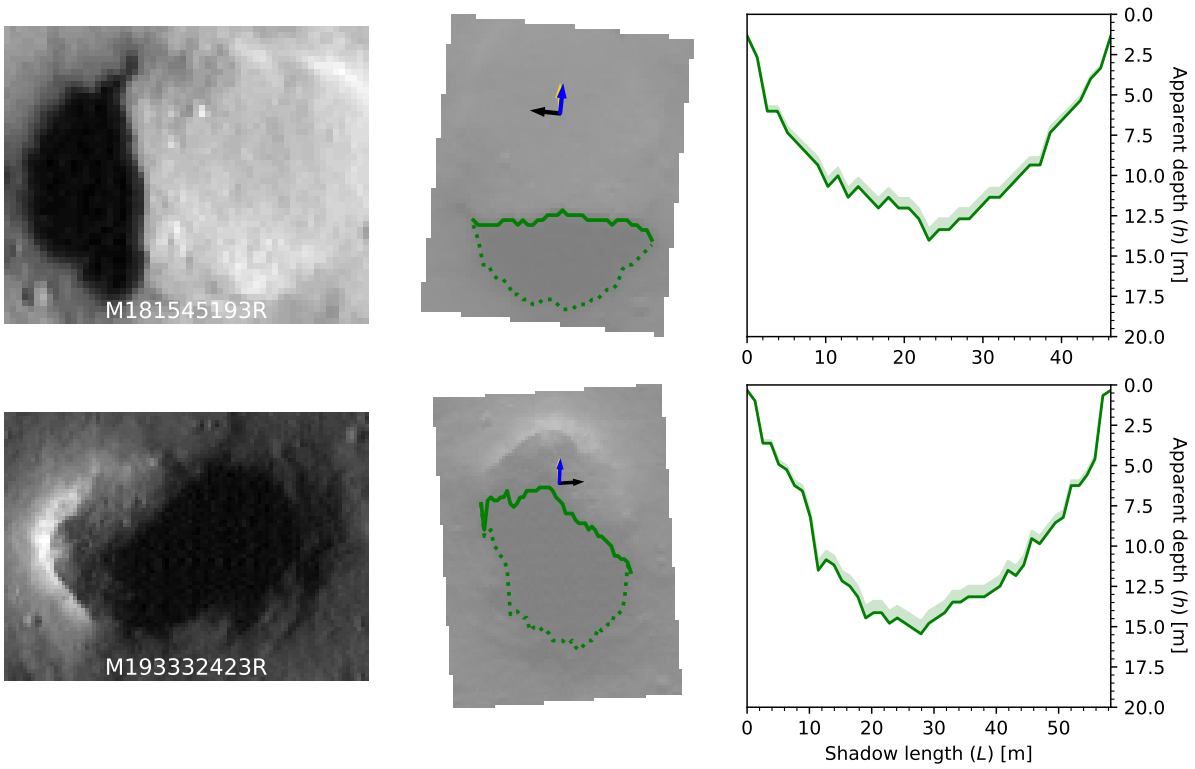

Figure 5.12: Apparent depth (h) profiles automatically calculated using the PITS tool upon two cropped LROC NAC images of the South Marius Hills Pit (SMHP). The arrows denote the directions of the Sun's illumination (yellow), LRO's sensing (blue), and north (black). M181545193R was cropped to SMHP's extents using the bounding box detection made by ESSA.

the periapsis of LRO's elliptical orbit now being at the south pole and the BAP being at a latitude where the sub-solar point at noon will always be to its south. Therefore, other data sources, such as infrared or radar imagery, may need to be employed.

Nonetheless, Figure 5.14 plots the PITS h profiles for two LROC NAC observations of the BAP. The first image (M1233723761R) has been cropped to the extent of the funnel surrounding the BAP, as the low Sun angle means that the shadow's edge falls within it. Whereas the second instance applies PITS to the image discussed above due to its high resolution. From this, it can be deduced that the combined h of the BAP and its surrounding funnel is ≈26 m. With consideration for the fact that the Sun is casting the shadow from different points along its length, the maximum h of BAP (minus the funnel) was $14.0^{+0.1}_{-0.7}$ m. There is also a steep slope (≈46°) from the eastern rim towards the possible overhang in the west.

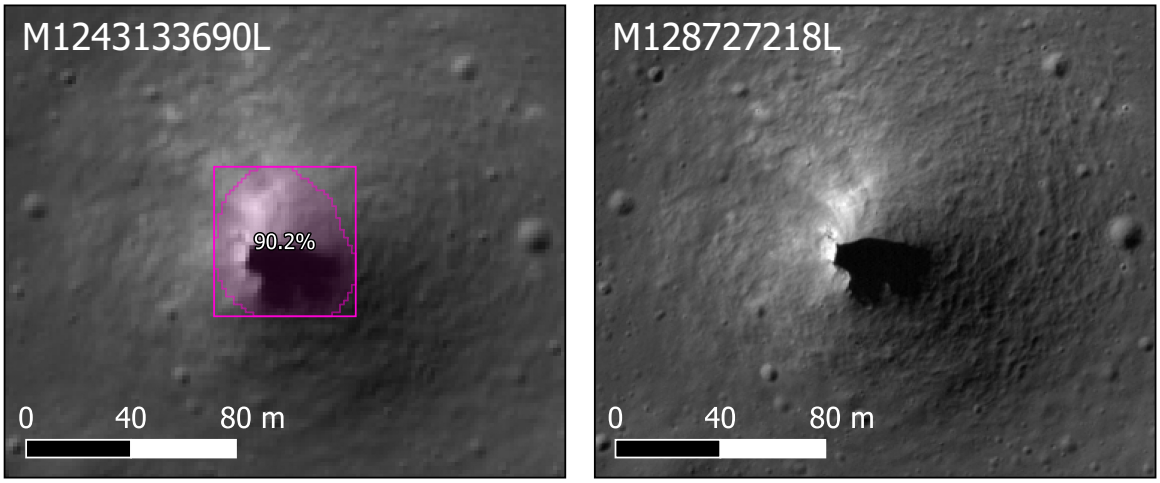

Figure 5.13: Two LROC NAC images showing the skylight and candidate cave entrance detected by ESSA within Bel'kovich A crater (nicknamed the Bel'kovich A Pit, BAP). (Left) The LROC NAC image M1243133690L (downscaled to 1.5 m/px), within which the feature was detected. (Right) The highest resolution LROC NAC image available for this feature, highlighting the possible overhanging rim.

A list of pit and skylight detections made by ESSA within the Lunar RoIs is given in Appendix C, along with their latitude-longitude coordinates, average confidence score, the LROC NAC images they were detected within, and a brief description. This table does not contain any of the rilles shown in Figure 5.10 since this mapping is available as a geo-referenced shapefile on Zenodo (Le Corre, 2025b).

5.4.4 Inference on HiRISE Reduced Data Record Version 1.1 January 2025 Release

Figure 5.15 presents a sample of the detections of pits (in green) and skylights (in pink) made by ESSA within the January 2025 PDS release of HiRISE RDRV11 red-band images. Not shown in Figure 5.15 are some of the erroneous detections, which were caused by dead pixels within the HiRISE sensor and the fact that the middle RED4 channel failed in 2023, meaning that valid data regions are now cut in two (McEwen et al., 2024).

Figure 5.15 (a) shows one of several instances where a section of the Cerberus fossae on Mars has been detected as a pit by ESSA (within ESP_085856_1895_RED). The sub-parallel

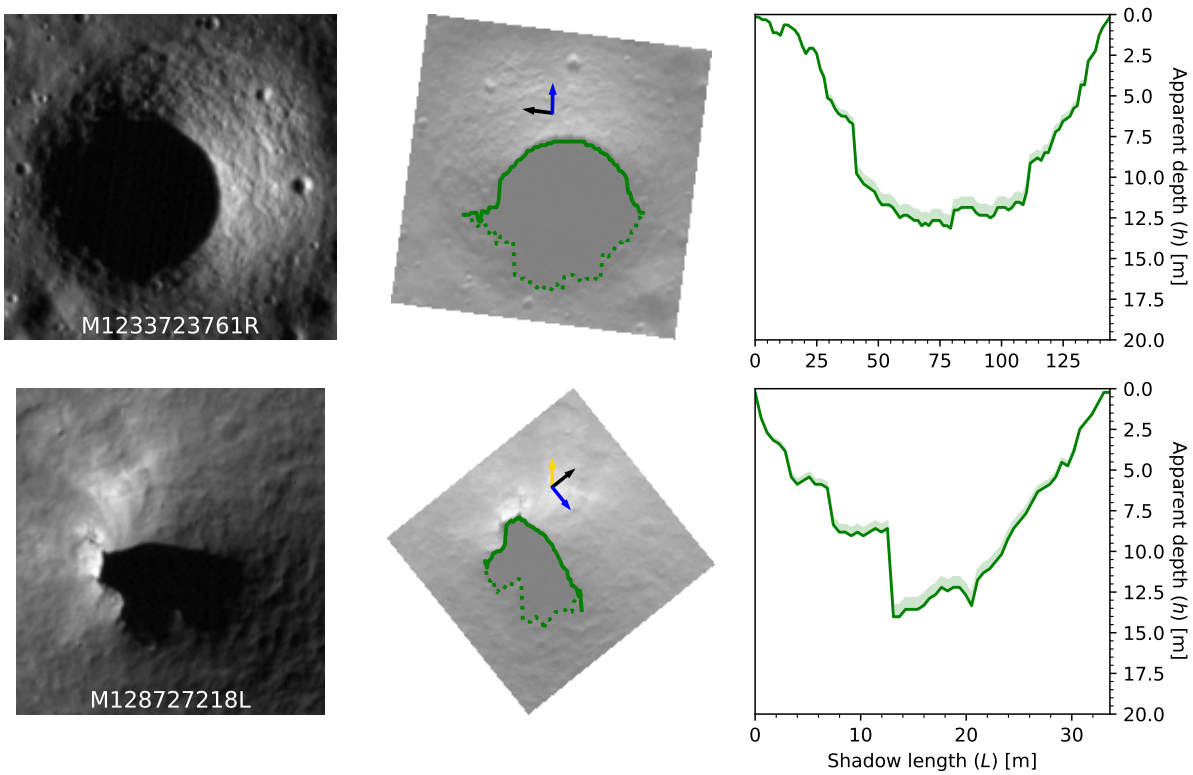

Figure 5.14: Apparent depth (h) profiles automatically calculated using the PITS tool upon two cropped LROC NAC images of the Bel'kovich A Pit (BAP). The symbology is the same as in Figure 5.12. The first image is a low Sun angle image where the shadow edge falls within the BAP's wide funnel, whilst the second is the highest resolution image available in which the BAP's floor is visible.

troughs are thought to be eruptive fissures, possibly opened up by extensional forces originating between the Tharsis and Elysium volcanic regions (Burr et al., 2002) and not related to the collapse of a sub-surface cavity. However, with its morphology of steep walls, prominent shadows and debris in the form of rockfalls, it is understandable why it has been detected in spite of the regional context.

Figure 5.15 (b) shows one of a series of elliptical depressions of unknown origin (found in the volcanic plains south of Olympus Mons within the image ESP_085801_1905_RED), which has been detected as a pit. Other similar features in the image are not aligned with each other, nor do they have steep walls despite the negligible ε. Therefore, these depressions may be genuine pits which have been buried due to aeolian processes or not related to sub-surface cavity collapses at all.

Figure 5.15 (c) and (d) present two examples of where low albedo material within the

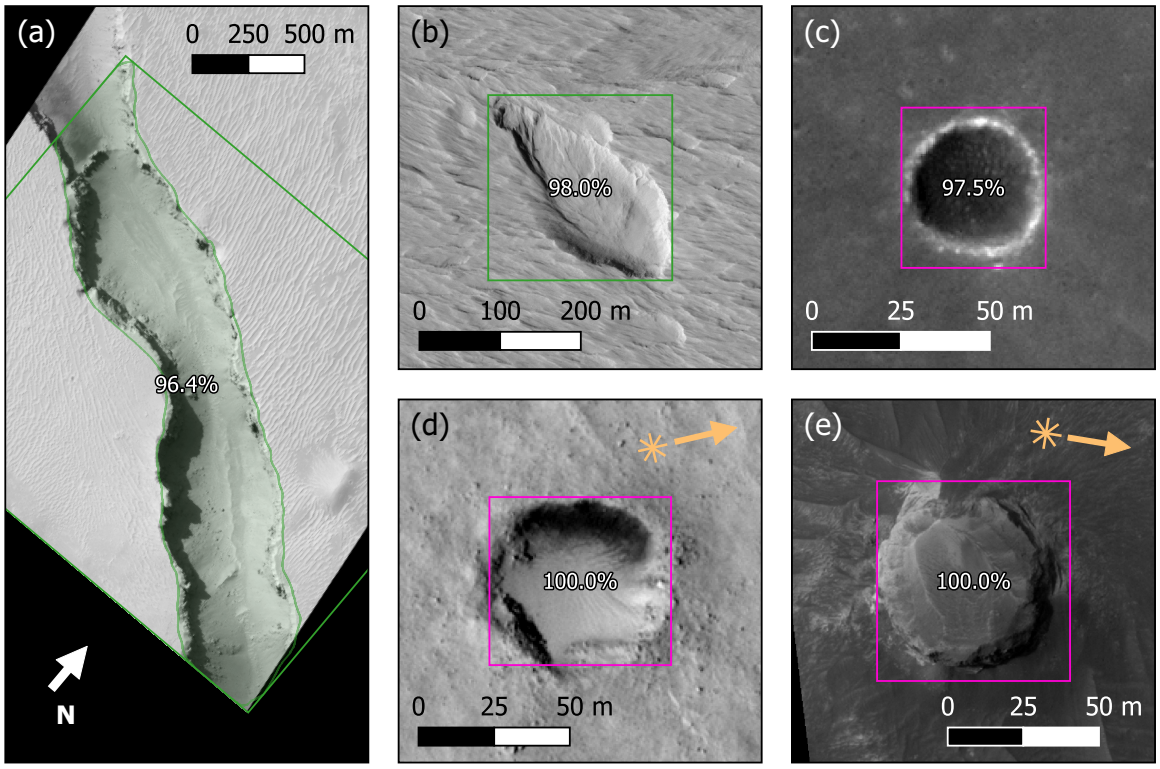

Figure 5.15: Detections of pits (green) and skylights (pink) made by ESSA upon the January 2025 PDS release of HiRISE RDRV11 images. These detections were made upon downscaled (1.5 m/px) versions of the full-resolution observations shown here (0.25–0.5 m/px).

floors of circular depressions has led to confident detections of skylights by ESSA (within the images ESP_085703_1730_RED and ESP_085964_1945_RED, respectively). Considering a α of 38°, the absence of a shadow in (c) means that it is too shallow to be a skylight. Meanwhile, plotting the direction of the Sun's illumination over (d) shows how the large low albedo region north of the feature's floor is not a shadow, meaning that it may be a weathered impact crater.

Figure 5.15 (e) shows what at first appears to be a circular depression with no raised rim, with a shadow being cast by steep walls within ESP_085942_1760_RED. However, the solar azimuth angle reveals that the feature is a positive elevation feature which is casting its own shadow onto the surrounding surface. As a result, it appears that (from the examples shown in Figure 5.15) ESSA may be more sensitive to morphologically confusing features without the context of the surrounding surface or sensing conditions on Mars than on the Moon—likely

due to the greater variety of features and changes in albedo. Consequently, ESSA would need retraining upon greater quantities of Martian data, possibly from other sensors, to be able to perform the same regional mapping of pits and skylights as was the case for the eight Lunar mare deposits in this work.

5.5 Key-Point Summary

- A new model has been developed, called ESSA (Entrances to Sub-Surface Areas), which is a Mask R-CNN instance segmentation deep convolutional neural network (DCNN) trained upon a combination of genuine and synthetic Lunar and Martian remote-sensing imagery and feature catalogues to detect Lunar pits and skylights.
- Between a range of Mask R-CNN models trained on different combinations of training data and with various backbones, the highest validation average F_1-score (F_1) of the bounding boxes (89.1% at epoch 57) was achieved by training all available training data and with a ResNet50 backbone.
- In testing upon LROC NAC imagery of the Mare Tranquillitatis Pit (MTP) and proposed lava tube collapses from Sauro et al. (2020), ESSA achieved an average F_1 of 82.4 and 93.7% for the bounding boxes and predicted masks, respectively.
- Having been applied to ≈0.23% of the total surface and ≈1.92% of the total Lunar maria, two candidate cave entrances have been found—with one being outside the latitude range of previous surveys and the other within the previously mapped Marius Hills.
- The South Marius Hills Pit (SMHP) is a 60–80 m-diameter skylight found within the Marius Hills region, and ≈4.3 km from the Marius A collapsed lava tube proposed in Sauro et al. (2020). Apparent depth (h) profiles calculated by the PITS tool found a maximum h of 15.4 m. SMHP may have been missed by previous investigations due to the low contrast between its floor and the immediate Lunar surface.
- The Bel'kovich A Pit (BAP) is a ≈40 m-diameter skylight situated within a wide funnel and found in the high-latitude Bel'kovich A crater. A maximum h of 14.0 m was returned by PITS, which increases to ≈26 m when including the funnel.
- Despite being the RoI with the largest volume of data, it took ESSA just 4.35 mins to survey all 4,905 image tiles within the Marius Hills region—far quicker than would have been possible via manual means.
- ESSA has been applied to the HiRISE RDRV11 images from the January 2025 PDS release, upon which preliminary results suggest that it is more sensitive to features which appear morphologically similar to pits/skylights without the knowledge of the surrounding regional context or the sensing conditions.

Chapter 6. Automated Impact Melt Fracture Mapping on the Moon with Weakly Supervised Deep Learning

6.1 Introduction

In the previous chapters, surface features that could potentially provide access to the sub-surface environment on planetary bodies have been established. Whether it be for the natural shelter that they would provide robotic or crewed space missions from micrometeorite impacts and harmful radiation, or the prospect of maintaining reserves of water ice leading to astrobiological interest, knowledge of the locations of these features is necessary for deciding which ones to target for such endeavours. Another form of surface feature with potential for being sub-surface entrances is impact melt fractures.

As the name suggests, impact melt fractures (herein referred to as IMFs) are cracks found in the melt deposits of impact craters, which can often occur in curvilinear or polygonal complexes. Figure 6.1 presents several patterns, scales and morphologies of IMFs as found on the Moon. The cracks are thought to form due to extensional stresses as the deposited shock-heated material (either found in the crater's floor or ejected from the site) cools and contracts after the event of an impact (Xiao et al., 2014). The presence of solid debris in a melt pond, such as material falling from the crater's walls, will affect the uniformity of the stresses experienced by the melt (Denevi et al., 2012)—and, by proxy, the patterns of IMFs observed.

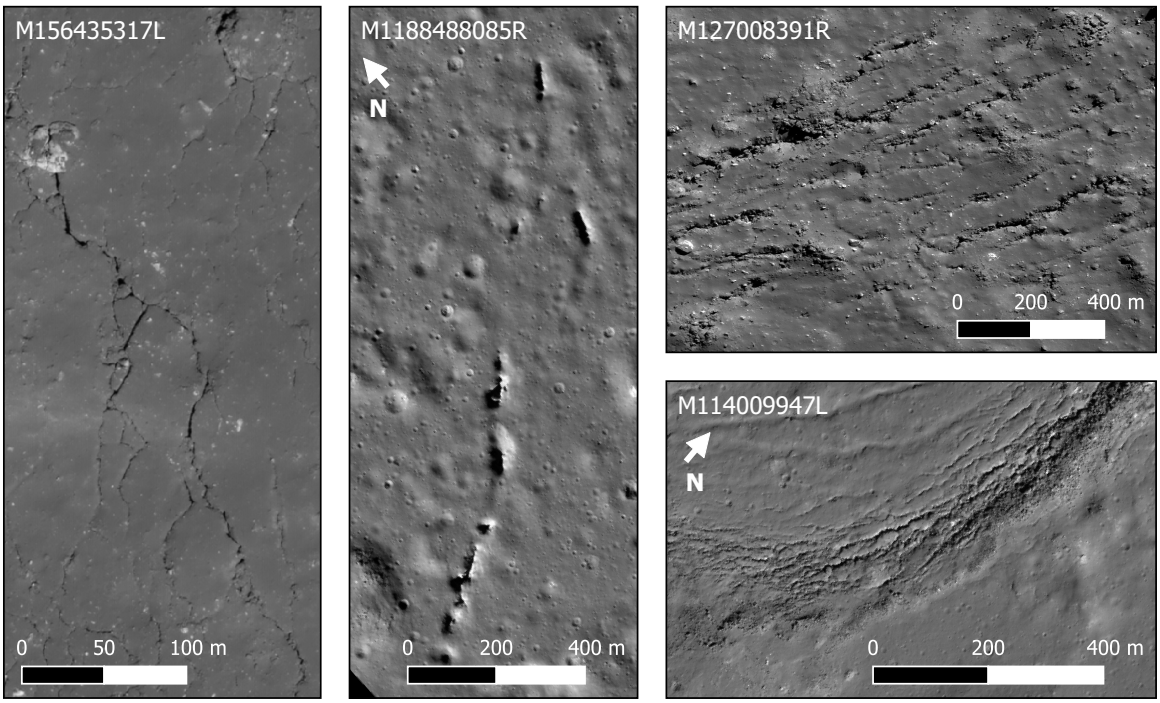

Figure 6.1: Examples of impact melt fractures (IMFs) found in Lunar impact craters. The Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) images shown here are taken of the (left) Virtanen F, (middle) Copernicus, (top right) Tycho and (bottom right) Lalande impact craters.

The rate of cooling of an impact melt deposit and, as a result, the likelihood of fracturing depend on several mechanisms: thermal radiation towards space, convection in the remaining molten material, conduction through the surrounding solid terrain or atmosphere and the latent heat released as the molten rock solidifies (Xiao et al., 2014). Due to the conditions under which these processes occur being similar on the two bodies, IMFs relating to the cooling of impact melt deposits have been observed on the Moon and Mercury (Xiao et al., 2014).

However, IMFs are not to be misconstrued for the scarp patterns found within fractured-floor craters (FFCs). FFCs are shallow complex impact craters which exhibit polygonal, radial or concentric patterns of linear troughs (Korteniemi, 2021) thought to be caused by shallow magmatic intrusion (Jozwiak et al., 2012), which have long been mapped on the Moon (Schultz, 1976) and Mars (Schultz, 1978) due to their larger size. An example of this is given in Figure 2.3.

Impact melt deposits (also known as melt ponds) and the IMFs found within them have

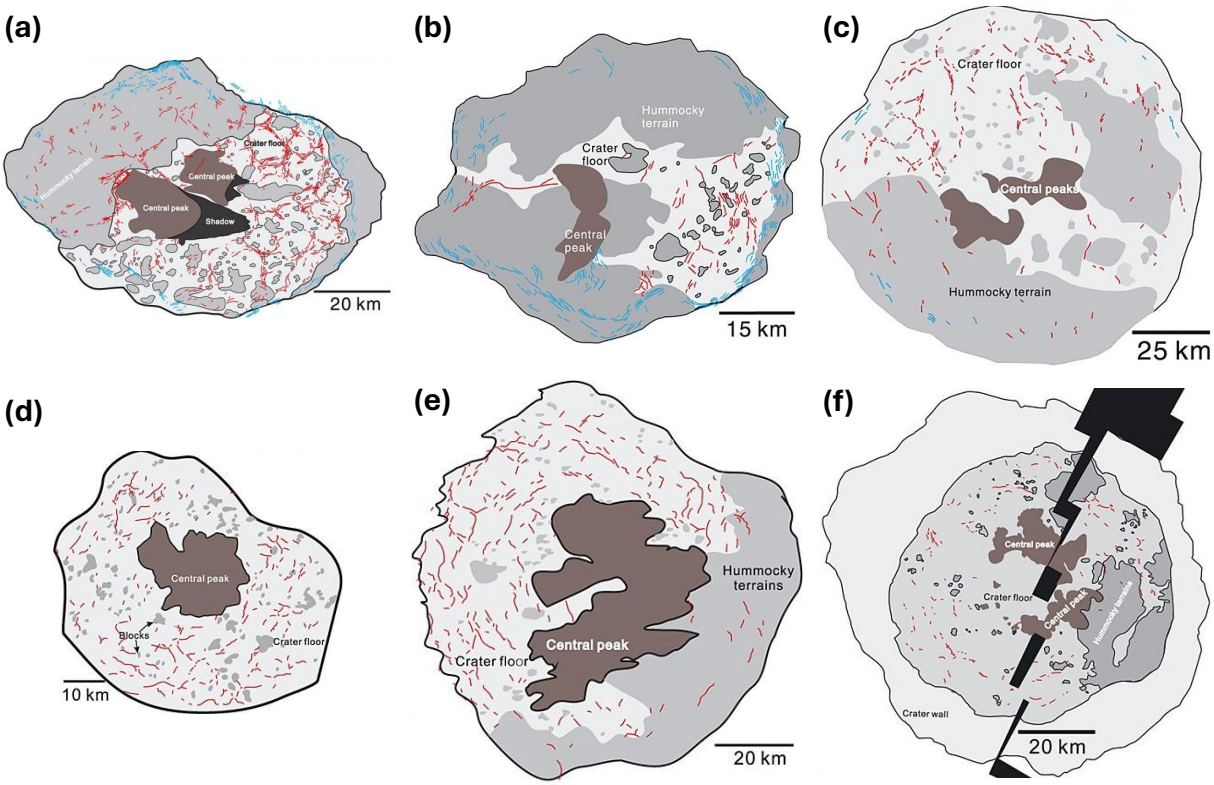

Figure 6.2: Mapping of IMFs by Xiao et al. (2014) within the melt deposits of the Lunar Copernican-aged craters: (a) Tycho, (b) Aristarchus, (c) Copernicus and the Kuiperian-aged (d) Degas, (e) Abedin and (f) Hokusai craters on Mercury. Red and blue lines correspond to interior and marginal IMFs, respectively. This figure has been adapted from Figures 1 to 6 of Xiao et al. (2014) with permission.

been the subject of several mapping exercises. Xiao et al. (2014) have carried out surveys of IMFs within the young impact melt deposits belonging to six complex impact craters on the Moon and Mercury. On the Moon, the Copernican-aged craters Tycho, Aristarchus and Copernicus were surveyed, whilst the Kuiperian-aged Degas, Abedin and Hokusai craters were studied on Mercury. In addition, Xiao et al. (2014) further categorised the delineated IMFs according to whether they occur close to the centre of the crater floor or at the margins between the floor and crater walls. Xiao et al. (2014) claim that i) parallel curvilinear fracture patterns occur at the border between crater wall and melt deposit, ii) IMFs tend to cluster around solid debris, iii) wider and more spaced-out IMFs tend to occur more centrally and/or in deeper melt deposits and iv) sub-parallel IMFs appear radially around local depressions caused by subsidence. While the maps are not publicly available in a GIS-ready format, they are displayed in Figure 6.2.

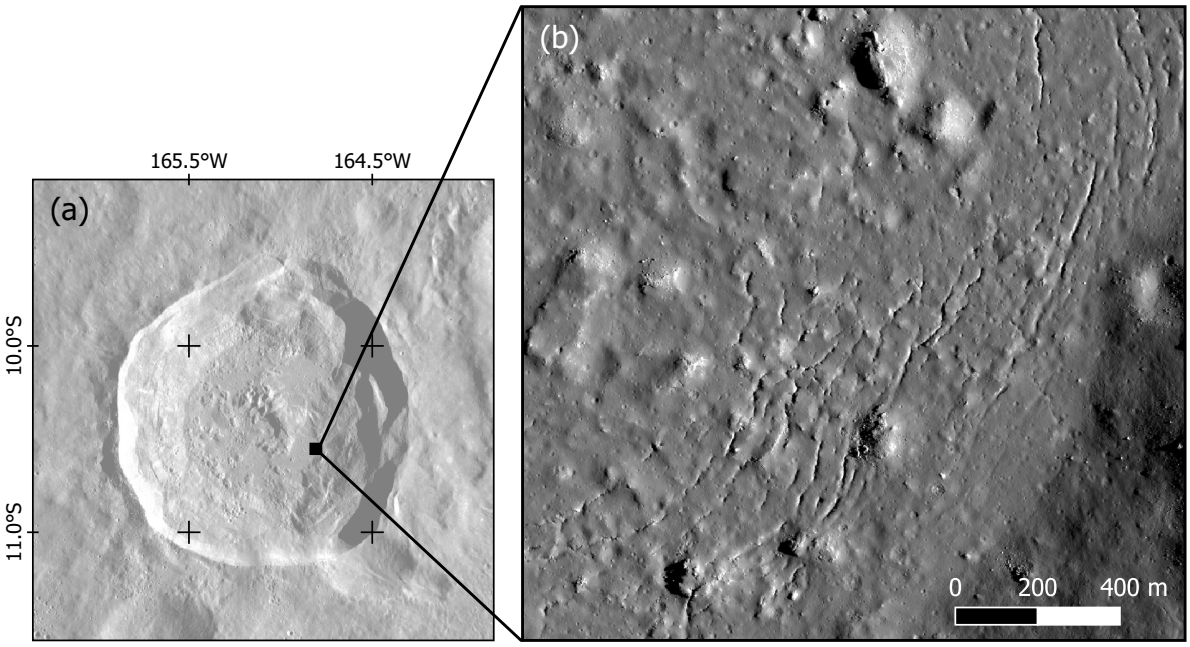

Figure 6.3: (a) Crookes crater within the Lunar Reconnaissance Orbiter Wide Angle Camera (LROC WAC; Robinson et al. 2010) global mosaic (Speyerer et al., 2011). (b) Sub-parallel IMFs near the boundary between the impact melt and the south-eastern wall of Crookes (shown within the LROC NAC feature mosaic NAC_ROI_CROOKES_LOB_E104S1949). (b) is the same region as displayed within B of Figure 7 in Thaker et al. (2020).

Dhingra et al. (2017) created geologic maps of the impact melt deposits within Tycho and Jackson craters containing 12 different morphological classes. However, they did not perform any cataloguing of IMFs—likely due to their use of 10 m/px resolution images from the Kaguya Terrain Camera (TC; Kato et al. 2008) for the mapping.

