Article

A Machine Learning Approach to Crater Classification from Topographic Data

Qiangyi Liu [1,2], Weiming Cheng [1,2,3,4,*] , Guangjian Yan [5,6] , Yunliang Zhao [7] and Jianzhong Liu [2,8]

1. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; liuqy.18b@igsnrr.ac.cn
2. University of Chinese Academy of Sciences, Beijing 100049, China; liujianzhong@mail.gyig.ac.cn
3. Jiangsu Centre for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
4. CAS Center for Excellence in Comparative Planetology, Hefei 230052, China
5. State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100088, China; gjyan@bnu.edu.cn
6. Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
7. School of Civil Engineering and Architecture, Southwest Petroleum University, Chengdu 610500, China; zhaoyl@lreis.ac.cn
8. Lunar and Planetary Science Research Center, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550002, China
* Correspondence: chengwm@lreis.ac.cn; Tel.: +86-010-6488-9777

check for updates

Abstract: Craters contain important information on geological history and have been widely used for dating absolute age and reconstructing impact history. The impact process results in a lot of ejected fragments and these fragments may form secondary craters. Studies on distinguishing primary craters from secondary craters are helpful in improving the accuracy of crater dating. However, previous studies about distinguishing primary craters from secondary craters were either conducted by manual identification or used approaches mainly concerning crater spatial distribution, which are time-consuming or have low accuracy. This paper presents a machine learning approach to distinguish primary craters from secondary craters. First, samples used for training and testing were identified and unified. The whole dataset contained 1032 primary craters and 4041 secondary craters. Then, considering the differences between primary and secondary craters, features mainly related to crater shape, depth, and density were calculated. Finally, a random forest classifier was trained and tested. This approach showed a favorable performance. The accuracy and F1-score for fivefold cross-validation were 0.939 and 0.839, respectively. The proposed machine learning approach enables an automated method of distinguishing primary craters from secondary craters, which results in better performance.

Keywords: moon; distinguish primary craters from secondary craters; machine learning; crater characteristics

1. Introduction

The current surface of a terrestrial planet is the result of geologic and geomorphologic processes, both having a significant effect on the landforms. Experiencing a continuous impact process, the terrestrial planet surface is covered by myriad craters. These craters contain important information on

geological history and have been widely used for dating absolute age, analyzing impact distribution, reconstructing impact history, and so on [1–3]. These applications are usually based on the assumption that all craters taken into consideration are primary craters produced by asteroids and comets [4,5]. However, the impact process results in a lot of ejected fragments, and these fragments may also form craters, which are considered secondary craters [6]. There is no doubt that research on distinguishing between primary and secondary craters is of great importance. On the one hand, it provides an opportunity to get a more precise geological age through the crater size-frequency distribution (CSFD) method. Making up a great percentage of small craters [7], secondary craters often lead to considerable uncertainty in the CSFD method [1,4,8–12]. Distinguishing primary from secondary craters is helpful in counting primary craters, hence improving the accuracy of crater dating. On the other hand, identification of primary and secondary craters also has a significant effect on projects concerning impact distribution, which may suggest diverse rotation. Besides, secondary craters also provide an approach to understand the impact characteristics of their parent craters.

Researchers have tried to distinguish primary craters from secondary craters, and identifying and learning the differences in their characteristics is the foundation of all related studies [1,13]. These differences play an important role in the distinguishing effort. Different from primary craters, some secondary craters occur in chains or clusters. These secondary craters are easy to identify, and this difference is the most commonly used one [5,14,15]. In addition, some secondary craters are associated with a "herringbone" ejecta pattern (V-shaped ridge), which indicates near-simultaneous formation during ejecta fragment deposition [16]. Oberbeck and Morrison [17] proved that the herringbone structure associated with lunar secondary craters can be accounted for simply by the interaction of ejecta plumes of secondary craters formed near one another by nearly simultaneous impact. Moreover, secondary craters have a more elliptical or irregular shape and are usually shallower than primary craters with the same diameter [16,18,19]. Researchers also found that primary and secondary craters also differ in rock size and center mound, which may be useful for distinguishing the two [20–22].

