The following publication Y. Wang and B. Wu, "Active Machine Learning Approach for Crater Detection From Planetary Imagery and Digital Elevation Models," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 8, pp. 5777-5789, Aug. 2019 is available at https://doi.org/10.1109/TGRS.2019.2902198.

Active Machine Learning Approach for Crater Detection from Planetary Imagery and Digital Elevation Models

Yiran Wang, Bo Wu*

***Abstract*— Craters are dominant geomorphological features on the surfaces of the Moon, Mars, and other planets. The distribution of craters provides valuable information on the planetary surface geology. Machine learning is a widely used approach to detect craters on planetary surface data. A critical step in machine learning is the determination of training samples. In previous studies, the training samples were mainly selected manually, which usually leads to insufficient numbers due to the high cost and unfavorable quality. Surface imagery and digital elevation models (DEMs) are now commonly available for planetary surfaces; this offers new opportunities for crater detection with better performance. This paper presents a novel active machine learning approach, in which the imagery and DEMs covering the same region are used for collecting training samples with more automation and better performance. In the training process, the approach actively asks for annotations for the 2D features derived from imagery with inputs from 3D features derived from the DEMs. Thus, the training pool can be updated accordingly, and the model can be retrained. This process can be conducted several times to obtain training samples in sufficient number and of favorable quality, from which a classifier with better performance can be generated, and it can then be used for automatic crater detection in other regions. The proposed approach highlights two advantages: 1) automatical generation of a large number of high-quality training samples, and 2) prioritization of training samples near the classification boundary so as to learn more quickly. Two sets of test data on the Moon and Mars were used for experimental validation. The performance of the proposed approach was superior to that of a regular machine learning method.**

***Index Terms*—Craters; Machine Learning; Imagery; Moon; Mars**

I. Introduction

CRATERS on planetary surfaces are among the most studied geomorphological features in planetary science. Crater distribution provides fundamental insights into geological processes and is used to reveal the ages of various geologic formations on a celestial body [1], [2]. The spatial distribution of craters also serves as an important consideration in planetary exploration missions when selecting candidate landing sites [3], [4].

Craters are normally detected from planetary surface images and/or digital elevation models (DEMs) with manual digitization [2] or computer-assisted methods. Manual detection is time-consuming and difficult to deal with large datasets, so numerous attempts have been made to create reliable methods for automatic crater detection. Salamunićcar et al. [5] tabulated 77 crater detection approaches from the literature and categorized them into three main types: DEM-based methods, image-based methods, and integrated methods based on both images and DEMs. DEM-based methods are based mainly on the analysis of slopes, horizontal and vertical structures, and the relationships among them. Techniques such as slope measurement [6]-[9], template matching [10], Hough transform [11], and machine learning [12] have been used for crater detection with DEMs. DEMs contain 3D information, and the morphological information of craters can be completely extracted. However, because DEMs usually have a relatively lower resolution and fewer DEM data are available than images, crater detection based on images has been more popular. Image-based crater detection approaches can be divided into rule-based methods and machine learning-based methods. Rule-based methods usually set a number of rules by experts and rely on pattern-recognition techniques to identify crater rims in images as circular or elliptical features. The original image is usually preprocessed to enhance the edges of the craters, and crater detection is achieved with methods such as Hough transform [13], [14], ellipse fitting [15], genetic algorithms [16], Gist features [17], watershed transform [18], pattern recognition [19], radial-consistency algorithms [20], or combinations of the above methods.

With the popularity in the field of computer vision, machine learning has been widely adopted in crater-detection tasks [21]-[24] and other similar feature detection tasks from remote sensing data [25]-[27]. Due to the variety of crater structures, machine learning based methods usually show more robust performance than the rule-based methods. The machine learning methods use samples to train a classifier to detect craters. Instead of relying on an expert's domain knowledge, machine learning methods learn the optimal filters and features based on a great number of training examples. In the learning phase, crater areas and non-crater areas are fed into a model to

Manuscript received March 31, 2018. The work described in this paper was funded by a grant from the National Natural Science Foundation of China (Project No.: 41671426) and grants from the Hong Kong Polytechnic University (Project No.: G-YBN8 and Project No.: 4-BCE5).

The authors are with the Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. (e-mail: yiran.wang@connect.polyu.hk; bo.wu@polyu.edu.hk)

form a classifier. In the detection phase, the previously trained classifier detects craters in a new set of images. The machine learning methods for crater detection are mainly based on strategies such as the scalable template model [28], [29], convolutional neural networks [24], boosting and modified boosting methods [22], [23], [30], and the least-squares support vector machine [21].

One of the most challenging tasks for machine learning is to obtain good labeled data. Training of a good model can require thousands of examples or even more, which results in huge and often prohibitive costs for many machine learning methods. The acquisition of training samples, which must be labeled by domain experts, is tedious and time consuming, so the number of training samples has been limited in most previous machine learning methods. For example, Jin and Zhang [22] used only 400 crater candidates and 400 non-crater candidates to train their modified AdaBoost-based model. Li et al. [31] tested the performance of their classifier with a training dataset of 200 to 1000. The detection accuracy improved noticeably as the number of training samples increased. However, the experiment stopped at 1000 training samples, and a larger dataset was not attempted. A limited number of training samples as in previous studies cannot cover the variety of crater features as they appear in the images due to variations in illumination, albedo, and surface roughness.

In addition to the number of training samples, their quality is also critical for successful model training. The manually selected training samples, which always focus on craters with distinct textures, are not a comprehensive representation of all possible cases. The lack of samples near the classification boundary will lead to an overfitting machine learning model. Martins et al. [30] noted that manual selection of samples is insufficient because it does not provide a representative collection of negative patterns. Thus, they enhanced their training dataset by adding false detections with a higher discriminant value. Their experimental analysis showed the effectiveness of the approach; however, this is again a manual process and requires great manual effort.

The idea of active learning [32] focused on the number and quality of training samples. The main content of active learning is that the machine learning method can actively query the domain experts for sample labels based on what has already been learned. By selecting the most targeted samples, active learning reduces the number of needed training samples and at the same time achieves better performances compared with the traditional machine learning methods. The concept of active learning was well-motivated in a variety of objection detection tasks [33], [34] to help collect training samples. However, in most of the traditional active learning methods, manual interaction is required for labeling the selected training samples. More automatic ways for generating high-quality training samples through active learning are in demand.

Crater detection based on the integration of images and DEMs may provide an opportunity to solve the problems mentioned above. Owing to the planetary exploration missions by different countries in recent years, high-resolution imagery and DEMs are now commonly available for planetary surfaces, such as the imagery and DEMs from the Chinese Chang'E-2 [35], Japanese SELENE [36], and NASA's Lunar Reconnaissance Orbiter [37] missions for the Moon, and the Mars Express [38] and Mars Reconnaissance Orbiter [39] missions for Mars. The DEMs can be generated from the stereo images via photogrammetric techniques or from other sources such as laser altimetry. This offers new opportunities for crater detection with better performance by allowing integration of 2D information from images and three-dimensional (3D) information from DEMs. Most previous studies of machine learning for crater detection have focused only on two-dimensional (2D) images. Few studies have reported the use of both images and DEMs for crater detection [40], [41]; these studies were conducted either by merely combining the detection results from the images and DEMs directly or by using the DEM data to confirm the results at the end. No existing study has used those two types of data in a more synergistic way for crater detection.

This paper presents a novel active machine learning approach in which both 2D information from imagery and 3D information from DEMs are used for more automatic collection of training samples. First, the training datasets including the imagery and the DEMs covering the same region are co-registered. Then in the training process the approach actively asks for annotations for the 2D features derived from imagery with inputs from the 3D features derived from the DEMs. Thus, the training pool can be updated accordingly, and the model can be retrained. This process can be conducted several times to obtain training samples in sufficient number and of favorable quality, from which a classifier with better performance can be obtained and used for automatic crater detection in other regions. Compared with previous studies, the active machine learning approach proposed in this paper highlights the following two novel aspects: 1) the more automatical generation of a large number of training samples can take the place of costly manually labeled data, and 2) instead of focusing on craters with distinct features, as in most existing machine learning methods, this approach prioritizes training samples near the classification boundary, which can help the machine learning process to focus on the most difficult cases and thus learn more quickly.

The remainder of this paper is organized as follows. Section II describes the active machine learning approach in detail. Its performance is evaluated using two sets of test data on the Moon and Mars, and the experimental results are presented in Section III. Finally, the concluding remarks are presented and discussed in Section IV.