On the other hand, Thaker et al. (2020) do produce IMF annotations within their morphological maps of the previously unsurveyed Lunar craters: Crookes and Ohm. Crookes crater is a Copernican-aged, complex impact crater found on the far side of the Moon with a diameter of ≈48 km (IAU, 2025). Whilst Thaker et al. (2020) found that all three craters exhibited sub-parallel or curvilinear IMFs (such as those given in Figure 6.3) within their crater floors, only Crookes and Ohm displayed IMFs within the melt ponds situated on the terraces of their crater walls. Figure 6.4 gives evidence of these IMFs found within the northern melt pond of Crookes. Moreover, the maps produced by Thaker et al. (2020) for the three craters appear to show Crookes having far fewer of such melt ponds. Although this may be due to Crookes

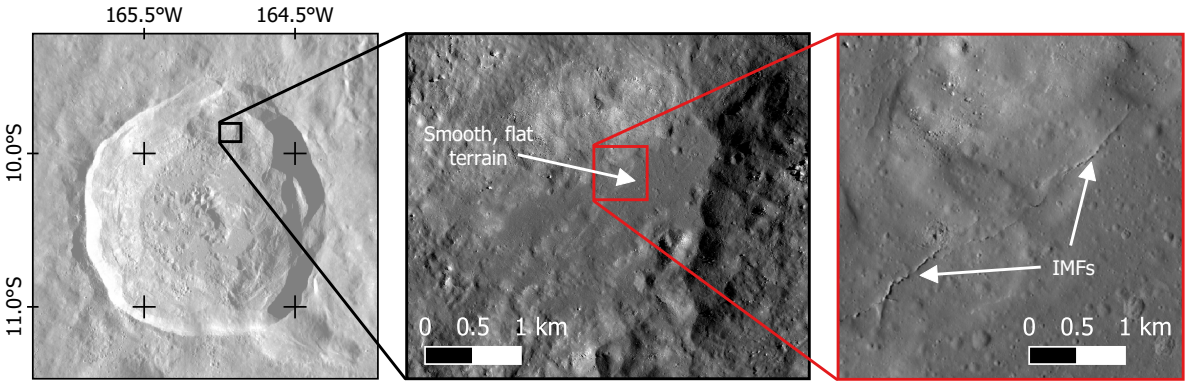

Figure 6.4: IMFs found by Thaker et al. (2020) within the melt pond situated within the northern wall of Crookes crater, as shown by the same LROC NAC mosaic as in Figure 6.3. The melt pond and constituent IMFs shown here are the same as those highlighted in Figure 8 of Thaker et al. (2020).

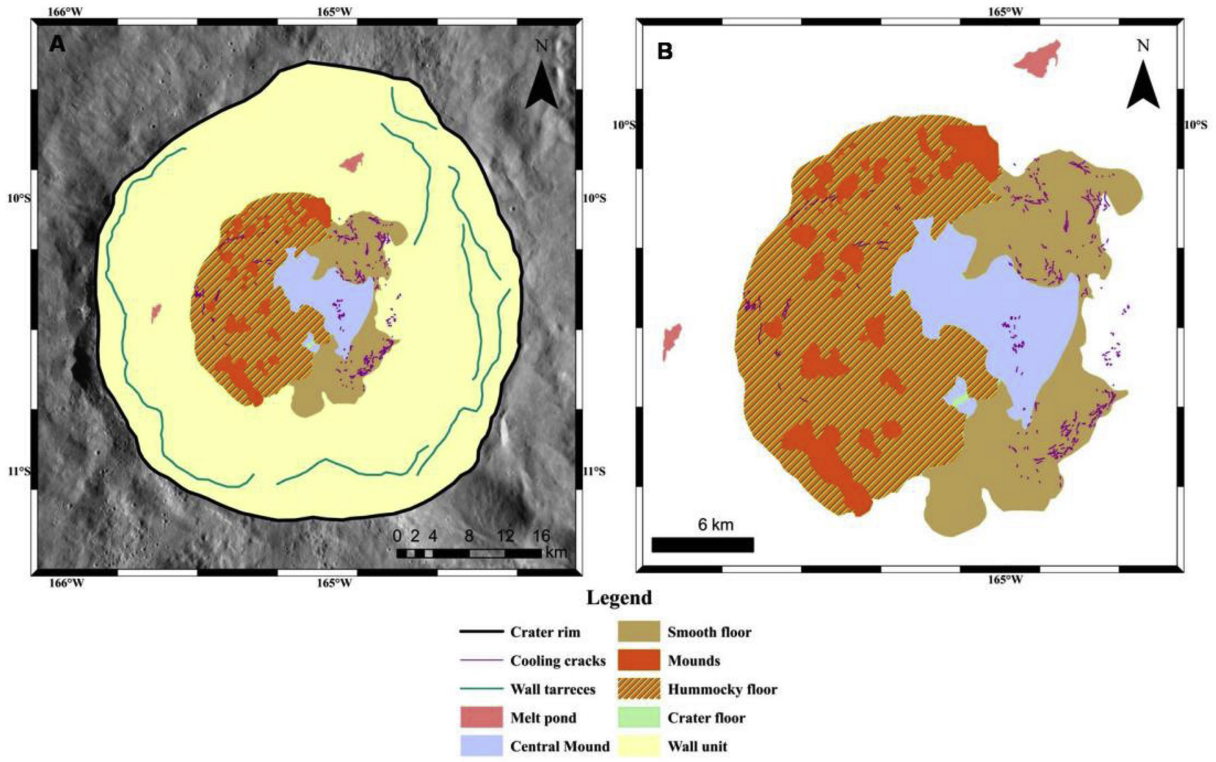

Figure 6.5: A morphological map of Crookes crater produced by Thaker et al. (2020). Crucially for this work, the map includes the mapping of IMFs (in purple) mostly within the various floor units and the identification of two melt ponds within the northern and western walls of Crookes crater. This is Figure 11 of Thaker et al. (2020), and is used with permission.

being the smallest of the three in terms of diameter, which generally results in less ejected material. The morphological map containing IMFs, as created by Thaker et al. (2020) for Crookes crater, is displayed in Figure 6.5.

Mapping of IMFs has also been inadvertently completed as part of the Lunar Pit Atlas (LPA). As previous chapters have discussed, the LPA category with the most features was the one describing pits found within impact melt deposits. Of the 256 impact melt pits, 8 have been specifically designated as 'fractures', with even more being associated with fracture systems. These 8 fractures within LPA can be found within Kepler, King, Messier A, Ohm and Tycho craters.

All of the studies mentioned previously have been concerned with impact melt deposits within multi-kilometre-scale complex craters. However, the deposition of impact melt is not solely reserved for such large impact events. For example, melt deposits have been found in the Lunar highlands within craters having diameters as small as 170 m (Plescia and Cintala, 2012). It is thought that a near-vertical impactor trajectory is required in order to generate and deposit sufficient shocked material, thus rendering these instances rare (Plescia and Cintala, 2012). IMFs, which were thought to only be indicative of larger volumes of melt material, were also observed in these small craters "in a few cases" (Plescia and Cintala, 2012).

Melt ponds have also been mapped by Gagnon et al. (2023) in the vicinity of five impact craters close to the Lunar south pole, which is the targeted landing site of the Artemis III mission. The mineral composition of these deposits was also probed using remote observations in preparation for potential future sample return missions. These five craters included three unnamed simple craters, along with De Forest and Schomberger A. Schomberger A was also the focus of Lee (2022), in which IMFs were identified within the visible, southern half of its melt deposit, as pictured in Figure 6.6. Lee (2022) draws comparisons between the IMFs in Schomberger A with terrestrial lava fractures found on Earth (such as Grjótagjá in Iceland) in order to suggest that they may be potential cave entrances with astrobiological interest. This is because at Schomberger A's latitude (78.61°S), IMFs within it with high depth-to-width ratios have the potential to act as permanently shadowed regions (PSRs). PSRs are topographic minima with sufficiently extreme latitudes such that they receive no sunlight, despite the Moon's $\approx 1.5^{\circ}$ obliquity to the ecliptic plane. As a result, PSRs can cold-trap volatiles such as water ice—with sub-surface deposits thought to be possible in regions even

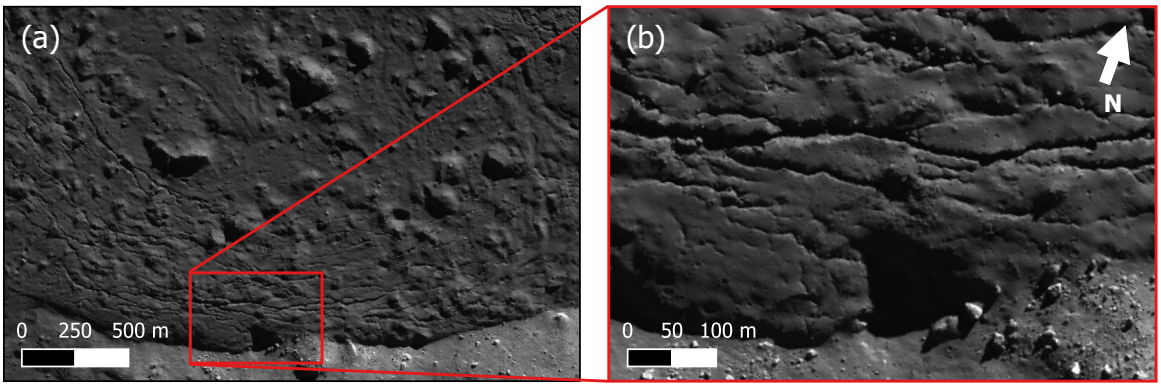

Figure 6.6: IMFs found within the visible floor of Schomberger A crater by Lee (2022) near the Lunar south pole. (a) and (b) show the same areas of Schomberger A as C and D, respectively, within Figure 2 of Lee (2022).

without surface frost (Brown et al., 2022), which IMFs may allow easier access to.

The studies above, of which the majority investigate melt deposits and/or IMFs on the Moon, have been enabled by the breadth of high-resolution imagery available of the Lunar surface. Of the data sources used, resolutions ranged from ≈0.5 m/px with the Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) to ≈10 m/px with the Kaguya TC (Kato et al., 2008). However, as was encountered by Xiao et al. (2014) in their mapping of IMFs within Degas, Abedin and Hokusai craters, the surface detail is more limited for Mercury since the best resolution available with images taken by MESSENGER's Mercury Dual Imaging System's Narrow Angle Camera (MDIS-NAC; Hawkins et al. 2007) is 12 m/px at best (Solomon et al., 2001).

Despite there having been many manual mapping efforts for IMFs, and some studies using machine and deep learning (ML/DL) to detect linear features in planetary remote-sensing imagery (dark slope streaks: Wang et al. 2017b; transverse aeolian ridges: Barrett et al. 2023; rockfalls: Bickel et al. 2019, 2020), there has yet to be any attempt at automatically mapping IMFs. This gap in the literature certainly does not mean that there are no reasons for doing so. For example, the fact that impact melt deposits are found in large complex impact craters and unnamed sub-kilometre-diameter simple craters simultaneously means that it is highly unlikely that the complete mapping of IMFs on the Moon will be a manually completable task.

This limits the ability to deduce the degree to which the higher ratio of ambient material to impact melt affects the dominant cooling processes in smaller craters. It is also perceivable that a tool for detecting the presence of IMFs on the Moon could be used as a proxy for the identification of previously uncatalogued impact melt deposits without the need for typically low-resolution datasets, such as radar or spectroscopy. The detection of more readily accessible deposits would be particularly relevant since one of the investigations within Science Objective 3 ("Interpreting the Impact History of the Earth-Moon System") of the Artemis III mission is the sample collection of impact melt (Weber et al., 2021).

In this chapter, a DL model is described that can perform semantic segmentation (SS) upon LROC NAC imagery of Lunar IMFs, meaning that individual pixels are classified to produce a binary mask of the predicted IMFs. This model, named IMFMapper, has been trained on LROC NAC observations of approximately 4,130 IMFs manually labelled by the author in a weak manner. As explained in further sections, IMFs have been weakly labelled (i.e. the labels are incomplete or imprecise) in order to make the labelling of such quantities more feasible and reduce annotator fatigue.

IMFMapper is tested upon a subset of the IMFs found within the impact craters Copernicus and Virtanen F. Copernicus, as one of those mapped by Xiao et al. (2014), contains numerous staple IMFs which, should they be detected, would prove a prospective model's efficacy when dealing with some of the clearest examples. Virtanen F is an 11 km diameter crater hosting an impact melt deposit that contains several complexes of IMFs with widths of approximately 5 m or less. Therefore, good performance in this instance would prove that some of the smallest examples of IMFs, which would be the most laborious for manual mappers, are detectable via automated means.

IMFMapper is then applied to self-processed mosaics of LROC NAC images of Crookes and Schomberger A craters. This will enable a qualitative comparison with the mapping of Crookes crater by Thaker et al. (2020), not just in whether IMFMapper can recall IMFs labelled by experts, but also whether it can detect smaller IMFs or find additional melt ponds. Applying IMFMapper to Schomberger A would also result in the first mapping of the IMFs previously

identified by Lee (2022), which may be of interest for the future exploration of cave entrances and PSRs.

The secondary objective of IMFMapper is to encourage the future mapping of IMFs on Mercury via automated means upon the arrival of the joint ESA and JAXA BepiColombo mission at Mercury. The SIMBIO-SYS (Spectrometers and Imagers for MPO Integrated Observatory System) aboard the ESA-provided Mercury Planetary Orbiter (MPO) aims to provide 6 m/px panchromatic imagery for ≈20% of the surface within the first year of operation (Benkhoff et al., 2021). Therefore, there is the potential for IMFs on the surface of Mercury to be revealed with greater detail than ever before, as well as in greater quantities. Thus, having a robust automated method for detecting IMFs will be a necessary tool for the study of cooling processes on Mercury and the search for melt deposits outside of large complex craters.

The following work presented in this chapter was published in the *'Journal of Geophysical Research (Planets)'* in November 2025 (Le Corre et al., 2025b).

6.2 Dataset

As was the case in Chapter 5, a significant amount of remote-sensing imagery and corresponding labels are required in order to train any supervised DL model. Thankfully, a large amount of the data containing IMFs had already been acquired and processed (in order to be viewable in GIS software) as part of the research in Chapter 5, wherein LROC NAC observations of impact melt pits from the LPA had been identified. Of this dataset of 111 LROC NAC images, manual inspection revealed that 64 included one or more IMFs. Since many of the 64 images were repeat observations of the same IMFs, a subset of 25 LROC NAC images was earmarked for labelling as training and validation data. These 25 images exhibited sufficient numbers of IMFs with a broad range of lengths, widths, morphologies and illumination conditions such that a high-quality training dataset could be built.

In order to gauge its real-world accuracy, a DL model should be tested on data that it had not 'seen' during the training process. Better yet, a DL model trained to detect IMFs should

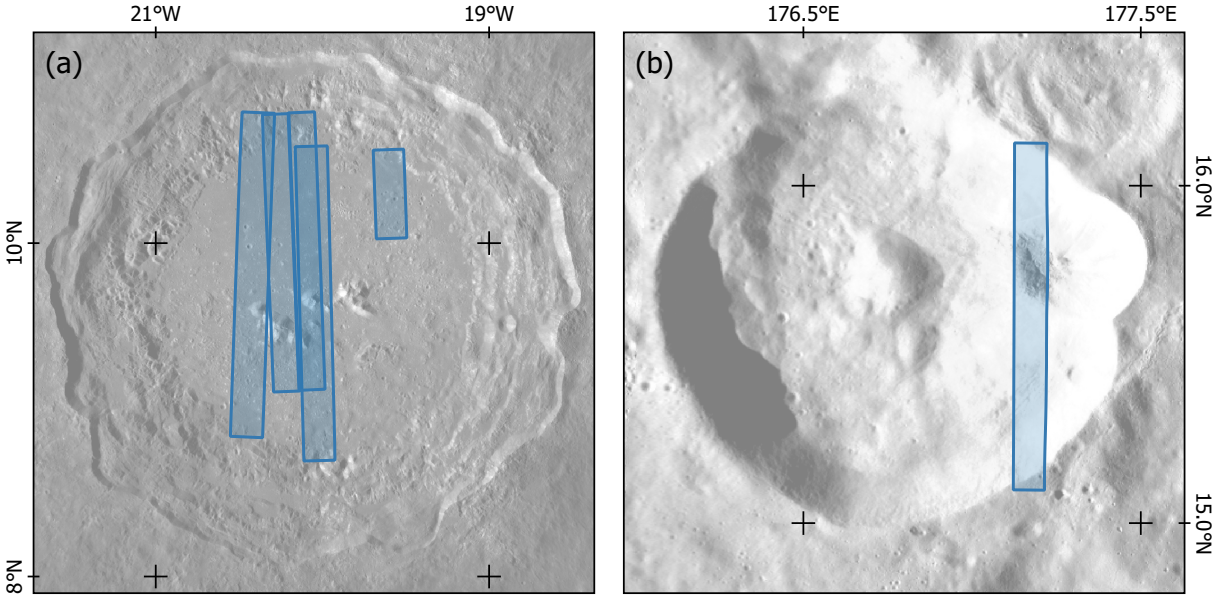

Figure 6.7: Footprints of the LROC NAC images used as testing data for DeepLabV3 models trained to detect Lunar IMFs. (a) and (b) show the footprints overlaying Copernicus and Virtanen craters, respectively, as seen in the LROC WAC global mosaic (Speyerer et al., 2011). Virtanen F is the crater superimposed upon the northeastern wall of Virtanen crater.

ideally be tested on imagery covering a completely different region (or impact crater in this instance). Therefore, none of the 25 training/validation images mentioned above contained the IMFs found within Copernicus or Virtanen F craters. Instead, 6 LROC NAC images covering these two craters (5 from Copernicus and 1 from Virtanen F) were reserved for independent testing. These 6 LROC NAC images are overlaid upon their respective craters in Figure 6.7.

As mentioned in Section 6.1, Crookes and Schomberger A are chosen as the sites for applying the trained model as examples of automatically mapping IMFs. In both cases, JMARS was used to find the LROC NAC images which overlap each crater. Then, the 'Mosaic Select Tool' within JMARS was employed to remove any images from the selection which did not improve the coverage of the crater. This resulted in a total of 56 and 55 LROC NAC images for Crookes and Schomberger A craters, respectively.

For the same reasons established in Chapter 5, the LROC NAC images described above had a range of resolutions which needed to be homogenised. As a result, after the raw LROC NAC image products were acquired from NASA's Planetary Data System (PDS; see Robinson 2009), the raw images were processed and map-projected to a common resolution of 1.5 m/px using

Crater Name	D [km]	Lat, Lon [°]	# IMFs	Used For
Aristarchus	40	23.73, -47.49	229	Tr+Va
Copernicus	96	9.62, -20.08	164	Te
Crookes	48	-10.40, -165.10		In
Das	36	-26.49, -137.05	14	Tr+Va
Jackson	71	22.05, -163.32	747	Tr+Va
Kepler	29	8.12, -38.01	45	Tr+Va
King	76	4.96, 120.49	439	Tr+Va
Klute W	31	37.98, -143.31	77	Tr+Va
Lalande	24	-4.46, -8.65	196	Tr+Va
Messier A	11	-2.03, 46.94	42	Tr+Va
Ohm	62	18.32, -113.78	50	Tr+Va
Philolaus	71	72.22, -32.88	13	Tr+Va
Rutherfurd	50	-61.15, -12.25	43	Tr+Va
Schomberger A	29	-78.61, 23.52		In
Tharp	13	-30.60, 145.63	77	Tr+Va
Tycho	85	-43.30, -11.22	2,118	Tr+Va
Virtanen F	11	15.79, 177.32	206	Te
Wiener F	45	41.19, 149.97	151	Tr+Va

Table 6.1: A list of the Lunar impact craters which were used for training and validation (Tr+Va), testing (Te), or inference (In) of a DL model tasked with detecting impact melt fractures (IMFs). The diameters (rounded to the nearest kilometre) and coordinates for each crater were retrieved from IAU (2025). The number of IMFs is the amount labelled for Tr+Va and Te, not necessarily the genuine total number of IMFs present in each crater.

the same ISIS (Integrated Software for Imagers and Spectrometers) and GDAL (Geospatial Data Abstraction Library) procedures as described in Section 5.2 (see Technical Note TN-2 in Appendix A). This resolution was chosen since the lowest resolution of the used images equalled 1.5 m/px when rounded to the nearest 0.1 m. However, in order to avoid repeat detections of the same IMF when mapping Crookes and Schomberger A craters, mosaics were created by stitching together the resulting processed and map-projected ISIS cubes.

Table 6.1 provides a list of the impact craters containing IMFs, which were used for either training and validation, testing, or inferring the model. Alongside the names of each crater is its location, the number of IMFs which were labelled in this work, and the stage of model development that it was responsible for. The number of labelled IMFs quoted for each crater are purely the amount of weak IMF labels that were produced in this work. Therefore, these

quantities should not be misconstrued as the absolute total of the IMFs present, since repeat observations of the same IMF were often labelled, and the training and testing images did not provide full coverage of each crater.

6.3 Methodology

6.3.1 Creating a Weakly-Labelled IMF Training Dataset

Several challenges are presented when attempting to create a high-quality training dataset of remote-sensing imagery containing IMFs. First is the linear morphology of these features, which can be kilometres-long and as small as a few metres in width, as shown in Figure 6.1. Therefore, a training image that fully encapsulates an IMF, from start to end, could become too large to pass through a DL model due to memory constraints. Also, if the resolution of such an image is lowered to counter the increasing memory load, then thinner IMFs would likely become unresolvable. The result of this balance between image size and resolution is that training images will frequently contain IMFs that have been partially cut off by the image's extent. The likelihood of a model seeing sections of IMFs is one of the motives for utilising a traditional SS model, whereby individual pixels are classified. The reasoning for this decision is that an object detection (OD) model, such as the R-CNN family discussed in Chapter 5, is learning the features that constitute an object, meaning that a cut-off IMF may be confused with other non-linear depressions.

Another issue arises in that supervised SS models also require pixel-level labels. As Table 6.1 suggests, the number of IMFs that were present in the imagery acquired in Section 6.2 was so high that manually labelling IMFs as polygons would be unachievable within available time frames. Pixel-level labels of some of the smallest IMFs, which could often be just a few pixels wide, would also be an unnecessary degree of labelling effort relative to the number of pixels being labelled. Therefore, the possibility of producing weak pixel-level labels of IMFs, which would enable a sufficiently large dataset to be built without a significant decrease in labels' accuracy, was explored.

As was discussed in Section 2.2.5, weakly supervised learning is the principle of training a supervised ML or DL model, but with the labelling of the data being incomplete, artificial or inaccurate. In this instance, instead of labelling IMFs as polygons at the first attempt, IMFs are annotated as lines along their entire path. This labelling was completed in QGIS since it has a facility for buffering lines by a certain map unit (i.e. metres or degrees) to produce a linear polygon. Therefore, from these loose labels of the paths of IMFs, weak pixel-level labels can be generated. However, IMFs are observed to have a range of widths, meaning that a buffer value that works for one IMF will not work for all. Therefore, line labels were given additional classifications according to their approximate width in pixels, which were: 'tiny' (0–5 px), 'small' (5–10), 'medium' (10–20), 'large' (20–30) and 'huge' (30 or more).

Figure 6.8 gives an example of the weak line labels and how they have been differentially buffered to generate pixel-level polygons ready for training a DL model. Figure 6.8 also shows two instances (which occurred for approximately 5% of features) where IMFs were labelled as polygons due to being LPA impact melt pits which resembled IMFs or an element of an IMF which was too difficult to label as a line.

In order to maximise the reliability when comparing model predictions against the labelled IMFs in the testing dataset, all features within the 5 LROC NAC images of Copernicus crater were labelled as polygons, whilst the thin IMFs in Virtanen F were still accurately approximated with the weak line labels.

Entire LROC NAC image products are far too large to pass through a DL model due to them being several 1,000s and 10,000s of pixels in width and height, respectively, even after downscaling to 1.5 m/px. As a result, all processed LROC NAC image products reserved for training and validation were tiled into a coordinate grid of 512×512 px crops with a 10% overlap between consecutive tiles (an example of this tiling is displayed in Figure 6.8). An overlap of 10% was chosen in order to increase the coverage of a single IMF within a given training tile while also minimising the risk of over-fitting. It was also decided that any resulting tiles that did not contain any IMFs would also be used for training and validation, with the aim of reducing the number of false positives. Nevertheless, where IMFs were present

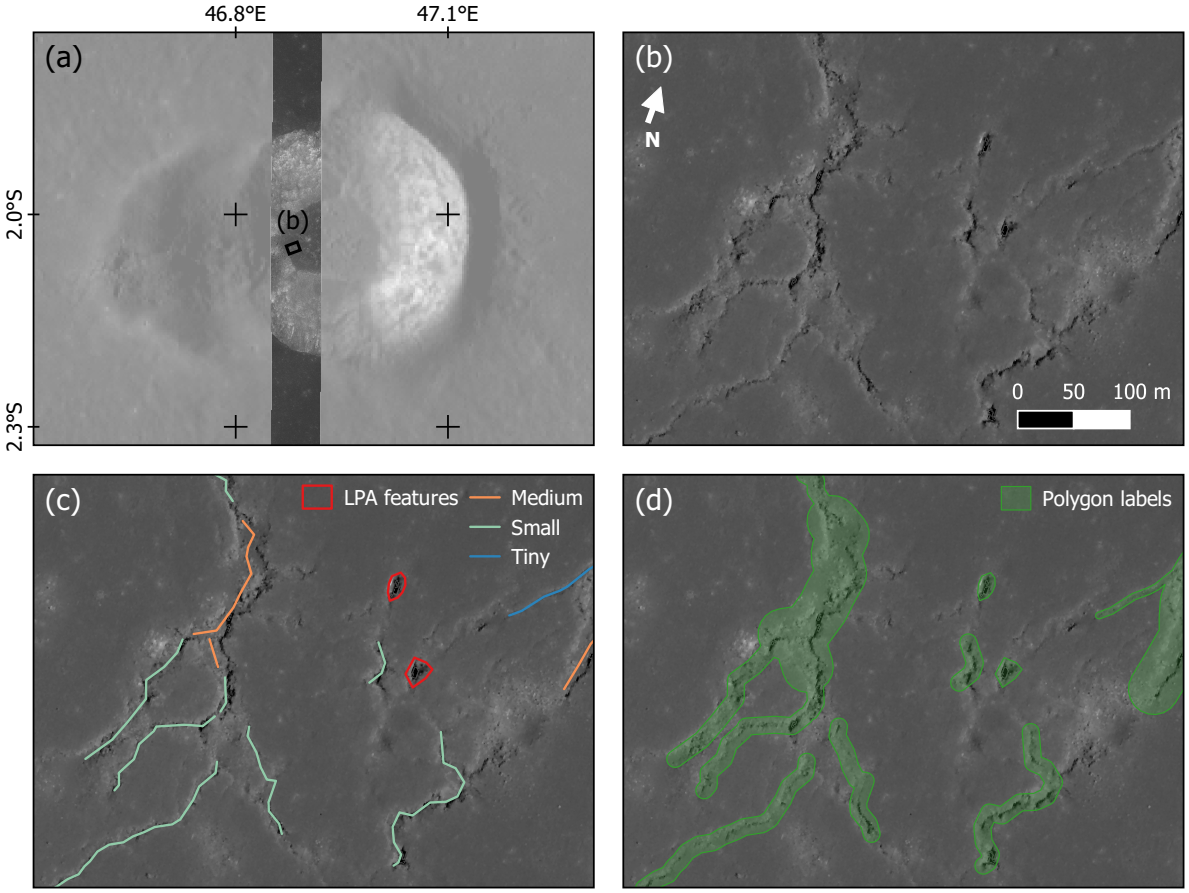

Figure 6.8: An example of the weak labelling of IMFs. (a) The impact crater 'Messier A' within the LROC WAC global mosaic (Speyerer et al., 2011), showing the LROC NAC image M126622485R and the location of the IMFs shown in (b). (c) The manually produced weak line labels and polygon annotations of the LPA features. (d) Pixel-level polygon labels created by buffering the lines in (c) and merging with the LPA polygons.

in a given tile, the weak polygon labels were rasterised using facilities within the GDAL in order to produce binary masks of their locations. Additionally, binary masks of the valid data regions within an image tile were also exported, which allowed for the removal of any adverse detections within the no-data regions after the loss calculation but before the performance metrics were calculated (see Section 6.3.2 for further details). Figure 6.9 gives an example of a training sample tiled from the image M113168034L, showing the image tile, as well as the corresponding label and valid data masks.

In order to bolster the size of the training and validation dataset while also reducing the risk of over-fitting to the training data, the same data augmentation steps that were performed in Chapter 5 were also used in this work. This involved the inclusion of rotated versions (by

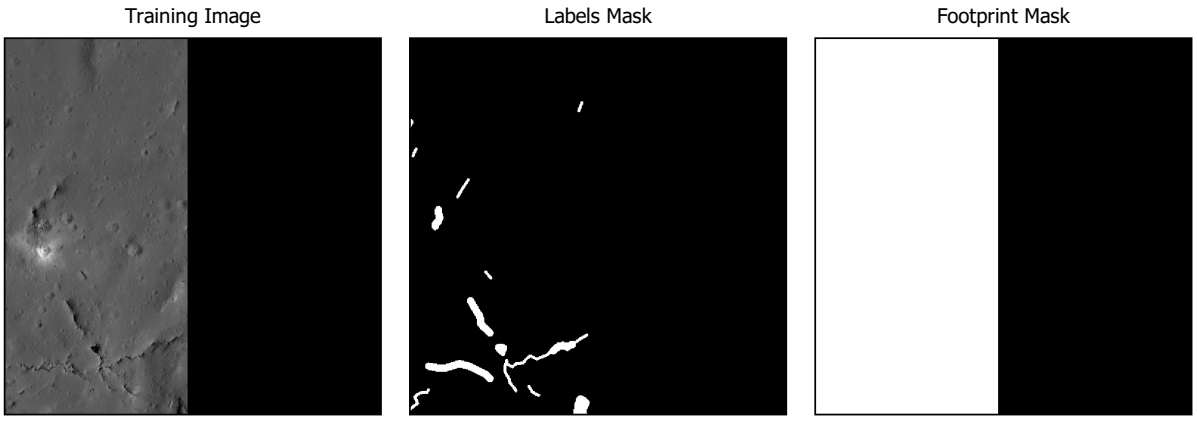

Figure 6.9: An example of a training sample containing IMFs and no-data regions. This shows the non-rotated image tile before any data augmentation has been applied. In the label and footprint masks, black and white pixels equal 0 and 1, respectively. The parent LROC NAC image for this tile is M113168034L and covers part of King crater.