Similar to crater detection, current methods aiming at distinguishing primary craters from secondary craters include manual and automatic techniques. Compared with the great progress made in automatic crater detection, however, manual identification is still the method used in the most recent works related to distinguishing primary craters from secondary craters [1,23–26], and very few works have tried to develop a useful automatic approach [5,10,27–30]. Distinguishing primary craters from secondary craters manually can give accurate results, but it requires a lot of time and expert knowledge. It is difficult to deal with large regions, and studies need to be repeated several times. As mentioned above, differences between primary and secondary craters are related to crater shape, depth, and density. Automatic methods of distinguishing primary craters from secondary craters are usually rule-based methods and have considered different features used for classification. According to the used classification basis, automatic methods can be divided into two types: those that only consider crater density and those that consider multiple aspects of the differences between primary and secondary craters. Among the few works related to distinguishing primary from secondary craters, most of them took crater density as the only basis, and proposed rule-based methods related to clustering.

Applying methods related to clustering to distinguish primary from secondary craters was first proposed by Bierhaus [5]. To estimate the primary crater population, Bierhaus developed a novel algorithm that removes the strongly clustered (secondary) craters, and the core idea of this method was calculating the probabilities of nonrandomness by comparing the cluster degree of a certain crater within the research area with that of a suite of random populations of craters that possess the same spatial density. This could be calculated using a single-linkage hierarchical clustering algorithm and Monte Carlo methods. However, in Bierhaus' method, there existed a problem of converting the probabilities of nonrandomness into crater types. On the one hand, using this method to distinguish or remove secondary craters depends largely on the selection of the threshold used to divide the probabilities of nonrandomness into two types, which is hard to decide and varies from region to

region. On the other hand, there still have some unclustered secondary craters and clustered primary craters located near secondary crater clusters. These craters are hard to distinguish by density but may show great differences in shape or depth. Kreslavsky [30] and Michael [10] also did some studies on Mars aimed at quantifying the spatial randomness and clustering of craters, and the methods in these works could also be used to distinguish primary craters from secondary craters. By inspecting clustering at different scales of crater diameter and introducing features related to crater distance such as the mean second-closest neighbor distance, the general idea of their works was the same as Bierhaus'. Honda [29] and Salih [27] proposed similar secondary candidate detection methods by replacing the single-linkage hierarchical clustering algorithm with Voronoi tessellation. All the methods mentioned above are based on the differences of spatial distribution patterns between primary and secondary craters, and are aimed at removing the influence of secondary craters in the CSFD method. The most important and difficult part of these methods is the selection of the threshold of the probability of nonrandomness, as the threshold is greatly affected by the research area, candidate crater radius, and so on, and the setting of a threshold directly affects the performance of these methods. As Bierhaus [5] pointed out in his work, a certain fraction of spatially random distributions are in fact secondary, so how to distinguish this part of craters and further improve the precision of distinguishing primary from secondary craters still needs to be solved. Until recently, researchers began considering using multiple aspects of differences between primary and secondary craters to distinguish them. Considering that a secondary crater may occur in a chain or cluster and has an elliptical shape, Wu et al. [28] proposed an automatic approach for detection. They converted the descriptions of crater chains, crater clusters, and the differences of crater shape into three criteria, and craters that met any of the criteria would be regarded as secondary craters. The whole process of parameter calculation and criteria judgment could be done automatically by a computer. Compared with previous algorithms, the algorithm developed by Wu et al. considers more characteristics describing differences between primary and secondary craters, but it needs more thresholds, which may also be difficult to define and will greatly affect the performance. The difficulty of threshold setting greatly increases the amount of work before running the algorithm and the uncertainty of results. Though researchers have tried to develop some methods to automatically distinguish primary craters from secondary craters, the number of similar works is relatively small and most of them are rule-base methods. There is still a lot of progress to be made in this field. Except for the weak points mentioned above, most of these works try to present an idea without providing statistical tests.