II. Active Machine Learning for Crater Detection

A. Overview of the Approach

For the proposed active machine learning approach for crater detection, the training datasets include images and DEMs that cover the same area. If the DEMs were generated from the images by photogrammetric techniques, they were already aligned. If the images and the DEMs were acquired from

different sources, co-registration between the two types of data would be necessary. The co-registration methods in the literature can be used for this purpose [42]-[44].

The active machine learning approach begins with a relatively small number of manually labeled training samples that serve as an initial input to train the classifier. Haar-like features are used to describe the samples. An Adaptive Boosting method is used to select and combine useful features, and a cascade structure is used to improve the processing efficiency. During the training process, the approach incorporates profiles derived from the DEM to judge the correctness of the detection results, from which the detection results will be automatically labeled as correct or incorrect and added to the positive or negative samples, respectively. The updated samples will be used to retrain and improve the classifier in machine learning. This cross-validation process will be conducted several times until saturation was reached, which means no improvement can be expected with the accumulation of the training samples. This is a process of active learning because the algorithm actively selects training samples and asks for annotations. But different from the traditional active learning, instead of interacted by domain experts, the 3D information from the DEM helps to select and label the positive and negative samples. The output classifier can be then applied on 2D images in other areas for crater detection. Fig. 1 shows the framework of the proposed approach.

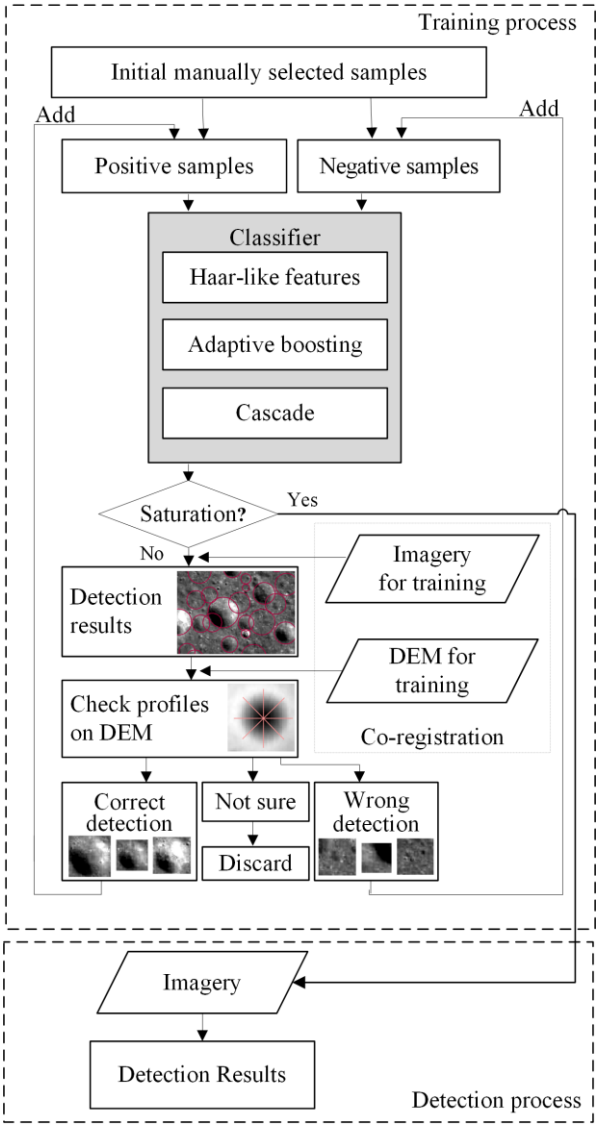


Fig. 1. Framework of the proposed active machine learning approach for crater detection.

B. Training Sample Determination

The approach begins with the selection of a small number of crater images (positive samples) and no-crater images (negative samples), as shown in Fig. 2. The crater samples, manually labeled on the satellite images, are cropped, rotated to the same illumination direction, and adjusted to the same size (e.g., 20×20 pixels). It is worthy of note that the positions of craters in positive samples are not strictly limited in previous related works. In our approach, however, positive areas are carefully selected to fit the craters. The strictly fixed position of the craters in the positive samples can bring about two types of benefits: 1) the features will be easier to describe, and 2) no further edge-detection process (which may cause additional errors) is needed because the largest inscribed circle of the detection window is the rim of the crater.

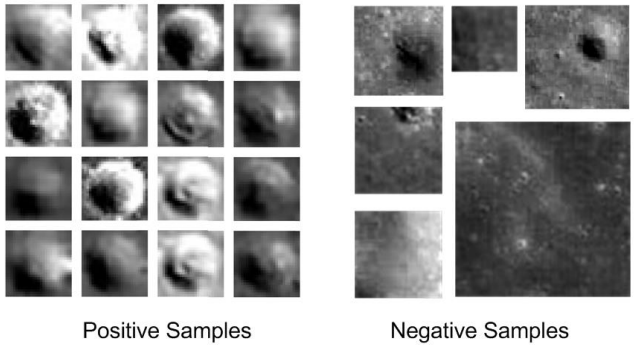


Fig. 2. Examples of positive and negative samples used in the training process.

C. Description of Classifier

The classifier in the approach includes three components: Haar-like features, adaptive boosting, and cascade.

The Haar-like features [45] are rectangle masks, which can be at any position and scale within the sample image. The feature value is the sum of the difference between the black and white image areas, which describes the change of the image's gray level. The Haar-like features were extended to tilted Haar-like features [46] by adding some tilted rectangles. Tilted Haar-like features are used in this study because their overall performance is better than that of the original Haar-like features. Fig. 3 shows examples of the tilted Haar-like features selected in the training phase. To accelerate the calculation of the feature values, the image is represented by an integral image [45]. The integral image can be calculated from the image at a low computational cost. The value of the integral image $II(x, y)$ at position (x, y) is defined by the summation of all pixels above and to the left:

$$II(x,y) = \sum_{x' \le x,\, y' \le y} i(x',\, y') \qquad (1)$$

where $i(x',\, y')$ is the pixel value of $(x',\, y')$. In the calculation,

it is not necessary to traverse each pixel and sum their values. The integral image value at a point (x, y) is calculated using only four references.

$$II(x,y) = II(x,y-1) + II(x-1,y) + I(x,y) - II(x-1,y-1) \quad (2)$$

With the integral image as a reference, for any of the tilted Haar-like features, it can be calculated with a few basic operations instead of calculating each involved pixel. For example, as Fig. 4 shows, the sum of the pixel values in rectangle A can be calculated as $II(d) - II(b) - II(c) + II(a)$.

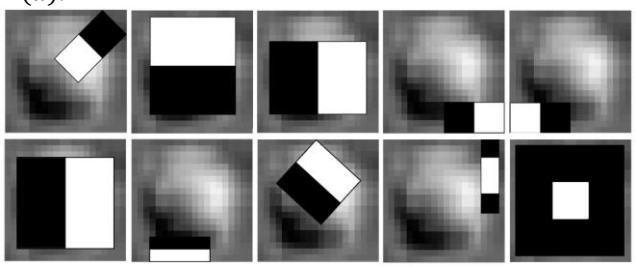


Fig. 3. Examples of tilted Haar-like features selected in the training phase.

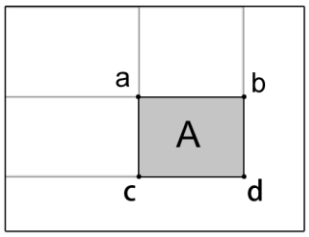


Fig. 4. Illustration of the calculation of Haar like features by referencing integral image.

Adaptive Boosting (AdaBoost) [47] is a machine learning algorithm used to combine weak classifiers into a strong classifier. Each tilted Haar-like feature can be regarded as a weak classifier $h_i(x)$:

$$h_i(x) = \begin{cases} 1 & if\ p_i f_i(x) < p_i \theta_i \\ 0 & otherwise \end{cases} \quad (3)$$

where f_i describe the type, size and location of the feature; θ_i is the optimal threshold and p_i is a polarity indicating the direction of the inequality sign. x is the sub-window of the training samples (e.g., 20×20 in this research). The AdaBoost training process combines the most useful weak classifiers to form a strong classifier. The features are combined into a weighted sum that represents the final output of the boosted classifier. A more useful feature will gain a higher weight, and the weight will be adjusted in the iteration. At each iteration of the training process, a weight is also assigned to each sample. Samples that are not correctly classified will have a higher weight in the next iteration. In this way, the AdaBoost training process selects only features that are known to improve the model's predictive power, thus reducing dimensionality and time.