90, 180 and 270° clockwise) of all training and validation tiles, except those which did not contain IMFs, thus quadrupling the number of IMFs seen by the model. This dataset was then split with a ratio of 4:1 between training and validation, before Contrast-limited Adaptive Histogram Equalisation (CLAHE), Gaussian noise, and random perspective transformations could then be applied (each with a probability of $p = 0.5$) to only training samples.

The same procedure as that described above was used to generate tiles for the 6 LROC NAC images used for testing, as well as the Crookes and Schomberger A mosaics, except that there was no overlap between consecutive tiles, nor any data rotation or augmentation.

Table 6.2 details the number of image tiles within each of the training, validation, testing and inference datasets. Where applicable, these quantities are given separately as to whether they contained labelled IMFs or not.

6.3.2 Model Architecture and Evaluation

In this work, a DeepLabV3 deep convolutional neural network (DCNN) DL model is chosen to perform SS on IMFs in tiled LROC NAC imagery (Chen et al., 2016). Some SS models (such as those based on U-Net) use an Encoder-Decoder structure, whereby images are downsampled to 'encode' a condensed representation of the data and then upsampled by the decoder to the original resolution. However, the foundations of the DeepLabV3 model are built on the principle

Dataset	Craters	# Image Tiles
Training	See Table 6.1	4,500 / 2,696
Validation		1,132 / 667
Testing	Copernicus	127 / 2,212
	Virtanen F	24 / 139
Inference	Crookes	3,332
	Schomberger A	1,341

Table 6.2: The number of 512×512 px image tiles used in the training, validation, testing or inference of IMFMapper. The quantities used for training, validation and testing have been split into those that contained annotated IMFs (left) and those that did not contain any IMFs (right). Of course, the number of tiles containing IMFs from Crookes and Schomberger A craters was not known a priori. Training and validation tiles had an overlap of 10%, whereas testing and inference tiles had none.

of atrous convolution. Deriving from the French for holes ('trous'), atrous convolution is where zeroes are embedded within a kernel so as to increase the receptive field of a convolutional layer without increasing the number of parameters that require learning (Chen et al., 2017).

DeepLabV3 implements atrous convolution to form an Atrous Spatial Pyramid Pooling (ASPP) block. Traditional Spatial Pyramid Pooling (SSP) structures attempt to solve the issue of detecting several scales of features by performing a range of resampling operations and pooling them together (He et al., 2014). Whereas ASPP replaces the image resampling with atrous convolution operations. The degree to which atrous convolution widens the receptive view is defined by the rate, which is equal to $N_{trous} + 1$ (where N_{trous} is the number of zeroes inserted between two consecutive kernel values in each dimension). This is shown (not to scale) in Figure 6.10.

The focus on learning feature maps at several scales is particularly advantageous in this work since IMFs can be observed with a wide range of different widths and lengths. Other SS structures, such as U-Net, are also tailored to multi-scale detection, but DeepLabV3 is more suited to this use-case due to the fewer parameters that require learning—thereby reducing the risk of over-fitting to what is a relatively small training dataset by DL standards. Another benefit to the use of DeepLabV3 is that it is available within PyTorch with pre-trained weights (for both the backbone and remaining DeepLab head layers) upon a subset of the COCO

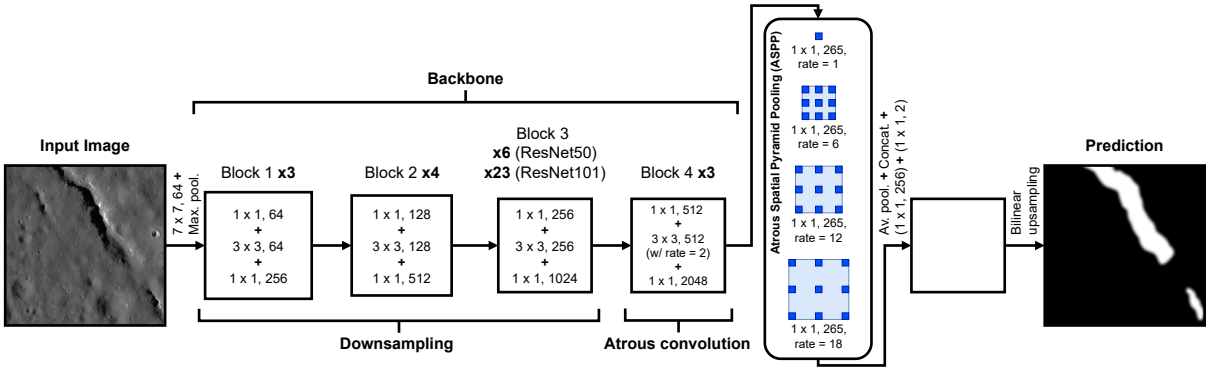

Figure 6.10: Diagram of the DeepLabV3 deep convolutional neural network (DCNN) with ResNet50 or ResNet101 backbones (Chen et al., 2016)—within the context of IMF semantic segmentation. Convolutional layers are represented by the kernel size and the number of output channels (e.g. '3×3', 126 for a kernel size of 3 and 126 output channels). The relative pixel and kernel sizes in the depiction of atrous convolution within the Atrous Spatial Pyramid Pooling (ASPP) block are not to scale.

(Common Objects in Context) dataset (Lin et al., 2014a). The pre-trained DeepLabV3, with ResNet50 and ResNet101 used as the backbones, achieved a mean Intersection over Union (IoU) upon the validation COCO dataset of 66.4 and 67.4%, respectively[1].

In this work, Binary Cross-Entropy (BCE) loss, also known as logarithmic loss, is used as the function whose objective it is to be minimised by the Adam optimiser (Kingma and Ba, 2017), which changes model weights with a standard learning rate of 0.0001. Ideally, a large batch size should be used to limit the effect that a single training sample has on the changing of weights during one pass through the model. Whilst a batch size of 32 could be used in conjunction with the ResNet50 backbone, memory errors occurred due to the addition of more convolutional layers when swapping in the ResNet101 backbone. Therefore, DeepLabV3 models with a batch size of 16 were also trained. All three of the following models were trained for 100 epochs: ResNet50 with a batch size of 32 (ResNet50-32), ResNet50 with a batch size of 16 (ResNet50-16) and ResNet101 with a batch size of 16 (ResNet101-16). A maximum of 100 epochs of training was selected as it was the point at which the training BCE loss began to plateau (i.e. no more notable improvement). As in Chapter 5, model training, testing and inference were performed using the specialist and high-performance computing

[1] The DeepLabV3 model is provided by the PyTorch framework with pre-trained DeepLab head and backbone (ResNet50 and ResNet101) weights. See https://pytorch.org/hub/pytorch_vision_deeplabv3_resnet101/ for more details

systems provided by Information Services at the University of Kent, of which a total of 13 nodes (sharing an NVIDIA A100 40 GB PCIE GPU and 256 GB of RAM) were available. The training of a single DeepLabV3 model took approximately 10–12 hrs.

For the purposes of assessing the model's performance upon the validation dataset during training, as well as the testing data of Copernicus and Virtanen F, the precision (P), recall (R) and F_1-score (F_1) were calculated. P and R, as previously defined in Table 2.1, are derived from the confusion matrix of true/false positives/negatives (TP, FP, TN and FN). TP, FP, TN and FN correspond to the number of pixels which have been correctly/falsely detected/missed, which was also the case for the masks predicted by the Mask R-CNN models in Chapter 5. To avoid division by zero errors, R is not calculated for tiles where there were no IMFs to recall. The same is true for P for tiles where no detections were made. The F_1, also defined in Table 2.1, is then calculated using the average P and R across all batches. Additionally, the Intersection over Union (IoU) between the detections made in Copernicus and Virtanen F with the associated labels was also calculated, as a means of assessing whether the performance differed between the two craters. In the context of SS, as opposed to assessing the accuracy of bounding box detections, IoU can be defined by Equation (6.1). During training, checkpoints for the model are saved after every 25 epochs, as well as a single checkpoint that is overwritten whenever the model reaches a new highest validation F_1.

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{6.1}$$

6.4 Results and Discussion

6.4.1 Training and Validation Performance

Figure 6.11 plots the BCE loss, P, R and F_1 averaged over all batches for the entire 100 epochs of the training process. These metrics have been calculated on the training and validation datasets separately in order to assess whether the DeepLabV3 models are over-fitting to the training data. As the metric chosen to determine the best performing of the three models

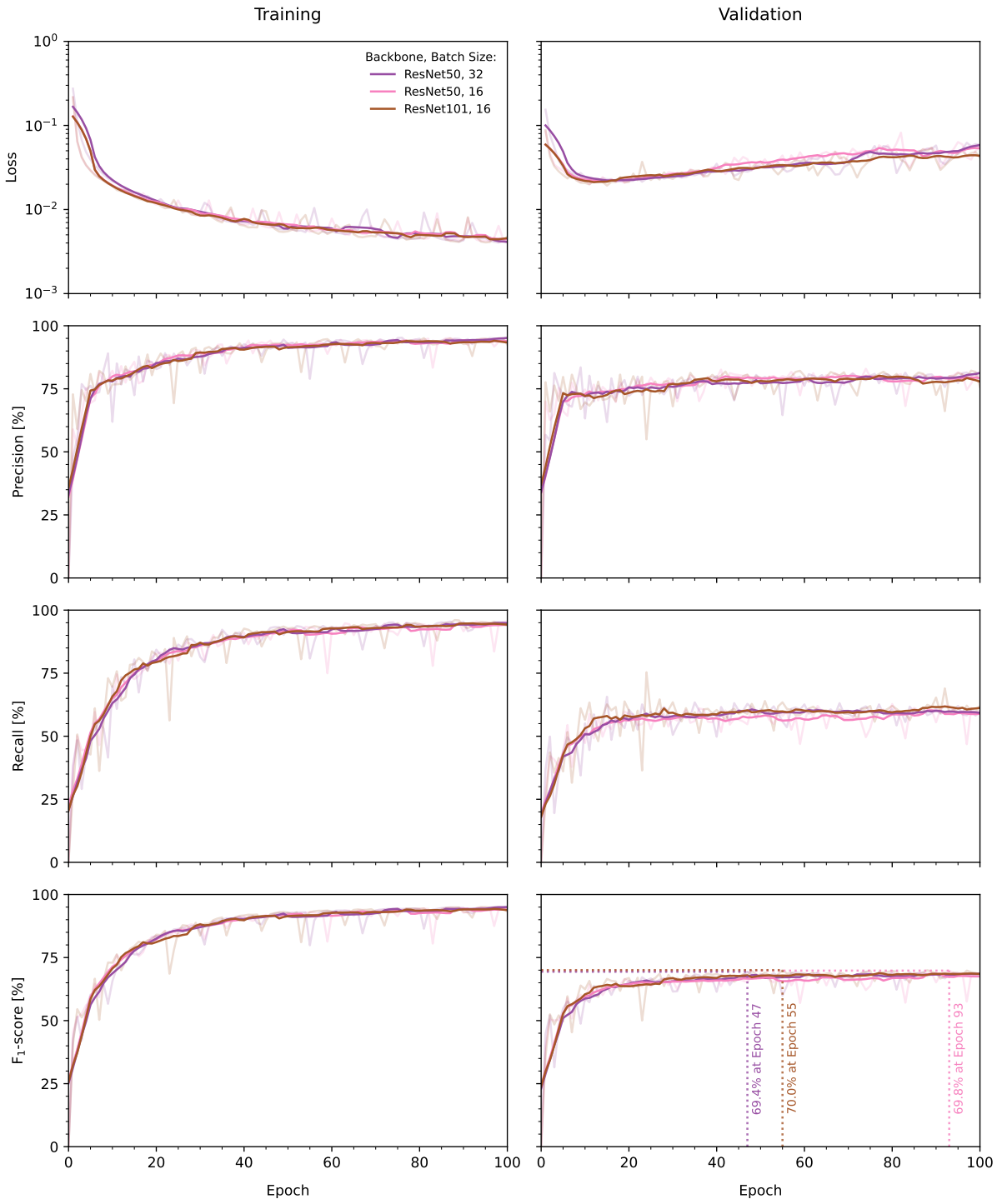

Figure 6.11: Performance metrics calculated upon the training and validation datasets when training DeepLabV3 models to detect IMFs for 100 epochs. The three models have different combinations of backbone and batch size. The solid lines represent the rolling average across 10 epochs, whereas the semi-transparent profiles are the genuine values. The highest F_1-scores achieved by each model are annotated with dotted lines.

during validation, the maximum F_1-scores achieved by each have been annotated along with the epoch at which it occurred. This shows that the models have achieved values with marginal difference (only 0.6%), with the ResNet101-16 model reaching the highest average F_1 with 70.0% at epoch 55. The ResNet50-16 and ResNet50-32 models each reached F_1-scores of 69.8 and 69.4% at epochs 93 and 57, respectively.

The profiles of F_1 calculated upon the training and validation datasets highlight that all three models exhibit a significant generalisation gap (i.e. performance upon training samples betters that of validation). This is driven in part by the fact that the training R reaches more than 90% by the end of the training, while the validation recall remains at ≈60%. This may be a sign that the data augmentation steps taken in Section 6.3.2 are collectively too strong. For example, the CLAHE applied to some samples may be hampering the ability of the models to detect IMFs in low-contrast scenarios within the validation dataset. With probabilities of each occurring of $p = 0.5$, it may also be that the DeepLabV3 models are seeing augmented data the majority of the time and are then learning to detect IMFs in augmented data as opposed to non-augmented.

The training and validation P, R and F_1 all appear to have plateaued within ≈60 epochs, suggesting that an early stopping condition could be implemented if no significant improvements are made within a user-defined number of epochs. However, the generalisation gap between the training and validation F_1 of all three DeepLabV3 models (≈25%) suggests that these models would benefit from more labelled data. It may also be necessary for future efforts at training DL models for IMF detection to sub-categorise them by morphology. This would allow for equal numbers of each morphology to be split across training and validation datasets, but also allow targeting of new data to label according to which morphologies are present, based upon which ones there is a higher deficiency in within the training dataset.

Consequently, as opposed to testing with the versions of the models reached at the end of the 100 epochs, in the following section, testing is conducted using the model checkpoints saved at the epoch which achieved the highest F_1 for that model.

Backbone	Batch Size	IoU [%]	P [%]	R [%]	F_1 [%]
Resnet50	32	**38.7**	55.7	**56.1**	**55.9**
Resnet50	16	36.2	**56.9**	50.1	53.3
Resnet101	16	34.0	50.0	51.6	50.8

Table 6.3: Summary of the average precision (P), recall (R), and F_1-score achieved by a series of DeepLabV3 models (trained with different combinations of backbone and batch size) upon 370 IMFs across 6 LROC NAC images of Copernicus and Virtanen F craters. The highest values for each metric are highlighted in bold.

6.4.2 Testing on Copernicus and Virtanen F Craters

Table 6.3 details the average performance metrics calculated across the entire testing dataset of IMFs in Copernicus and Virtanen F craters for the same models plotted in Figure 6.11. This shows that, despite the ResNet101-16 model performing marginally better during validation, the ResNet50-32 model achieved the highest F_1 of all with 55.9% during testing upon independent data. ResNet50-32 also returned a significantly higher R than the other models with 56.1%.

The fact that training with a larger batch size has led to a better testing performance adds weight to the idea that the ResNet50-16 and ResNet101-16 models were more prone to over-fitting to the training data. A possible explanation for why this same discrepancy between models was not seen in the validation metrics is that the rotated versions of a single tile could be split between the training and validation datasets. This means that the model could be trained on a tile containing IMFs and then validated on that same image rotated by 90, 180 or 270°. Therefore, it may be that the ResNet50-16 and ResNet101-16 models have preferentially learnt to detect those features within the training dataset, as opposed to the ResNet50-32 model, which has learnt those features but is also more readily transferred to unseen regions. As a result, the ResNet50-32 model is selected for use upon the mosaics of Crookes and Schomberger A craters—and herein nickname the model 'IMFMapper'.

Figures 6.12 and 6.13 present examples of the visual performance of IMFMapper upon the testing data of Copernicus and Virtanen F craters, respectively. Figure 6.12 highlights some of the TP, FP and FN detections made by IMFMapper, showing that IMFs with prominent shadows and enclosed debris were most frequently detected, with more ambiguous sections

being missed. Figure 6.13 suggests that IMFMapper is able to detect some of the thinnest IMFs, with a sufficient recall for finding IMF complexes. However, thin IMFs were occasionally detected with sparse extents, leading to incorrectly predicted widths and/or connected sections. Figure 6.13 also gives an example of a common FP across both regions—in ridges or boulders casting shadows due to low Sun angles.

As the values in Figure 6.11 and Table 6.3 suggest, there is some discrepancy between the average F_1 of IMFMapper upon the validation and testing datasets (69.4 and 55.9%, respectively). This discrepancy is a consequence of two factors, with the first being that there is a greater proportion of image tiles hosting no IMFs in the testing dataset than in the training and validation datasets. Therefore, there is a larger number of empty tiles which, when one or more false IMF detections are made within them, will result in a P of 0%.

Secondly, the testing data deliberately represents a more difficult distribution of IMFs. For example, features similar to the large bowl pits catalogued in the LPA within Copernicus, and shown in M1136683665L of Figure 6.12, had not been seen during training. The IMFs in Virtanen F would also be significantly trickier to detect, due to their small widths, than the majority of those seen during training/validation. The increased difficulty with detecting thin IMFs is reflected in the fact that the IoU calculated between the detections and labels made by IMFMapper within Virtanen F was 21.2%, as opposed to 40.1% for Copernicus.

As previously mentioned in Section 6.2, the IMFs in Copernicus crater were labelled as polygons, as opposed to the weak line labelling of Virtanen F (as was the case for the training and validation data). The purpose of labelling the IMFs in Copernicus with polygons was to be able to visualise whether the weak labels have influenced the geometries of the detections made by IMFs. Crucially, the shapes of the detections within Copernicus crater indicate that the weakly-supervised training of IMFMapper has succeeded in developing a model which

Figure 6.12: (Following page) Visual testing performance of IMFMapper within Copernicus crater. Examples of correct and false detections, as well as missed IMFs, are shown for each of the 5 LROC NAC images. M1175545925L is shown twice to highlight the detection of the LPA feature 'Copernicus 4' (bottom left).

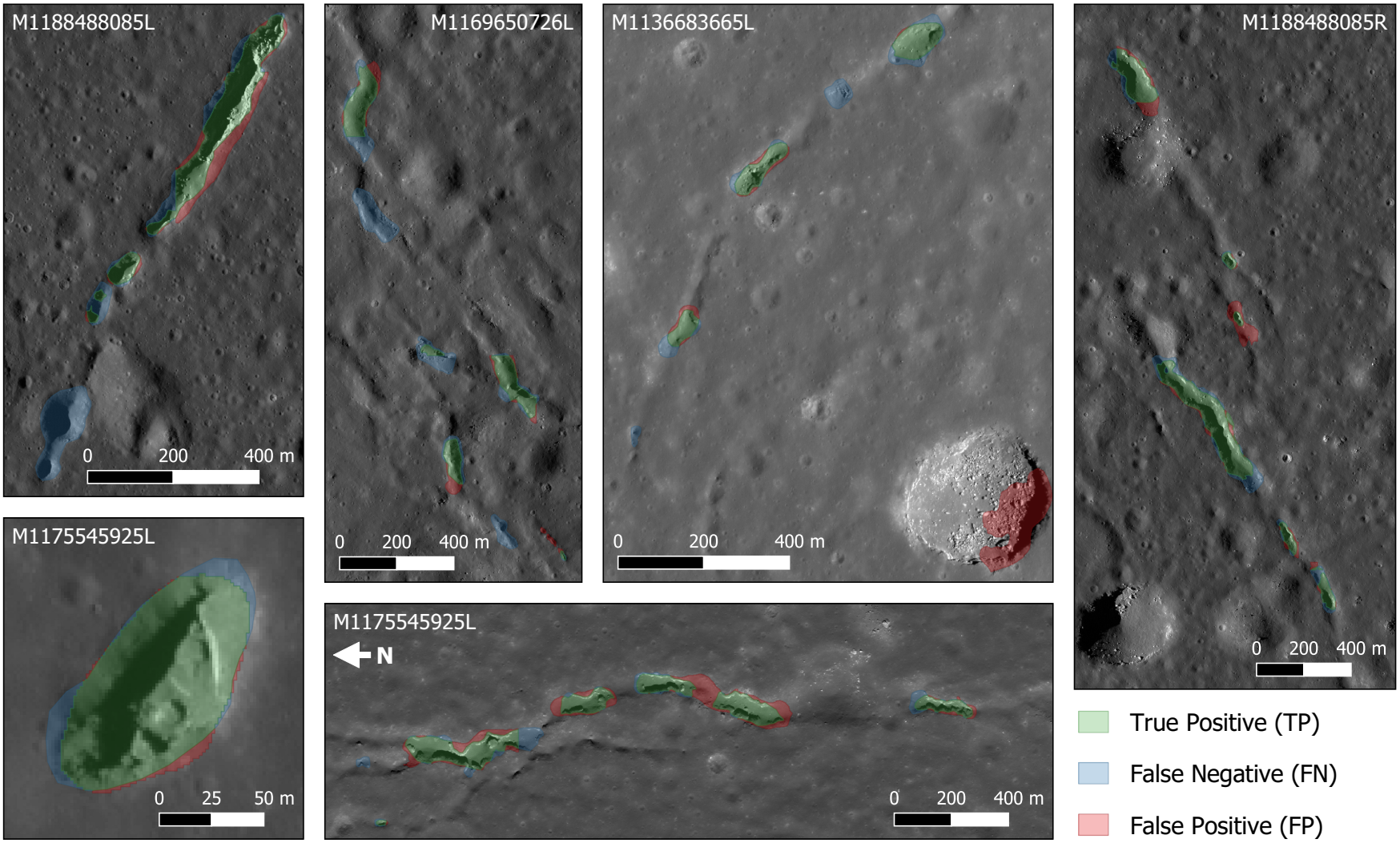

Figure 6.12: (Caption on previous page)

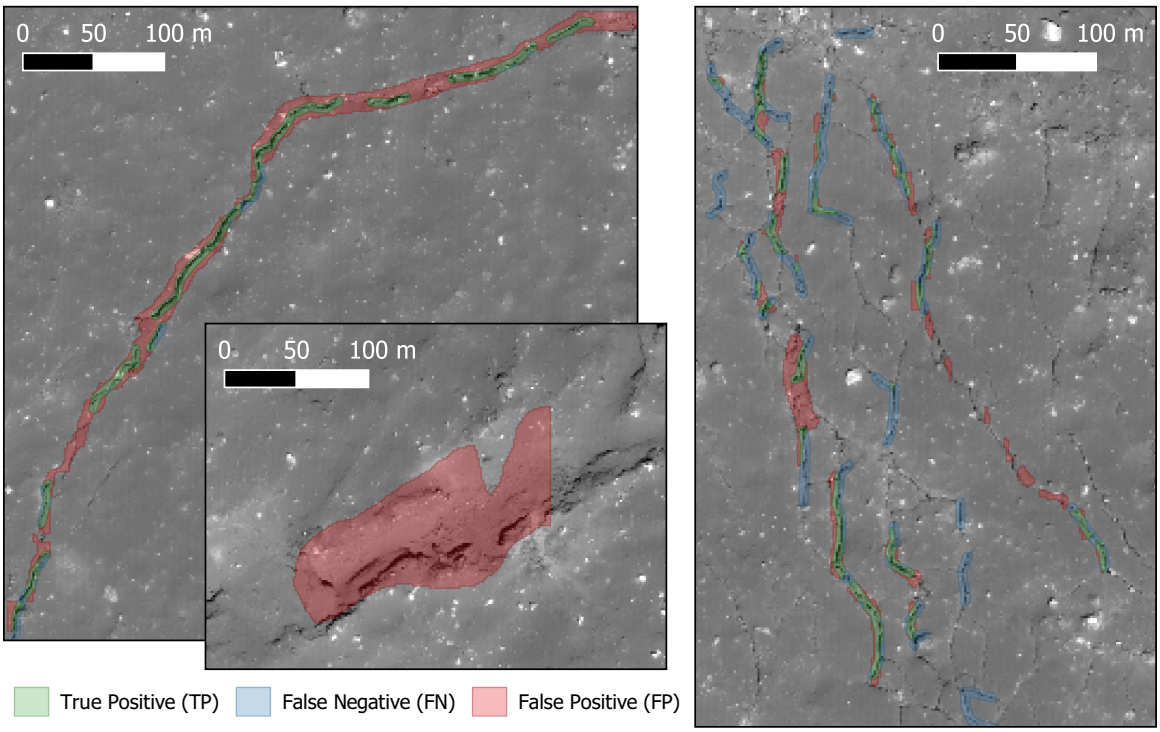

Figure 6.13: Visual testing performance of IMFMapper within Virtanen F crater. The examples shown here (all within LROC NAC image M156435317L) present the performance of IMFMapper upon some of the thinnest IMFs, as well as a common FP in ridges.

can delineate the pixel-level extents of IMFs without the need for equally detailed labelling. Whereas, the detections in Virtanen F imply that, while the principle of weakly-supervised training has worked, the values by which the lines for the 'tiny' IMFs were buffered may require revision or that occasionally such IMFs were mislabelled as 'small'.

6.4.3 Mapping of IMFs in Crookes and Schomberger A Craters

Now that IMFMapper has been tested upon IMFs within previously unseen craters, it can now be applied to the self-processed mosaics of Crookes and Schomberger A craters. As is stated in Table 6.2, the tiling of the mosaics of Crookes and Schomberger A produced 3,332 and 1,341 image tiles, respectively, to be fed through IMFMapper. Inferring IMFMapper on these quantities of image tiles took just 1.01 and 0.70 mins, far surpassing what would be achievable by humans.

The detections made by IMFMapper were converted from raster format (i.e. binary images

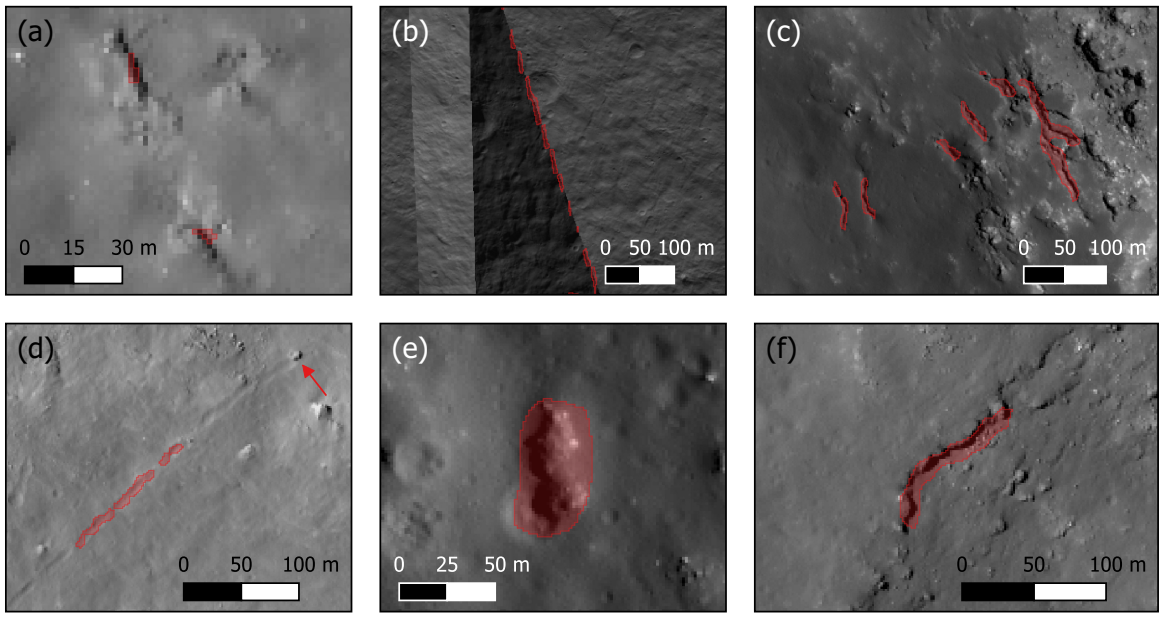

Figure 6.14: Examples of some of the detections made by IMFMapper within Crookes and Schomberger A craters that were filtered out due to (a) insufficient area, (b) mosaic seaming or (c) being located on the crater's rim. (d to f) Examples of some of the false positives which were observed: (d) A rockfall trail left by the falling boulder marked with a red arrow, as well as prominent shadows cast by an isolated boulder (e) and a ridge (f).

where 1=IMF and 0=background) to ESRI shapefiles using the GDAL Python API—wherein detections made at the edges of adjacent tiles could be merged. Since the resolution of the tiles remains constant, the surface area (in m^2) of each merged IMF could also be recorded. This allowed for the filtering out of any detections (smaller than 150 m^2 in area) which are likely too small to be resolvable and identified as IMFs. Similarly, adverse detections made along the seams of the mosaics were removed since these were not indicative of IMFMapper's performance but of the quality of the data to which it had been applied. Furthermore, only the IMF detections which fell within the craters' floors and, in the case of Crookes, any melt ponds found within the terraces of its walls were considered for inclusion in the maps presented in Figures 6.15 and 6.17.

Figure 6.14 gives examples of some of the detections which were filtered out, either due to insufficient size, mosaic seams, or location—as well as some common false positives which remained in the maps. For example, despite being filtered out due to their locations on the

crater wall, the majority of the trails left by rockfalls within Crookes crater were misclassified as IMFs. Therefore, there is potential for IMFMapper to be transfer-learnt to satellite imagery of rockfalls and compared to previous efforts in literature (Bickel et al., 2019, 2020). This would also be the first use of SS for Lunar rockfalls since Bickel et al. (2019) trained a RetinaNet OD DCNN model.