A machine learning approach has already been introduced in lunar study, especially in crater identification and remote sensing image classification [31–37]. Due to the variety of crater structures, machine learning–based methods usually show more robust performance that rule-based methods [37]. As primary craters also differ from secondary craters in various aspects and cannot be distinguished according to one aspect, a machine learning–based method may have better performance that the previous rule-based method too. A machine learning method learns the optimal filters and features based on a great number of training examples. Compared with previous simple rule-based methods, it could better imitate and learn the complex judgment rules contained in visual recognition. Also, a machine learning method has better generalization ability and data adaptability [38]. Besides, a machine learning method needs fewer predefined thresholds and can give results from a comprehensive perspective. In the learning phase, features of primary craters and secondary craters are fed into a model to form a classifier. In the detection phase, the previously trained classifier distinguishes primary and secondary craters in a new set of candidate craters.

Based on a public crater database [39], this paper presents a machine learning approach to distinguish primary craters from secondary craters. First, samples used for training and testing were identified and unified. Then, features good at distinguishing primary craters from secondary craters were calculated and used as training features. Finally, a random forest classifier was trained and tested. The training process using different features was conducted several times, from which we selected a classifier that had the best performance, such as the highest accuracy or sensitivity, and this classifier

was used for automatic testing in other regions. Compared with previous studies, the approach developed in this paper is based on machine learning and emphasizes the following two innovations: (1) this method uses two groups of features to quantify a crater chain or a crater cluster, and (2) instead of simply focusing on density and using a rule-based method, as in most secondary crater automatic identification methods, this approach takes features related to shape, depth, and density into consideration to develop a machine learning-based method, which may improve its performance.

2. Data

2.1. Reference Data

This approach is mainly aimed at distinguishing primary and secondary craters based on an existing crater database. The lunar crater database used in this research was presented by Robbins [39], estimated to be a complete census of all craters with diameters larger than 1–2 km. The identification and feature extraction of training and testing samples are based primarily on the examination of the following data:

Lunar Reconnaissance Orbiter (LRO) and Kaguya merged digital elevation model (DEM), which spans 60° in latitude and has a resolution of 59 m/pixel [40];

- 1024 pixel per degree Lunar Orbiter Laser Altimeter topography data [41];
- 100 m/pixel Lunar Reconnaissance Orbiter Camera (LROC) wide-angle images [42].

2.2. Sample Data

The precision of the sample inventory greatly affects the reliability of the training classifier, and the first step of most machine learning–based methods is to prepare a set of positive and negative samples. One of the most used applications of the proposed method is to help get a more precise geological age through the CSFD method, and this means that it is important to make sure that craters identified as primary craters by this method are actual primary craters. Besides, there are usually more secondary craters than primary craters, and secondary craters account for most small craters. Based on these two reasons, in this method, primary craters are regarded as positive and secondary craters are regarded as negative and the number of negative samples is larger than the number of positive samples. It is commonly accepted that the diameter of the crater needs to be larger than 10 pixels when identifying and extracting attributes, otherwise the uncertainty resulting from artifacts and data accuracy may lead to unreliable conclusions [43]. Constrained by the resolution of DEM and remote sensing images, only craters with a diameter greater than 1 km are taken into consideration.

Though some studies have manually identified secondary craters [1,20,44], their databases only contain secondary craters and lack primary craters. Head et al. [3] proved that craters with diameters larger than 20 km are usually primary craters by statistically searching the density of significantly increased craters (>20 km) in annular zones of the Imbrium Basin and South Pole–Aitken Basin. Thus, combining existing secondary crater databases and selected craters with diameters larger than 20 km can create a new database that meets the basic demand for distinguishing secondary craters from primary craters. However, this new database is less representative, as it lacks small primary craters. A useful way to improve the representativity is adding additional samples to this database. To fully use previous study results and enhance data accuracy, the research region was set near Orientale Basin, covering an area centered at the basin and extended out to a radial distance of 6 radius (Figure 1). Samples used in this paper contain the following parts:

- Primary and secondary craters identified in this research, located within Orientale Basin (manual identification);
- Secondary craters identified by Guo et al. [1], located within the green circle and outside Orientale Basin;
- Randomly selected primary craters.
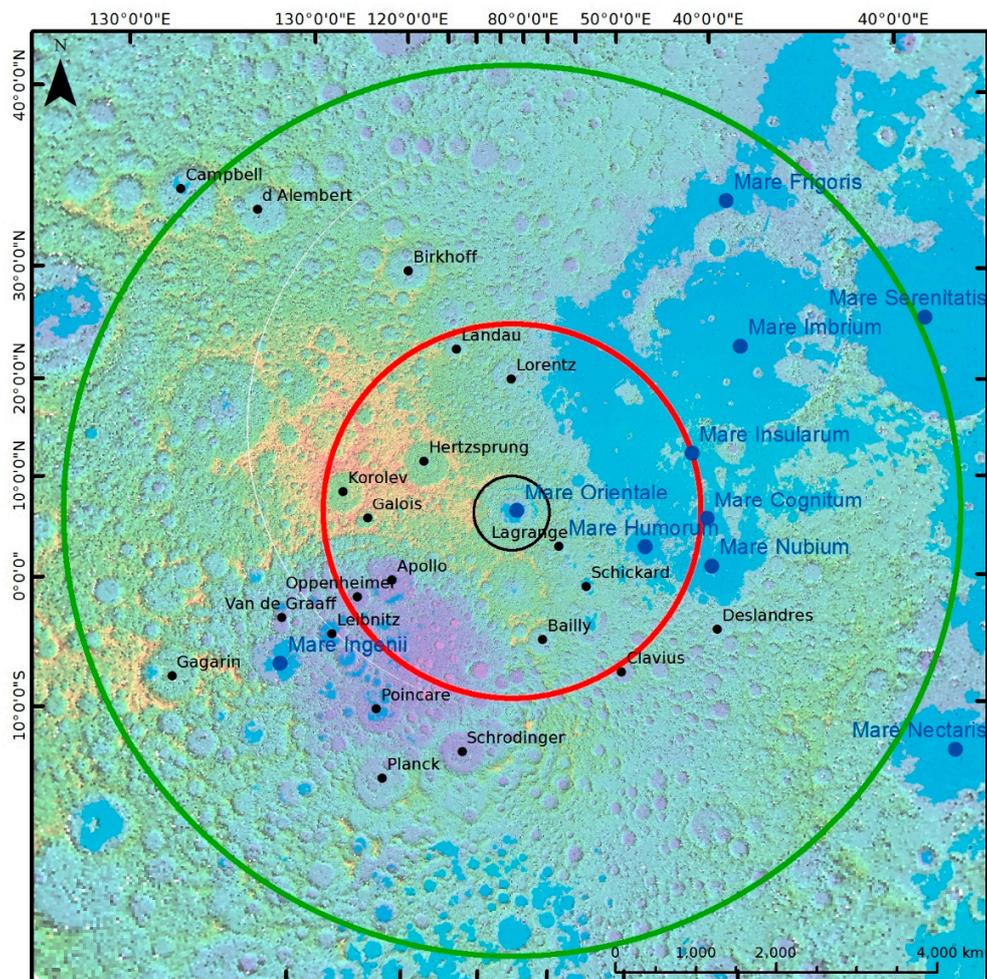
Figure 1. Sketch map showing the geomorphological setting of the area for machine learning training and testing. Background map is a hillshade image from Lunar Orbiter Laser Altimeter elevation data. Dark blue dots mark the crater and crater basin center with names labeled in dark blue. Black circle depicts the boundary of the Orientale Basin.

To keep data consistency, all the samples were obtained based on or unified with the lunar crater database published by Robbins [39].

Manual sample identification was conducted within the Orientale Basin. According to the definition, secondary craters are irregular, shallow, and elongate impact craters formed by fragments. Secondary and primary craters were distinguished based on the following five criteria, and finally 554 primary craters and 1420 secondary craters were identified:

- Secondary craters occur in chains (lines of regularly spaced rows of three or more with similar sizes) or clusters of 10 or more [1,14–16,45].
- Secondary craters are associated with a "herringbone" ejecta pattern (V-shaped ridge), which indicates their near-simultaneous formation during ejecta fragment deposition and points toward the parent crater [12,16–18,46].
- Secondary craters have an elliptical or irregular shape [15,16].
- Secondary craters show interference features such as septa and mounds [20].
- Secondary craters are usually shallower than primary impact craters with the same diameter [1,16,18,19].