Cascade is a mechanism to accelerate processing. The cascade structure [45] is formed by a series stage of classifiers, and each stage either rejects the image window or passes it to the next stage. The first classifier eliminates many non-crater images with very little processing and leaves the more difficult ones for the following classifier. The image region that passes the last stage is finally classified as a crater. The image regions that are easy to classify (e.g., homogeneous flat areas) are rejected in the early stages, whereas the difficult image regions (e.g., large rocks and hills) are classified as non-craters in the later stages.

Fig. 5 illustrates the structure of the classifier. Given the training example $(x_i, y_i), \ldots, (x_{m+n}, y_{m+n})$, $y_i \in \{+1, -1\}$ for the crater and non-crater samples, for each feature i, initial weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ are assigned for the crater and non-crater samples, respectively. m and l are the number of crater and non-crater samples. For each sample $x_1, x_2, \ldots, x_i$, $x_i = \{h_1(x)\alpha_1, h_2(x)\alpha_2, \ldots, h_l(x)\alpha_l \ldots, h_L(x)\alpha_L\}$, where l is the lth tilted Haar-like feature and assuming the total number of tilted features is L.

For t=1,…,T (T is the number of selected weak classifier):

1. Normalize the weights, $\alpha_{t,i} \leftarrow \frac{\alpha_{t,i}}{\sum_{j=1}^{n} \alpha_{t,j}}$
2. For each classifier $h_j(x)$ of feature j, the error is evaluated by $\varepsilon_t = \sum_i \alpha_i |h_j(x_i) - y_i|$. A weak classifier is selected from all tilted Haar-like classifiers when it has the lowest error rate.
3. The weight is updated as $\alpha_{t+1,i} = \alpha_{t,i}\beta_t^{1-e_i}$, where $e_i \in \{0, 1\}$ for correct and wrong classification, respectively, and $\beta_t = \varepsilon_t/(1-\varepsilon_t)$.

Finally, a weighted combination of the weak classifier $h_l(x)$ forms a strong classifier $H(x)$:

$$H(x) = sign(\sum_{t=1}^{T} \alpha_t h_t(x)) \quad (4)$$

Through several rounds of the process, a strong classifier is generated in each round and they are combined to form the final cascade classifier.

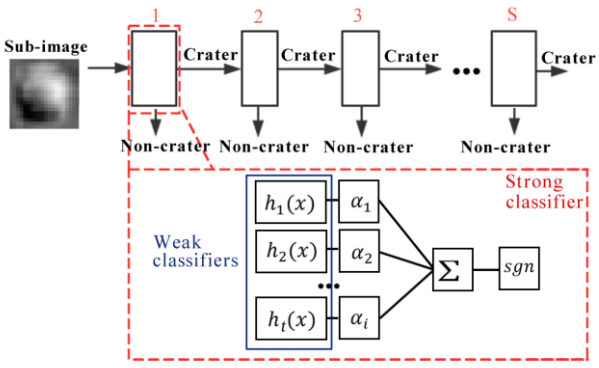


Fig. 5. The structure of the classifier using Haar-like features, AdaBoost and cascade technic.

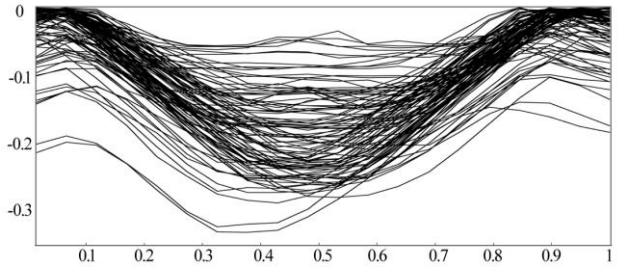
D. *Active Learning Process*

The active learning process consists of three steps: 1) initial classifier training, 2) crater detection, and 3) training sample updating and retraining. First, an initial classifier can be trained by a small number of samples. Second, the initial classifier is used for crater detection, and a large number of crater

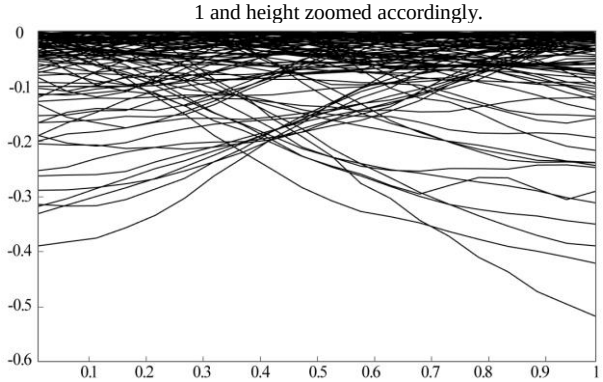
candidates can be detected. Third, the corresponding DEM is used to identify false and true craters from the detected candidates, and the training pool is updated simultaneously. The updated training pool is then used to retrain the model, and the process can be repeated several times until favorable results are achieved.

The first round of training and detection with a small number of training samples will inevitably produce many wrong detections. The method will actively ask the DEM for annotations. With the co-registered DEM, verification based on profiles can be conducted.

To explore the profiles derived from different terrain types, 100 crater profiles and 100 non-crater profiles were selected, as shown in Fig. 6. The profiles have different scales. For the sake of comparison, the horizontal distances were normalized, and the heights were zoomed accordingly. The X-axis shows the distance and the Y-axis displays the elevation after the normalization. It can be seen that the crater profiles (Fig. 6 (a)) have concave shapes while the non-crater profiles (Fig. 6 (b)) are just random lines. Therefore, it is possible to classify the crater profiles and non-crater profiles automatically.



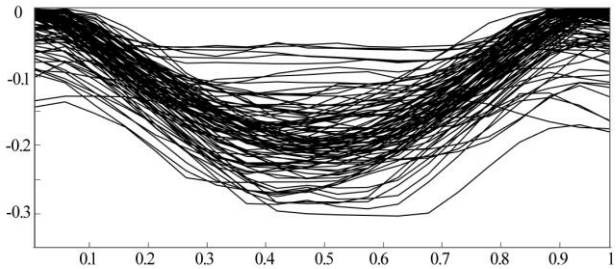
(a) 100 crater profiles with the horizontal distance normalized to 0-1 and height zoomed accordingly.



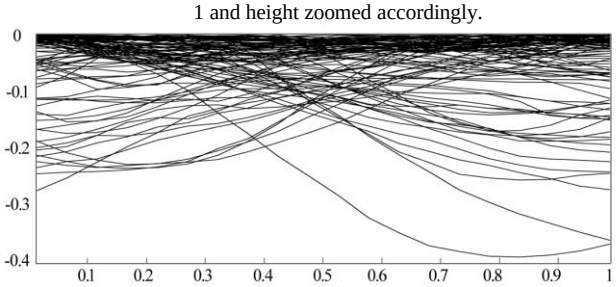
(b) 100 non-crater profiles with the horizontal distance normalized to 0-1 and height zoomed accordingly.

Fig. 6. Crater profiles and non-crater profiles for training the SVM classifier.

A Support-Vector Machine (SVM) classifier was used for the profile classification. Initially conceived by Cortes and Vapnik [48], SVM is used to solve binary classification problems. The goal of SVM is to determine an optimal hyperplane, which can not only separate two classes but also in a way that would make it as far as possible from the closest members of both classes. To prepare the training data for the SVM classifier, the 200 profiles (as shown in Fig. 6) were sampled by 20 points. Each profile was labeled as 1 for crater profiles and -1 for non-crater profiles. Thus the input of the training process was a $20 \times n$ matrix of profile data and a $1 \times n$ matrix of label data. n is the total number of profiles. After training, a mathematical depiction of the separating hyperplane training was obtained. To evaluate the performance of the well-trained SVM classifier for profile classification, another 200 profiles were manually selected and labeled, as shown in Fig. 7. They were normalized and resampled in the same way as previously described. The 100 crater profiles (Fig. 7. (a)) were extracted from the craters with various size and shapes. The 100 non-crater profiles (Fig.7. (b)) were randomly selected in the non-crater areas. For the 200 testing data, the prediction made by the SVM classifier is 100% consistent with the manual label.



(a) 100 crater profiles with the horizontal distance normalized to 0-1 and height zoomed accordingly.



(b) 100 non-crater profiles with the horizontal distance normalized to 0-1 and height zoomed accordingly.

Fig. 7. Crater profiles and non-crater profiles for evaluating the performance of the SVM classifier.