6.4.3.1 Crookes Crater

Figure 6.15 presents the automated mapping of IMFs within Crookes crater by IMFMapper. The centre map of the figure plots the coordinates of the detected IMFs (in green) against the LROC WAC global mosaic with 50% transparency (Speyerer et al., 2011). Within this map, the crater's floor is outlined in red, which is also overlaid upon the mosaic of 56 LROC NAC images produced as part of this work.

The melt ponds identified by Thaker et al. (2020) in the northern and western walls of Crookes are also highlighted in blue. However, the presence of IMF detections within smooth, low albedo terrains suggests the discovery of additional melt ponds: two smaller melt ponds in the western wall, as well as another in the south. Maps (a) and (c) of Figure 6.15 show IMF detections being made within the melt ponds identified by Thaker et al. (2020), with the IMFs in (a) being the same ones spotted in Figure 6.4. Whereas, map (b) exhibits parallel IMFs in the candidate melt pond situated approximately 2 km north of the western melt pond mapped by Thaker et al. (2020). Map (d) displays the largest of the IMFs detected in the proposed southern melt pond in Crookes crater, which is ≈600 m-long and between 15–40 m in width. In total, 70 IMFs were detected by IMFMapper within the melt ponds of Crookes

Figure 6.15: (Following page) Detections made by IMFMapper within Crookes crater. (Centre) IMF detections within the floor (red) and melt ponds (blue) of Crookes, overlaid upon the LROC WAC global mosaic (Speyerer et al., 2011); (a) The IMFs shown in Figure 6.4 being detected by IMFMapper; (c) IMFs being detected within the western melt pond presented in Figure 6.5; (b and d) Discovery of candidate melt ponds in the western and southern walls due to detected IMFs; (e) the largest IMF detection made by IMFMapper in Crookes; (f) IMF detections within the region shown in Figure 6.3.

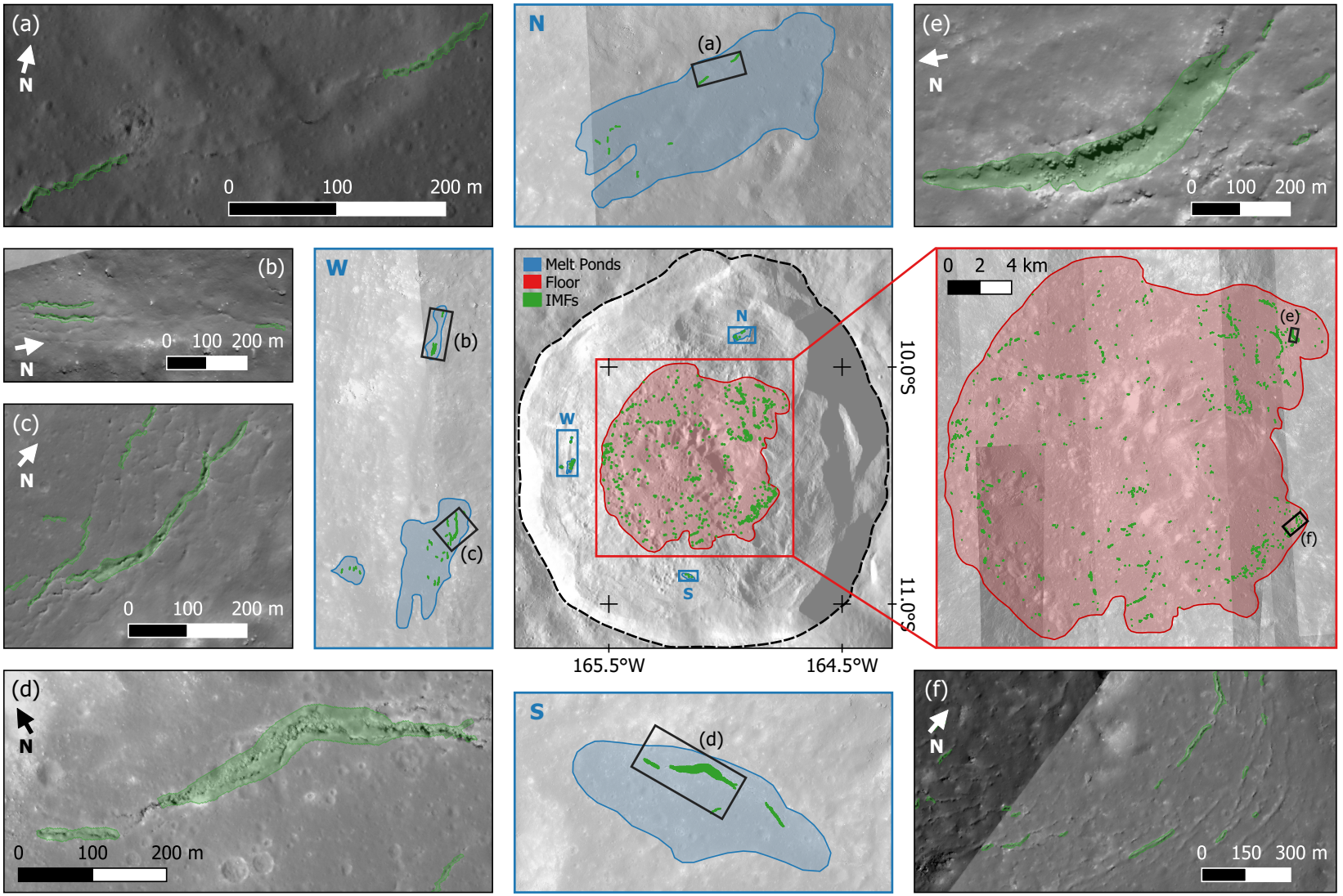

Figure 6.15: (Caption on previous page)

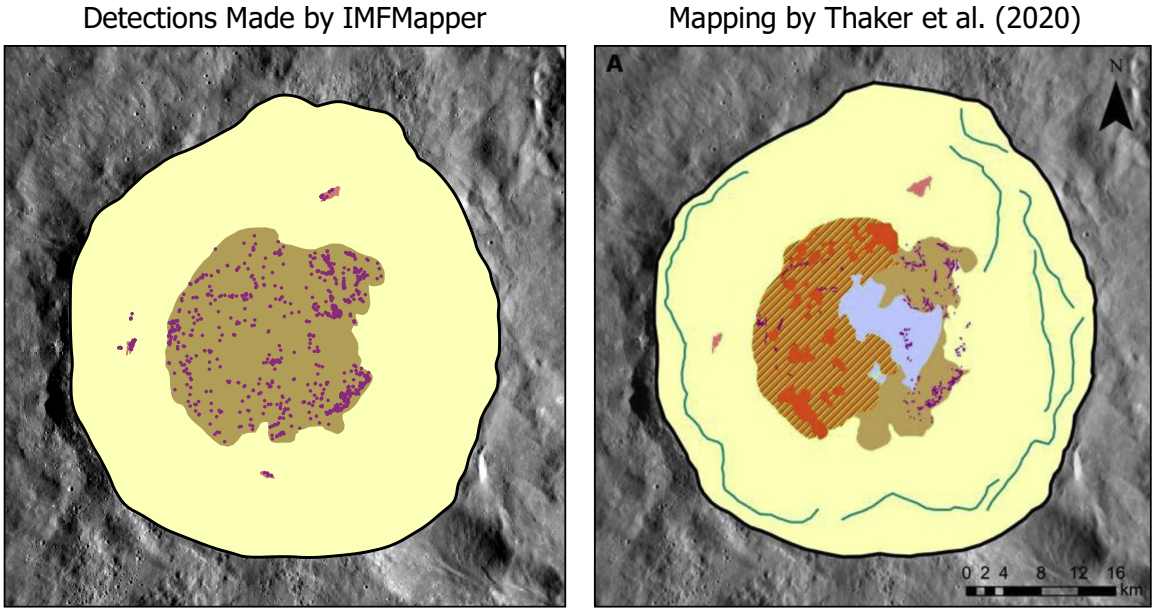

Figure 6.16: Direct comparison between the results of applying IMFMapper to a mosaic of LROC NAC images of Crookes crater and the map produced by Thaker et al. (2020), the latter of which has been adapted for this figure with permission. The symbology in this figure is identical to that in Figure 6.5.

crater, of which most were found within the western melt pond shown in (c).

With regard to the IMFs mapped within the crater floor, map (e) of Figure 6.15 contains the largest IMF found anywhere in Crookes due to it possessing a surface area, length and maximum width of more than 0.05 km^2, 700 m and 90 m, respectively. Map (f) also presents the IMFs detected by IMFMapper within the region towards the top right of that shown in Figure 6.3, proving that it can detect complexes of sub-parallel IMFs.

The examples in Figure 6.15 reveal the fact that features of various scales and morphologies have been recalled by IMFMapper in regions of Crookes crater previously highlighted by experts. However, in order to better aid the comparison between the results of IMFMapper and the map by Thaker et al. (2020) of Crookes crater, Figure 6.16 re-plots the mapping shown in Figure 6.15 with the same symbology as in Figure 6.5. This shows that many of the denser patterns of IMFs mapped by Thaker et al. (2020) are also visible within the mapping by IMFMapper. Despite this agreement, there are noticeably more IMFs mapped by IMFMapper, particularly in the hummocky interior of the crater floor. Upon visual inspection, the rates

of FPs are sufficiently low that it can be said that IMFMapper is detecting IMFs which have been previously missed. Detections of previously uncatalogued IMFs have also aided in the discovery of candidate impact melt ponds. All of these factors help towards the case of using DL to automatically map IMFs.

6.4.3.2 Schomberger A Crater

Figure 6.17 is the first mapping of the IMFs found within Schomberger A crater, as completed automatically by IMFMapper. The top left map overlays the coordinates of the IMF detections (in green) within the visible portion of Schomberger A's floor (in red) upon the mosaic of 55 LROC NAC images created in this work (with 50% transparency). Maps (a to c) display the patterns of IMF detections at the 'complex' scale, and maps (d to f) zoom in to show individual IMFs in greater detail.

Maps (a) and (e), in particular, correspond to the same impact melt flow shown in Figure 2 of Lee (2022), showing how the spacing between the sub-parallel IMFs appears to get closer towards the margin between melt and wall. Maps (b) and (e) focus on some of the largest IMF detections (by area) with lengths between 350–400 m and widths of ≈15 m. Maps (c) and (f) show how IMFs are increasingly missed by IMFMapper due to the shadow cast by the crater's rim encroaching upon its floor.

The promising visual performance of IMFMapper within Schomberger A crater, despite the extreme latitudes and sensing conditions, means that there is potential for IMFMapper to be applied to further examples of south-polar craters harbouring impact melt deposits. Such an exercise could generate a regional mapping of IMFs, which may be of interest for future space exploration. For example, knowledge of the distribution of IMFs across the south pole

Figure 6.17: (Following page) Detections made by IMFMapper within Schomberger A crater. (Top left) IMF detections within the visible floor (in red) of Schomberger A, overlaying the self-processed mosaic of 55 LROC NAC images; (a and d) IMF detections made within the same regions as those shown in Figure 2 of Lee (2022), respectively; (b and e) The largest IMF detections made by IMFMapper in terms of area, shown at the 'complex' and pixel-level scales; (c and f) IMFs being detected close to the terminus of the visible portion of the crater floor.

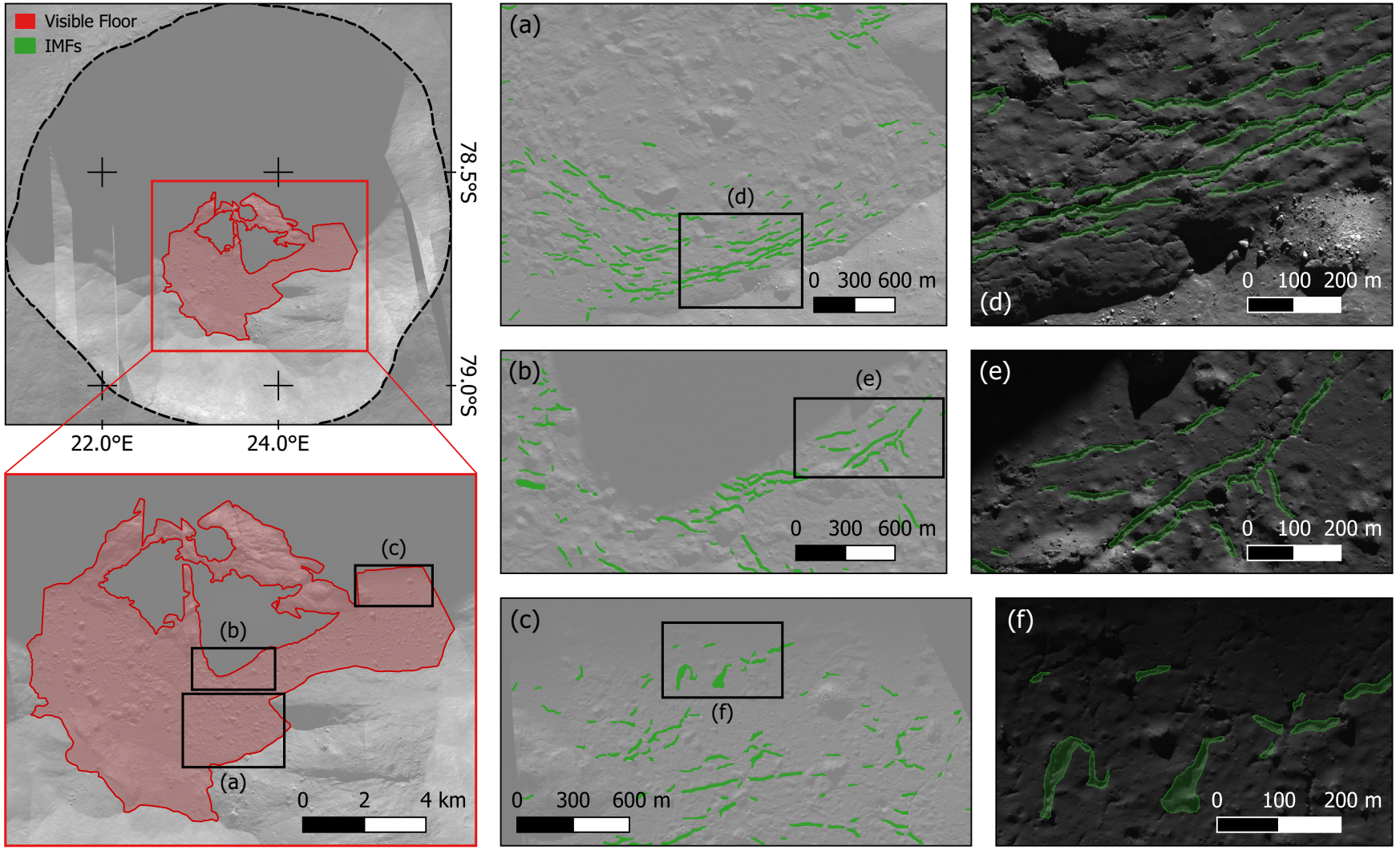

Figure 6.17: (Caption on previous page)

will surely be an important factor in the decision of which impact melt to target for an impact melt sampling mission, since any such effort may be able to append a secondary objective of exploring their potential as cave entrances and/or PSRs containing water ice.

6.5 Key-Point Summary

- IMFMapper is a DeepLabV3 semantic segmentation DCNN trained upon over 4,000 weakly labelled impact melt fractures (IMFs) within 25 LROC NAC images of Lunar impact craters.
- Due to memory constraints when using backbones with more network layers, three different DeepLabV3 models were trained: ResNet50 backbone with a batch size of 32, ResNet50 with a batch size of 16, and ResNet101 with a batch size of 16. At the end of training, the models achieved their highest average F_1-scores of 69.4, 69.8 and 70.0%, respectively.
- However, during testing upon IMFs within Copernicus and Virtanen F craters, it was found that the ResNet50-32 configuration (which is called 'IMFMapper') returned the highest average F_1 and IoU of 55.9 and 38.7%, respectively. The IoUs, when averaged separately across all testing samples from Copernicus or Virtanen F, were 40.1 and 21.2%, respectively, showing that Copernicus contains more obvious IMFs.
- Applying IMFMapper to a mosaic of LROC NAC imagery of Crookes crater achieved good visual agreement between the distributions of IMFs within the maps by Thaker et al. (2020) and IMFMapper. However, additional candidate melt ponds have been found in the western and southern walls of Crookes due to the smooth, low albedo terrain and the presence of IMFs detected by IMFMapper.
- In Schomberger A, IMFMapper has produced the first mapping of IMFs in a crater so close to the Lunar south pole, where these features have the potential to be permanently shadowed regions (PSRs). With the planned arrival of the Artemis missions, whose objective is to sample an impact melt deposit (Weber et al., 2021), IMFMapper could be used as a proxy to search for uncatalogued melt deposits. IMFMapper may also be of use for transfer learning to remote-sensing imagery of Lunar rockfalls, since they were a common false positive.
- Inferring IMFMapper on the mosaics of Crookes and Schomberger A craters took just 1.01 and 0.70 mins, respectively, far surpassing what would be achievable by humans.
- IMFMapper has laid the foundations for future work concerning the use of automated techniques, such as deep learning, to map IMFs within SIMBIO-SYS high-resolution imagery upon the arrival of BepiColombo at Mercury.

Chapter 7. Conclusions

7.1 Summary

7.1.1 Summary of Chapter 4: PITS

The Pit Topography from Shadows (PITS) tool described in Chapter 4 has been tested with a range of image segmentation methods in order to automatically extract shadows from planetary remote-sensing imagery of pits. Detecting shadow pixels allows for the measuring of the shadow's width along its entire length, which, along with knowledge of the Sun's position, returns a profile of the pit's apparent depth (h)—the relative depth at the edge of the shadow. PITS is an automated approach to calculating h, which has been performed manually before in literature (Wyrick et al., 2004; Cushing et al., 2015). PITS develops the work further by implementing a generalisation of the correction applied by Cushing et al. (2015) such that PITS can correct the observed shadow widths for all values of satellite emission (ε), solar azimuth (ϕ) and satellite azimuth (ω) angles.

Otsu's method and k-means clustering with silhouette analysis were adapted and tested for automated shadow extraction from shadow-labelled Mars Reconnaissance Orbiter (MRO) High Resolution Imaging Science Experiment (HiRISE; McEwen et al. 2007) imagery of Mars Global Cave Candidate Catalog (MGC^3) atypical pit craters (APCs)—19 red-band and 12 colour images in total. Otsu's method generated highly variable results since only two classes (foreground and background) were often insufficient to delineate the shadow without including non-shadow pixels within the detection. k-means clustering with silhouette analysis was also tested when maximising the overall silhouette score of all clusters and when only maximising the silhouette coefficient of the darkest cluster ($\bar{s_0}$). The latter achieved the highest and least

variable F_1-score (F_1) of 97.1% (with a standard deviation of 2.6%) upon the 19 red-band images across all methods of automated shadow extraction tested in this work. Therefore, k-means clustering with silhouette analysis when maximising $\bar{s_0}$ is the method that has been implemented into the PITS algorithm. A marginal improvement in F_1 was observed when applying k-means clustering when maximising $\bar{s_0}$ to the 12 colour images compared to the performance on the corresponding red versions. Yet, the increase in F_1 was not sufficiently large to merit the significant increase in average run-time.

PITS has also been applied to 123 HiRISE red-band observations of 88 MGC^3 APCs. Across the entire dataset, PITS exhibited a minimum and maximum run-time of ≈5 and 500 s, respectively. The variation in run-time was a result of the tool being applied to a range of image sizes. These are much shorter time-scales than can be achieved manually, considering the range of products that PITS can output. Applying PITS' ε-correction significantly reduced the relative standard deviation in the h calculated by PITS (σ_h/h) for the majority of APCs that had multiple HiRISE observations. In situations where σ_h/h did increase after ε-correction, this was explained by the shadow's edge being cast to a location in the pit floor at a different depth, or the pit's floor being fully covered in shadow over multiple images with ranging solar incidence angles (α).

The h values calculated by PITS for these 88 APCs have been compared to the depths provided in the MGC^3 catalogue. According to the relevant documentation, the quoted h values in MGC^3 were "generally" calculated using the approach in Cushing et al. (2015), with a correction being applied to all images with $\varepsilon > 5°$ (Cushing, 2015). Assuming that PITS and MGC^3 have measured shadow widths from the same HiRISE images, the h_c and h_m calculated by PITS upon images that are the only HiRISE RDRV11 observation of a given APC were plotted against their corresponding MGC^3 h. Figure 4.13 showed that PITS and MGC^3 calculate comparable h in many cases. Although for the features where they disagreed most, there were no visible errors with the extracted shadows. The lack of an error in the shadow extraction suggests that some of the MGC^3 APC depths are instead estimates of the maximum depth of the pit, as opposed to h. This explanation is supported by the above

statement in the MGC^3 documentation, as well as the fact that the h_c and h_m calculated by PITS are less than those in MGC^3 in the majority of instances.

The outputs of these 123 HiRISE images have been used to identify 11 examples (taken of 10 APCs) whose h profiles, as well as the context provided by the image, suggest that they could be possible cave entrances. The wider context came in the form of using PITS' geo-referenced shapefiles of the detected shadows to enhance the contrast of the shadow pixels, whereby any deeply shaded regions may be due to steep or overhanging rims. These 10 APCs were considered to be possible cave entrances since their h profiles exhibited a decline towards one of these deeply shadowed regions. It was also considered whether wider shadows were, in fact, being cast by an uneven pit rim as opposed to being a result of the pit's topography. Whilst it is impossible to definitively confirm that these features are indeed cave entrances without exploring them on the surface, tools such as PITS will help to narrow down the features with the highest priority for investigating Martian cave entrances for astrobiological significance and habitability feasibility.

Given that the approach of PITS is highly applicable to data from other sensors, the tool has been applied to 10 Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) images covering the Mare Tranquillitatis Pit (MTP). When considering the uncertainty in the h profiles, the values retrieved by PITS agreed with a series of manual measurements within the literature, such as its sloping towards its eastern rim. This first use of PITS on a Lunar example has returned promising results. However, the same testing procedure that was carried out for Martian pits is also needed here to assess the shadow extraction performance within LROC NAC data.

The work presented in Chapter 4 (prior to Section 4.4.3) has been published in the journal *'Royal Astronomy Society's Techniques and Instruments (RASTI)'* (Le Corre et al., 2023a). The front page of the published article is appended to the end of this thesis. See Appendix D for a full list of conference presentations and published or in-review research articles that have been produced as part of this thesis.

7.1.2 Summary of Chapter 5: ESSA

In Chapter 5, a series of Mask R-CNN (region-based convolutional neural network) instance segmentation deep learning (DL) models have been trained to detect Lunar pits and skylights. The purpose of which was to find candidate cave entrances that may provide access to intact lava tubes. The various Mask R-CNN models were trained upon different combinations of LROC NAC and MRO HiRISE imagery of the Moon and Mars, and a range of backbone models—from the ResNet, DenseNet, VGG and MobileNet families. Due to the discrepancy in the best resolutions achievable with LROC NAC and HiRISE imagery, as well as the range of resolutions seen within each dataset, all imagery was downscaled to a common resolution of 1.5 m/px. Images were labelled using known pits and skylights from various catalogues, such as the Lunar Pit Atlas (LPA; Wagner and Robinson 2021), MGC^3 (Cushing et al., 2015), lava-tube-related rilles mapped by Hurwitz et al. (2013), and the potential cave entrances found by Watson and Baldini (2024). After tiling, synthetic Moon training samples were also generated by implanting warped HiRISE pixels covering Martian pits and skylights into empty LROC NAC tiles.

During validation, the highest average F_1-scores of the bounding boxes and predicted masks (89.1 and 95.6%, respectively) were achieved by the Mask R-CNN that was trained on Lunar, Martian and synthetic Lunar data with a ResNet50 backbone. This particular model, named ESSA (Entrances to Sub-Surface Areas), then underwent testing upon self-produced mosaics of LROC NAC observations of the Gruithuisen, Marius A and Marius B lava tube candidates proposed by Sauro et al. (2020), as well as five LROC NAC images containing the Mare Tranquillitatis Pit (MTP). This testing led to an average F_1 of 82.1 and 93.7% for the bounding boxes and predicted masks, respectively, with the MTP being detected in every image with high confidence. The testing results on the lava tube mosaics also highlighted that some features that were not seen during training, such as kilometre-scale impact craters, commonly led to false positive detections.

Following testing, ESSA was applied to LROC NAC imagery covering eight regions of the Moon's mare deposits (six being impact craters or basins and two being Lunar maria them-

selves). Considering the rarity of skylights on the Moon, and the fact that ESSA was applied to just 1.92% of the total maria, two previously uncatalogued skylights have been detected. The first, dubbed the South Marius Hills Pit, was detected in a region which had already been surveyed by the LPA and is within ≈4.3 km of the Marius A proposed lava tube collapse analysed by Sauro et al. (2020). Therefore, this particular area of Marius Hills represents a potential site of interest for space missions, such as Moon Diver, which are tasked with investigating cave entrances. The second, situated within the satellite crater Bel'kovich A, appears as though the prominent shadow cast within is adjacent to an overhanging rim. However, due to the high latitude of the feature, the Sun's illumination will only cast northwards, meaning that observations with other data types may be required to confirm the presence of a cave entrance. The detection of a skylight so close to the north pole shows that the use of DL in this work has extended the latitude range that the search for cave entrances can continue within, but also that there is scope for repeat surveys of areas which had already been thought to have been completed.

ESSA has also been inferred upon the January 2025 release of HiRISE RDRV11 images in order to assess its preliminary performance compared to the Lunar surface. This resulted in several detections which appeared morphologically similar to pits or skylights until further inspection of the wider context of the image and the sensing conditions when the image was taken. Confusion was particularly caused by circular depressions, which also exhibited low albedo material. Since ESSA has no concept of α or ϕ, it is understandable why some of these features were falsely detected.

The work presented in Chapter 5 (prior to Section 5.2.3.2) has been published in the journal *'Icarus'* (Le Corre et al., 2025a). The front page of the published article is appended to the end of this thesis.

7.1.3 Summary of Chapter 6: IMFMapper

IMFMapper is a DeepLabV3 DCNN with a ResNet50 backbone trained to perform the semantic segmentation of impact melt fractures (IMFs) within LROC NAC imagery of Lunar impact melt

deposits. Due to the quantities of IMFs available to label and the difficulty with creating pixel-level annotations for linear features, IMFMapper was trained in a weakly-supervised manner (i.e. upon labels which were produced from a series of buffered line labels). Similar to in Chapter 5, all LROC NAC imagery used for training, testing or inference required downscaling to a common resolution of 1.5 m/px. During validation, it was found that IMFMapper achieved an average F_1 of 69.4% after just 47 out of the total 100 epochs of training. This was marginally less than the highest average F_1 reached by a DeepLabV3 model with a ResNet101 backbone and a smaller batch size (70.0%). However, during testing of the three combinations of model backbone and batch size upon IMFs within Copernicus and Virtanen F craters, it was found that the configuration ultimately called 'IMFMapper' returned the highest average F_1 of 56.8%.

By applying IMFMapper to two self-processed mosaics of LROC NAC imagery, two maps have been created of the IMFs found within Crookes and Schomberger A craters on the Moon. The former has been the subject of a previous manual mapping exercise (Thaker et al., 2020), whilst the latter is confirmed to host IMFs (Lee, 2022) that may act as permanently shadowed regions (PSRs). In the absence of the ability to quantitatively compare the two, the map of Crookes crater generated by IMFMapper visually agrees with the distributions of IMFs in Thaker et al. (2020). However, additional candidate melt ponds have been found in the western and southern walls of Crookes due to the smooth, low albedo terrain and the presence of IMFs detected by IMFMapper. In Schomberger A, IMFMapper has produced the first mapping of IMFs in a crater so close to the Lunar south pole. With the planned arrival of the Artemis missions, whose objective it is to sample an impact melt deposit (Weber et al., 2021), the ability to map IMFs at these latitudes and at scales unachievable by humans (both spatially and temporally) will be a useful tool—either to assess their potential as PSRs or to use this as a proxy to search for uncatalogued melt deposits.

Even a common false positive detection made by IMFMapper presents an interesting result, in that the majority of Lunar rockfalls were classified as IMFs. Therefore, the learnt model weights that make up IMFMapper may prove advantageous for the transfer learning of a

DeepLabV3 model trained to detect rockfalls.

The work presented in Chapter 6 has been published in the *'Journal of Geophysical Research (Planets)'* (Le Corre et al., 2025b). The front page of the published article is appended to the end of this thesis.

7.2 Relating to Thesis Objectives

As discussed in Chapter 1, this PhD project was established with several aims and objectives, which the three constituent research efforts presented in Chapters 4 to 6 should collectively and individually satisfy. This section will determine, based upon the evidence in the respective chapters, whether these goals have been met.

The primary objective was that the projects should all involve the use of ML or DL techniques to automatically detect or analyse planetary surface features. Moreover, they should quantify the performance and justify the use of the techniques used in order to compare with previous manual or automated literature. As such, Chapter 4 assessed the shadow extraction performance by manually labelling the shadows of HiRISE red-band and colour images of Martian pits. Chapter 5 calculated performance metrics for the various Mask R-CNNs during training and validation upon different combinations of Lunar and Martian pits/skylights, as well as upon testing data containing known candidate lava tube collapses. The same performance monitoring was also performed in Chapter 6, with the testing data coming from two impact craters which were not seen during the training process. The use of P, R, and F_1 was constant throughout these projects as a means of depicting the relative accuracy and difficulty of the tasks being performed in each. Meanwhile, Chapters 4 to 6 presented the many advantages that justify the use of these automated methods as opposed to the manual alternative. Firstly, PITS' automated depth estimation produces a profile along the entire length of the shadow, as well as the geo-referenced shapefiles of the detected shadows. These outputs would take considerable time to be manually produced, particularly as one would need to measure and correct the shadow width at every pixel. In the case of Marius Hills, ESSA was able to review

4,905 image tiles in only 4.35 mins, within which time it detected the previously uncatalogued East Marius Hills skylight and recalled both of the residing LPA features (Marius Hills and West Marius Hills skylights). Similarly, the good visual agreement with the map of Crookes crater by Thaker et al. (2020) and the detection of IMFs in new candidate melt ponds were achieved by IMFMapper within only 1.01 mins. Applying IMFMapper to the mosaic of Schomberger A was also completed even more quickly (in just 0.70 mins). However, the unique advantage of IMFMapper is that it was trained using weak line labels, rendering the labelling less laborious and prone to errors. As a result, the detailed performance assessments, the levels of outputs generated, and the timescales that these were achievable within, all serve to prove the reasons for using ML/DL in each of the use cases within this work—and in future.