In addition, 2632 secondary craters identified by Guo et al. [1] were added to the sample inventory. Guo et al. [1] identified a total of 2728 secondary craters of the Orientale Basin. These craters were

unified with Robbins' database and included in this inventory after examination. Only craters with similar diameters and locations in both databases were used in this study. Taking the total ratio into consideration, 478 craters with diameters ranging from 20 to 40 km were selected as a supplement. These craters were used after a careful check and excluded secondary craters. It should be noted that craters located on ramps were excluded from the catalog, as slope has a great influence on crater ellipticity and depth-to-diameter ratio. Also, this inventory may not contain the self-secondary craters, which have circular rims and dispersed spatial distribution.

A total of 5073 craters were identified and used in this study (Table 1). The statistics of the whole crater inventory are described in Table 2. The diameters of the primary and secondary craters range from 1.68 to 63.06 km and from 1.18 to 27.74 km, respectively. In order to fully evaluate the accuracy of the model, the crater inventory was partitioned into three subsets used for training and testing (Table 2). Craters located in lunar mare may show different characteristics compared with those in highland, so we allocated mare craters to testing data, aiming at testing whether the classification approach trained with highland craters can be applied to mare craters. One hundred and fifty-one craters located in maria regions were selected as Testing Dataset I, and they were obtained by manual identification and large primary crater selection. To test the performance of the proposed model on well-accepted data, we further divided the remaining part into two subsets, a training dataset obtained by all three ways and a testing dataset obtained by combining results from Guo et al. [1] and large primary craters (Figure 2). Testing Dataset II can be regarded as true data accepted by the public, and this dataset can minimize the errors caused by our personal identification. The training dataset contained 606 primary craters and 3037 secondary craters. To avoid overfitting, the portion of the two classes used for training should be similar. As secondary craters usually make up a large part of small craters in a certain area, the model performance on a dataset with a biased portion of secondary craters is much more representative than that on a dataset with a similar portion. A modified fivefold cross-validation was used in this paper, which ensured that the proportion of the two classes used for training was similar and that used for testing was different.

Table 1. Statistics of crater inventory.

Parameter	Value		
	Primary Craters	**Secondary Craters**	**Craters**
Count	1032	4041	5073
Mean	14.77	6.25	7.98
Standard deviation	12.93	4.55	7.89
Minimum	1.18	1.18	1.18
25th percentile	2.31	1.90	2.01
Median	8.50	5.66	5.72
75th percentile	25.80	8.94	10.08
Maximum	63.06	27.74	63.06

Table 2. Classification of crater inventory.

