

# Lunar Landing Site Selection using Machine Learning

Daison Darlan

*Department of Artificial Intelligence  
Kyungpook National University,  
Daegu, Republic of Korea  
daisondarlan33@gmail.com*

Oladayo S. Ajani

*Department of Artificial Intelligence  
Kyungpook National University,  
Daegu, Republic of Korea  
oladayosolomon@gmail.com*

Rammohan Mallipeddi

*Department of Artificial Intelligence  
Kyungpook National University,  
Daegu, Republic of Korea  
mallipeddi.ram@gmail.com*

**Abstract**—The value of a planetary landing mission is critically dependent on the choice of landing sites which are generally influenced by several scientific and engineering constraints. The conventional method for Landing Site (LS) selection requires the manual analyses of individual sites involving a large talent force of multi-domain experts hence rendering the process cumbersome and expensive. Simultaneously, the currently employed methodology of recycling previously selected LSs for future missions would not be realized in the case of hitherto unexplored planets. As a consequence, machine learning algorithms are being actively explored in aiding effective space exploration. However, their application to selection of LSs that satisfy the scientific and engineering requirements of a mission have not yet been explored. In this paper, we propose an end-to-end Landing Site selection methodology using Moon as a case study and employing Hierarchical clustering of regions based on their altitude, expandable to other planets. Furthermore, we enforce commonly used constraints on the potential sites and select final sites for landing based on the user provided scientific and engineering constraints. With this approach, the LS selection process is simplified and the temporal requirement is reduced.

**Index Terms**—Landing Site Selection, Clustering.

## I. INTRODUCTION

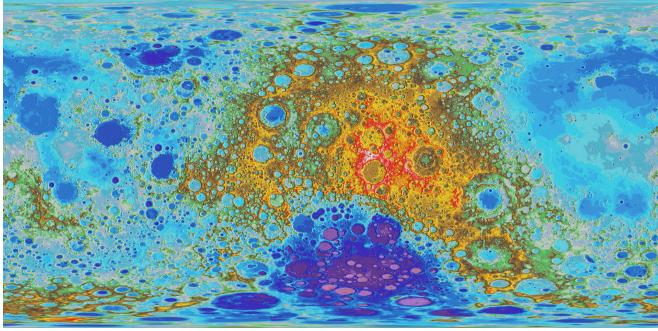
Since the beginning of the 21st century, space agencies of world leading nations have carried out several Lunar exploration missions paving the way for scientific discovery and innovations that extend beyond the domain of space exploration. These missions can be widely classified as being in-situ and ex-situ. In-situ exploration refers to extraction and analysis of local resources on the surface of the moon with the help of a Lunar Lander [1]. The acquisition of planetary surface and topographical data using Orbiters falls under the scope of ex-situ exploration [2]. Some of the recent ex-situ missions include those from NASA (ARTEMIS, Lunar Reconnaissance Orbiter, LCROSS, GRAIL,CAPSTONE), European Space Agency (SMART-1), ISRO (Chandrayaan-1, Chandrayaan-2), CNSA (CE-1,CE-2,CE-3,CE-4) [3] and KARI (Danuri) [4]. The successful Lunar landing missions in comparison to Orbiter missions are far fewer in number and are undertaken majorly by CNSA, NASA and NPO Lavochkin [5], [6]. The success of these missions were heavily influenced by the site selected for landing. Therefore, research related to developing the next generation of landing sites has seen a great deal of attention lately [3].

There are several factors that influence the selection of the landing site and can be divided into those arising from scientific or engineering objectives. Some of the constraints used for current and previous Moon and Mars LS selection include i) Scientific Constraints: Geological Characteristics, In-Situ Resources, Space Environment Characteristics, Special Regions and ii) Engineering constraints: topographic slope, terrain obstacles, communication ability, temperature, latitude, elevation, ellipse dimension [3], [7].