Additionally, there is the aim of ensuring that the thesis would be approachable by researchers unfamiliar with either computer or planetary science. Working towards this, Chapter 2 simultaneously introduced readers to the various categories of planetary surface features, as well as the necessary ML/DL principles and models. Chapter 3 also detailed which ML/DL techniques and architectures have been previously utilised and the surface features that they have been applied to, such that a researcher coming from either background can more readily see where gaps in the literature could be filled in future. With this aim in mind, Chapters 4 to 6 reported how planetary remote-sensing datasets have been acquired and processed, as well as where ML/DL algorithms were sourced from and how they were trained and assessed—whilst being explicit with considerations such as confidence thresholds, for example. Appendix E also lists the primary remote-sensing datasets and feature catalogues that were exploited in this thesis, which future researchers may find useful. For these reasons, this thesis will hopefully act as an enabler for planetary or computer science researchers with limited experience in the opposite field to consider using planetary remote-sensing datasets and use ML/DL architectures to exploit them.

Chapter 3 provided evidence for the need for the work in this thesis to move away from impact craters and crater detection algorithms (CDAs) and instead focus on different planetary surface features of equal significance. It just so happens that in trying to achieve this, the

surface features which have received ML/DL treatment within this thesis are all touted as entrances to underground cavities. Whilst there had been some previous work in detecting pits and skylights within satellite imagery of Mars (CTX: Watson and Baldini 2024; HiRISE: Nodjoumi et al. 2021, 2023), no automated techniques had yet been applied to the Moon in the search for more Lunar examples. Chapter 6 presented a first venture into using DL to automatically map IMFs on any planetary surface. Therefore, this thesis has expanded the number of surface features, as well as the number of surfaces themselves, which have been the subject of comprehensive treatment from ML/DL.

Finally, the tools in this thesis should be made publicly available in order to foster and accelerate future work in these areas. This includes the algorithms or models themselves, but also the frameworks to be able to use them for future research. For this reason, PITS is available to download from Zenodo as a containerised Docker image, such that it is disseminated with all of the necessary Python packages (Le Corre et al., 2023b). The instructions for how to set up PITS locally from this Docker image can be found on its GitHub repository (Le Corre, 2025e). The PyTorch model checkpoint files for both ESSA and IMFMapper can be downloaded from Zenodo (ESSA; Le Corre 2025b, IMFMapper; Le Corre 2025d), from which they can be transfer-learnt on new training datasets or inferred on previously unseen data. As such, Python scripts have been made available on the GitHub repositories for ESSA (Le Corre, 2025a) and IMFMapper (Le Corre, 2025c), which allow users to infer each model on their own tiled (2,048x2,048 for ESSA and 512x512 for IMFMapper) and downscaled (to 1.5 m/px) LROC NAC data. In terms of the content of this thesis, Technical Notes have been given in Appendix A as examples of some of the code and processes used in this work. Technical Note TN-1 provides the functionality of filtering cumulative index files for use with the PITS tool. Technical Note TN-2 outlines code using the ISIS and GDAL suites for the batch processing of raw LROC NAC imagery. Lastly, Technical Note TN-3 presents the method for correcting the map-projection error in some HiRISE RDRV11 images. Consequently, with this suite of publicly available tools, frameworks and scripts, there is an abundance of materials to encourage future work.

7.3 Significance

Whilst the work in this thesis has satisfied the objectives originally set in Chapter 1, there is also a wider significance for the areas of planetary surface feature detection, future space exploration, astrobiology and planetary geology.

The PITS tool is relevant to all three of the fields mentioned above, as a result of it deriving h profiles via k-means clustering from just a single cropped satellite image of a pit or skylight. PITS has filled a gap in the literature since previous studies typically only used ML/DL to detect features, not analyse their morphometry. With PITS, researchers can also gain an estimate of the depth, volume and topography of a pit or skylight in scenarios where there is no stereo imagery available for calculating DEMs or the elevation across the feature is not resolvable in global DEMs. Therefore, where there is a lack of other data sources, PITS allows for the search to continue for potential cave entrances, which mission proposals, such as Moon Diver (Nesnas et al., 2023), may want to consider as candidate landing sites. The discovery of 10 MGC^3 APCs exhibiting characteristics of cave entrances (i.e. a downward sloping in the h profile towards an overhanging rim or deeply shaded region, as revealed by using the shadow shapefile to increase contrast) is an example of PITS being used to narrow down which features are of most interest for future exploration. As discussed in Chapter 4, planetary caves are considered one of the "best solutions" for a habitat on Mars for mitigating the effects of radiation exposure (Atri et al., 2022) and have been simulated to host stable reserves of ice water (Williams et al., 2010).

As was mentioned in Section 7.2, ESSA represents the first effort to use ML or DL methods to detect Lunar pits and skylights—whilst PitScan was the first attempt to search for pits and skylights on the Moon by any automated means (Wagner and Robinson, 2014). However, where PitScan was limited to the $\pm 50^\circ$ latitude range, ESSA is not restricted in the same way (albeit FPs will increase with latitude due to smaller elevation changes casting larger shadows). ESSA's robustness to extreme latitudes has been proven by the detection of the candidate skylight in Bel'kovich A crater. Therefore, ESSA has expanded the areas in which Lunar pits and skylights can be searched for, with the relevance for this being the same

as stated in the paragraph above. In fact, between PITS and ESSA, two tools have been developed that, when combined, could accelerate the discoveries of pits and skylights on the Moon and Mars, including those that exhibit signs of a cave entrance.

Where ESSA was the first use of DL to detect pits and skylights on the Moon, IMFMapper is the first DL model to be trained on and inferred to detect IMFs on the Moon. As was the focus for Xiao et al. (2014), the mapping of IMFs within impact melt deposits can give an indication of the more dominant mechanisms as the melt cools after the impact. Therefore, IMFMapper opens the door to scaling up this investigation from the three Lunar and three Mercurian impact craters in Xiao et al. (2014) to as many as have corresponding high-resolution imagery. As was the case in Crookes crater, the detection of IMFs within its walls led to the discovery of new candidate melt ponds. Thus, if IMFMapper were to be applied to the surface surrounding impact craters near the south pole, this may result in the identification of smaller and more reachable melt deposits for the sample return objective of the Artemis III mission (Weber et al., 2021). The detection of IMFs in reachable areas of the Lunar surface would also be useful for any future mission which may want to investigate whether they could provide access to underground cavities similar to terrestrial analogues (Lemaire et al., 2024). The same is also true for the search for PSRs that may cold-trap volatiles, such as water (Brown et al., 2022), which would be critical for prolonged crewed missions on the Lunar surface.

7.4 Limitations

The work shown in this thesis, and any future work which develops upon it, comes with some limitations which require consideration. The main limitation that requires clarification is that none of the surface features concerning Chapters 4 to 6 (pits, skylights, sub-surface cavities, lava tubes, IMFs, nor PSRs) have been visited by robotic or crewed missions on the surface of the Moon or Mars to date. Therefore, any mention of these features being what they are is on the basis of remote observations and comparisons with known terrestrial analogues here on Earth. An example of this would be the comparisons made by Sauro et al. (2020)

between known terrestrial lava tube systems and Lunar/Martian candidate lava tube collapses using a host of remote-sensing datasets. For this reason, the words 'candidate', 'proposed', or 'possible' are frequently used to try to make this distinction, except where features from existing catalogues are being referred to.

Secondly, as is true for any attempt to train any supervised ML/DL model to perform a given task, the performance of the model is a factor of the quality of the training data—in terms of the size of the dataset and the accuracy of the labelling. Therefore, any confusion with labelling the features is likely to lead to incorrect detections or missed targets when applying the model to unseen data. The accuracy of labelling was only a concern in Chapters 5 and 6, rather than Chapter 4, since PITS uses unsupervised ML clustering algorithms. However, decisions often had to be made in the former as to whether certain features should be labelled as pits, skylights or IMFs. For example, due to the downscaling of LROC NAC and HiRISE data in Chapter 5 to 1.5 m/px, catalogued features from the MGC[3] or the LPA that were less than 10 m in diameter were barely resolvable, meaning that labelling them could lead to small clusters of pixels being falsely detected. Therefore, due to the non-uniformity of resolutions found within planetary remote-sensing datasets, there is often a trade-off between spatial detail and the ability to produce more training data.

Whilst one of the main aims of the thesis was to focus on non-impact-related features, impact craters are the most prolific surface features in terms of the existing catalogues and available planetary remote-sensing datasets. Any such researcher interested in planetary surface feature detection aside from impact craters may notice a lack of coverage of high-resolution datasets, which means that there is comparatively little data for these features if they are below 100 m-scales. This was the case for Lunar and Martian pits and skylights in Chapter 5, hence the need for creating synthetic Lunar data by implanting augmented Martian examples into empty LROC NAC tiles.

Another limitation which researchers may experience is the bottleneck in processing planetary remote-sensing data. As detailed in Chapters 4 to 6, the resulting tools can all achieve results in impressive timescales considering the size of the dataset they are applied to and/or

the time it would take for a researcher to manually generate the same output. However, within the entire workflow of going from the identification of a region of interest to inspecting any detections made by ESSA or IMFMapper, the vast majority of time taken was in the processing of remote-sensing data. For example, LROC NAC image products are available through NASA's Planetary Data System (PDS) as Experimental (i.e. raw) or Calibrated (i.e. echo- and radiometrically-corrected) Data Records (EDR/CDR). However, neither EDR nor CDR versions are map-projected, meaning that loading these images into GIS software will not place them in the correct location on the surface. In this thesis, the choice to make the outputs of all tools completely geo-referenced means that LROC NAC images had to be processed in-house from raw image products to a map-projected, GIS-ready GeoTiff file format. Thus, whilst it only took 1.01 mins for IMFMapper to detect the IMFs within Figure 6.15, it took approximately 2 hrs to process, map-project and then mosaic the 56 LROC NAC images of Crookes crater. The prospect of a bottleneck due to map-projection is not a reason to avoid using automated tools, such as IMFMapper, since performing the same task manually will require similar processing.

7.5 Future Work

There are several avenues of future work regarding the tools presented in this thesis, such as making further improvements to the tools to increase their accuracy or adapting them to operate on other planetary bodies. First of all, having achieved good visual results when being applied to the MTP and the skylights detected in Chapter 5, PITS should also be applied to the remaining LPA mare and highland features. Two opportunities for further work are created due to the fact that a single Lunar pit/skylight is typically imaged far more by LROC NAC than a Martian equivalent would be by HiRISE. The first is that several h profiles (derived from images in which the floors are not completely covered in shadow) could be exploited to produce a coarse 3D topographic map. Additionally, the multiple stereo LROC NAC observations could be used to calculate a DEM, upon which PITS' results could be validated. However, due to

changing satellite emission angles, both of these developments would require co-registration between the images used for 3D modelling or the location of the shadow's edge with the DEM.

Due to limited memory and storage allocations, it was not possible to deploy ESSA to the entirety of the Lunar maria to search for uncatalogued pits and skylights. This is why a select few mare RoIs were chosen, typically having well-defined boundaries (such as the mare unit inside Plato crater). However, with sufficient resources and time, LROC NAC imagery of the Lunar maria could be processed and passed through ESSA in increments of latitude and longitude. Given that two candidate cave entrances have been found by ESSA after only being applied to less than 2% of the total Lunar maria, the expectation is that this endeavour would result in further discoveries. Furthermore, ESSA should continue to be applied to archival and future periodic releases of MRO RDRV11 HiRISE data in the search for new Martian pits and skylights. Prior to this, retraining ESSA with the addition of HiRISE image tiles not containing any target features would help in removing some of the false positives made within the January 2025 release. The highest resolution (≈1.4 m/px) Mars Orbiter Camera narrow-angle (MOC-NA) images may also be a useful source of training/inference data in this instance (Albee et al., 2001). Unfortunately, due to the limited global coverage of current high-resolution datasets (HiRISE: ≈4.5%; McEwen et al. 2024, MOC-NA: ≈5.4%; Sidiropoulos and Muller 2015), the same regional or global mapping that is achievable for the Moon is not possible at the same resolution (1.5 m/px) on Mars.

As discussed in Chapter 6, IMFMapper was partly developed in anticipation of IMFs on Mercury being revealed in greater detail upon the beginning of science operations with Bepi-Colombo. Due to MESSENGER's elliptical orbit, resolving IMFs is only currently possible in Mercury's northern hemisphere, with a best resolution of 12 m/px (Hawkins et al., 2007; Xiao et al., 2014). Whereas, the SIMBIO-SYS instrument will provide 6 m/px imagery for roughly 20% of Mercury's surface within its first year of operation (Benkhoff et al., 2021). Given the difference in the pixel scales between LROC NAC and what SIMBIO-SYS is expected to achieve, IMFMapper may need to be retrained upon further downscaled LROC NAC imagery. Of course, downscaling will result in a loss of currently labelled IMFs whose widths fall short

of ≈30 m. Although the loss of training data caused by downscaling could be mitigated by labelling the large amount of LROC NAC data containing yet unannotated IMFs, such as those within Lowell crater (Srivastava and Varatharajan, 2016), as well as those already annotated but under different sensing conditions. Applying a retrained IMFMapper to SIMBIO-SYS high-resolution images would open up Mercury's southern hemisphere to the automated mapping of IMFs, as well as the detection of smaller, marginal IMFs within the northern hemisphere examples already mapped by Xiao et al. (2014). On the other hand, it would be interesting to observe if the current state of IMFMapper could transfer directly to SIMBIO-SYS data, as this alone would justify the use of DeepLabV3, which is tailored to the detection of multiple target feature scales.

Aside from making further improvements, the tools presented in this thesis all contain principles that may be applicable to other planetary surface features. PITS' methodology for calculating depths could be applied to remote-sensing imagery of simple impact craters with sufficiently low Sun angles, which is necessary to cast prominent shadows despite the back-reflection of visible light caused by a crater's bowl-shaped topography. The ability to detect shadows cast within impact craters with PITS would enable the automatic production of h profiles with greater spatial detail than the elevation profiles typically taken as cross-sections from DEMs produced from radar or laser altimetry. In Chapter 5, the training on labelled remote-sensing data taken of two different bodies is also relevant for other surface features. Previous studies have trained models to detect rockfalls in Martian and Lunar data (Bickel et al., 2019, 2020), but there are still many surface features (namely impact craters) that are present across two or more bodies for which there is sufficient remote-sensing imagery at similar scales to use for ML/DL purposes.

7.6 Implications for Other Fields

There are several concepts and methodologies utilised throughout the work presented in this thesis that could be relevant for fields other than planetary remote-sensing, such as Earth

Observation (EO). Firstly, there is scope for planetary remote-sensing data to be used in training EO. With the large amounts of publicly available high-resolution data, EO DL models trained to monitor dune migration on Earth could feasibly use satellite data containing Martian analogues.

The principle of producing synthetic data to aid in the training of DL models could also prove useful for EO, as it was in Chapter 5. For example, cloud cover is a common problem in EO, which means that a large percentage of acquired data has little use due to the surface being obscured. DL models can be trained to super-sample satellite data in order to remove cloud cover (Ebel et al., 2021). Such models may benefit from synthetic 'cloudy' images in which cloud pixels have been artificially implanted, as the objective of a 'cloudless' image was already known.

The weakly supervised learning implemented in Chapter 6 has several EO applications. Linear features in EO data that are of such abundance that it would take too much time to produce pixel-level labels would also benefit from the same buffering of line labels. For example, roads, field boundaries and rivers are all examples of features found within EO data which could be labelled in a similar manner.

DL techniques were used within the winning solution of the 2022 Machine Learning Lunar Data Challenge hosted by the EXPLORE project (Lovell et al., 2024). EXPLORE (see `https://explore-platform.eu`) is a consortium of academic and professional institutions whose main goal is to develop platforms that promote the exploration and exploitation of the vast amounts of space science data using Artificial Intelligence (AI) and visual tools. The work of this thesis was closely linked with the EXPLORE Lunar Data Challenge, which proves that the use of ML/DL in planetary remote-sensing has had an impact on science outreach and collaboration.

The work in this thesis also aligns with several of the mission aims of ACRI-ST, an SME based in France, with which the project was a collaboration. For example, the several different ML and DL methodologies utilised throughout this thesis (clustering: k-means, instance segmentation: Mask R-CNN, semantic segmentation: DeepLabV3) serve to expand and deepen

expertise in automated techniques and AI. ACRI-ST also holds a strong interest in these methodologies and applications as part of its activities in operating and exploiting current and future EO and space missions. Of particular focus is feature classification and detection within data retrieved by the EO satellites within the European Space Agency's (ESA) Copernicus program, such as the Sentinel-2 and 3 missions. The same is also true for two upcoming ESA EO missions: the Copernicus Hyper-spectral Imaging Mission for the Environment (CHIME; Celesti et al. 2022) and Land Surface Temperature Monitoring (LSTM; Bernard et al. 2023) satellites. The former will perform agricultural and biodiversity monitoring, including soil characterisation, via its hyperspectral (visible to near-infrared) imaging suite. LSTM will provide the agricultural community with land-surface temperatures at high spatial and temporal resolutions using its thermal infrared instrument.

With a long tradition in EO, ACRI-ST is developing new activities on big data management in astrophysics and planetary sciences. This thesis bolsters the scientific expertise of ACRI-ST in the domain of planetary sciences, in particular of the Moon and Mars and expands ACRI-ST's network of scientific partners.

References

Albee, A. L., Arvidson, R. E., Palluconi, F., and Thorpe, T. (2001). Overview of the Mars Global Surveyor mission. *Journal of Geophysical Research*, 106(E10):23291–23316. doi: 10.1029/2000JE001306.

Alemanno, G., Orofino, V., and Mancarella, F. (2018). Global Map of Martian Fluvial Systems: Age and Total Eroded Volume Estimations. *Earth and Space Science*, 5(10):560–577. doi: 10.1029/2018EA000362.

Ali-Dib, M. (2022). A machine-generated catalogue of Charon's craters and implications for the Kuiper belt. *Icarus*, 386:115142. doi: 10.1016/j.icarus.2022.115142.

Ali-Dib, M., Besse, S., and Kueppers, M. (2022). Automatic boulders identification on 67P/C-G with machine learning. In *Planetary Science Informatics and Data Analytics conference (PSIDA) 2022*.

Ali-Dib, M., Menou, K., Jackson, A. P., Zhu, C., and Hammond, N. (2020). Automated crater shape retrieval using weakly-supervised deep learning. *Icarus*, 345:113749. doi: 10.1016/j.icarus.2020.113749.

Alshehhi, R. and Gebhardt, C. (2022). Detection of Martian dust storms using mask regional convolutional neural networks. *Progress in Earth and Planetary Science*, 9(1):4. doi: 10.1186/s40645-021-00464-1.

Alvarez, L. W., Alvarez, W., Asaro, F., and Michel, H. V. (1980). Extraterrestrial Cause for the Cretaceous-Tertiary Extinction. *Science*, 208(4448):1095–1108. doi: 10.1126/science.208.4448.1095.

Atri, D. et al. (2022). Crewed Missions to Mars: Modeling the Impact of Astrophysical Charged PARTICLEs on Astronauts and Their Health. *arXiv e-prints*. doi: 10.48550/arXiv.2208.00892.

Azari, A. et al. (2021). Integrating Machine Learning for Planetary Science: Perspectives for the Next Decade. *Bulletin of the AAS*, 53(4). https://baas.aas.org/pub/2021n4i128.

Baker, V. R., Strom, R. G., Gulick, V. C., Kargel, J. S., and Komatsu, G. (1991). Ancient oceans, ice sheets and the hydrological cycle on Mars. *Nature*, 352(6336):589–594. doi: 10.1038/352589a0.

Balme, M. R., Whelley, P. L., and Greeley, R. (2003). Mars: Dust devil track survey in Argyre Planitia and Hellas Basin. *Journal of Geophysical Research (Planets)*, 108(E8):5086. doi: 10.1029/2003JE002096.

Barker, M. K., Mazarico, E., Neumann, G. A., Zuber, M. T., Haruyama, J., and Smith, D. E. (2016). A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera. *Icarus*, 273:346–355. doi: 10.1016/j.icarus.2015.07.039.

Barrett, A. M. et al. (2022). NOAH-H, a deep-learning, terrain classification system for Mars: Results for the ExoMars Rover candidate landing sites. *Icarus*, 371:114701. doi: 10.1016/j.icarus.2021.114701.

Barrett, A. M., Favaro, E. A., and Balme, M. R. (2023). Detecting and Measuring Transverse Aeolian Ridges (TARs) on Mars Using Deep Learning. In *54th Lunar and Planetary Science Conference*, volume 2806 of *LPI Contributions*, page 1534.

Becker, K. J. et al. (2016). First Global Digital Elevation Model of Mercury. In *47th Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 2959.

Benedix, G. Â. K. et al. (2020). Deriving Surface Ages on Mars Using Automated Crater Counting. *Earth and Space Science*, 7(3):e01005. doi: 10.1029/2019EA001005.

Benkhoff, J. et al. (2021). BepiColombo - Mission Overview and Science Goals. *Space Science Reviews*, 217(8):90. doi: 10.1007/s11214-021-00861-4.

Bernard, F. et al. (2023). The Copernicus land surface temperature monitoring (LSTM) mission: design, technology and status. In Babu, S. R., Hélière, A., and Kimura, T., editors, *Sensors, Systems, and Next-Generation Satellites XXVII*, volume 12729 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 127290A. doi: 10.1117/12.2679705.

Bickel, V. T., Conway, S. J., Tesson, P.-A., Manconi, A., Loew, S., and 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. doi: 10.1109/JSTARS.2020.2991588.

Bickel, V. T., Lanaras, C., Manconi, A., Loew, S., and Mall, U. (2019). Automated Detection of Lunar Rockfalls Using a Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, 57(6):3501–3511. doi: 10.1109/TGRS.2018.2885280.

Blair, D. M. et al. (2017). The structural stability of lunar lava tubes. *Icarus*, 282:47–55. doi: 10.1016/j.icarus.2016.10.008.

Boston, P. J. et al. (2004). Extraterrestrial Subsurface Technology Test Bed: Human Use and Scientific Value of Martian Caves. In El-Genk, M. S., editor, *Space Technology and Applications*, volume 699 of *American Institute of Physics Conference Series*, pages 1007–1018. AIP. doi: 10.1063/1.1649667.

Bottke, W. F., Love, S. G., Tytell, D., and Glotch, T. (2000). Interpreting the Elliptical Crater Populations on Mars, Venus, and the Moon. *Icarus*, 145(1):108–121. doi: 10.1006/icar.1999.6323.

Brown, H. M. et al. (2022). Resource potential of lunar permanently shadowed regions. *Icarus*, 377:114874. doi: 10.1016/j.icarus.2021.114874.

Burr, D. M., Grier, J. A., McEwen, A. S., and Keszthelyi, L. P. (2002). Repeated Aqueous Flooding from the Cerberus Fossae: Evidence for Very Recently Extant, Deep Groundwater on Mars. *Icarus*, 159(1):53–73. doi: 10.1006/icar.2002.6921.

Calvari, S. and Pinkerton, H. (1999). Lava tube morphology on Etna and evidence for lava flow emplacement mechanisms. *Journal of Volcanology and Geothermal Research*, 90(3):263–280. doi: 10.1016/S0377-0273(99)00024-4.

Cameron, A. G. W. and Ward, W. R. (1976). The Origin of the Moon. In *Lunar and Planetary Science Conference*, volume 7 of *Lunar and Planetary Science Conference*, page 120.

Cardinale, M., Hargitai, H., and Fitzsimmons, K. E. (2021). Dune. In *Encyclopedia of Planetary Landforms*, pages 1–18. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_163-1.

Carr, M. H. and Head, J. W. (2003). Oceans on Mars: An assessment of the observational evidence and possible fate. *Journal of Geophysical Research (Planets)*, 108(E5):5042. doi: 10.1029/2002JE001963.

Carrer, L., Pozzobon, R., Sauro, F., Castelletti, D., Patterson, G. W., and Bruzzone, L. (2024). Radar evidence of an accessible cave conduit on the Moon below the Mare Tranquillitatis pit. *Nature Astronomy*, 8:1119–1126. doi: 10.1038/s41550-024-02302-y.

Celesti, M., Rast, M., Adams, J., Boccia, V., Gascon, F., Isola, C., and Nieke, J. (2022). The Copernicus Hyperspectral Imaging Mission for the Environment (Chime): Status and Planning. In *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, pages 5011–5014. 10.1109/IGARSS46834.2022.9883592.

Chaini, C. and Jha, V. K. (2024). A review on deep learning-based automated lunar crater detection. *Earth Science Informatics*, 17(5):3863–3898. doi: 10.1007/s12145-024-01396-2.

Chappaz, L. et al. (2017). Evidence of large empty lava tubes on the Moon using GRAIL gravity. *Geophysical Research Letters*, 44(1):105–112. doi: 10.1002/2016GL071588.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., and Yuille, A. L. (2016). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *arXiv e-prints*. doi: 10.48550/arXiv.1606.00915.

Chen, L.-C., Papandreou, G., Schroff, F., and Adam, H. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv e-prints*. doi: 10.48550/arXiv.1706.05587.

Christensen, P. R. et al. (2004). The Thermal Emission Imaging System (THEMIS) for the Mars 2001 Odyssey Mission. *Space Science Reviews*, 110(1):85–130. doi: 10.1023/B:SPAC.0000021008.16305.94.

Christensen, P. R. et al. (2009). JMARS - A Planetary GIS. In *AGU Fall Meeting Abstracts*, volume 2009, pages IN22A–06. https://jmars.asu.edu/.

Christoff, N., Jorda, L., Viseur, S., Bouley, S., Manolova, A., and Mari, J.-L. (2020). Automated Extraction of Crater Rims on 3D Meshes Combining Artificial Neural Network and Discrete Curvature Labeling. *Earth Moon and Planets*, 124(3-4):51–72. doi: 10.1007/s11038-020-09535-7.

Collins, G. S., Melosh, H. J., Morgan, J. V., and Warner, M. R. (2002). Hydrocode Simulations of Chicxulub Crater Collapse and Peak-Ring Formation. *Icarus*, 157(1):24–33. doi: 10.1006/icar.2002.6822.

Coombs, C. R. and Hawke, B. R. (1992). A Search for Intact Lava Tubes on the Moon: Possible Lunar Base Habitats. In Mendell, W. W. et al., editors, *Lunar Bases and Space Activities of the 21st Century*, page 219.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Crown, D. A., Scheidt, S. P., and Berman, D. C. (2022). Distribution and Morphology of Lava Tube Systems on the Western Flank of Alba Mons, Mars. *Journal of Geophysical Research (Planets)*, 127(6):e07263. doi: 10.1029/2022JE007263.

Cushing, G. E. (2015). Mars Global Cave Candidate Catalog PDS4 Archive Bundle. PDS Cartography and Imaging Sciences Node (IMG). doi: 10.17189/1519222. https://astrogeology.usgs.gov/search/map/mars_global_cave_candidate_catalog_v1_cushing.

Cushing, G. E., Okubo, C. H., and Titus, T. N. (2015). Atypical pit craters on Mars: New insights from THEMIS, CTX, and HiRISE observations. *Journal of Geophysical Research (Planets)*, 120(6):1023–1043. doi: 10.1002/2014JE004735.

Cushing, G. E., Titus, T. N., Wynne, J. J., and Christensen, P. R. (2007). THEMIS observes possible cave skylights on Mars. *Geophysical Research Letters*, 34:17201. doi: 10.1029/2007GL030709.

Dai, J., Li, Y., He, K., and Sun, J. (2016). R-FCN: Object Detection via Region-based Fully Convolutional Networks. *arXiv e-prints*. doi: 10.48550/arXiv.1605.06409.

Davey, S. C., Ernst, R. E., Samson, C., Grosfils, E. B., and Willamson, M.-C. (2013). Hierarchical clustering of pit crater chains on Venus. *Canadian Journal of Earth Sciences*, 50(1):109–126. doi: 10.1139/cjes-2012-0054.

Dawson, H. L., Dubrule, O., and John, C. M. (2023). Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. *Computers and Geosciences*, 171:105284. doi: 10.1016/j.cageo.2022.105284.

De Hon, R. (2021). Ghost Crater. In *Encyclopedia of Planetary Landforms*, pages 1–6. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_172-2.

De Toffoli, B., Plesa, A. C., Hauber, E., and Breuer, D. (2021). Delta Deposits on Mars: A Global Perspective. *Geophysical Research Letters*, 48(17):e94271. doi: 10.1029/2021GL094271.

DeLatte, D. M., Crites, S. T., Guttenberg, N., Tasker, E. J., and Yairi, T. (2019a). Segmentation Convolutional Neural Networks for Automatic Crater Detection on Mars. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8):2944–2957. doi: 10.1109/JSTARS.2019.2918302.

DeLatte, D. M., Crites, S. T., Guttenberg, N., and Yairi, T. (2019b). Automated crater detection algorithms from a machine learning perspective in the convolutional neural network era. *Advances in Space Research*, 64(8):1615–1628. doi: 10.1016/j.asr.2019.07.017.

Denevi, B. W. et al. (2012). Physical constraints on impact melt properties from Lunar Reconnaissance Orbiter Camera images. *Icarus*, 219(2):665–675. doi: 10.1016/j.icarus.2012.03.020.

Deng, J., Dong, W., Socher, R., Li, L.-J., Kai, L., and Li, F.-F. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255. doi: 10.1109/CVPR.2009.5206848.

Dhingra, D., Head, J. W., and Pieters, C. M. (2017). Geological mapping of impact melt deposits at lunar complex craters Jackson and Tycho: Morphologic and topographic diversity and relation to the cratering process. *Icarus*, 283:268–281. doi: 10.1016/j.icarus.2016.05.004.

Di, K., Li, W., Yue, Z., Sun, Y., and Liu, Y. (2014). A machine learning approach to crater detection from topographic data. *Advances in Space Research*, 54(11):2419–2429. doi: 10.1016/j.asr.2014.08.018.