Classification		Training Dataset	Testing Dataset I	Testing Dataset II 2	Sum
Way 1	Primary craters	510	44	0	554
	Secondary craters	1331	89	0	1420
Way II	Primary craters	0	0	0	0
	Secondary craters	1706	0	915	2621
Way III	Primary craters	96	18	364	478
	Secondary craters	0	0	0	0
Sum	Primary craters	606	62	364	1032
	Secondary craters	3037	89	915	4041
	Craters	3643	151	1279	5073
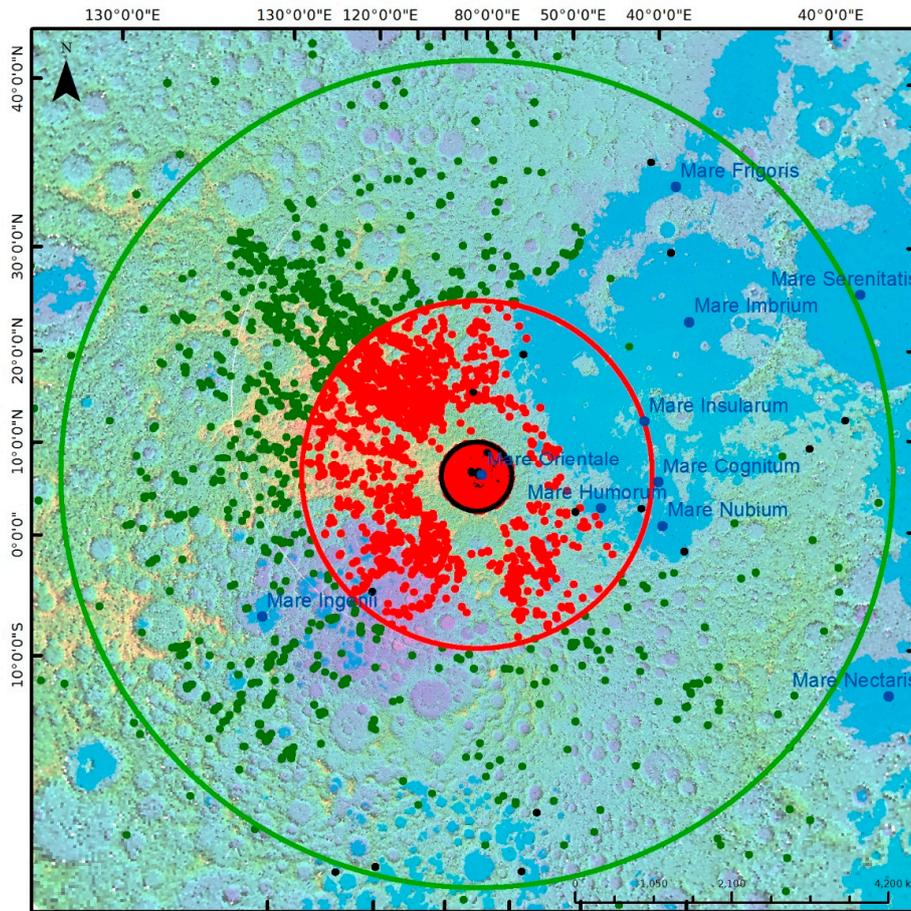
Figure 2. Sketch map showing samples of different datasets. Red points represent samples in the training dataset, black points show samples in Testing Dataset I, and green points mark samples in Testing Dataset II. Black circle depicts the boundary of the Orientale Basin. Red and green circles denote the region at the center of the basin and extending out to a radial distance of 3.5 radius (**R**) and 6 R, respectively.

The number of primary craters decreases with increased diameter, and the distribution of secondary craters is the same. However, for primary craters, the number decreases in two intervals, from 1 km to 20 km and from 20 km to 64 km. This is because we manually added a few large primary craters (diameter larger 20 km). This inconstant distribution of primary craters does not affect the precision of the proposed method very much, as these craters are distributed all over the moon surface and the number in the training area is not so large (Figure 3).
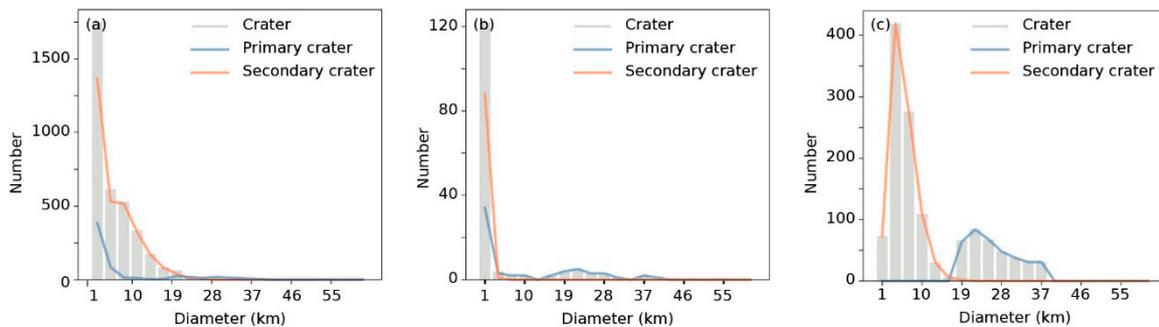
Figure 3. Diameter distribution of craters in three datasets: (**a**) training dataset, (**b**) Testing Dataset I, (**c**) Testing Dataset II. Geometric scale of bins is the fourth root of 2, and the leftmost edge represents the smallest diameter, 1 km. This figure uses the same geometric scale of bins when involving diameter.