The classification results of profiles can serve as a principle to select the candidate crater areas and non-crater areas for updating the training samples. Due to the complicity of the planetary surface, it is possible for a non-crater area to have a crater-like profile (as Fig.8 (a) shows) and also for a crater area, not all the profiles have the concave shape, especially for connected craters (as Fig.8 (b) shows). In this approach, four profiles were derived for each candidate crater area, namely a vertical profile, a horizontal profile, and two diagonal profiles. Three rules were set to select crater and non-crater areas for sample updating based on the four derived profiles:

1) If three or more profiles are classified as crater profiles by the SVM classifier, this area is regarded as a crater area and is added to the positive-sample dataset;

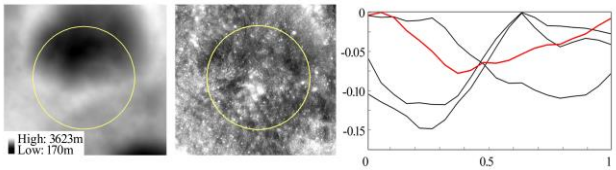
2) If less than two profiles are classified as crater profile, the area is judged as non-crater area, and is added into the negative-sample dataset; and

3) For other situations, it will be regarded as not sure and will be excluded from the training sample datasets.

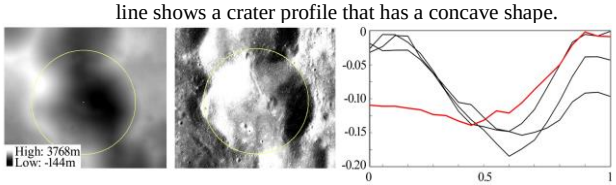
To verify the three rules for selecting the training samples, 100 patches of crater areas and another 100 patches of non-

crater areas were manually selected. Four profiles were derived from each area and classification was conducted by the SVM classifier on the profiles. Table I shows the numbers of the crater or non-crater areas with a certain number of profiles classified as crater profiles (among the four in total) for each area. 97 out of the 100 crater areas have three or more crater profiles, and the other three crater areas have two crater profiles. For the non-crater areas, 97 out of 100 of them have zero crater profile, and only three non-crater areas have been identified with one crater profile (out of the four profiles), which could happen by coincidence.

By adding the detection results, the training samples will accumulate automatically, which can free manual efforts from collecting training samples. In addition, this approach can make the negative samples more targeted because inaccurate detection is the difficult case in the training and classification process. This active learning process can be conducted several rounds. It should be noted that in each round, 30% of the samples will be reserved for the testing purpose. If the testing shows that the performance improvement is less than 1% between two rounds (a saturation status), the sample updating process will be stopped.



(a) Example of a crater area. Left: SLDEM. Middle: SELENE imagery. Right: four profiles derived from SLDEM, where the red line shows a crater profile that has a concave shape.



(b) Example of a crater area. Left: SLDEM. Middle: SELENE imagery. Right: four profiles derived from SLDEM, where the red line shows a non-crater profile.

Fig. 8. Examples of the situation that a crater area has a non-crater profile and a non-crater area has a concave shape crater profile.

TABLE I
VALIDATION OF THE SVM CLASSIFIER FOR SELECTING CRATER AND NON-CRATER SAMPLES

	Number of crater profiles in each area					Total
	0	1	2	3	4	
Crater area	0	0	3	8	89	100
Non-crater area	97	3	0	0	0	100

E. Crater Detection

With the final output classifier from the above steps, crater detection can be conducted on images in other regions. According to the sun azimuth of the imagery, each input image is rotated to ensure that the illumination direction is the same as in the training samples. To detect impact craters of various sizes, an image pyramid is then built for multiple-scales and traversed by the classifier. The detection results from each layer are merged into the final output. The output of the detection is a set of squares with information for location and size. The largest inscribed circle of the square describes the rim of the detected craters.

It should be noted that no DEM is used in the detection step. For each dataset, the training process need only be conducted once, and the classifier can be applied to other untrained imagery.

III. EXPERIMENTAL ANALYSIS

A. Evaluation Criterions

Two test datasets were selected representing the general characteristics of the planetary surface on the Moon and Mars for this study. Both the detection and evaluation areas contain diverse sizes and types of craters.

The ground truth crater catalog manually digitized from the images was used for elevation purposes. To guarantee the validation performance, we focused on craters with diameters greater than 20 pixels on imagery. Manual labeling was processed with the assistance of a 20-pixel reference grid overlaid on the image as a reference for the crater size. To guarantee the reliability, the ground truth data were cataloged separately by two operators, and the results were combined in the end. Craters identified by only one operator were further checked.

To evaluate the performance of the proposed approach, the true detection rate (TDR), the false detection rate (FDR), and the detection rate (DR) were computed as given by the following equations:

$$TDR(\%) = \frac{TP}{TP+FN} \times 100 \quad (5)$$

$$FDR(\%) = \frac{FP}{TP+FP} \times 100 \quad (6)$$

$$DR(\%) = \frac{TP}{TP+FN+FP} \times 100 \quad (7)$$

where true positive (TP) is the number of correct detections; false negative (FN) is the number of the missed detections; and false positive (FP) is the number of incorrect detections. The rules used in Salamunićcar et al. [49] were used to decide whether the two records belong to the same crater. The TDR, FDR, and DR were used to evaluate the performance of the classifier generated from each round of the process. In our experiments, the threshold parameters were carefully determined to reach a balance between TDR and FDR.

Another traditional machine learning approach [30] was involved in the experimental analysis for comparison purposes. The results from our approach and those from the traditional approach are compared with the ground truth data, in terms of the TDR, FDR, and DR.

B. Experiment for the Dataset on the Moon

The experiment for the dataset on the Moon used the SELENE TC images [36] and the SLDEM [50]. The SLDEM

was produced by co-registering the DEM from SELENE TC images and data from the Lunar Orbiter Laser Altimeter, giving an effective spatial resolution of 60 m. Two sets of data in different locations were selected for training and evaluation purposes, respectively. Fig. 9 (a) shows a 10 m/pixel SELENE TC image, and Fig. 9 (b) shows the co-registered SLDEM. Both were used for training purposes. Another SELENE TC image (Fig. 9 (c)) was used for crater detection and evaluation. It is a heavily cratered terrain, in which craters adjoin or even overlap with each other. Detailed information of the Moon dataset are shown in Table II.

TABLE II
INFORMATION OF THE MOON DATASET

	Longitude	Latitude	Data Size (pixels)
Fig. 9. (a)	55.3°W-61.9°W	41.3°N-44.8°N	22190×16839
Fig. 9. (b)	55.3°W-61.9°W	41.3°N-44.8°N	2303×1748
Fig. 9. (c)	53.4°W-50.0°W	41.2°N-42.2°N	11747×4780

The active machine learning approach started with 130 positive and 300 negative samples that were used as inputs to train the classifier; this process was regarded as round one. The classifier was then used to detect craters on the image for training purposes. For each detection result, the profiles were extracted from the co-registered DEM. According to the judgment made by the profiles, the result was automatically labeled as correct or incorrect and added to positive or negative datasets, respectively. The newly generated positive samples and negative samples were resized and added in the second-round training datasets. The result of the renewed datasets was used to retrain the classifiers of the second round. This cross-validation process was conducted six times until testing shows the performance improvement is less than 1% between two iterations. For comparison, the experiments were also conducted using a traditional boosting approach [30]. The training processes of the traditional boosting approach and the proposed approach were conducted in the same way, except for the way of collecting training samples. For the experiments using the traditional boosting approach, we followed the regular way of manually selecting the training samples of the same number for each round of training. We stopped at round 3 due to the time-consuming manual process.