Ebel, P., Meraner, A., Schmitt, M., and Zhu, X. X. (2021). Multisensor Data Fusion for Cloud Removal in Global and All-Season Sentinel-2 Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7):5866–5878. doi: 10.1109/TGRS.2020.3024744.

Edwards, C. S., Christensen, P. R., and Hill, J. (2011a). Mosaicking of global planetary image datasets: 2. Modeling of wind streak thicknesses observed in Thermal Emission Imaging System (THEMIS) daytime and nighttime infrared data. *Journal of Geophysical Research (Planets)*, 116(E10):E10005. doi: 10.1029/2011JE003857.

Edwards, C. S., Nowicki, K. J., Christensen, P. R., Hill, J., Gorelick, N., and Murray, K. (2011b). Mosaicking of global planetary image datasets: 1. Techniques and data processing for Thermal Emission Imaging System (THEMIS) multi-spectral data. *Journal of Geophysical Research (Planets)*, 116(E10):E10008. doi: 10.1029/2010JE003755.

Emami, E., Ahmad, T., Bebis, G., Nefian, A., and Fong, T. (2018). On Crater Classification Using Deep Convolutional Neural Networks. In *49th Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 2566.

Emami, E., Ahmad, T., Bebis, G., Nefian, A., and Fong, T. (2019). Crater Detection Using Unsupervised Algorithms and Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8):5373–5383. doi: 10.1109/TGRS.2019.2899122.

Erkeling, G. and Carling, P. A. (2021). Streamlined Island. In *Encyclopedia of Planetary Landforms*, pages 1–5. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_547-1.

Fassett, C. I., Kadish, S. J., Head, J. W., Solomon, S. C., and Strom, R. G. (2011). The global population of large craters on Mercury and comparison with the Moon. *Geophysical Research Letters*, 38(10):L10202. doi: 10.1029/2011GL047294.

Fergason, R. L., Hare, T. M., and Laura, J. (2018). HRSC and MOLA Blended Digital Elevation Model at 200m v2. Astrogeology PDS Annex, U.S. Geological Survey. https://astrogeology.usgs.gov/search/map/mars_mgs_mola_mex_hrsc_blended_dem_global_200m.

Ferrill, D. A. (2015). Pit. In Hargitai, H. and Kereszturi, Á., editors, *Encyclopedia of Planetary Landforms*, pages 1566–1570. Springer New York, New York, NY. doi: 10.1007/978-1-4614-3134-3_267.

Ferrill, D. A. and Morris, A. P. (2003). Dilational normal faults. *Journal of Structural Geology*, 25(2):183–196. doi: 10.1016/S0191-8141(02)00029-9.

Fortezzo, C. M., Spudis, P. D., and Harrel, S. L. (2020). Release of the Digital Unified Global Geologic Map of the Moon at 1:5,000,000-Scale. In *51st Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 2760.

Francis, A. et al. (2020). A Multi-Annotator Survey of Sub-km Craters on Mars. *Data*, 5(3). doi: 10.3390/data5030070.

Freund, Y. and Schapire, R. E. (1996). Experiments with a new boosting algorithm. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, ICML'96, page 148–156, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Gadányi, P. and van der Bogert, C. H. (2015). Lava Tube. In Hargitai, H. and Kereszturi, Á., editors, *Encyclopedia of Planetary Landforms*, pages 1181–1186. Springer New York, New York, NY. doi: 10.1007/978-1-4614-3134-3_474.

Gagnon, S., Lemelin, M., Neish, C., and Diotte, F. (2023). Localisation and Study of Impact Melts in the South Polar Region of the Moon. In *54th Lunar and Planetary Science Conference*, volume 2806 of *LPI Contributions*, page 2655.

Galgana, G. A. (2021). Caldera. In *Encyclopedia of Planetary Landforms*, pages 1–15. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_29-2.

GDAL/OGR contributors (2025). GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation. doi: 10.5281/zenodo.5884351. https://gdal.org.

Gichu, R. and Ogohara, K. (2019). Segmentation of dust storm areas on Mars images using principal component analysis and neural network. *Progress in Earth and Planetary Science*, 6(1):19. doi: 10.1186/s40645-019-0266-1.

Gillis-Davis, J. J. et al. (2009). Pit-floor craters on Mercury: Evidence of near-surface igneous activity. *Earth and Planetary Science Letters*, 285(3-4):243–250. doi: 10.1016/j.epsl.2009.05.023.

Girshick, R. (2015). Fast R-CNN. *arXiv e-prints*. doi: 10.48550/arXiv.1504.08083.

Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2013). Rich feature hierarchies for accurate object detection and semantic segmentation. *arXiv e-prints*. doi: 10.48550/arXiv.1311.2524.

Goodfellow, I. J. et al. (2014). Generative Adversarial Networks. *arXiv e-prints*. doi: 10.48550/arXiv.1406.2661.

Greeley, R. (1971). Lava Tubes and Channels in the Lunar Marius Hills. *Moon*, 3(3):289–314. doi: 10.1007/BF00561842.

Greeley, R., Bender, K., Thomas, P. E., Schubert, G., Limonadi, D., and Weitz, C. M. (1995). Wind-related features and processes on Venus: Summary of Magellan results. *Icarus*, 115(2):399–420. doi: 10.1006/icar.1995.1107.

Greeley, R. and Gault, D. E. (1979). Endogenic craters on basaltic lava flows: size frequency distributions. *Lunar and Planetary Science Conference Proceedings*, 3:2919–2933.

Gregg, T. K. P. and Greeley, R. (1992). Formational Constraints on Venusian "Canali". In *Lunar and Planetary Science Conference*, volume 23 of *Lunar and Planetary Science Conference*, page 449.

Haase, I., Wählisch, M., Gläser, P., Oberst, J., and Robinson, M. S. (2019). Coordinates and Maps of the Apollo 17 Landing Site. *Earth and Space Science*, 6(1):59–95. doi: 10.1029/2018EA000408.

Hagen, T. H. (2015). Pit Crater Chain, Pit Chain. In Hargitai, H. and Kereszturi, Á., editors, *Encyclopedia of Planetary Landforms*, pages 1575–1579. Springer New York, New York, NY. doi: 10.1007/978-1-4614-3134-3_261.

Hahn, R. M. and Byrne, P. K. (2023). A Morphological and Spatial Analysis of Volcanoes on Venus. *Journal of Geophysical Research (Planets)*, 128(4):e2023JE007753. doi: 10.1029/2023JE007753.

Halliday, A. N. (2000). Terrestrial accretion rates and the origin of the Moon. *Earth and Planetary Science Letters*, 176(1):17–30. doi: 10.1016/S0012-821X(99)00317-9.

Handmer, C. (2021). Digital Elevation Model enhancement using Deep Learning. *arXiv e-prints*. doi: 10.48550/arXiv.2101.04812.

Hansen, C. J. et al. (2006). Enceladus' Water Vapor Plume. *Science*, 311(5766):1422–1425. doi: 10.1126/science.1121254.

Hargitai, H. (2021a). Lava Channel. In *Encyclopedia of Planetary Landforms*, pages 1–10. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_221-1.

Hargitai, H. (2021b). Pedestal Crater (Mars). In *Encyclopedia of Planetary Landforms*, pages 1–8. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_257-1.

Hargitai, H. and Kereszturi, A. (2015). *Encyclopedia of Planetary Landforms*. Springer New York, NY. doi: 10.1007/978-1-4614-3134-3.

Hargitai, H., Reimold, W. U., and Bray, V. J. (2021). Impact Structure. In *Encyclopedia of Planetary Landforms*, pages 1–39. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_194-1.

Hartmann, W. K. and Kuiper, G. P. (1962). Concentric Structures Surrounding Lunar Basins. *Communications of the Lunar and Planetary Laboratory*, 1:51–66.

Hartmann, W. K. and Neukum, G. (2001). Cratering Chronology and the Evolution of Mars. *Space Science Reviews*, 96:165–194. doi: 10.1023/A:1011945222010.

Haruyama, J. et al. (2009). Possible lunar lava tube skylight observed by SELENE cameras. *Geophysical Research Letters*, 36(21):L21206. doi: 10.1029/2009GL040635.

Haruyama, J. et al. (2010). New Discoveries of Lunar Holes in Mare Tranquillitatis and Mare Ingenii. In *41st Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 1285.

Hashimoto, S. and Mori, K. (2019). Lunar Crater Detection based on Grid Partition using Deep Learning. In *2019 IEEE 13th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pages 75–80. doi: 10.1109/SACI46893.2019.9111474.

Hawkins, S. E. et al. (2007). The Mercury Dual Imaging System on the MESSENGER Spacecraft. *Space Science Reviews*, 131(1-4):247–338. doi: 10.1007/s11214-007-9266-3.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask R-CNN. *arXiv e-prints*. doi: 10.48550/arXiv.1703.06870.

He, K., Zhang, X., Ren, S., and Sun, J. (2014). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *arXiv e-prints*. doi: 10.48550/arXiv.1406.4729.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, page 1. doi: 10.1109/CVPR.2016.90.

Head, J. W., I. (1976). Lunar volcanism in space and time. *Reviews of Geophysics and Space Physics*, 14:265–300. doi: 10.1029/RG014i002p00265.

Head, J. W. et al. (2010). Global Distribution of Large Lunar Craters: Implications for Resurfacing and Impactor Populations. *Science*, 329(5998):1504. doi: 10.1126/science.1195050.

Heldmann, J. L., Diniega, S., Kereszturi, Á., and Conway, S. J. (2014). Gully. In *Encyclopedia of Planetary Landforms*, pages 1–11. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_184-1.

Herrick, R. R., Bateman, E. M., Crumpacker, W. G., and Bates, D. (2018). Observations From a Global Database of Impact Craters on Mercury With Diameters Greater than 5 km. *Journal of Geophysical Research (Planets)*, 123(8):2089–2109. doi: 10.1029/2017JE005516.

Herrick, R. R., Sharpton, V. L., Malin, M. C., Lyons, S. N., and Feely, K. (1997). Morphology and Morphometry of Impact Craters. In Bougher, S. W., Hunten, D. M., and Phillips, R. J., editors, *Venus II: Geology, Geophysics, Atmosphere, and Solar Wind Environment*, page 1015.

Hill, J., Edwards, C. S., and Christensen, P. R. (2014). Mapping the Martian Surface with THEMIS Global Infrared Mosaics. In LPI Editorial Board, editor, *Eighth International Conference on Mars*, volume 1791 of *LPI Contributions*, page 1141.

Hinton, G. E. and Zemel, R. S. (1993). Autoencoders, minimum description length and Helmholtz free energy. In *Proceedings of the 7th International Conference on Neural Information Processing Systems*, NIPS'93, page 3–10, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Horstman, K. C. and Melosh, H. J. (1989). Drainage pits in cohesionless materials: implications for the surface of Phobos. *Journal of Geophysical Research*, 94:12433–12441. doi: 10.1029/JB094iB09p12433.

Horvath, T., Hayne, P. O., and Paige, D. A. (2022). Thermal and Illumination Environments of Lunar Pits and Caves: Models and Observations From the Diviner Lunar Radiometer Experiment. *Geophysical Research Letters*, 49(14):e99710. doi: 10.1029/2022GL099710.

Hsu, C.-Y., Li, W., and Wang, S. (2021). Knowledge-Driven GeoAI: Integrating Spatial Knowledge into Multi-Scale Deep Learning for Mars Crater Detection. *Remote Sensing*, 13(11):2116. doi: 10.3390/rs13112116.

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2016). Densely Connected Convolutional Networks. *arXiv e-prints*. doi: 10.48550/arXiv.1608.06993.

Hurwitz, D. M., Head, J. W., and Hiesinger, H. (2013). Lunar sinuous rilles: Distribution, characteristics, and implications for their origin. *Planetary and Space Science*, 79:1–38. doi: 10.1016/j.pss.2012.10.019.

IAU (2025). Gazetteer of Planetary Nomenclature. International Astronomical Union (IAU) Working Group for Planetary System Nomenclature. http://planetarynames.wr.usgs.gov/.

Jagert, F. (2021). Volcanic Plain. In *Encyclopedia of Planetary Landforms*, pages 1–6. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_399-1.

Jaumann, R. et al. (2007). The high-resolution stereo camera (HRSC) experiment on Mars Express: Instrument aspects and experiment conduct from interplanetary cruise through the nominal mission. *Planetary and Space Science*, 55(7-8):928–952. doi: 10.1016/j.pss.2006.12.003.

Jozwiak, L. M., Head, J. W., Zuber, M. T., Smith, D. E., and Neumann, G. A. (2012). Lunar floor-fractured craters: Classification, distribution, origin and implications for magmatism and shallow crustal structure. *Journal of Geophysical Research (Planets)*, 117(E11):E11005. doi: 10.1029/2012JE004134.

Kaku, T. et al. (2017). Detection of Intact Lava Tubes at Marius Hills on the Moon by SELENE (Kaguya) Lunar Radar Sounder. *Geophysical Research Letters*, 44(20):10,155–10,161. doi: 10.1002/2017GL074998.

Kato, M., Sasaki, S., Tanaka, K., Iijima, Y., and Takizawa, Y. (2008). The Japanese lunar mission SELENE: Science goals and present status. *Advances in Space Research*, 42(2):294–300. doi: 10.1016/j.asr.2007.03.049.

Kerner, H. R., Wagstaff, K. L., Bue, B. D., Gray, P. C., Bell, James F., I., and Ben Amor, H. (2019). Toward Generalized Change Detection on Planetary Surfaces With Convolutional Autoencoders and Transfer Learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(10):3900–3918. doi: 10.1109/JSTARS.2019.2936771.

Kim, J.-R., Lin, S.-Y., and Oh, J.-W. (2022). The Survey of Lava Tube Distribution in Jeju Island by Multi-Source Data Fusion. *Remote Sensing*, 14(3):443. doi: 10.3390/rs14030443.

Kingma, D. P. and Ba, J. (2017). Adam: A Method for Stochastic Optimization. *arXiv e-prints*. doi: 10.48550/arXiv.1412.6980.

Kirk, R. L. et al. (2008). Ultrahigh resolution topographic mapping of Mars with MRO HiRISE stereo images: Meter-scale slopes of candidate Phoenix landing sites. *Journal of Geophysical Research (Planets)*, 113(E12):E00A24. doi: 10.1029/2007JE003000.

Klem, S. M., Henriksen, M. R., Stopar, J., Boyd, A., Robinson, M. S., and LROC Science Team (2014). Controlled LROC Narrow Angle Camera High Resolution Mosaics. In *45th Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 2885.

Korteniemi, J. (2021). Fractured-Floor Crater. In *Encyclopedia of Planetary Landforms*, pages 1–9. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_160-2.

Korteniemi, J., Eldridge, D. L., Lough, T., Werblin, L., Singer, K., and Kring, D. (2010). Assessment of Lunar Volcanic Morphological Diversity: Distribution of Floor-fractured Craters. In *41st Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 1335.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.

La Grassa, R., Cremonese, G., Gallo, I., Re, C., and Martellato, E. (2023). YOLOLens: A Deep Learning Model Based on Super-Resolution to Enhance the Crater Detection of the Planetary Surfaces. *Remote Sensing*, 15(5):1171. doi: 10.3390/rs15051171.

Lagain, A. et al. (2021a). The Tharsis mantle source of depleted shergottites revealed by 90 million impact craters. *Nature Communications*, 12:6352. doi: 10.1038/s41467-021-26648-3.

Lagain, A., Servis, K., Benedix, G. K., Norman, C., Anderson, S., and Bland, P. A. (2021b). Model Age Derivation of Large Martian Impact Craters, Using Automatic Crater Counting Methods. *Earth and Space Science*, 8(2):e01598. doi: 10.1029/2020EA001598.

Laura, J. et al. (2023). Integrated Software for Imagers and Spectrometers. Zenodo. doi: 10.5281/zenodo.7644616. https://doi.org/10.5281/zenodo.7644616.

Le Corre, D. (2025a). Entrances to Sub-Surface Areas (ESSA) GitHub Repository. Zenodo. doi: 10.5281/zenodo.17593909. https://github.com/dlecorre387/Entrances-to-Sub-Surface-Areas.

Le Corre, D. (2025b). ESSA PyTorch Model Checkpoint and Shapefiles Containing Pit and Skylight Detections on the Moon. Zenodo. doi: 10.5281/zenodo.15095750. https://doi.org/10.5281/zenodo.15095750.

Le Corre, D. (2025c). IMFMapper GitHub Repository. Zenodo. doi: 10.5281/zenodo.17460144. https://github.com/dlecorre387/IMFMapper.

Le Corre, D. (2025d). IMFMapper PyTorch Model Checkpoint and Shapefiles of IMFs in Crookes and Schomberger A Craters. Zenodo. doi: 10.5281/zenodo.15101175. https://doi.org/10.5281/zenodo.15101175.

Le Corre, D. (2025e). Pit Topography from Shadows (PITS) GitHub Repository Release. Zenodo. doi: 10.5281/zenodo.17593928. https://github.com/dlecorre387/Pit-Topography-from-Shadows.

Le Corre, D. (2025f). Planetary Image Processing GitHub Repository. Zenodo. doi: 10.5281/zenodo.17215911. https://github.com/dlecorre387/Planetary-Image-Processing.

Le Corre, D., Mary, D., Mason, N., Bernard-Salas, J., and Cox, N. (2023a). Automatically calculating the apparent depths of pits using the Pit Topography from Shadows (PITS) tool. *RAS Techniques and Instruments*, 2(1):492–509. doi: 10.1093/rasti/rzad037.

Le Corre, D., Mary, D., Mason, N., Bernard-Salas, J., and Cox, N. (2023b). Docker Image for the Pit Topography from Shadows (PITS) Tool. Zenodo. doi: 10.5281/zenodo.8192849. https://doi.org/10.5281/zenodo.8192849.

Le Corre, D., Mason, N., Bernard-Salas, J., Cox, N., and Mary, D. (2023c). Automatically Calculating Depths of Martian and Lunar Pits with Satellite Imagery. In *EGU General Assembly Conference Abstracts*, EGU General Assembly Conference Abstracts, pages EGU–14927. doi: 10.5194/egusphere-egu23-14927.

Le Corre, D., Mason, N., Bernard-Salas, J., Mary, D., and Cox, N. (2022a). Martian Pit Shadow Extractor (MAPS): Determining the Apparent Depths of Martian Pits from the Morphology of their Shadows. In *10th International Conference on Geomorphology*, pages ICG2022–186. doi: 10.5194/icg2022-186.

Le Corre, D., Mason, N., Bernard-Salas, J., Mary, D., and Cox, N. (2022b). Planetary Surface Feature Detection Using Machine Learning. In *European Planetary Science Congress*, pages EPSC2022–997. doi: 10.5194/epsc2022-997.

Le Corre, D., Mason, N., Bernard-Salas, J., Mary, D., and Cox, N. (2025a). New candidate cave entrances on the Moon found using deep learning. *Icarus*, 441:116675. doi: 10.1016/j.icarus.2025.116675.

Le Corre, D., Mason, N., Bernard-Salas, J., Mary, D., and Cox, N. L. J. (2025b). Automated Impact Melt Fracture Mapping on the Moon With Weakly Supervised Deep Learning. *Journal of Geophysical Research (Planets)*, 130(12):e2025JE009145. doi: 10.1029/2025JE009145.

Lee, C. (2019). Automated crater detection on Mars using deep learning. *Planetary and Space Science*, 170:16–28. doi: 10.1016/j.pss.2019.03.008.

Lee, C. and Hogan, J. (2021). Automated crater detection with human level performance. *Computers and Geosciences*, 147:104645. doi: 10.1016/j.cageo.2020.104645.

Lee, P. (2022). Impact Melt Lava Inflation Fractures in Schomberger-A Crater, Moon: New Permanently Shadowed Regions and Possible Caves Near the Lunar South Pole. In *53rd Lunar and Planetary Science Conference*, volume 2678 of *LPI Contributions*, page 2685.

Lemaire, T. et al. (2024). Lava flow field development and lava tube formation during the 1858–1861 eruption of Vesuvius (Italy), unravelled by historical documentation, lidar data and 3D mapping. *Journal of Volcanology and Geothermal Research*, 455:108197. doi: 10.1016/j.jvolgeores.2024.108197.

Léveillé, R. J. and Datta, S. (2010). Lava tubes and basaltic caves as astrobiological targets on Earth and Mars: A review. *Planetary and Space Science*, 58(4):592–598. doi: 10.1016/j.pss.2009.06.004.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2016). Feature Pyramid Networks for Object Detection. *arXiv e-prints*. doi: 10.48550/arXiv.1612.03144.

Lin, T.-Y. et al. (2014a). Microsoft COCO: Common Objects in Context. *arXiv e-prints*. doi: 10.48550/arXiv.1405.0312.

Lin, Y., Bunte, M., Saripalli, S., Bell, J., and Greeley, R. (2014b). Autonomous volcanic plume detection on planetary bodies. *Acta Astronautica*, 97:151–163. doi: 10.1016/j.actaastro.2013.11.029.

Lindholm, A., Wahlström, N., Lindsten, F., and Schön, T. B. (2022). *Machine Learning - A First Course for Engineers and Scientists*. Cambridge University Press.

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137. doi: 10.1109/TIT.1982.1056489.

Lorenz, R. D. et al. (2006). The Sand Seas of Titan: Cassini RADAR Observations of Longitudinal Dunes. *Science*, 312(5774):724–727. doi: 10.1126/science.1123257.

Loshchilov, I. and Hutter, F. (2017). Decoupled Weight Decay Regularization. *arXiv e-prints*. doi: 10.48550/arXiv.1711.05101.

Lovell, L., Adriani, I., Nodjoumi, G., Suarez-Valencia, J., Le Corre, D., Heward, A., Rossi, A., and Cox, N. (2024). Design of robotic traverses on the Archytas Dome on the Moon. *Open Research Europe*, 4(116). doi: 10.12688/openreseurope.17424.1.

Maeda, K., Ogawa, T., and Haseyama, M. (2015). Automatic Martian Dust Storm Detection from Multiple Wavelength Data Based on Decision Level Fusion. *IPSJ Transactions on Computer Vision and Applications*, 7:79–83. doi: 10.2197/ipsjtcva.7.79.

Malin, M. C. et al. (1998). Early Views of the Martian Surface from the Mars Orbiter Camera of Mars Global Surveyor. *Science*, 279:1681. doi: 10.1126/science.279.5357.1681.

Malin, M. C. et al. (2007). Context Camera Investigation on board the Mars Reconnaissance Orbiter. *Journal of Geophysical Research (Planets)*, 112(E5):E05S04. doi: 10.1029/2006JE002808.

Mao, A., Mohri, M., and Zhong, Y. (2023). Cross-Entropy Loss Functions: Theoretical Analysis and Applications. *arXiv e-prints*, page arXiv:2304.07288. doi: 10.48550/arXiv.2304.07288.

Martin, R. P. and Benaroya, H. (2023). Pressurized lunar lava tubes for habitation. *Acta Astronautica*, 204:157–174. doi: 10.1016/j.actaastro.2022.12.013.

McCauley, J. F. (1973). Mariner 9 Evidence for Wind Erosion in the Equatorial and Mid-Latitude Regions of Mars. *Journal of Geophysical Research*, 78(20):4123–4137. doi: 10.1029/JB078i020p04123.

McEwen, A. (2007). Mars Reconnaissance Orbiter High Resolution Imaging Science Experiment, Reduced Data Record, MRO-M-HIRISE-3-RDR-V1.1. NASA Planetary Data System. doi: 10.17189/1520303. https://hirise-pds.lpl.arizona.edu/PDS/RDR/.

McEwen, A. S. et al. (2007). Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE). *Journal of Geophysical Research (Planets)*, 112(E5):E05S02. doi: 10.1029/2005JE002605.

McEwen, A. S. et al. (2024). The high-resolution imaging science experiment (HiRISE) in the MRO extended science phases (2009–2023). *Icarus*, 419:115795. doi: 10.1016/j.icarus.2023.115795.

McSween, Jr, H. Y. et al. (2019). *Planetary Geoscience*. Cambridge University Press.

Melosh, H. J. (1989). *Impact cratering : a geologic process*. Oxford University Press, New York.

Michalski, J. R., Goudge, T. A., Crowe, S. A., Cuadros, J., Mustard, J. F., and Johnson, S. S. (2022). Geological diversity and microbiological potential of lakes on Mars. *Nature Astronomy*, 6:1133–1141. doi: 10.1038/s41550-022-01743-7.

Morabito, L. A., Synnott, S. P., Kupferman, P. N., and Collins, S. A. (1979). Discovery of Currently Active Extraterrestrial Volcanism. *Science*, 204(4396):972. doi: 10.1126/science.204.4396.972.

Mouginis-Mark, P. (2001). The geomorphology of planetary calderas. *Geomorphology*, 37(3-4):201–223. doi: 10.1016/S0169-555X(00)00083-0.

Nagle-McNaughton, T. P., Scuderi, L. A., and Erickson, N. (2022). Squeezing Data from a Rock: Machine Learning for Martian Science. *Geosciences*, 12(6):248. doi: 10.3390/geosciences12060248.

Negi, K., Kiran, B. S., and Singh, S. K. (2019). Mission Design and Analysis for Mars Orbiter Mission. *Journal of the Astronautical Sciences*, 67(3):932–949. doi: 10.1007/s40295-019-00199-8.

Nesnas, I. A. D. et al. (2023). Moon Diver: Exploring a pit's exposed strata to understand lunar volcanism. *Acta Astronautica*, 211:163–176. doi: 10.1016/j.actaastro.2023.05.042.

Neukum, G., Ivanov, B. A., and Hartmann, W. K. (2001). Cratering Records in the Inner Solar System in Relation to the Lunar Reference System. *Space Science Reviews*, 96:55–86. doi: 10.1023/A:1011989004263.

Nodjoumi, G., Pozzobon, R., and Rossi, A. P. (2021). Deep Learning Object Detection for Mapping Cave Candidates on Mars: Building Up the Mars Global Cave Candidate Catalog (MGC^3). In *52nd Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 1316.

Nodjoumi, G., Pozzobon, R., Sauro, F., and Rossi, A. P. (2023). DeepLandforms: A Deep Learning Computer Vision Toolset Applied to a Prime Use Case for Mapping Planetary Skylights. *Earth and Space Science*, 10(1):e2022EA002278. doi: 10.1029/2022EA002278.

Norman, C. J., Paxman, J., Benedix, G. K., Tan, T., Bland, P. A., and Towner, M. (2018). Automated Detection of Craters in Martian Satellite Imagery Using Convolutional Neural Networks. In *Planetary Science Informatics and Data Analytics Conference*, volume 2082 of *LPI Contributions*, page 6004.

Okubo, C. H. and Martel, S. J. (1998). Pit crater formation on Kilauea volcano, Hawaii. *Journal of Volcanology and Geothermal Research*, 86(1):1–18. doi: 10.1016/S0377-0273(98)00070-5.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

Palafox, L. F., Hamilton, C. W., Scheidt, S. P., and Alvarez, A. M. (2017). Automated detection of geological landforms on Mars using Convolutional Neural Networks. *Computers and Geosciences*, 101:48–56. doi: 10.1016/j.cageo.2016.12.015.

Patil, J. and Kini, S. N. (2015). A Survey Paper on Crater Detection. *International Journal of Science and Research*, 4:81–85.

Persello, C. et al. (2022). Deep Learning and Earth Observation to Support the Sustainable Development Goals: Current approaches, open challenges, and future opportunities. *IEEE Geoscience and Remote Sensing Magazine*, 10(2):172–200. doi: 10.1109/MGRS.2021.3136100.

Pizer, S. M. et al. (1987). Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3):355–368. doi: 10.1016/S0734-189X(87)80186-X.

Platt, J. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. Technical Report MSR-TR-98-14, Microsoft.

Plescia, J. B. and Cintala, M. J. (2012). Impact melt in small lunar highland craters. *Journal of Geophysical Research (Planets)*, 117:E00H12. doi: 10.1029/2011JE003941.

Popova, O., Nemtchinov, I., and Hartmann, W. K. (2003). Bolides in the present and past martian atmosphere and effects on cratering processes. *Meteoritics and Planetary Science*, 38(6):905–925. doi: 10.1111/j.1945-5100.2003.tb00287.x.

Povilaitis, R. Z., Robinson, M. S., van der Bogert, C. H., Hiesinger, H., Meyer, H. M., and Ostrach, L. R. (2018). Crater density differences: Exploring regional resurfacing, secondary crater populations, and crater saturation equilibrium on the moon. *Planetary and Space Science*, 162:41–51. doi: 10.1016/j.pss.2017.05.006.

Prieur, N. C. et al. (2023). Automatic Characterization of Boulders on Planetary Surfaces From High-Resolution Satellite Images. *Journal of Geophysical Research (Planets)*, 128(11):e2023JE008013. doi: 10.1029/2023JE008013.

QGIS (2025). QGIS Geographic Information System. QGIS Development Team. https://www.qgis.org.

Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv e-prints*. doi: 10.48550/arXiv.1511.06434.

Rajaneesh, A., Vishnu, C. L., Oommen, T., Rajesh, V. J., and Sajinkumar, K. S. (2022). Machine learning as a tool to classify extra-terrestrial landslides: A dossier from Valles Marineris, Mars. *Icarus*, 376:114886. doi: 10.1016/j.icarus.2022.114886.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2015). You Only Look Once: Unified, Real-Time Object Detection. *arXiv e-prints*. doi: 10.48550/arXiv.1506.02640.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv e-prints*. doi: 10.48550/arXiv.1506.01497.

RGCPS (2025). Lava Tube Database. Ronald Greeley Center for Planetary Studies (RGCPS), Arizona State University. https://rgcps.asu.edu/ltdb/.

Richardson, D. C. and Hagen, T. H. (2014). Crater Chain (Impact, Primary). In *Encyclopedia of Planetary Landforms*, pages 1–4. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_73-2.

Robbins, S. J. (2019). A New Global Database of Lunar Impact Craters >1-2 km: 1. Crater Locations and Sizes, Comparisons With Published Databases, and Global Analysis. *Journal of Geophysical Research (Planets)*, 124(4):871–892. doi: 10.1029/2018JE005592.

Robbins, S. J. and Hynek, B. M. (2012a). A new global database of Mars impact craters ≥1 km: 1. Database creation, properties, and parameters. *Journal of Geophysical Research (Planets)*, 117(E5):E05004. doi: 10.1029/2011JE003966.