3. Method

3.1. Overview

Figure 4 presents a flowchart of the machine learning approach. Since secondary craters are generally elliptical and shallower than primary craters, ellipticity and depth–diameter ratio are mostly used in distinguishing the two. In order to fully detect and use these differences, we developed a machine learning approach containing several kinds of features using random forest that needs to be trained before use. The approach begins extracting features. Then, a training process is conducted. Finally, each crater is recorded with a number that describes exactly how it can be a primary crater. Accuracy assessment is conducted with manually labeled results. The pseudocode of the whole algorithm could be found in Appendix A.
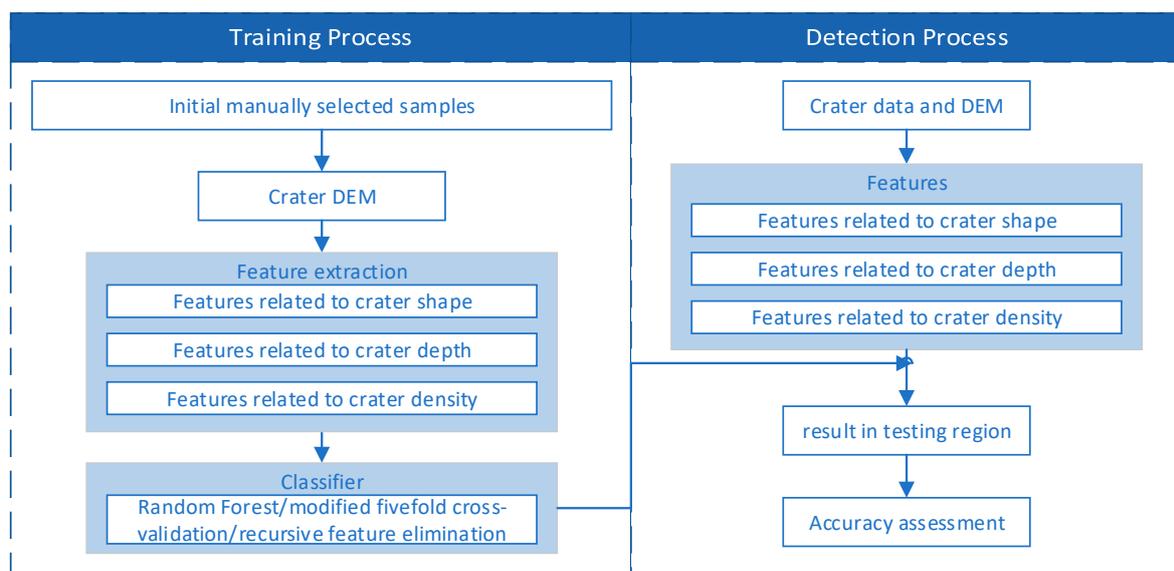
Figure 4. Workflow of this study. DEM, digital elevation model.

3.2. Features

To distinguish primary craters from secondary craters, we analyzed features that can be used. Features such as irregularity and eccentricity describe crater shape and are commonly used to quantify the differences between primary and secondary craters. Besides, different incident velocities result in different crater depths, thus features related to crater depth could well express the difference too. Primary and secondary craters differ not only in their own features but also in density, as the existence of chains or clusters affects crater density a lot. For this reason, features characteristics by crater density should also be considered. Samples for this study consist of 32 features, reflecting crater shape, depth, and density.