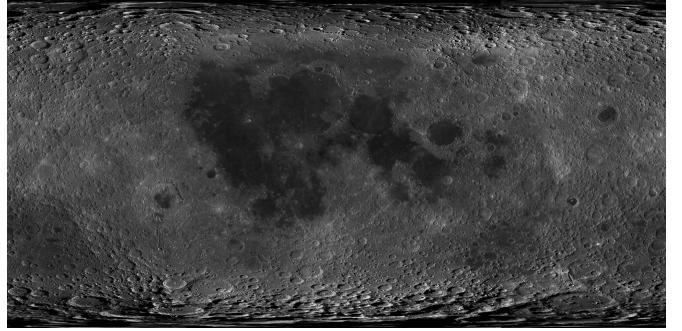
The recent increment of interest in space exploration has motivated researchers to explore efficient and economical machine learning (ML) algorithms that aid activities for space exploration [8]. Although several works have explored ML algorithms for landing site selection, they are limited to atmospheric entry or powered descent [9] stages that occur from the parking orbit around the planet. This indicates the absence of investigation into the use of ML for selecting LSs in the preliminary stages of mission planning.

The traditional method of selecting a set of LSs involves manual analysis of each individual LS by a team of multi-domain experts and assigning a score of feasibility to all potential sites [10]. For example, the selection of Gale crater as the LS for the Curiosity rover was a result of global scientific effort, that spanned a period of 5 years [11]. However, this process is extremely time consuming and requires a large talent force [12]. Furthermore, the subsequent missions recycle the previous regions that were considered potentially safe due to the cumbersome task of re-selecting regions that satisfy the requirements of scientific exploration and safety constraints. While this methodology is being used for currently proposed missions on Moon and Mars, the expansion of space exploration to hitherto unexplored planets would render it ineffective and expensive.

To address these limitations, we propose an end-to-end Landing Site selection methodology choosing the Moon as a case study by employing Hierarchical clustering of regions based on their altitude. Furthermore, we enforce commonly used constraints on the potential sites and select final sites for landing based on the user provided scientific and engineering constraints. With this approach, the LS selection process is simplified and the time complexity is reduced. Additionally, the proposed framework is expandable to include other infor-



(a) Global Topographic map.



(b) Global Morphology Mosaic

Fig. 1: Maps used as Datasets

mation such as mineral concentration in the form of global morphological maps. The incorporation of maps will allow the enforcing of additional mission or safety constraints, hence improving the quality of the site selected by the algorithm.

The rest of the manuscript is arranged as follows. Section II discusses the characteristics of the real-world data and provides a brief discourse on the data preparation steps undertaken on the acquired datasets. Section III goes into details of the clustering algorithm followed by discussion of results in Section IV. The final section V concludes the paper and proposes future scope of the research.

## II. DATA ACQUISITION AND PREPROCESSING

### A. Data Acquisition

In this paper, two global maps are generated using data captured by a Wide Angle Camera (WAC) aboard the Lunar Reconnaissance Orbiter (LRO) that are subjected to clustering and visual analysis. The two maps dubbed Auxiliary (Aux) Map and Base Map are Global Topographic map and Global Morphology Mosaic of the moon respectively as seen in Fig. 1. Clustering algorithm is applied on the Aux map containing encoded information regarding the topology of the Lunar surface. While, the main purpose of the Base map is to demonstrate the functionality of adding multiple maps containing encoded information as well as to provide an actual view of the surface. Specifically, the WAC is designed to generate global images at a scale of 100m/pixel using 7 narrow-band interference filters, 2 of which are in ultraviolet and the rest in visible. Further in depth information regarding the maps and the data acquisition modalities can be found in [8], [13], [14].

### B. Data Preprocessing

For the purpose of clustering, the two maps are discretized into sub-images with horizontal and vertical divisions that can be expressed as:

$$S = hdiv \times vdiv \quad (1)$$

where  $S$  is the total number of sub-images which the map is divided into, where  $hdiv$  and  $vdiv$  represent the number of horizontal and vertical divisions respectively. The values of  $hdiv$  and  $vdiv$  determine the shape and size of the individual cells. Consequently, the height and width of the maps are

extracted, which is then used to compute the coordinate centers of the sub-images that are to be generated. This is so that every sub-image that encompasses a region of the Lunar surface has a unique label with which it is referred. This step also provides the encoding of the latitudinal and longitudinal information of the region to the image which expedites further processing of the images.

To generate the sub-images, the coordinates of the nodes of the individual cells in the discretized grid is calculated. Figure 2 depicts a sample discretized cell obtained from the Aux map. Each cell is initially identified with a set of  $x$  and  $y$  coordinates that denote the top-left node and bottom-right nodes of the cell. This set of coordinates for the entire set of discretized cells can be represented as finite arithmetic series.