Robbins, S. J. and Hynek, B. M. (2012b). A new global database of Mars impact craters ≥1 km: 2. Global crater properties and regional variations of the simple-to-complex transition diameter. *Journal of Geophysical Research (Planets)*, 117(E6):E06001. doi: 10.1029/2011JE003967.

Robbins, S. J., Kirchoff, M. R., and Hoover, R. H. (2023). Fully Controlled 6 Meters per Pixel Equatorial Mosaic of Mars From Mars Reconnaissance Orbiter Context Camera Images, Version 1. *Earth and Space Science*, 10(3):e2022EA002443. doi: 10.1029/2022EA002443.

Robinson, M. (2009). Lunar Reconnaissance Orbiter Narrow Angle Camera, Experiment Data Record, LRO-L-LROC-2-EDR-V1.0. NASA Planetary Data System (PDS). doi: 10.17189/1520250. https://pds.lroc.asu.edu/data/LRO-L-LROC-2-EDR-V1.0/.

Robinson, M. S. et al. (2010). Lunar Reconnaissance Orbiter Camera (LROC) Instrument Overview. *Space Science Reviews*, 150(1-4):81–124. doi: 10.1007/s11214-010-9634-2.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv e-prints*. doi: 10.48550/arXiv.1505.04597.

Rouseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65. doi: 10.1016/0377-0427(87)90125-7.

Rubanenko, L., Perez-Lopez, S., Schull, J., and Lapotre, M. G. A. (2021). Automatic Detection and Segmentation of Barchan Dunes on Mars and Earth Using a Convolutional Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:9364–9371. doi: 10.1109/JSTARS.2021.3109900.

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv e-prints*. doi: 10.48550/arXiv.1609.04747.

Salamunićcar, G., Lončarić, S., and Mazarico, E. (2012). LU60645GT and MA132843GT catalogues of Lunar and Martian impact craters developed using a Crater Shape-based interpolation crater detection algorithm for topography data. *Planetary and Space Science*, 60(1):236–247. doi: 10.1016/j.pss.2011.09.003.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *arXiv e-prints*. doi: 10.48550/arXiv.1801.04381.

Sauro, F. et al. (2019). Volcanic Caves of Lanzarote: A Natural Laboratory for Understanding Volcano-Speleogenetic Processes and Planetary Caves. In Mateo, E., Martínez-Frías, J., and Vegas, J., editors, *Lanzarote and Chinijo Islands Geopark: From Earth to Space*, pages 125–142. Springer International Publishing, Cham. doi: 10.1007/978-3-030-13130-2_9.

Sauro, F., Pozzobon, R., Massironi, M., De Berardinis, P., Santagata, T., and De Waele, J. (2020). Lava tubes on Earth, Moon and Mars: A review on their size and morphology revealed by comparative planetology. *Earth Science Reviews*, 209:103288. doi: 10.1016/j.earscirev.2020.103288.

Schörghofer, N. (2015). Slope Streak (Mars). In Hargitai, H. and Kereszturi, Á., editors, *Encyclopedia of Planetary Landforms*, pages 1980–1986. Springer New York, New York, NY. doi: 10.1007/978-1-4614-3134-3_344.

Schultz, P. H. (1976). Floor-Fractured Lunar Craters. *Moon*, 15(3-4):241–273. doi: 10.1007/BF00562240.

Schultz, P. H. (1978). Martian intrusions: Possible sites and implications. *Geophysical Research Letters*, 5(6):457–460. doi: 10.1029/GL005i006p00457.

Settles, B. (2012). Active Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 6(1):1–114. doi: 10.2200/S00429ED1V01Y201207AIM018.

Sidiropoulos, P. and Muller, J. P. (2015). On the status of orbital high-resolution repeat imaging of Mars for the observation of dynamic surface processes. *Planetary and Space Science*, 117:207–222. doi: 10.1016/j.pss.2015.06.017.

Silburt, A. et al. (2019). Lunar crater identification via deep learning. *Icarus*, 317:27–38. doi: 10.1016/j.icarus.2018.06.022.

Simonyan, K. and Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv e-prints*. doi: 10.48550/arXiv.1409.1556.

Smith, D. E. et al. (2001). Mars Orbiter Laser Altimeter: Experiment summary after the first year of global mapping of Mars. *Journal of Geophysical Research*, 106(E10):23689–23722. doi: 10.1029/2000JE001364.

Smith, D. E. et al. (2010). The Lunar Orbiter Laser Altimeter Investigation on the Lunar Reconnaissance Orbiter Mission. *Space Science Reviews*, 150(1-4):209–241. doi: 10.1007/s11214-009-9512-y.

Solomon, S. C. et al. (2001). The MESSENGER mission to Mercury: scientific objectives and implementation. *Planetary and Space Science*, 49(14-15):1445–1465. doi: 10.1016/S0032-0633(01)00085-X.

Som, S. M., Greenberg, H. M., and Montgomery, D. R. (2008). The Mars Orbiter Laser Altimeter dataset: Limitations and improvements. *International Journal of Mars Science and Exploration*, 4:14–26. doi: 10.1555/mars.2008.0002.

Spencer, J. R. and Fanale, F. P. (1990). New models for the origin of Valles Marineris closed depressions. *Journal of Geophysical Research*, 95:14301–14313. doi: 10.1029/JB095iB09p14301.

Speyerer, E. J., Robinson, M. S., Denevi, B. W., and LROC Science Team (2011). Lunar Reconnaissance Orbiter Camera Global Morphological Map of the Moon. In *42nd Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 2387. https://astrogeology.usgs.gov/search/map/moon_lro_lroc_wac_global_morphology_mosaic_100m.

Srivastava, N. and Varatharajan, I. (2016). Geomorphology of Lowell crater region on the Moon. *Icarus*, 266:44–56. doi: 10.1016/j.icarus.2015.11.013.

Stepinski, T. F., Ding, W., and Vilalta, R. (2012). Detecting impact craters in planetary images using machine learning. In Magdalena-Benedito, R., Martinez-Sober, M., Martinez-Martinez, J., Escandell-Moreno, P., and Vila-Frances, J., editors, *Intelligent Data Analysis for Real-Life Applications: Theory and Practice*, pages 146–159. IGI Global. doi: 10.4018/978-1-4666-1806-0.ch008.

Sun, M. and Chen, S. (2022). Deep Learning-Based Super-Resolution Reconstruction and Algorithm Acceleration of Mars Hyperspectral CRISM Data. *Remote Sensing*, 14(13):3062. doi: 10.3390/rs14133062.

Szegedy, C. et al. (2014). Going Deeper with Convolutions. *arXiv e-prints*. doi: 10.48550/arXiv.1409.4842.

Tewari, A., Verma, V., Srivastava, P., Jain, V., and Khanna, N. (2022). Automated Crater detection from Co-registered optical images, elevation maps and slope maps using deep learning. *Planetary and Space Science*, 218:105500. doi: 10.1016/j.pss.2022.105500.

Thaker, A. D., Patel, S. M., and Solanki, P. M. (2020). Morphological analysis and mapping of complex craters of Copernican age: Crookes, Ohm and Stevinus. *Planetary and Space Science*, 184:104856. doi: 10.1016/j.pss.2020.104856.

Theinat, A. K. et al. (2020). Lunar lava tubes: Morphology to structural stability. *Icarus*, 338:113442. doi: 10.1016/j.icarus.2019.113442.

van der Bogert, C. H. and Ashley, J. W. (2015). Skylight. In Hargitai, H. and Kereszturi, Á., editors, *Encyclopedia of Planetary Landforms*, pages 1964–1969. Springer New York, New York, NY. doi: 10.1007/978-1-4614-3134-3_342.

van der Bogert, C. H., Ashley, J. W., and Ferrill, D. (2014). Pit Crater. In *Encyclopedia of Planetary Landforms*, pages 1–7. Springer New York, New York, NY. doi: 10.1007/978-1-4614-9213-9_260-1.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–I. doi: 10.1109/CVPR.2001.990517.

Wagner, R. and LROC Team (2025). Lunar Reconnaissance Orbiter Narrow Angle Camera Processing Guide. Zenodo. doi: 10.5281/zenodo.15392665. https://doi.org/10.5281/zenodo.15392665.

Wagner, R. V. and Robinson, M. S. (2014). Distribution, formation mechanisms, and significance of lunar pits. *Icarus*, 237:52–60. doi: 10.1016/j.icarus.2014.04.002.

Wagner, R. V. and Robinson, M. S. (2021). Occurrence and Origin of Lunar Pits: Observations from a New Catalog. In *52nd Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 2530. https://www.lroc.asu.edu/atlases/pits.

Wagner, R. V. and Robinson, M. S. (2022). Lunar Pit Morphology: Implications for Exploration. *Journal of Geophysical Research (Planets)*, 127(8):e07328. doi: 10.1029/2022JE007328.

Wallace, I. et al. (2017). LabelMars.net: Crowd-Sourcing an Extremely Large High Quality Martian Image Dataset. In *48th Annual Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 1170.

Wang, C., Zhang, Y., Zhang, Y., Tian, R., and Ding, M. (2021). Mars Image Super-Resolution Based on Generative Adversarial Network. *IEEE Access*, 9:108889–108898. doi: 10.1109/ACCESS.2021.3101858.

Wang, H., Jiang, J., and Zhang, G. (2018). CraterIDNet: An End-to-End Fully Convolutional Neural Network for Crater Detection and Identification in Remotely Sensed Planetary Images. *Remote Sensing*, 10(7):1067. doi: 10.3390/rs10071067.

Wang, J., Cheng, W., Zhou, C., and Zheng, X. (2017a). Automatic mapping of lunar landforms using DEM-derived geomorphometric parameters. *Journal of Geographical Sciences*, 27(11):1413–1427. doi: 10.1007/s11442-017-1443-z.

Wang, S., Fan, Z., Li, Z., Zhang, H., and Wei, C. (2020). An Effective Lunar Crater Recognition Algorithm Based on Convolutional Neural Network. *Remote Sensing*, 12(17):2694. doi: 10.3390/rs12172694.

Wang, Y., Di, K., Xin, X., and Wan, W. (2017b). Automatic detection of Martian dark slope streaks by machine learning using HiRISE images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129:12–20. doi: 10.1016/j.isprsjprs.2017.04.014.

Wang, Y. and Wu, B. (2019). Active Machine Learning Approach for Crater Detection From Planetary Imagery and Digital Elevation Models. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8):5777–5789. doi: 10.1109/TGRS.2019.2902198.

Wang, Y., Wu, B., Xue, H., Li, X., and Ma, J. (2021). An Improved Global Catalog of Lunar Impact Craters ($\geq$1 km) With 3D Morphometric Information and Updates on Global Crater Analysis. *Journal of Geophysical Research (Planets)*, 126(9):e06728. doi: 10.1029/2020JE006728.

Watson, T. H. and Baldini, J. U. L. (2024). Martian cave detection via machine learning coupled with visible light imagery. *Icarus*, 411:115952. doi: 10.1016/j.icarus.2024.115952.

Weber, R. C. et al. (2021). The Artemis III Science Definition Team Report. In *52nd Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 1261.

Wilhelm, T. et al. (2020). DoMars16k: A Diverse Dataset for Weakly Supervised Geomorphologic Analysis on Mars. *Remote Sensing*, 12(23):3981. doi: 10.3390/rs12233981.

Williams, D. A. et al. (2011). Volcanism on Io: New insights from global geologic mapping. *Icarus*, 214(1):91–112. doi: 10.1016/j.icarus.2011.05.007.

Williams, K. E., McKay, C. P., Toon, O. B., and Head, J. W. (2010). Do ice caves exist on Mars? *Icarus*, 209(2):358–368. doi: 10.1016/j.icarus.2010.03.039.

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., and Girshick, R. (2019). Detectron2. https://github.com/facebookresearch/detectron2.

Wu, Y., Wan, G., Liu, L., Wei, Z., and Wang, S. (2021). Intelligent Crater Detection on Planetary Surface Using Convolutional Neural Network. In *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, volume 5, pages 1229–1234. doi: 10.1109/IAEAC50856.2021.9391002.

Wyrick, D., Ferrill, D. A., Morris, A. P., Colton, S. L., and Sims, D. W. (2004). Distribution, morphology, and origins of Martian pit crater chains. *Journal of Geophysical Research (Planets)*, 109(E6):E06005. doi: 10.1029/2004JE002240.

Xiao, Z., Zeng, Z., Li, Z., Blair, D. M., and Xiao, L. (2014). Cooling fractures in impact melt deposits on the Moon and Mercury: Implications for cooling solely by thermal radiation. *Journal of Geophysical Research (Planets)*, 119(7):1496–1515. doi: 10.1002/2013JE004560.

Yang, C. et al. (2020). Lunar impact crater identification and age estimation with Chang'E data by deep and transfer learning. *Nature Communications*, 11:6358. doi: 10.1038/s41467-020-20215-y.

Zhang, S. et al. (2024). Detecting Lunar Linear Structures Based on Multimodal Semantic Segmentation: The Case of Sinuous Rilles. *Remote Sensing*, 16(9):1602. doi: 10.3390/rs16091602.

Zhang, Z., Liu, Q., and Wang, Y. (2018). Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5):749–753. doi: 10.1109/LGRS.2018.2802944.

Zhao, Y. and Ye, H. (2024). Crater Detection and Population Statistics in Tianwen-1 Landing Area Based on Segment Anything Model (SAM). *Remote Sensing*, 16(10):1743. doi: 10.3390/rs16101743.

Zhou, M. et al. (2024). High-Resolution Morphology of Lunar Lava Tube Pits Using Photogrammetric Modeling of Multiple Stereo Images. *Earth and Space Science*, 11(11):2024EA003532. doi: 10.1029/2024EA003532.

Zuo, W., Li, C., Zhang, Z., Zeng, X., Liu, Y., and Xiong, Y. (2021). China's Lunar and Planetary Data System: Preserve and Present Reliable Chang'e Project and Tianwen-1 Scientific Data Sets. *Space Science Reviews*, 217(8):88. doi: 10.1007/s11214-021-00862-3.

Appendices

A Technical Notes

TN-1 BASH Code for Filtering PDS Index Files

Below is a piece of BASH code which will filter the cumulative index files found on NASA's Planetary Data System (PDS) for all Mars Reconnaissance Orbiter (MRO) High Resolution Imaging Science Experiment (HiRISE; McEwen et al. 2007) Reduced Data Record Version 1.1 (RDRV11) or Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC; Robinson et al. 2010) images. Since there are over 2.6 million LROC NAC images on PDS at the time of writing, it may be necessary to download the individual index files for each volume, as opposed to the entire cumulative file.

```
# Get product names from input directory
images=`ls $PWD`

# For HiRISE RDRV11 images
old=$PWD"RDRCUMINDEX.TAB"

# For LROC NAC EDR images
old=$PWD"CUMINDEX.TAB"

# Define the name of the filtered index file
new=$PWD"CUMINDEX_filtered.TAB"

# Loop through every image in the directory
for image in $images; do

    # Get the product name (minus the extension)
    name=`basename $image $ext`

    # Get the relevant line from the index file
    grep $name $old >> $new

done
```

TN-2 BASH Code for Processing Lunar Reconnaissance Orbiter Narrow Angle Camera Experimental Data Record Products

Below is a piece of BASH code which will process a batch of raw LROC NAC Experimental Data Record (EDR) image products (IMG files). The code converts the raw files into ISIS cubes, initialises their SPICE kernels using the USGS web service, performs radiometric and echo correction, projects the image onto the Lunar surface, and converts from an ISIS cube to a GIS-ready Uint8 GeoTiff. The code will also generate an accurate footprint (in ESRI shapefile format) of the location of the image. If the user inputs the minimum/maximum latitude-longitude extent for their particular region of interest, the code will crop the ISIS cubes during the map-projection stage in order to save storage space. The code assumes that the user has two Conda environments created, with each having one of ISIS or GDAL installed. Instructions for doing so can be found on the Planetary Image Processing GitHub repository (Le Corre, 2025f).

```
# Loop through each raw PDS file in the directory
for i in *.IMG; do

    # Activate the ISIS conda environment
    source /home/username/anaconda3/bin/activate isis

    # Get the base name of the image product
    base=`basename $i .IMG`

    # Convert the raw PDS file to an ISIS cube
    lronac2isis from=$i to=$base.cub

    # Initialise the spice kernels for the raw ISIS cube
    spiceinit from=$base.cub web=yes
    ↪  url=https://astrogeology.usgs.gov/apis/ale/v0.9.1/spiceserver/

    # Radiometrically calibrate the spice-initialised ISIS cube
    lronaccal from=$base.cub to=$base.cal.cub

    # Echo-correct the calibrated ISIS cube
    lronacecho from=$base.cal.cub to=$base.echo.cub

    # Produce a .map file for the echo-corrected ISIS cube
    list=$base.lis
    map=$base.map
    echo $base.echo.cub > $list
    mosrange fromlist=$list projection=equirectangular to=$map

    # Map-project the processed ISIS cube without cropping to a specific RoI
    if [ $roi == "no" ]; then

        # Do or do not up/downscale the image
        if [ "$res" == "0" ]; then
            cam2map from=$base.echo.cub to=$base.map map=$map
            ↪  warpalgorithm=forwardpatch patchsize=$patch matchmap=true
        elif [ "$res" != "0" ]; then
            cam2map from=$base.echo.cub to=$base.map map=$map interp=cubicconvolution
            ↪  warpalgorithm=forwardpatch patchsize=$patch pixres=mpp
            ↪  resolution="$res"
        fi

    # Map-project the processed ISIS cube and crop the images to a specific RoI
    else
```

```
        # Get the min/max latitude and longitude of the image from the .map file
        camminlat=`getkey from=$map grpname=Mapping keyword=MinimumLatitude`
        camminlon=`getkey from=$map grpname=Mapping keyword=MinimumLongitude`
        cammaxlat=`getkey from=$map grpname=Mapping keyword=MaximumLatitude`
        cammaxlon=`getkey from=$map grpname=Mapping keyword=MaximumLongitude`

        # Check if the extents of the image fall outside of the RoI
        if (( $(echo "$camminlat < $roiminlat" | bc -l) )); then
            camminlat=$roiminlat
        fi
        if (( $(echo "$camminlon < $roiminlon" | bc -l) )); then
            camminlon=$roiminlon
        fi
        if (( $(echo "$cammaxlat > $roimaxlat" | bc -l) )); then
            cammaxlat=$roimaxlat
        fi
        if (( $(echo "$cammaxlon > $roimaxlon" | bc -l) )); then
            cammaxlon=$roimaxlon
        fi

        # Do or do not up/downscale the image
        if [ "$res" == "0" ]; then
            cam2map from=$base.echo.cub to=$base.map map=$map
            ↪  warpalgorithm=forwardpatch patchsize=$patch matchmap=true
            ↪  defaultrange=map minlat="$camminlat" maxlat="$cammaxlat"
            ↪  minlon="$camminlon" maxlon="$cammaxlon"
        elif [ "$res" != "0" ]; then
            cam2map from=$base.echo.cub to=$base.map map=$map interp=cubicconvolution
            ↪  warpalgorithm=forwardpatch patchsize=$patch pixres=mpp
            ↪  resolution="$res" defaultrange=map minlat="$camminlat"
            ↪  maxlat="$cammaxlat" minlon="$camminlon" maxlon="$cammaxlon"
        fi
    fi

    # Deactivate the ISIS conda environment
    conda deactivate

    # Activate the GDAL conda environment
    source /home/username/anaconda3/bin/activate gdal

    # Convert the mapped ISIS cube to a 0-255 GeoTiff
    gdal_translate -a_nodata none -of GTiff -ot Byte -scale $base.map.cub $base.tif

    # Produce a footprint for the image as an ESRI shapefile
    gdal_calc.py -A $base.tif --quiet --calc="numpy.where(A >= 0, 1, 0)"
    ↪  --format=GTiff --type=UInt16 --NoDataValue=0 --A_band=1
    ↪  --outfile=$base.binary.tif
    mv $base.tif.msk $base.binary.tif.msk
    gdal_polygonize.py $base.binary.tif -f "ESRI Shapefile" $base.shp

    # Deactivate the GDAL conda environment
    conda deactivate

done
```

TN-3 BASH Code for Correcting Mars Reconnaissance Orbiter HiRISE Map-Projection Errors

Below is some BASH code which will correct the latitudinal error in the embedded map-projection found in some MRO HiRISE RDRV11 images. The code will loop through all JP2 files (the format of HiRISE RDRV11 images on PDS) and switch the 'standard parallel' and 'centre latitude' parameters within the Proj4 string used to map-project the image products to the Martian surface. If there is no error found in the map projection, then the image is simply converted from a UInt16 JP2 file to a UInt8 GeoTiff. This code requires the GDAL software suite to be installed.

```
# Loop through all JP2 files
for i in *.JP2; do

    # Get the name of the image
    base=`basename $i ".JP2"`

    # Get the Proj4 string for the spatial reference
    str=`gdalsrsinfo -o proj4 $i`
    str_list=($str)

    # Find the standard parallel (lat_ts) and the centre latitude (lat_0)
    lat_ts=${str_list[1]:8}
    lat_0=${str_list[2]:7}

    # Find if lat_0 is non-zero
    if [ $lat_0 -ne 0 ]; then

        # Reassign parameter values
        var1="+lat_ts=${lat_0}"
        var2="+lat_0=${lat_ts}"

        # Input new parameter values into Proj4 string
        str_list[1]=$var1
        str_list[2]=$var2
        srs=$(IFS=" " ; echo "${str_list[*]}")

        # Translate .JP2 to GeoTiff and use new projection
        gdal_translate -ot Byte -a_srs "${srs[@]}" -scale $i $base.tif

    else

        # If lat_0 is zero, just convert to UInt8 GeoTiff
        gdal_translate -ot Byte -a_srs "$str" -scale $i $base.tif

    fi
done
```

B Appendix for Chapter 4: PITS

B.1 Derivation of Uncertainty in True Shadow Width (ΔS_{true})

As discussed in Section 4.3.3, the uncertainty in the true shadow width (ΔS_{true}) depends only on the uncertainties in the observed shadow width (ΔS_{obs}) and the obliquity of the satellite parallel to the Sun's line of sight ($\Delta\varepsilon_{\shortparallel}$). Hence, ΔS_{true} is given below.

$$\Delta S_{\text{true}} = \sqrt{{\Delta S_{\text{obs}}}^2 \left(\frac{\partial S_{\text{true}}}{\partial S_{\text{obs}}}\right)^2 + {\Delta\varepsilon_{\shortparallel}}^2 \left(\frac{\partial S_{\text{true}}}{\partial \varepsilon_{\shortparallel}}\right)^2}$$

$\Delta\varepsilon_{\shortparallel}$ can be found by knowing the range of possible ε values given the size of the image, and is defined by the equation below.

$$\begin{aligned}
\Delta\varepsilon_{\shortparallel} &= \Delta\varepsilon \left(\frac{\partial \varepsilon_{\shortparallel}}{\partial \varepsilon}\right) \\
&= \Delta\varepsilon \left(\frac{\partial}{\partial \varepsilon}\left[\arctan\left(v/d_h\right)\right]\right) \\
&= \Delta\varepsilon \left(\frac{\partial}{\partial \varepsilon}\left[\arctan\Big(|\cos\gamma|\tan\varepsilon\Big)\right]\right) \\
&= \Delta\varepsilon \left(\frac{|\cos\gamma|\sec^2\varepsilon}{\cos^2\gamma\tan^2\varepsilon + 1}\right)
\end{aligned}$$

In case A), where $0° < \gamma < 90°$, S_{true} is given by Equation (4.14) and the uncertainty is derived by the following.

$$\begin{aligned}
\frac{\partial S_{\text{true}}}{\partial S_{\text{obs}}} &= \frac{\partial}{\partial S_{\text{obs}}}\left[\frac{S_{\text{obs}}}{\cos\varepsilon_{\shortparallel}} + \frac{S_{\text{obs}}\tan\varepsilon_{\shortparallel}}{\cos\varepsilon_{\shortparallel}(\tan\alpha - \tan\varepsilon_{\shortparallel})}\right] \\
&= \frac{1}{\cos\varepsilon_{\shortparallel}}\left(1 + \frac{\tan\varepsilon_{\shortparallel}}{\cos\varepsilon_{\shortparallel}(\tan\alpha - \tan\varepsilon_{\shortparallel})}\right) \\
\frac{\partial S_{\text{true}}}{\partial \varepsilon_{\shortparallel}} &= \frac{\partial}{\partial \varepsilon_{\shortparallel}}\left[\frac{S_{\text{obs}}}{\cos\varepsilon_{\shortparallel}} + \frac{S_{\text{obs}}\tan\varepsilon_{\shortparallel}}{\cos\varepsilon_{\shortparallel}(\tan\alpha - \tan\varepsilon_{\shortparallel})}\right] \\
&= S_{\text{obs}}\left(\frac{\partial}{\partial \varepsilon_{\shortparallel}}\left[\frac{1}{\cos\varepsilon_{\shortparallel}}\right] + \frac{\partial}{\partial \varepsilon_{\shortparallel}}\left[\frac{\tan\varepsilon_{\shortparallel}}{\cos\varepsilon_{\shortparallel}(\tan\alpha - \tan\varepsilon_{\shortparallel})}\right]\right) \\
&= \frac{S_{\text{obs}}}{\cos\varepsilon_{\shortparallel}}\left(\frac{\sin\varepsilon_{\shortparallel}}{\cos\varepsilon_{\shortparallel}} + \frac{\sec^2\varepsilon_{\shortparallel}}{\cos\varepsilon_{\shortparallel}(\tan\alpha - \tan\varepsilon_{\shortparallel})^2}\right. \\
&\quad \left. + \frac{\sin\varepsilon_{\shortparallel}\tan\varepsilon_{\shortparallel}}{\tan\alpha - \tan\varepsilon_{\shortparallel}} + \frac{\sec^2\varepsilon_{\shortparallel}\tan\varepsilon_{\shortparallel}}{(\tan\alpha - \tan\varepsilon_{\shortparallel})^2}\right)
\end{aligned}$$

In case B), where $90° < \gamma < 180°$, S_{true} is given by Equation (4.16) and the uncertainty

is derived by the following:

$$\begin{aligned}
\frac{\partial S_{\text{true}}}{\partial S_{\text{obs}}} &= \frac{\partial}{\partial S_{\text{obs}}}\left[\frac{S_{\text{obs}}\sin\alpha}{\cos\left(90^\circ-\alpha-\varepsilon_{\shortparallel}\right)}\right] \\
&= \frac{\sin\alpha}{\cos\left(90^\circ-\alpha-\varepsilon_{\shortparallel}\right)} \\
\frac{\partial S_{\text{true}}}{\partial \varepsilon_{\shortparallel}} &= \frac{\partial}{\partial \varepsilon_{\shortparallel}}\left[\frac{S_{\text{obs}}\sin\alpha}{\cos\left(90^\circ-\alpha-\varepsilon_{\shortparallel}\right)}\right] \\
&= S_{\text{obs}}\sin\alpha\left(\frac{\partial}{\partial \varepsilon_{\shortparallel}}\left[\frac{1}{\cos\left(90^\circ-\alpha-\varepsilon_{\shortparallel}\right)}\right]\right) \\
&= S_{\text{obs}}\sin\alpha\left(\frac{\sin\left(90^\circ-\alpha-\varepsilon_{\shortparallel}\right)}{\cos^2\left(90^\circ-\alpha-\varepsilon_{\shortparallel}\right)}\right)
\end{aligned}$$

It is possible, albeit highly unlikely, that γ could equal exactly 0, 90, or 180°. If $\gamma = 0$ or 180°, $\varepsilon_{\perp}$ will be zero, meaning that $\varepsilon_{\shortparallel} = \varepsilon$ and $\Delta\varepsilon_{\shortparallel} = \Delta\varepsilon$. This means that Equations (4.14) and (4.16) would reduce to those used by Cushing et al. (2015) if $\gamma = 0$ or 180°, respectively. However, $\varepsilon_{\shortparallel}$ will be zero if $\gamma = 90^\circ$, meaning that there would be no distortion in the shadow width (i.e. $S_{\text{true}} = S_{\text{obs}}$ and $\Delta S_{\text{true}} = \Delta S_{\text{obs}}$).

B.2 PITS Apparent Depths (h) for 123 HiRISE Images of 88 Atypical Pit Craters (APCs)

Below is a table of the apparent depths (h) taken at the centre (h_c) and the maximum (h_m) of the profiles calculated by the Pit Topography from Shadows (PITS) tool upon 123 HiRISE images of 88 Mars Global Cave Candidate Catalog (MGC^3) atypical pit craters (APCs). This table is available as a CSV file within the supporting material of Le Corre et al. (2023a), which also gives the sensing information for each HiRISE image.