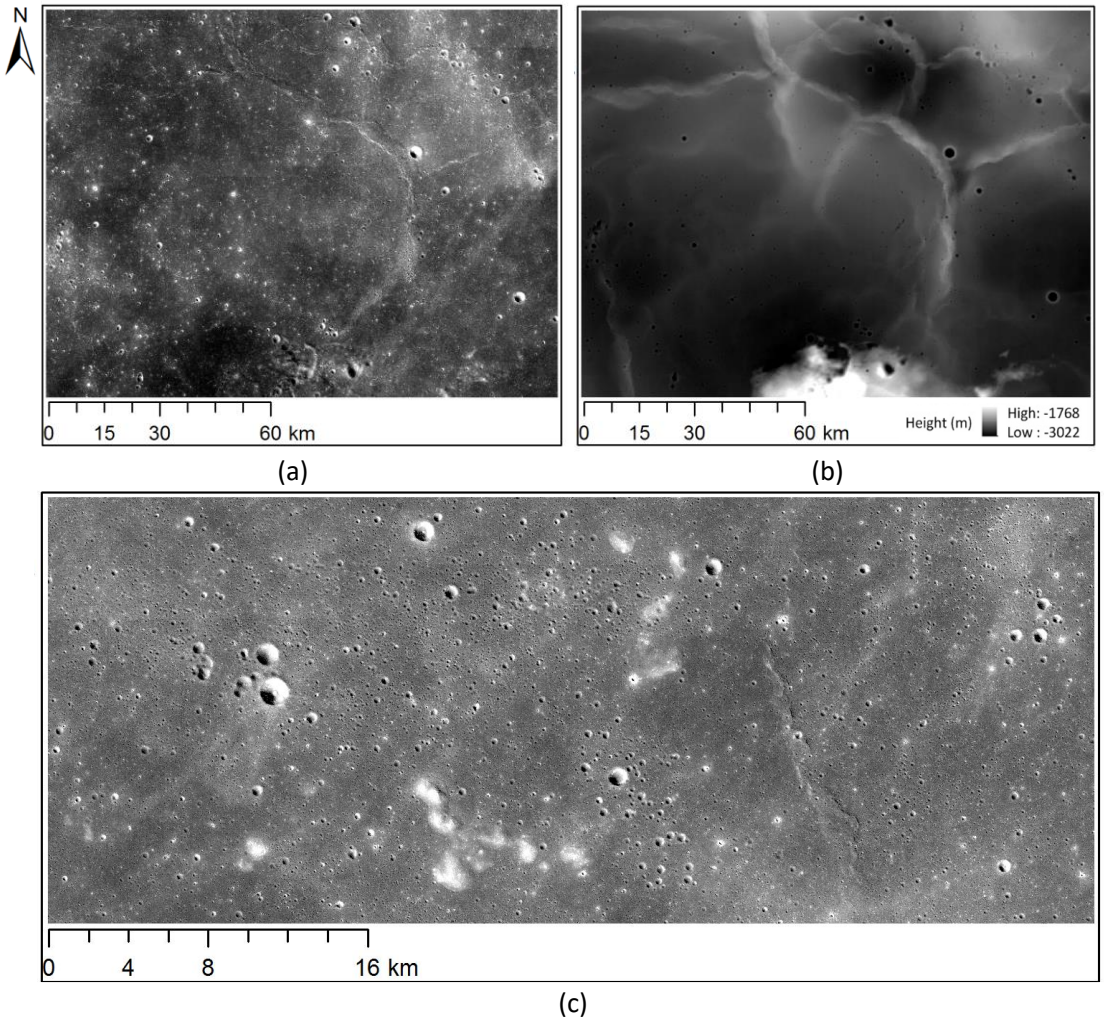


Fig. 9. (a) SELENE TC image (10 m/pixel) used for training. (b) SLDEM (60 m/pixel) used for training. (c) SELENE TC image (10 m/pixel) used for evaluation.

TABLE III
COMPARISON RESULTS FOR THE MOON DATASET

			Active machine learning approach			Traditional boosting approach [30]		
	Number of Positive Samples	Number of Negative Samples	TDR	FDR	DR	TDR	FDR	DR
Round 2	250	1250	80.78%	52.06%	43.03%	60.90%	58.25%	32.93%
Round 3	400	2000	82.88%	48.60%	46.47%	72.54%	64.02%	31.66%
Round 4	1000	2500	85.26%	23.59%	67.49%	/	/	/
Round 5	2000	4500	93.63%	10.74%	84.15%	/	/	/
Round 6	2600	5500	94.01%	11.09%	84.18%	/	/	/

The evaluation results (TDR, FDR, and DR) of each round for the active machine learning method and the traditional method are listed in Table III. For the active machine learning method, the performance of the classifier gradually improved as the training samples accumulated until saturation was reached after round 5, with a TDR of 93.63% and an FDR of 10.74%. With the same number of training samples, the classifier trained by the proposed active machine learning performed better than that trained with the traditional method. This finding proves that prioritizing the most confusing samples in the training process can greatly reduce the amount of labeled data required to train a model.

The ground truth (Fig. 10 (a)) has 6701 craters, and the detection result (Fig. 10 (b)) contains 6864 craters. Fig. 10 (c) presents the comparisons between the ground truth and the detection results using the classifier from round 5. Yellow circles stand for TP, blue circles stand for FN, and orange circles stand for FP.

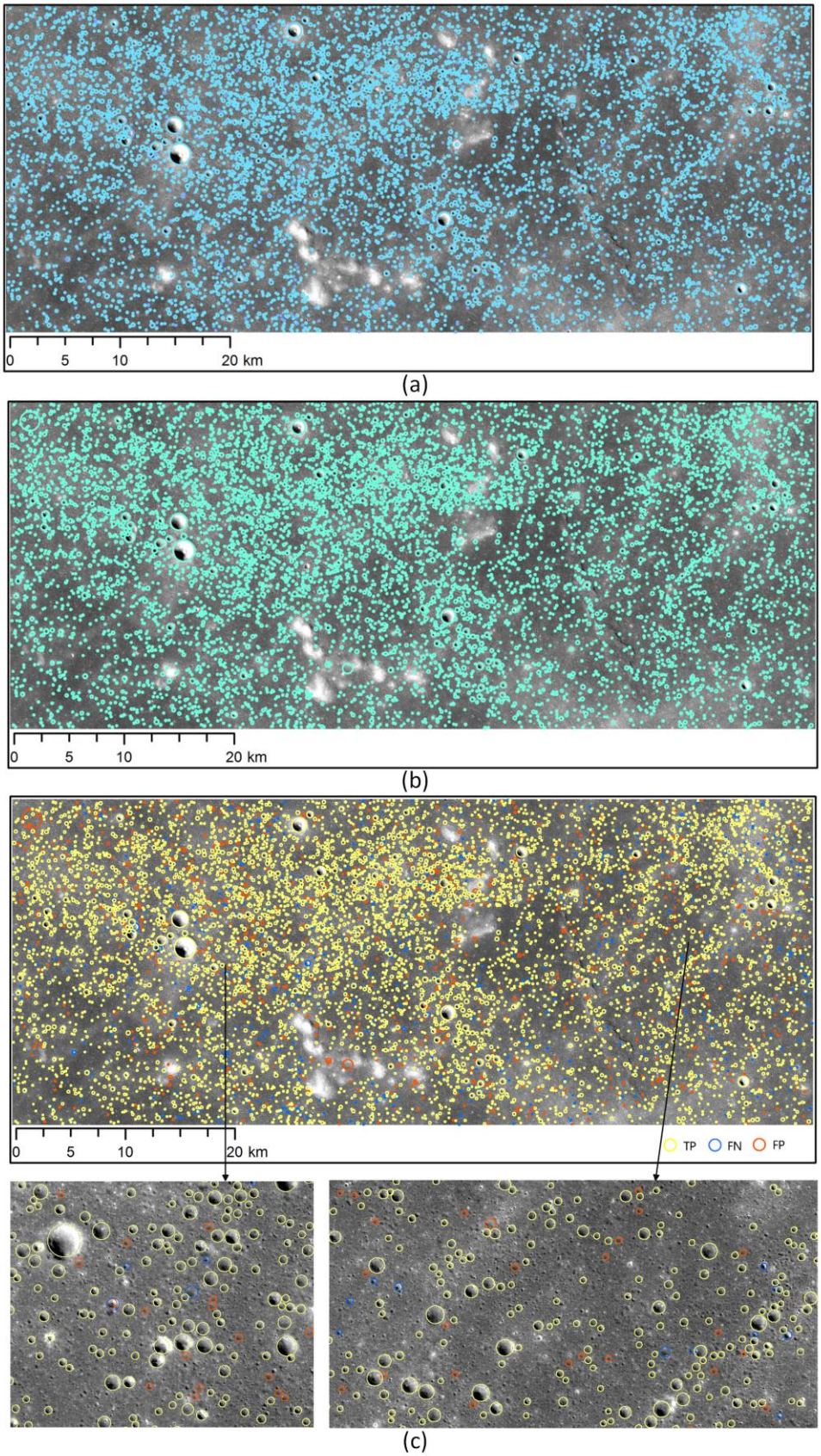


Fig. 10. Experimental results for the Moon dataset. (a) Ground truth data, (b) Detection results from round 5, and (c) comparison between detection results and ground truth.

C. Experiment for the Dataset on Mars

The experiments for the dataset on Mars used the High-Resolution Stereo Camera (HRSC) [38] images and DEMs derived from HRSC image. Both the HRSC image and the DEM were downloaded from http://pds-geosciences.wustl.edu/missions/mars_express/hrsc.htm. Like the experiments conducted on Lunar data, two sets of data in different locations were selected for training and evaluation purposes, respectively. Fig. 11 (a) shows a 12.5-m/pixel HRSC image, and Fig. 11 (b) shows the co-registered 50-m/pixel HRSC DEM. Both are used for training purposes. Fig. 11 (c) isa 12.5-m/pixel HRSC image used for crater detection and evaluation. Detailed information of the Mars dataset are shown in Table IV.