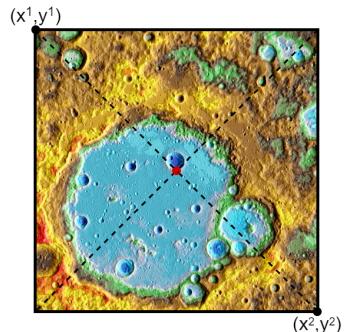


Fig. 2: A sample discretized region from the Aux Map. The region is centered at  $(142.5^\circ, 7.5^\circ)$  denoted by the red circle.

Let  $(x^1, y^1)$  and  $(x^2, y^2)$  be the Cartesian coordinates of the top-left node and bottom-right node of a cell respectively, such that

$$x^1, x^2 \in x \quad (2)$$

$$y^1, y^2 \in y \quad (3)$$

and

$$\{x | x \subseteq \mathbb{R}, x \leq w\} \quad (4)$$

$$\{y | y \subseteq \mathbb{R}, y \leq h\} \quad (5)$$

where,  $w$  and  $h$  are the height and width respectively that extracted from the image of the map.

$x^1, x^2, y^1$  and  $y^2$  can be formulated as an arithmetic series of the form

$$(a_n)_{n=0}^f = (a_0, a_1, a_2, \dots, a_f) \quad (6)$$

where,

$$(a_n)_{n=0}^m = (a_0, a_1, a_2, \dots, a_m) \quad (7)$$

$$: a \in \{x, y\} \quad (8)$$

Specifically,

$$a_k^1 = k \times \left( \frac{m}{hdiv} \right) \quad (9)$$

$$a_k^2 = (k+1) \times \left( \frac{m}{hdiv} \right) \quad (10)$$

where,

$$k \in [0, w], m = w | a \Leftrightarrow x \quad (11)$$

$$k \in [0, h], m = h | a \Leftrightarrow y \quad (12)$$

The coordinate system for the resulting cells is generated by iterating through the Eqs. 9 and 10. Furthermore, every cell is identified by the coordinates at the location of its geometric center. Once the entire map is discretized into sub-images, the resulting images are converted to an array of their pixel values and scaled according to the formula

$$z = (x - u)/s \quad (13)$$

where,  $u$  and  $s$  represent the mean and standard deviation respectively, of the array of pixel values  $x$ . Finally, the clustering algorithm is applied on the 1 dimensional array that is obtained after flattening out the 3 dimensional array containing information from the Red, Green and Blue channels.

### III. CLUSTERING

An Agglomerative Hierarchical Clustering (AHC) algorithm is employed to cluster the discretized sub-images. Specifically, a bottom-top approach is followed for identifying clusters in the data in which, all individual data points are considered as a separate cluster at the beginning of clustering operation. The AHC algorithm then proceeds to merge clusters together based on their similarity until only a single cluster is left. The formation of clusters is dependent on 2 factors: affinity metric used and linkage method [15]. Affinity is a metric that is used to compute the distance between two clusters. In this work, the Euclidean distance is used as the affinity due to its simplicity. On the other hand, the linkage method determines the proximity between two clusters. In this, the Minimal Increase of Sum Squares is used that computes proximity between two clusters as the difference between summed square of distance of the joint cluster and the summation of individual summed square of distance with each cluster. [16].

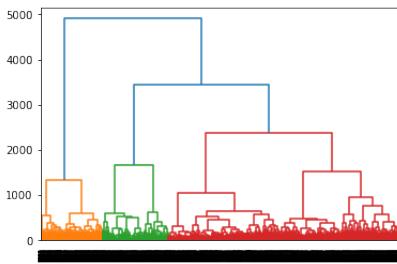


Fig. 3: Dendrogram generated for the data extracted from the Auxiliary Image.

The formation of clusters and their optimal number can be observed from the dendrogram depicted in Fig. 3. Additionally, the number of clusters can also be determined by plotting the Within Cluster Sum-of-Squares(WCSS) with the number of clusters also known as the Elbow method. The formula to calculate WCSS is given by

$$WCSS = \sum_{C_k}^{C_n} \left( \sum_{d_i \in C_k}^{d_m} \text{distance}(d_i, C_k)^2 \right) \quad (14)$$

where,  $C$  is centroid of a cluster and  $d$  is the data point present in each cluster. From these two analyses, the number of clusters was decided to be 4.