Image Name	MGC^3 Name	h_c [m]	+	-	h_m [m]	+	-
ESP_011386_2065	APC095	175.3	0.8	9.2	177.6	0.8	9.3
ESP_011531_2065	APC095	176.4	0.8	9.2	177.9	0.8	9.3
ESP_011677_1655	APC079	69.7	0.3	3.7	85.9	0.4	4.5
ESP_011756_1735	APC091	131.5	0.6	6.9	132.1	0.6	6.9
ESP_012600_1655	APC079	56.4	0.2	3.0	75.7	0.3	4.0
ESP_013167_1785	APC042	87.0	2.8	7.0	90.2	0.4	4.7
ESP_013589_1785	APC042	88.2	3.1	5.7	119.0	3.1	7.6
ESP_014077_1660	APC117	39.3	0.2	2.1	43.3	0.2	2.3
ESP_016411_1605	APC096	115.2	0.5	6.0	121.5	2.9	8.7
ESP_016622_1660	APC034	47.6	0.2	2.5	48.7	0.2	2.5
ESP_016978_1730	APC119	67.3	3.6	6.9	131.2	12.9	19.2
ESP_017189_1730	APC119	67.8	2.1	5.4	137.8	1.1	7.7
ESP_017544_2055	APC121	94.8	0.4	5.0	96.7	0.4	5.1
ESP_019351_1795	APC084	88.1	0.4	4.6	90.0	0.4	4.7

Table continues on following page

Image Name	MGC^3 Name	h_c [m]	+	-	h_m [m]	+	-
ESP_019984_1795	APC084	83.5	2.1	6.2	90.9	0.4	4.8
ESP_019997_1975	APC085	72.7	20.7	24.2	88.5	0.4	4.6
ESP_019997_1975	APC092	161.5	0.7	8.4	163.6	0.7	8.6
ESP_021738_1625	APC099	147.7	0.6	7.7	150.8	0.7	7.9
ESP_024481_1605	APC096	94.8	0.4	5.1	176.7	0.8	9.4
ESP_025892_1780	APC129	104.5	0.8	5.5	105.9	0.8	5.6
ESP_028450_1730	APC061	51.8	0.2	2.7	54.2	0.2	2.8
ESP_028793_1655	APC078	68.4	0.3	3.6	68.7	0.3	3.6
ESP_029043_1975	APC085	82.6	1.3	5.3	82.6	1.3	5.3
ESP_029043_1975	APC092	139.7	0.6	7.3	140.3	0.6	7.3
ESP_030995_1610	APC089	58.6	0.3	3.1	85.5	0.4	4.5
ESP_031285_1785	APC052	22.6	0.1	1.2	26.2	0.1	1.4
ESP_033342_1660	APC071	83.6	0.4	4.4	88.8	0.4	4.6
ESP_033355_1635	APC098	51.1	8.6	11.1	74.7	0.3	3.9
ESP_033618_2050	APC094	54.7	0.2	2.9	68.2	0.3	3.6
ESP_033935_1805	APC054	62.7	0.5	3.3	63.6	0.5	3.4
ESP_036454_1660	APC074	43.6	0.3	2.3	45.3	0.3	2.4
ESP_036770_1805	APC054	56.3	0.2	2.9	57.1	0.2	3.0
ESP_038076_1795	APC056	67.0	0.3	3.5	67.6	0.3	3.5
ESP_041162_1665	APC131	46.8	0.2	2.5	49.6	0.2	2.6
ESP_041900_1805	APC054	54.4	0.3	3.0	124.6	0.6	6.9
ESP_042085_1795	APC050	78.1	1.7	4.4	78.8	1.7	4.5
ESP_042085_1795	APC056	74.8	0.3	3.9	84.1	0.4	4.4
ESP_042128_2090	APC106	95.5	0.4	5.0	98.7	0.4	5.2
ESP_042629_2090	APC108	96.6	0.4	5.1	99.2	0.4	5.2
ESP_042721_2045	APC114	15.3	0.1	0.8	25.1	0.1	1.3
ESP_042932_2050	APC112	62.1	0.3	3.2	67.6	0.3	3.5
ESP_042985_2100	APC110	40.5	0.2	2.1	45.1	0.2	2.4
ESP_043222_2035	APC115	61.3	0.3	3.2	63.9	0.3	3.3
ESP_043763_2100	APC109	48.5	0.2	2.5	70.1	0.3	3.7
ESP_043974_2090	APC107	25.6	0.1	1.3	50.5	0.2	2.6
ESP_044748_2110	APC139	87.7	0.4	4.6	92.4	0.4	4.8
ESP_045830_1735	APC104	56.0	0.2	2.9	56.0	0.2	2.9
ESP_046753_1610	APC029	16.6	0.1	0.9	18.1	0.1	0.9
ESP_049812_1735	APC082	78.4	0.3	4.1	86.8	5.2	9.4
ESP_050089_1660	APC071	79.1	0.4	4.1	90.2	0.4	4.7
ESP_050234_1735	APC082	74.4	0.3	3.9	76.7	0.3	4.0
ESP_051829_1755	APC140	93.1	0.4	4.9	96.6	0.4	5.1
ESP_052638_2020	APC130	44.3	0.2	2.3	57.2	0.2	3.0

Table continues on following page

Image Name	MGC^3 Name	h_c [m]	+	-	h_m [m]	+	-
ESP_053547_2095	APC111	54.3	0.2	2.8	64.1	0.3	3.4
ESP_053781_1660	APC074	42.5	0.3	2.2	50.8	0.3	2.7
ESP_055196_2015	APC130	44.8	0.2	2.3	50.1	0.2	2.6
ESP_055614_1605	APC040	19.8	0.1	1.0	38.1	0.2	2.0
ESP_056867_1695	APC090	51.0	1.4	3.8	57.2	0.2	3.0
ESP_056946_1665	APC049	42.6	0.2	2.2	46.4	0.2	2.4
ESP_057025_1640	APC048	26.8	0.1	1.4	26.8	0.1	1.4
ESP_057434_1790	APC001	9.2	0.0	0.5	10.2	0.0	0.5
ESP_057434_1790	APC033	42.0	0.2	2.2	43.0	0.2	2.2
ESP_057645_1770	APC032	21.4	0.2	1.1	28.6	0.2	1.5
ESP_057750_1870	APC039	27.1	0.1	1.4	33.7	0.1	1.8
ESP_057764_1775	APC135	15.3	0.1	0.8	16.0	0.1	0.8
ESP_057764_1775	APC136	34.2	2.1	3.8	36.5	0.2	1.9
ESP_057764_1775	APC137	40.2	0.2	2.1	42.2	0.2	2.2
ESP_057777_1790	APC041	31.9	0.1	1.7	35.1	0.2	1.8
ESP_057830_1820	APC059	34.3	0.1	1.8	37.5	0.2	2.0
ESP_057975_1810	APC076	69.1	0.3	3.6	72.3	0.3	3.8
ESP_057988_1735	APC008	32.8	0.1	1.7	33.8	0.1	1.8
ESP_058133_1770	APC022	18.8	0.3	1.0	22.4	0.3	1.2
ESP_058186_1660	APC117	32.5	3.2	4.8	40.1	0.2	2.1
ESP_058199_1785	APC043	29.8	0.1	1.6	34.8	0.1	1.8
ESP_058489_1785	APC007	17.8	0.1	0.9	18.2	0.1	1.0
ESP_058542_1625	APC062	62.8	0.3	3.3	77.0	0.3	4.0
ESP_058687_1810	APC027	39.1	0.2	2.0	43.9	0.2	2.3
ESP_058700_1785	APC006	31.3	1.2	2.0	38.0	1.2	2.5
ESP_058845_1780	APC038	35.4	0.2	1.9	43.2	0.2	2.3
ESP_058898_1775	APC134	20.4	0.1	1.1	22.5	0.1	1.2
ESP_059544_1650	APC125	22.5	0.1	1.2	25.5	0.1	1.3
ESP_059544_1650	APC126	22.2	0.1	1.2	28.5	0.1	1.5
ESP_059557_1740	APC017	62.0	0.3	3.2	63.0	0.3	3.3
ESP_059702_1790	APC012	45.9	0.2	2.4	46.6	0.2	2.4
ESP_059702_1790	APC013	46.9	0.2	2.5	48.7	0.2	2.5
ESP_060379_2020	APC130	68.8	0.4	3.9	82.6	0.4	4.7
ESP_061772_2020	APC055	55.9	0.2	2.9	57.2	0.2	3.0
ESP_062924_2020	APC130	50.7	0.2	2.7	51.7	0.2	2.8
ESP_063262_1755	APC140	105.2	0.5	5.5	105.4	0.5	5.5
ESP_063434_1780	APC045	55.1	0.2	2.9	55.2	0.2	2.9
ESP_063477_2090	APC108	112.3	0.5	5.9	116.9	0.5	6.1
ESP_063829_1755	APC140	96.8	0.4	5.1	97.0	0.4	5.1

Table continues on following page

Image Name	MGC^3 Name	h_c [m]	+	-	h_m [m]	+	-
ESP_065887_1660	APC014	41.1	0.2	2.2	41.6	0.2	2.2
ESP_066005_2110	APC133	22.3	0.1	1.2	27.0	0.1	1.4
ESP_066507_1610	APC138	33.2	0.1	1.7	43.9	0.2	2.3
ESP_066599_1755	APC011	15.4	0.1	0.8	15.4	0.1	0.8
ESP_066863_1660	APC065	15.0	0.1	0.8	30.2	0.1	1.6
ESP_066876_1785	APC009	28.2	0.1	1.5	31.0	0.1	1.6
ESP_066929_1780	APC045	54.1	0.2	2.8	55.4	0.2	2.9
ESP_066942_1735	APC015	16.6	0.1	0.9	33.2	0.1	1.8
ESP_067707_1840	APC025	35.9	0.9	2.1	40.7	0.9	2.4
ESP_071022_2100	APC109	60.9	1.7	4.7	86.2	0.4	4.5
PSP_003317_1975	APC085	64.4	0.3	3.4	69.1	0.3	3.6
PSP_003317_1975	APC092	123.0	0.5	6.4	128.2	0.5	6.7
PSP_003647_1745	APC075	120.0	0.5	6.3	120.8	0.5	6.3
PSP_004847_1745	APC075	221.1	1.0	11.7	227.4	1.0	12.1
PSP_004913_1735	APC082	97.8	0.4	5.1	99.7	0.4	5.2
PSP_005058_1720	APC072	68.7	5.8	9.1	73.5	2.2	5.7
PSP_005203_1730	APC068	53.7	0.2	2.8	58.1	0.2	3.0
PSP_005203_1730	APC086	116.8	1.1	6.7	123.4	0.5	6.4
PSP_005414_1735	APC091	99.4	0.4	5.2	103.1	0.4	5.4
PSP_005625_1730	APC068	56.7	1.4	4.1	67.8	0.3	3.6
PSP_005625_1730	APC086	119.5	0.6	6.3	122.3	0.6	6.5
PSP_005770_1745	APC075	278.3	1.2	14.6	280.6	1.2	14.7
PSP_005836_1735	APC082	80.4	0.3	4.2	88.5	0.4	4.6
PSP_006693_1755	APC093	110.2	0.5	5.8	122.2	0.5	6.4
PSP_006904_1755	APC093	124.1	6.7	12.7	131.3	0.6	6.9
PSP_007774_1745	APC023	33.0	0.1	1.7	34.4	0.1	1.8
PSP_008130_1745	APC023	34.4	0.1	1.8	35.2	0.2	1.8
PSP_009488_1745	APC005	12.3	0.1	0.6	12.6	0.1	0.7
PSP_009488_1745	APC030	29.0	0.1	1.5	29.3	0.1	1.5
PSP_009765_1780	APC045	52.6	0.2	2.8	54.7	0.2	2.9
PSP_009989_1760	APC016	33.4	0.1	1.7	33.8	0.1	1.8

C Appendix for Chapter 5: ESSA

Below is a summary table of the pit and skylight detections made by the ESSA model (described in Chapter 5) within eight different Regions of Interest (RoIs) on the Moon. The table gives the RoI and LROC NAC images within which each feature was detected, as well as the corresponding classification (1: skylight and 2: pit) and confidence score. The latitude-longitude coordinates (in a 180° domain) for each detection have also been given, along with a brief description of the feature. This table does not contain any of the Lunar rilles shown in Figure 5.10, as this mapping is available on Zenodo as an ESRI shapefile (Le Corre, 2025b).

RoI (Name)	Lat, Lon [°]	Class	Score [%]	Image ID(s)	Description
Antoniadi	-69.3310, -172.9860	2	99.9	M1286001151R, M107994247L	Elongated
Antoniadi	-69.3610, -170.5190	2	80.8	M174042488R	Elongated
Antoniadi	-68.9480, -174.9590	2	99.8	M1265918336L	Pit
Bel'kovich A (BAP)	58.6974, 87.5999	1	90.2	M1243133690L	Skylight
Endymion	55.0756, 56.3904	2	95.6	M1146806616R	Elongated
Mare Humboldtianum	57.0034, 76.4213	2	96.2	M1406405110R	Elongated
Mare Humboldtianum	56.6937, 76.5045	2	98.2	M170042063R, M1149033839R, M1210254180R	Elongated
Mare Humboldtianum	56.5497, 76.2684	2	99.1	M1266722478R	Elongated
Mare Humboldtianum	56.4141, 76.8732	2	98.9	M1399363573L	Elongated
Mare Humboldtianum	56.1603, 77.3487	2	97.4	M1223119197L, M1327837599L	Elongated
Marius Hills (MHP)	14.0898, -56.7705	1	100.0	M1341583189L, M183904057R	Skylight
Marius Hills (SMHP)	13.0184, -56.9323	1	98.5	M1132212699L, M181545193R, M183904057R	Skylight
Marius Hills (WMHP)	13.5504, -58.1725	1	97.8	M1346290850R	Skylight
Plato	51.8214, -11.2123	2	96.5	M1384775946R	Elongated
Poincaré	-56.5120, 160.6730	2	86.9	M1429422566R	Rille section
Poincaré	-58.4780, 162.8877	2	99.8	M1314350129R	Elongated

D Presentations and Publications

Below is a chronological list (most recent first) of all of the presentations (oral and posters) that were given upon the work in this thesis, as well as any papers published or due to be submitted for review. The front pages of the manuscripts for these published and draft research articles are also appended to the end of this thesis.

Title	Journal/Conference	Date
Automated Impact Melt Fracture Mapping on the Moon with Weakly-Supervised Deep Learning (Le Corre et al., 2025b)	JGR Planets	11/25
New Candidate Cave Entrances on the Moon Found Using Deep Learning (Le Corre et al., 2025a)	Icarus	06/25
Searching for Intact Lunar Lava Tubes with the Help of Deep Learning	BPSC (Poster)	06/24
Searching for Intact Lunar Lava Tubes with Deep Learning	UKPF EC (Oral)	01/24
Automatically calculating the apparent depths of pits using the Pit Topography from Shadows (PITS) tool (Le Corre et al., 2023a)	RASTI	08/23
Pit Topography from Shadows (PITS) tool	CryoMars (Poster)	06/23
Automatically Calculating Depths of Martian and Lunar Pits with Satellite Imagery (Le Corre et al., 2023c)	EGU (Oral)	04/23
Planetary Surface Feature Detection Using Machine Learning (Le Corre et al., 2022b)	EPSC (Oral)	09/22
Martian Pit Shadow Extractor (MAPS): Determining the Apparent Depths of Martian Pits from the Morphology of their Shadows (Le Corre et al., 2022a)	ICG (Oral)	09/22
Planetary Surface Feature Detection Using Machine Learning	RAS NAM (Poster)	07/22
Determining the Apparent Depths of Martian Pits Using Machine Learning	EAS (Poster)	06/22

E Datasets and Feature Catalogues

Below is a summary of the main planetary remote-sensing imagery datasets and planetary surface feature catalogues that were used within this thesis. This is not an exhaustive list of all imagery datasets and catalogues that were discussed in this thesis, but is instead those which were primarily used to bring about the results of this work. MRO/LRO = Mars/Lunar Reconnaissance Orbiter.

Name	Description	Reference
Planetary Remote-Sensing Imagery Datasets		
MRO High Resolution Imaging Science Experiment (HiRISE)	0.25–0.5 m/px images, with red and often-times colour versions	McEwen (2007)
LRO Narrow Angle Camera (LROC NAC)	Panchromatic images at 0.5–1.0 m/px during circular orbit, now 0.4–2.0 m/px	Robinson (2009)
LRO Wide Angle Camera (LROC WAC)	100 m/px global morphological mosaic of the Moon's surface	Speyerer et al. (2011)
Planetary Surface Feature Catalogues		
Lunar Pit Atlas (LPA)	281 Lunar pits: 256 found in impact melt (8 of which were labelled as 'fractures'), 16 in mare, 5 in highlands	Wagner and Robinson (2021)
Mars Global Cave Candidate Catalog (MGC^3)	1,062 possible cave entrances on Mars, with 132 atypical pit craters (APCs) and 354 skylights	Cushing (2015)

RAS Techniques and Instruments

RASTAI **2,** 492–509 (2023) https://doi.org/10.1093/rasti/rzad037
Advance Access publication 2023 August 11

Automatically calculating the apparent depths of pits using the Pit Topography from Shadows (PITS) tool

Daniel Le Corre [1,2]★ David Mary,[3] Nigel Mason,[1] Jeronimo Bernard-Salas[2] and Nick Cox[2]

[1]*Centre for Astrophysics and Planetary Science, University of Kent, Giles Lane, Canterbury CT2 7NH, UK*
[2]*Centre d'Etudes et de Recherche de Grasse, ACRI-ST, Av. Nicolas Copernic, F-06130 Grasse, France*
[3]*Lagrange UMR 7293, Université Côte d'Azur, Observatoire de la Côte d'Azur, Bd de l'Observatoire CS 34229, F-06304 Nice, France*

Accepted 2023 July 28. Received 2023 July 12; in original form 2023 January 29

ABSTRACT

Pits, or pit craters, are near-circular depressions found in planetary surfaces, which are generally formed through gravitational collapse. Pits will be primary targets for future space exploration and habitability for their presence on most rocky Solar System surfaces and their potential to be entrances to sub-surface cavities. This is particularly true on Mars, where caves have been simulated to harbour stable reserves of ice water across much of the surface. Caves can also provide natural shelter from the high radiation dosages experienced at the surface. Since pits are rarely found to have corresponding high-resolution elevation data, tools are required for approximating their depths in order to find those which are the ideal candidates for follow-up remote investigation and future exploration. The Pit Topography from Shadows (PITS) tool has been developed to automatically calculate the apparent depth of a pit (h) by measuring the width of its shadow as it appears in satellite imagery. The tool requires just one cropped single- or multiband image of a pit to calculate a profile of h along the length of the shadow, thus allowing for depth calculation where altimetry or stereo image data is not available. We also present a method for correcting shadow width measurements made in non-nadir observations for all possible values of emission and solar/satellite azimuth angles. Shadows are extracted using image segmentation in the form of k-means clustering and silhouette analysis. Across 19 shadow-labelled Mars Reconnaissance Orbiter red-band HiRISE images of atypical pit craters (APCs) from the Mars global cave candidate catalogue (MGC[3]), PITS detected 99.6 per cent of all shadow pixels (with 94.8 per cent of all detections being true shadow pixels). Following this testing, PITS has been applied to 123 red-band HiRISE images containing 88 APCs, which revealed an improvement in the variation of the calculated h due to emission angle correction, and also found 10 APCs that could be good candidates for cave entrances on Mars due to their h profiles.

Key words: Machine Learning – Algorithms – Mars – Planetary surfaces – Pits – Cave entrances.

1 INTRODUCTION

Pits are circular-to-elliptical depressions on rocky planetary surfaces which are generally formed by the gravitational collapse of surface material (Ferrill, Hargitai & Kereszturi 2015). Pits can often be misconstrued as impact craters due to their circular morphology and negative elevation relative to the surrounding surface. None the less, pits are most distinguishable from impact craters by the clear lack of a raised rim or any ejecta rays – formed by the impact ejecta settling on the surface. Pits can also range from being bowl-shaped – similar to the interior of impact craters – to being either more regularly conical or cylindrical in shape (Ferrill et al. 2015).

There are several proposed mechanisms for how pits are formed on planetary surfaces. 'Pit crater' is a commonly used term for defining a pit that has been created through volcanic, tectonic, or faulting-related processes (van der Bogert, Ashley & Ferrill 2014). Pits can also be formed due to the loss of subsurface volatiles through chemical dissolution, impact melting, or sublimation (Ferrill et al. 2015). Different formation processes, as well as the gravitational attraction of the host planet/moon, can lead to pits of a range of diameters and volumes, though an extraterrestrial pit exceeding 10 km in diameter will instead be known as a 'caldera' (Mouginis-Mark & Rowland 2001). Many of these mechanisms can also result in several pits occurring in linear or sinuous series, often called 'pit chains' (Hagen 2014). For example, a survey of Martian pit chains with diameters greater than 930 m suggested that dilational normal faulting was likely to be their most prominent formation pathway (Wyrick et al. 2004). This is where normal faults are refracted as they pass through unconsolidated layers, causing dilation under the surface (Ferrill & Morris 2003) through which surface material may drain (see Wyrick et al. 2004, fig. 2). By comparing with known terrestrial examples, Lunar and Martian pit chains that indicate the presence of partially intact lava tubes have also been identified (Sauro et al. 2020). Lava tubes are underground conduits that currently or formerly held flowing or solid lava (Gadányi & van der Bogert 2014), which can form pit chains as their roofs collapse. Collapses along a lava tube are commonly called 'skylights' when they appear to have lateral entrances into the conduit (Gadányi & van der Bogert 2014).

★ E-mail: dl387@kent.ac.uk

Icarus 441 (2025) 116675

Contents lists available at ScienceDirect

Icarus

journal homepage: www.elsevier.com/locate/icarus

Research Paper

New candidate cave entrances on the Moon found using deep learning

Daniel Le Corre [a,b,*], Nigel Mason [a], Jeronimo Bernard-Salas [b], David Mary [c], Nick Cox [b]

[a] Centre for Astrophysics and Planetary Science (CAPS), University of Kent, Park Wood Road, Canterbury, CT2 7NH, Kent, United Kingdom
[b] Centre d'Etudes et de Recherche de Grasse (CERGA), ACRI-ST, Avenue Nicolas Copernic, Grasse, 06130, Provence-Alpes-Côte-d'Azur, France
[c] Lagrange UMR 7293, Université Côte d'Azur, Observatoire de la Côte d'Azur, Boulevard de l'Observatoire CS 34229, Nice, F-06304, Provence-Alpes-Côte-d'Azur, France

ARTICLE INFO

Keywords:
The Moon
Mars
Planetary surfaces
Remote sensing
Deep learning

ABSTRACT

Pits and skylights are circular to elliptical, rimless, steep-sided depressions on planetary surfaces formed through gravitational collapse, which are of interest for astrobiological investigation and future space exploration. This is due to their ability to signify the presence of, or allow access to, underground cave systems such as lava tubes. The Lunar Pit Atlas contains 16 such features situated within mare regions that were partly discovered via the automated PitScan tool, which was limited by searchable latitudes and data coverage. In order to search for pits and skylights within these unmapped regions, we have trained a series of Mask R-CNN (Region-based Convolutional Neural Network) models on various combinations of Lunar and Martian remote-sensing imagery to detect Lunar pits and skylights. The best-performing model, named ESSA (Entrances to Sub-Surface Areas), was trained upon all available training data with a ResNet50 backbone. During testing on imagery of the famous Mare Tranquillitatis Pit and self-produced mosaics of proposed lava tube collapses, ESSA achieved average F_1-scores of 82.4 and 93.7% for the bounding boxes and predicted masks, respectively. Despite only having surveyed ≈1.92% of the Lunar maria, ESSA has detected two previously uncatalogued skylights: the South Marius Hills and Bel'kovich A Pits (SMHP and BAP) - which are possible candidates for cave entrances on the Moon.

1. Introduction

Pits (also known as pit craters, collapse pits, or sinkholes) are circular-to-elliptical depressions on the surface of rocky planetary bodies that are generally formed through gravitational collapse (Ferrill, 2015). Unlike impact craters, which are formed through the collision between a planetary surface and a meteorite or ejecta material, pits will lack a raised rim or any visible ejecta rays. Pits can have a range of topographies and sizes, from having steep walls and relatively flat floors (Cushing et al., 2015; Wagner and Robinson, 2022) to being more conical or bowl-shaped (Wyrick et al., 2004) and from being on the order of tens of metres to several kilometres in diameter. Several genetic mechanisms have been suggested for the formation of pits, such as the subsidence of an underground evacuated magma chamber (Wagner and Robinson, 2022), sublimation of sub-surface volatiles (Spencer and Fanale, 1990), or dilational normal faulting (Wyrick et al., 2004). These mechanisms can also lead to pits occurring in linear or sinuous alignment — called pit chains (Hagen, 2015). Pits and pit chains can be found on Earth, such as those surrounding the Kilauea volcano site on the island of Hawai'i (Okubo and Martel, 1998), but have also been observed on the Moon (Greeley and Gault, 1979), Mars (Wyrick et al., 2004), Mercury (Gillis-Davis et al., 2009), Venus (Davey et al., 2013) and Phobos (Horstman and Melosh, 1989).

Pits are of particular focus for their potential to provide access to underground cavities such as lava tubes. Lava tubes are sub-surface conduits found sub-parallel to terrestrial surfaces that have been formed by the current or historic flowing of lava (Gadányi and van der Bogert, 2015). Intact lava tube systems can reveal themselves to the surface when their roofs become unstable and collapse to form circular, coalesced or elongated pits that can appear along the path of the tube. These pits are known as 'skylights' in cases where they exhibit steep walls and floors which are flat or sloping towards what may be a roof overhanging an underground cavity (van der Bogert and Ashley, 2015). Infrared images taken by Mars Odyssey's Thermal Emission Imaging System (THEMIS) have suggested that the thermal properties of seven Martian skylights in nighttime images appear to resemble what would be expected of sub-surface materials (Cushing et al., 2007). More recently, radar imagery taken by the Lunar Reconnaissance Orbiter (LRO) MiniRF instrument of the renowned Mare Tranquillitatis Pit

* Corresponding author at: Centre for Astrophysics and Planetary Science (CAPS), University of Kent, Park Wood Road, Canterbury, CT2 7NH, Kent, United Kingdom.
E-mail address: dl387@kent.ac.uk (D. Le Corre).

https://doi.org/10.1016/j.icarus.2025.116675
Received 11 April 2025; Received in revised form 29 May 2025; Accepted 29 May 2025
Available online 13 June 2025

JGR Planets

RESEARCH ARTICLE

10.1029/2025JE009145

Automated Impact Melt Fracture Mapping on the Moon With Weakly Supervised Deep Learning

Daniel Le Corre[1,2], Nigel Mason[1], Jeronimo Bernard-Salas[2], David Mary[3], and Nick L. J. Cox[2]

[1]Centre for Astrophysics and Planetary Science (CAPS), University of Kent, Canterbury, UK, [2]Centre d'Etudes et de Recherche de Grasse (CERGA), ACRI-ST, Grasse, France, [3]Lagrange UMR 7293, Observatoire de la Côte d'Azur (OCA), Université Côte d'Azur, Nice, France

Key Points:

- IMFMapper is a DeepLabV3 deep learning model trained to detect impact melt fractures (IMFs) within LROC NAC imagery of the Moon
- We propose new candidate melt ponds within the western and southern walls of Crookes crater, thanks to the detection of IMFs
- In Schomberger A crater, we present the first mapping via any means of IMFs so close to the Lunar South Pole

Correspondence to:

D. Le Corre,
dl387@kent.ac.uk

Citation:

Le Corre, D., Mason, N., Bernard-Salas, J., Mary, D., & Cox, N. L. J. (2025). Automated Impact melt fracture mapping on the Moon with weakly-supervised Deep Learning. *Journal of Geophysical Research: Planets*, *130*, e2025JE009145. https://doi.org/10.1029/2025JE009145

Received 13 APR 2025
Accepted 13 NOV 2025

Author Contributions:

Conceptualization: Daniel Le Corre
Data curation: Daniel Le Corre
Formal analysis: Daniel Le Corre
Funding acquisition: Nigel Mason, Jeronimo Bernard-Salas
Investigation: Daniel Le Corre
Methodology: Daniel Le Corre
Project administration: Daniel Le Corre, Nigel Mason, Jeronimo Bernard-Salas
Resources: Daniel Le Corre
Software: Daniel Le Corre
Supervision: Nigel Mason, Jeronimo Bernard-Salas, David Mary
Validation: Daniel Le Corre
Visualization: Daniel Le Corre
Writing – original draft: Daniel Le Corre
Writing – review & editing: Daniel Le Corre, Nigel Mason, Jeronimo Bernard-Salas, David Mary

Abstract Cooling fractures found within impact melt deposits have been manually mapped within several craters on the Moon and Mercury, as their distribution can indicate which heat-loss processes were most significant in the periods after impact. However, due to the discovery of melt deposits in Lunar impact craters with sub-km diameters, it is unlikely that the complete mapping of these impact melt fractures (IMFs) on the Moon will be achievable without automation. As such, we have trained a DeepLabV3 semantic segmentation deep convolutional neural network, called IMFMapper, to detect IMFs within Lunar Reconnaissance Orbiter Narrow Angle Camera (LROC NAC) satellite imagery. As a means of maximizing the size of the training data set, "weak" pixel-level labels were generated by buffering line annotations. In testing upon the IMFs found within Ohm crater, IMFMapper achieved an average F_1-score of 69.3%. IMFMapper has also been deployed to map IMFs within the previously surveyed Crookes crater, where we have found new candidate melt deposits within the crater's western and southern walls. In addition, IMFMapper has produced the first map of IMFs within Schomberger A crater, in which IMFs may act as permanently shadowed regions due to the crater's proximity to the Lunar South Pole. The successful mapping of IMFs in Schomberger A also signifies IMFMapper's robustness to extreme solar incidence angles. We also demonstrate that IMFMapper could be implemented for automated mapping of IMFs on Mercury upon the commencement of BepiColombo's science operations.

Plain Language Summary After a meteorite hits a planetary body at hyper-velocity speeds, the surface material is melted into molten rock under the immense energies of an impact event. As the crater forms underneath, this pool of molten rock–often called an impact melt deposit - contracts as it cools over time. This contraction causes linear fractures to pop open on the surface of the impact melt deposit. Previously, these impact melt fractures (IMFs) have been manually mapped by researchers on the Moon and Mercury in order to investigate which cooling mechanisms may be most influential on airless bodies such as these. In this work, we use a deep learning model to automate this process. Our model, called IMFMapper, builds upon the previous manual mapping of Crookes crater on the Moon and finds new impact melt deposits within its western and southern walls. We also use IMFMapper to catalog the IMFs found within the Moon's Schomberger A crater, which could potentially trap reserves of water ice within them due to being deprived of sunlight at such high latitudes.

1. Introduction

Impact melt fractures (herein referred to as IMFs) are cracks found in the melt deposits of impact craters, which can often occur in curvilinear or polygonal complexes. Figure 1 presents several patterns, scales and morphologies of IMFs as found on the Moon. The cracks are thought to form due to extensional stresses as the deposited shock-heated material (either found in the crater's floor or ejected from the site) cools and contracts after the event of an impact (Xiao et al., 2014). The presence of solid debris in a melt pond, such as material falling from the crater's walls, will affect the uniformity of the stresses experienced by the melt (Denevi et al., 2012) - and, by proxy, the patterns of IMFs observed. Due to the conditions under which these processes occur being similar on the two bodies, fractures relating to the cooling of impact melt deposits have been observed on both the Moon and Mercury (Xiao et al., 2014).

Impact melt deposits (also known as melt ponds) and the IMFs found within them have been the subject of several mapping exercises. Xiao et al. (2014) have carried out surveys of IMFs within the young impact melt deposits