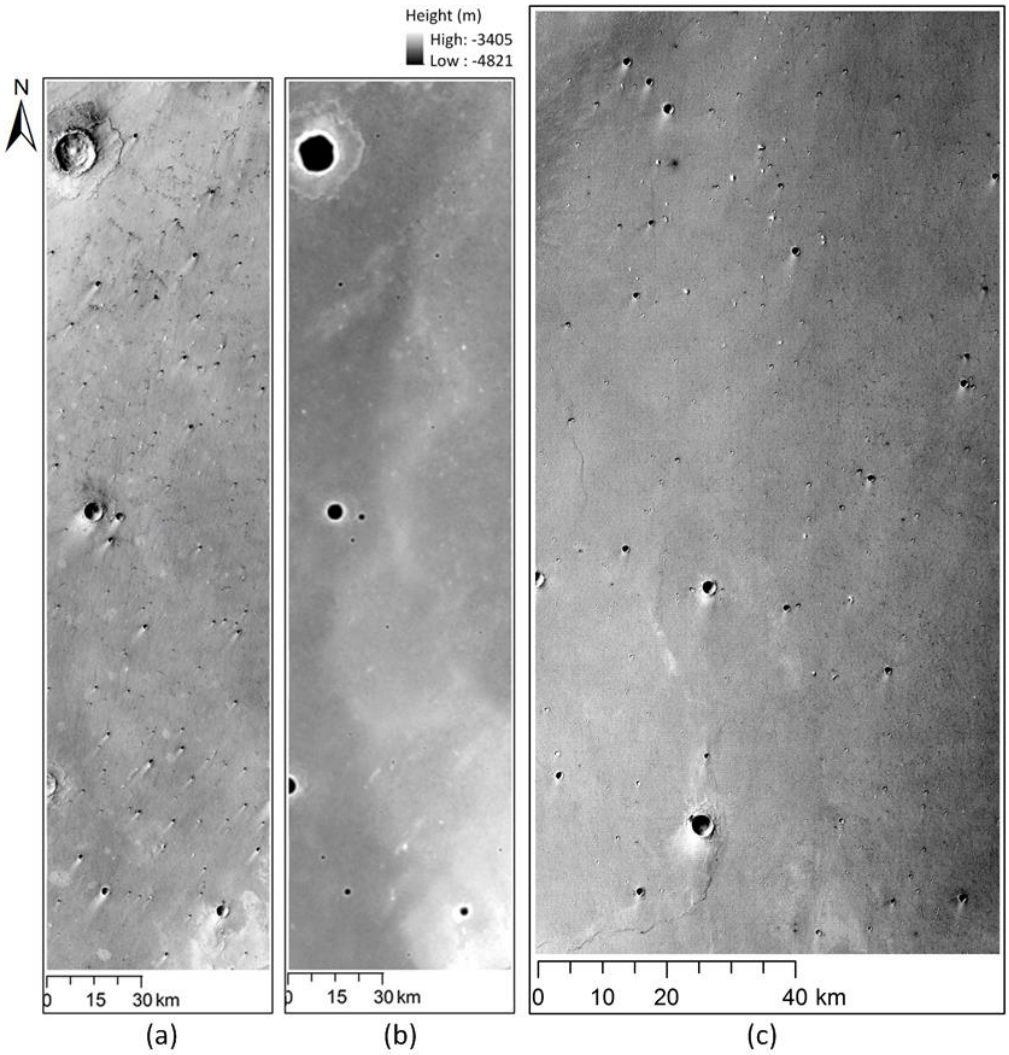


Fig. 11. (a) HRSC image (12.5 m/pixel) used for training. (b) HRSC DEM (50 m/pixel) used for training. (c) HRSC image (12.5 m/pixel) used for evaluation.

TABLE IV
INFORMATION OF THE MARS DATASET

	Longitude	Latitude	Data Size (pixels)
Fig. 11. (a)	37.9°W-39.1°W	21.3°N-25.7°N	5264×20864
Fig. 11. (b)	37.9°W-39.1°W	21.3°N-25.7°N	1317×5216
Fig. 11. (c)	43.8°W-45.0°W	20.6°N-22.9°N	6816×45680

The experiment conducted on the Mars data followed the similar process as that on the Moon data. The active machine learning started with 150 positive samples and 700 negative samples and ended up with about 6,000 samples in total. The performance of the detection classifier gradually improved as the training samples accumulated until saturation was reached after round 5, with a TDR of 92.27% and an FDR of 3.83%, as shown in Table V. It should be noted that saturation was reached with a smaller dataset than the experiment on the Moon, possibly because the Mars has fewer craters than the Moon, which makes detection easier. Compared with the traditional method (i.e., training by the same number of manually selected samples), the active machine learning method shows an apparently better result.

Fig. 12 (a) shows the 1306 ground truth craters, and Fig. 12 (b) shows the 1273 detected craters. Fig. 12 (c) compares the ground truth and the detection results using the classifier from round 5. Yellow circles stand for TP, blue circles stand for FN, and orange circles stand for FP.

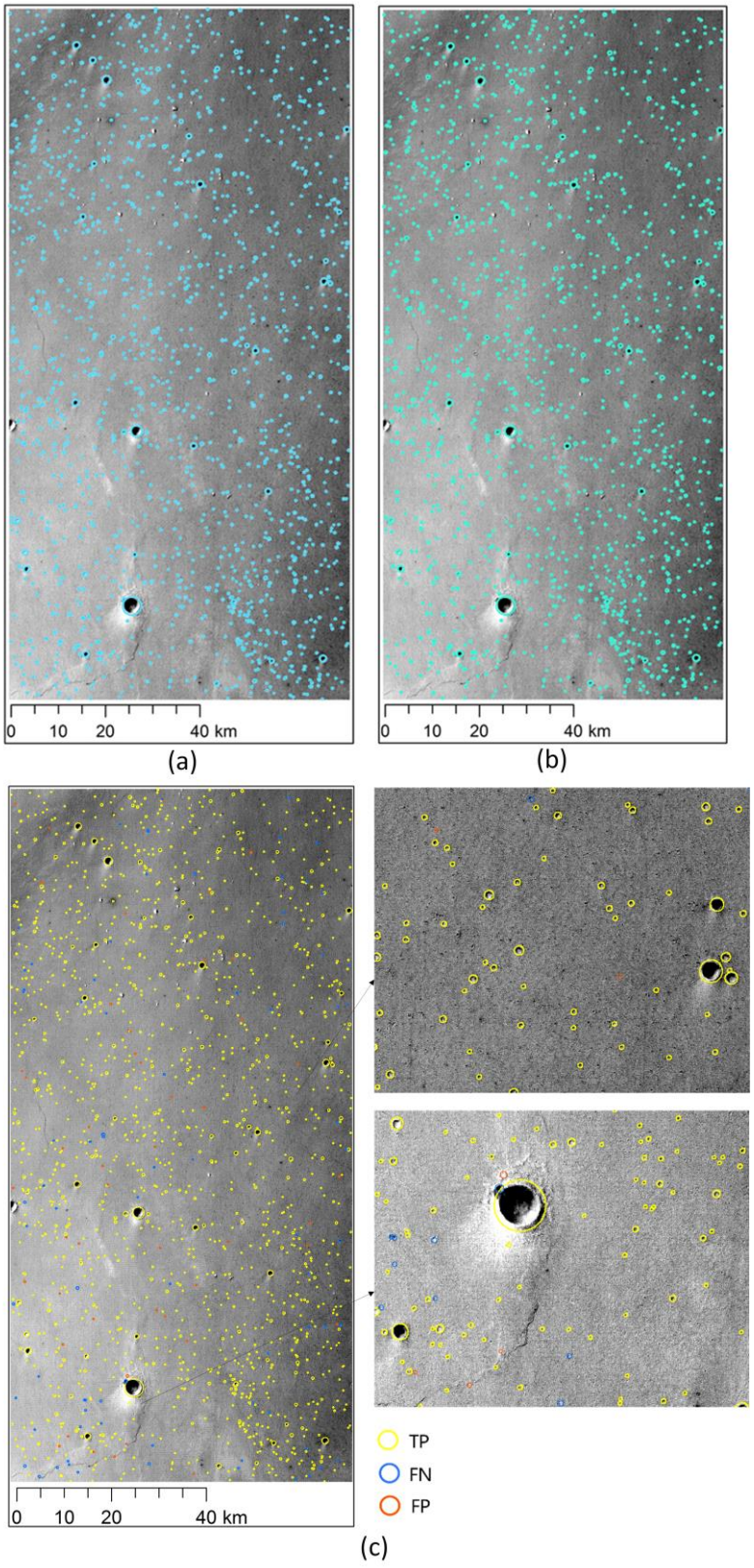


Fig. 12. Experimental results for Mars dataset: (a) ground truth data, (b) detection results from round 5, and (c) comparison between detection results and ground truth.