Subsequently, the array of pixel values of the discretized sub-images from the Aux map are clustered. The AHC algorithm attempts to create clusters based on the altitude of the region present in the image. Since it is the information about the altitude that is color encoded in the map and consequently reflected in the pixel values, the AHC algorithm attempts to generate clusters depending on the altitude of the region present in the image. Although the base map is discretized, it is not subjected to clustering. This is because the primary motivation for including the Base map is to provide a reference for the landing sites that is not obscured with any information. A trajectory constraint is then applied to narrow down the list of potential landing sites.

### IV. RESULTS

18432 image cells created by the 192 and 96 horizontal and vertical divisions respectively are used for the experiments in this work. The simulations are run on a NVIDIA GTX 1650 Ti GPU with a compute capability of 7.5 and a base clock speed of 1530 MHz. The algorithm is implemented in Python and leverages the predictive data analysis functionality.

Figure 4 (b) is a visual representation of the 4 clusters obtained where the sizes of the clusters are 7569, 3637, 3848 and 3378 respectively. The AHC algorithm groups the images based on their pixel values and which is evident in the generated clusters. The significance of this result is that the algorithm divides the Lunar surface based on the altitude of the topography. This information can be used during the decision making process for LS selection. Additionally, it is to be noted that for this study just the altitude of the Lunar surface is considered and the addition of maps will provide us with more accurate site selection that includes more factors. For example, if a map that contains the density of mineral content is added in conjecture to the already existing map, the different objectives of altitude and mineral density can be combined to provide new LSs.

The visualization of the clusters with images as data points can be seen in Fig. 4 (a). To generate this graph, Principal Component Analysis (PCA) is used to reduce the dimensionality of the data points to two dimensions. Here, the different clusters are assigned different colors to be differentiable from each other. Finally, a trajectory constraint is applied to the cluster LSs. This constraint is such that the sites that exist within 10 Kilometers of the 90° latitudinal line are deemed

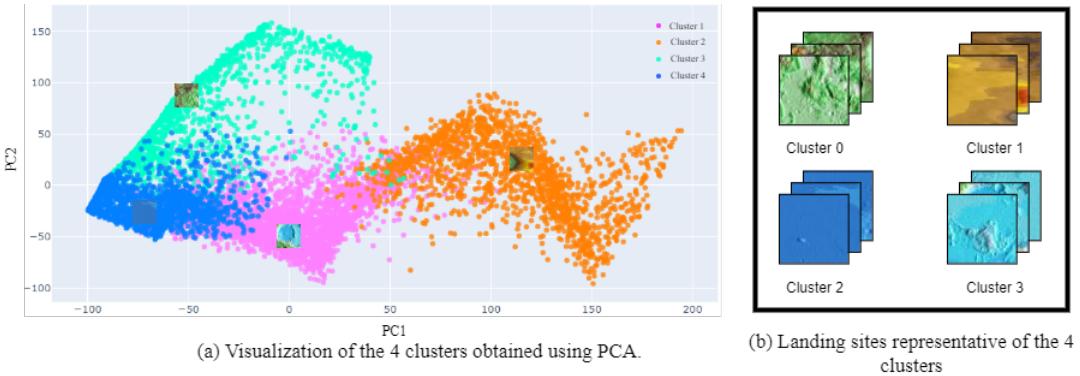


Fig. 4: Analysis of generated clusters.

to be suitable for landing. The application of the constraint selects a total of 104 landing sites that are marked feasible of which 68, 3, 16 and 17 belong to the clusters 0, 1, 2, 3 and 4 respectively. To elaborate further, cluster 0 that contains shallow valleys accounts for over 41 % of the total image cells which is reflected in the number of selected sites. Cluster 1 that contains elevated regions such as mountains are sparsely located especially in the regions specified by the constraint and hence only 3 viable sites are found. Furthermore, clusters 2 and 3 contain large and medium sized craters respectively that are widely interspersed on the Lunar surface. Consequently, depending upon the mission requirements a site from one of the 4 clusters can be chosen as the final site of landing.