TABLE V
COMPARISON RESULTS FOR THE MARS DATASET

			Active machine learning approach			Traditional boosting approach [30]		
	Positive Samples	Negative Samples	TDR	FDR	DR	TDR	FDR	DR
Round 2	250	1250	65.05%	36.32%	47.45%	57.53%	46.50%	38.35%
Round 3	400	2000	72.04%	21.18%	60.36%	67.74%	41.40%	45.82%
Round 4	650	3000	85.99%	21.08%	69.93%	/	/	/
Round 5	800	4000	92.27%	3.83%	89.00%	/	/	/
Round 6	1100	5000	92.80%	4.72%	88.73%	/	/	/

IV. CONCLUSIONS AND DISCUSSION

An active machine learning approach is proposed for crater detection with greater automation and better performance. The developed approach was evaluated with actual datasets collected on the Moon and on Mars. The theoretical analysis and experimental validation convey the following conclusions.

1) The evaluation process is conducted with manually labeled ground truth data, and the proposed active machine learning shows a favorable performance. The final TDR and FDR were 93.63% and 10.74%, respectively, for the Moon dataset and 92.27% and 3.83%, respectively, for the Mars dataset.
2) The proposed method iteratively updates and optimizes the training datasets with a more automatic active-learning mechanism, which greatly reduces the labor cost of labeling the training samples. The performance of the detection classifier also improves as the training samples accumulate until favorable results are obtained.
3) The experimental results on the Moon and Mars show that the proposed machine learning method produces more accurate crater detection than the traditional machine learning method. This means that the samples generated from the active learning process provided much useful information during the training process by prioritizing the most confusing examples.

It should be noted that the automatic training process must be conducted with both the image and DEM data available; however, once the classifier obtained it can be applied to other areas with only images available. It should also be noted that this paper selected a diameter threshold of 20 pixels on imagery when detecting craters, which is mainly for evaluation convenience. The proposed approach is able to detect craters with diameters less than 20 pixels on the image, however, the detection rate might be decreased. Fig. 13 shows examples of using the proposed approach for crater detection without setting a specific diameter threshold. The detected smallest craters reach to 7 pixels in size, however, the TDR and FDR are degraded to 85.01% and 15.61%, respectively. Our future work will investigate strategies to improve the crater detection rate for small craters by incorporating other feature information.

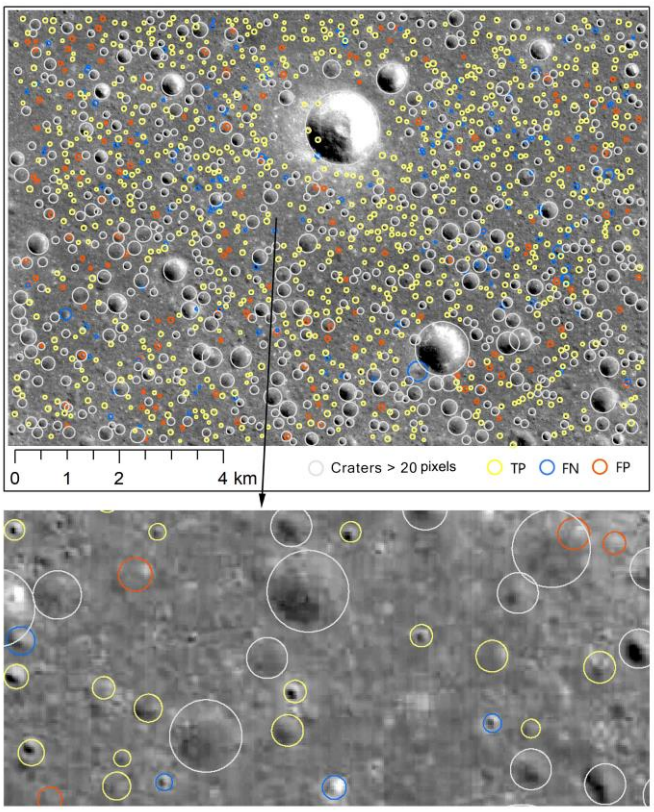


Fig. 13. Examples of crater detection results without setting a diameter threshold. The bottom row shows an enlarged view of a local region marked by the arrow.

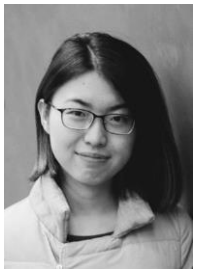
The proposed active machine learning approach enables a more automated method of collecting training samples with a favorable number and quality, which results in better performance in crater detection. The proposed method provides new insights in the field of machine learning for various applications.

REFERENCES

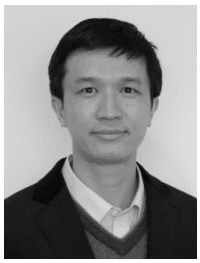
[1] D. E. Wilhelms, F. John, N. J. Trask, "The geologic history of the Moon," U.S. Geol. Surv. Prof. Pap. 1348, 1987. (U.S. Government Printing Office, Washington, DC).
[2] J. W. Head, C. I. Fassett, S. J. Kadish, D. E. Smith, M. T. Zuber, G. A. Neumann, E. Mazarico, "Global distribution of large lunar craters:

Implications for resurfacing and impactor populations," Science 329, 1504-1507, 2010.
[3] D. De Rosa, B. Bussey, J. T. Cahill, T. Lutz, I. A. Crawford, T. Hackwill, S. van Gasselt, G. Neukum, L. Witte, A. McGovern, "Characterisation of potential landing sites for the European Space Agency's Lunar Lander project," Planet. Space Sci., 74, 224-246, 2012.
[4] B. Wu, F. Li, L. Ye, S. Qiao, J. Huang, X. Wu, H. Zhang, "Topographic modeling and analysis of the landing site of Chang'E-3 on the Moon," Earth Planet. Sci. Lett 405, 257-273, 2014.
[5] G. Salamunićcar, S. Lončarić, P. Pina, L. Bandeira, J. Saraiva, "MA130301GT catalogue of Martian impact craters and advanced evaluation of crater detection algorithms using diverse topography and image datasets," Planet. Space Sci., 59, 111-131, 2011.
[6] B. D. Bue, T. F. Stepinski, "Machine detection of Martian impact craters from digital topography data," IEEE Transactions on Geoscience and Remote Sensing 45, 265-274, 2007.
[7] J. R. Kim, J. P. Muller, S. van Gasselt, J. G. Morley, G. Neukum," Automated crater detection, a new tool for Mars cartography and chronology," PE&RS, 71, 1205-1217, 2005.
[8] L. Luo, L. Mu, X. Wang, C. Li, W. Ji, J. Zhao, H. Cai, "Global detection of large lunar craters based on the CE-1 digital elevation model," Front Earth Sci., 7, 456-464, 2013.
[9] G. Michael, "Coordinate registration by automated crater recognition," Planet. Space Sci., 51, 563-568, 2003.
[10] A. L. Salih, M. Mühlbauer, A. Grumpe, J. Pasckert, C. Wöhler, H. Hiesinger, "Mapping of planetary surface age based on crater statistics obtained by an automatic detection algorithm," ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 479-486, 2016.
[11] N. Harada, T. Hayashi, N. Hirata, H. Demura, N. Asada, "Recognition algorithm for topographic features, Computer and Information Technology," CIT 2007. 7th IEEE International Conference on. IEEE, 2007, pp. 685-689.
[12] K. Di, W. Li, Z. Yue, Y. Sun, Y. Liu, "A machine learning approach to crater detection from topographic data," Adv. Space Res, 54, 2419-2429, 2014.
[13] M. Galloway, J. Paxman, G. Benedix, T. Tan, M. Towner, P. Bland, "Automated crater detection and counting using the Hough transform and canny edge detection," Workshop on Issues in Crater Studies and the Dating of Planetary Surfaces, p. 9024, 2015.
[14] M. J. Galloway, G. K. Benedix, P. A. Bland, J. Paxman, M. C. Towner, T. Tan, "Automated crater detection and counting using the Hough transform," Image Processing (ICIP), IEEE International Conference on. IEEE, 2014, pp. 1579-1583.
[15] Y. Cheng, A. E. Johnson, L. H. Matthies, C. F. Olson, "Optical landmark detection for spacecraft navigation," Proceedings of the 13th AAS/AIAA Space Flight Mechanics Meeting, Ponce, Puerto Rico, February, 2003, AAS 03–235.
[16] R. Honda, Y. Iijima, O. Konishi, "Mining of topographic feature from heterogeneous imagery and its application to lunar craters," Progress in Discovery Science, 27-44, 2002.
[17] J. Yin, H. Li, X. Jia, "Crater detection based on gist features," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8, 23-29, 2015.
[18] T. Barata, E. I. Alves, J. Saraiva, P. Pina, "Automatic recognition of impact craters on the surface of Mars," International Conference Image Analysis and Recognition. Springer, 2004, pp. 489-496.
[19] S. Vijayan, K. Vani, S. Sanjeevi, "Crater detection, classification and contextual information extraction in lunar images using a novel algorithm," Icarus 226, 798-815, 2013.
[20] J. Earl, A. Chicarro, C. Koeberl, P. G. Marchetti, M. Milnes, "Automatic recognition of crater-like structures in terrestrial and planetary images," 36th Annual Lunar and Planetary Science Conference, no. 1319, 2005.
[21] M. Ding, Y. Cao, Q. Wu, "Novel approach of crater detection by crater candidate region selection and matrix-pattern-oriented least squares support vector machine," Chin. J. Aeronaut., 26, 385-393, 2013.
[22] S. Jin, T. Zhang, "Automatic detection of impact craters on Mars using a modified adaboosting method," Planet. Space Sci., 99, 112-117, 2014.
[23] Y. Wang, G. Yang, L. Guo, "A novel sparse boosting method for crater detection in the high resolution planetary image," Adv. Space Res. 56, 982-991, 2015.
[24] J. P. Cohen, H. Z. Lo, T. Lu, W. Ding, "Crater detection via convolutional neural networks," arXiv preprint arXiv:1601.00978, 2016.
[25] X. Bai, H. Zhang, J. Zhou, "VHR object detection based on structural feature extraction and query expansion." IEEE Transactions on Geoscience and Remote Sensing. Oct;52(10):6508-20, 2014.
[26] H. Zhang, X. Bai, J. Zhou, J. Cheng, H. Zhao. "Object detection via structural feature selection and shape model." IEEE transactions on image processing. Dec;22(12):4984-95, 2013.
[27] C. Wang, X. Bai, S. Wang, J. Zhou, P. Ren. Multiscale Visual Attention Networks for Object Detection in VHR Remote Sensing Images. IEEE Geoscience and Remote Sensing Letters, 2018.
[28] M. C. Burl, T. Stough, W. Colwell, E. Bierhaus, W. Merline, C. Chapman, "Automated detection of craters and other geological features," In: Proceedings of the 6th International Symposium on Artificial Intelligence, Robotics, and Automation in Space, AM118, 2001.
[29] T. Vinogradova, M. Burl, E. Mjolsness, "Training of a crater detection algorithm for Mars crater imagery," Aerospace Conference Proceedings, IEEE, 2002, pp. 7-7.
[30] R. Martins, P. Pina, J. S. Marques, Silveira, M., "Crater detection by a boosting approach," IEEE Geosci. Remote Sens. Lett. 6, 127-131, 2009.
[31] H. Li, J. Yin, Z. Gu. "Crater detection based on local non-negative matrix factorization," IGARSS, IEEE, pp. 521-524, 2014.
[32] B. Settles, "Active learning". Morgan & Claypool, 2012.
[33] S. Sivaraman, M.M. Trivedi, "Active learning for on-road vehicle detection: a comparative study", Mach. Vis. Appl., 25, 3, 599-611, 2014.
[34] E. Dura, Y. Zhang, X. Liao, G. J. Dobeck, L. Carin, "Active learning for detection of mine-like objects in side-scan sonar imagery," IEEE J. Ocean. Eng., 30, 2, 360-371, 2005.
[35] B. Wu, H. Hu, J. Guo, "Integration of Chang'E-2 imagery and LRO laser altimeter data with a combined block adjustment for precision lunar topographic modeling," Earth Planet. Sci. Lett., 391, 1-15, 2014.
[36] M. Kato, S. Sasaki, K. Tanaka, Y. Iijima, Y. Takizawa, "The Japanese lunar mission SELENE: Science goals and present status," Adv. Space Res., 42, 294-300, 2008.
[37] G. Chin, S. Brylow, M. Foote, J. Garvin, J. Kasper, J. Keller, M. Litvak, I. Mitrofanov, D. Paige, K. Raney, "Lunar Reconnaissance Orbiter overview: The instrument suite and mission," Space Sci. Rev., 129, 391, 2007.
[38] G. Neukum, R. Jaumann, "HRSC: The high-resolution stereo camera of Mars Express," Mars Express: The Scientific Payload, pp. 17-35, 2004.
[39] A. S. McEwen, E. M. Eliason, J. W. Bergstrom, N. T. Bridges, C. J. Hansen, W. A. Delamere, J. A. Grant, V. C. Gulick, K. E. Herkenhoff, L. Keszthelyi, "Mars Reconnaissance Orbiter's high-resolution imaging science experiment (HiRISE)," J. Geophys. Res.: Planets, 112, 2007.
[40] G. Salamunićcar, S. Lončarić, P. Pina, L. Bandeira, J. Saraiva, "Integrated method for crater detection from topography and optical images and the new PH9224GT catalogue of Phobos impact craters," Adv. Space Res., 53, 1798-1809, 2014.
[41] J. Simpson, J. Kim, J. Muller, "3D crater database production on Mars by automated crater detection and data fusion," In: 21st ISPRS Congress, 2008, pp. 1049–1054.
[42] B. Wu, J. Guo, H. Hu, Z. Li, Y. Chen, "Co-registration of lunar topographic models derived from Chang'E-1, SELENE, and LRO laser altimeter data based on a novel surface matching method," Earth Planet. Sci. Lett., 364, 68-84, 2013.
[43] B. Wu, H. Hu, J. Guo, "Integration of Chang'E-2 imagery and LRO laser altimeter data with a combined block adjustment for precision lunar topographic modeling," Earth Planet. Sci. Lett., 391, 1-15, 2014.
[44] Y. Wang, B. Wu, "Investigation of boresight offsets and co-registration of HiRISE and CTX imagery for precision Mars topographic mapping," Planet. Space Sci., 139, 18-30, 2017.
[45] P. Viola, M. Jones, 2001. "Rapid object detection using a boosted cascade of simple features," CVPR, Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001, pp. I-I.
[46] R. Lienhart, J. Maydt, "An extended set of Haar-like features for rapid object detection," Image Processing. Proceedings. International Conference on. IEEE, 2002, pp. I-I.

[47] Y. Freund, R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” J. Comput. Syst. Sci., 55, 119-139, 1997.
[48] C. Cortes, V. Vapnik, Support-vector networks. Machine Learning 1995, 20, 273-297.
[49] G. Salamunićcar, S. Lončarić, “Open framework for objective evaluation of crater detection algorithms with first test-field subsystem based on MOLA data,” Adv. Space Res., 42, 6-19, 2008.
[50] M. Barker, E. Mazarico, G. Neumann, M. Zuber, J. Haruyama, D. Smith, “A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera,” Icarus 273, 346-355, 2016.



Yiran Wang received the B.S. degree in Marine Technology from the Ocean University of China, Qingdao, China. She is currently working toward the Ph.D. degree with a major in photogrammetry and remote sensing at The Hong Kong Polytechnic University, Kowloon, Hong Kong. Her research interests include planetary mapping, automatic crater detection and machine learning. She is now working on projects for landing site selection of the future Chinese Mars and Lunar exploration missions.



Bo Wu is an Associate Professor with the Department of Land Surveying & Geo-Informatics of the Hong Kong Polytechnic University, mainly working on Photogrammetry and Planetary Mapping. He served as the Vice President of the Hong Kong Geographic Information System Association, the Co-Chair of the Working Group II/6 - Geo-Visualization and Virtual Reality of the International Society for Photogrammetry and Remote Sensing (ISPRS), and currently is serving as the Secretary of the Working Group IC III/IV - Planetary Remote Sensing and Mapping of ISPRS. He is currently the Associate Editor of international journal “Photogrammetric Engineering & Remote Sensing” and an Editorial Board member of the “ISPRS Journal of Photogrammetry and Remote Sensing” and “The Photogrammetric Record”.