## V. CONCLUSION

In this paper, an end-to-end clustering based framework for Lunar Landing Site Selection is proposed. Two maps obtained from NASA's Lunar Reconnaissance Orbiter is used in which one map contains the topographical information encoded in it while the other is a Mosaic map of the surface. Agglomerative Hierarchical Clustering is used to obtain different regions in the map based on the altitude. With this methodology, four clusters of landing sites are obtained. Subsequently, a trajectory constraint is applied to narrow the list of potential landing sites.

The framework proposed is exoplanet-agnostic. This means that the algorithm can be applied for any planet or moon as the decision making process depends on the Maps provided and the user and the scientific and engineering constraints. The methodology adopted is robust to expansion by including additional maps hence providing specific LSs tailored to the mission requirements.

## ACKNOWLEDGEMENTS

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1I1A3049810).

## REFERENCES

- [1] K. Sacksteder and G. Sanders, "In-situ resource utilization for lunar and mars exploration," in *45th AIAA Aerospace Sciences Meeting and Exhibit*, 2007, p. 345.
- [2] D. E. Smith, M. T. Zuber, G. B. Jackson, J. F. Cavanaugh, G. A. Neumann, H. Riris, X. Sun, R. S. Zellar, C. Coltharp, J. Connally *et al.*, "The lunar orbiter laser altimeter investigation on the lunar reconnaissance orbiter mission," *Space science reviews*, vol. 150, no. 1, pp. 209–241, 2010.
- [3] J. Liu, X. Zeng, C. Li, X. Ren, W. Yan, X. Tan, X. Zhang, W. Chen, W. Zuo, Y. Liu *et al.*, "Landing site selection and overview of china's lunar landing missions," *Space Science Reviews*, vol. 217, no. 1, pp. 1–25, 2021.
- [4] K. Kim, K. W. Min, R. Elphic, R. Choi, N. Hasebe, H. Nagaoka, J. Park, E. S. Yi, S. Lee, Y.-K. Yeon *et al.*, "Introduction to the lunar gamma-ray spectrometer for korea pathfinder lunar orbiter," *42nd COSPAR Scientific Assembly*, vol. 42, pp. B3–1, 2018.
- [5] A. L. Turkevich, "Comparison of the analytical results from the surveyor, apollo, and luna missions," in *Lunar and Planetary Science Conference Proceedings*, vol. 2, 1971, p. 1209.
- [6] Y. Zheng, Z. Ouyang, C. Li, J. Liu, and Y. Zou, "China's lunar exploration program: present and future," *Planetary and Space Science*, vol. 56, no. 7, pp. 881–886, 2008.
- [7] M. Golombek, D. Kipp, N. Warner, I. J. Daubar, R. Fergason, R. L. Kirk, R. Beyer, A. Huertas, S. Piqueux, N. Putzig *et al.*, "Selection of the insight landing site," *Space Science Reviews*, vol. 211, no. 1, pp. 5–95, 2017.
- [8] S. Chen, Y. Li, T. Zhang, X. Zhu, S. Sun, and X. Gao, "Lunar features detection for energy discovery via deep learning," *Applied Energy*, vol. 296, p. 117085, 2021.
- [9] K. Iiyama, K. Tomita, B. A. Jagatia, T. Nakagawa, and K. Ho, "Deep reinforcement learning for safe landing site selection with concurrent consideration of divert maneuvers," *arXiv preprint arXiv:2102.12432*, 2021.
- [10] A. Binder and D. Roberts, "Criteria for lunar site selection," Tech. Rep., 1970.
- [11] NASA. Landing site selection.
- [12] M. Golombek, J. Grant, T. Parker, D. Kass, J. Crisp, S. Squyres, M. Carr, M. Adler, R. Zurek, A. Haldemann *et al.*, "Selection of the final four landing sites for the mars exploration rovers," in *Lunar and Planetary Science Conference*, 2003, p. 1754.
- [13] "Iro camera team releases high resolution global topographic map of moon'2011."
- [14] E. Speyerer, M. Robinson, B. Denevi *et al.*, "Lunar reconnaissance orbiter camera global morphological map of the moon," in *42nd Annual Lunar and Planetary Science Conference*, no. 1608, 2011, p. 2387.
- [15] F. Nielsen, "Hierarchical clustering," in *Introduction to HPC with MPI for Data Science*. Springer, 2016, pp. 195–211.
- [16] J. H. Ward Jr, "Hierarchical grouping to optimize an objective function," *Journal of the American statistical association*, vol. 58, no. 301, pp. 236–244, 1963.