3.2.1. Features Related to Crater Shape

Features related to crater shape used in this study include irregularity (Irr), eccentricity (Ecc), and rim integrity (R_i). Irregularity and eccentricity have proven to be useful in distinguishing primary craters from secondary craters. Irregularity is usually defined as the ratio of the crater perimeter to the perimeter of a circle whose area is the same as the crater [47]. For the convenience of calculation, irregularity in this paper is defined as the difference between the boundary and the fit circle of a certain crater. In Robbins' crater database, a feature named DIAM_CIRC_SD_IMG was calculated and defined as the standard deviation of kilometers of the fit residuals [39]. Each manual rim point's distance from the crater center is calculated and subtracted from the best-fit radius, and this value is the

standard deviation of those differences [39]. Irregularity used in this study is derived from Robbins and defined as

$$\text{Irr} = \frac{\sqrt[2]{\frac{\sum_{i=1}^{n}(r_i-R)^2}{n}}}{R}, \tag{1}$$

where R is the radius from a circle fit, r_i is the distance between a manual rim point and the crater center, and n is the number of manual rim points. The closer the irregularity is to 0, the more likely the crater is a primary crater.

Eccentricity used in this paper is provided by Robbins in his [39]. It is calculated as

$$\text{Ecc} = \sqrt[2]{\left(1 - \frac{b^2}{a^2}\right)}, \tag{2}$$

where a is the major axis from an ellipse fit and b is the minor axis. The closer the eccentricity is to 1, the closer the crater shape is to an ellipse, and the more likely the crater is a secondary crater.

Rim integrity (R_i) is an estimation of the fraction of the complete rim that was traced, and can also be obtained directly from Robbins' lunar crater database [39]. Though this feature seems to have no direct relationship to the difference between primary and secondary craters, it implies the degree of boundary destruction and, further, the accuracy of other features related to the crater rim. Thus, we consider rim integrity as a feature included in this method.

3.2.2. Features Related to Crater Depth

The selected features in this study related to crater depth include features describing the standard deviation of rim elevation and depth-to-diameter ratio. Unlike the features mentioned above, features related to crater depth usually describe craters from a three-dimensional perspective and can be derived from DEM data. In a depth-to-diameter ratio, diameter refers to the distance from the crater center to the rim, and depth refers to the difference between the minimum elevation within the crater and the average rim height [48]. For ease of batch calculation, the calculation of the depth-to-diameter ratio can be simplified by taking the diameter from a circle fit as the diameter and the difference in elevation between the average fitted circle height and the deepest point within the crater as depth [2]. This simplified calculation is easy to apply, as the output of most crater databases or crater detection approaches is a set of circles describing the rim of the crater. However, this simplified calculation also results in uncertain accuracy of the depth-to-diameter ratio, as the fitted circle and crater rim do not coincide completely and the rim may be destroyed at different levels. To reduce the uncertainty caused by the precision of the fitted circle, we further consider calculating the depth-to-diameter ratio based on an ellipse fit. The standard deviation of the height of the fitted line can also serve as a supplement, as it represents the height difference caused by the ruined rim.

Features related to crater depth used in this study are the standard deviation of fitted circle height (STD_d_c), the standard deviation of fitted ellipse height (STD_d_e), the fitted circle depth-to-circle diameter (d_c/D), the fitted ellipse depth-to-ellipse major axis (d_e/A_{maj}), and the fitted ellipse depth-to-ellipse minor axis (d_e/A_{min}). It should be noted that the merged LRO and Kaguya DEM spans $60°$ in latitude, thus for craters within that latitude, their features are calculated based on the merged DEM, and the calculation of features of remaining craters is based on Lunar Orbiter Laser Altimeter topography data. The standard deviation of a fitted circle/ellipse can be obtained by overlapping DEM data and is formulated as follows:

$$STD\_d_c = \sqrt[2]{\frac{\sum_{i=1}^{n}\left(hc_i - \overline{hc}\right)^2}{n_c}}, \tag{3}$$

$$STD\_d_e = \sqrt[2]{\frac{\sum_{i=1}^{n}\left(he_i - \overline{he}\right)^2}{n_e}}, \tag{4}$$

where hc_i represents the DEM value of each pixel falling on the fitted circle; n_c represents the number of these pixels; $\overline{hc}$ represents the average of hc_i; he_i represents the DEM value of each pixel falling on the fitted ellipse; n_e represents the number of these pixels; and $\overline{he}$ represents the average of he_i.

The fitted circle's depth-to-circle diameter (d_c/D), depth-to-ellipse major axis (d_e/A_{maj}), and depth-to-ellipse minor axis (d_e/A_{min}) can be calculated as

$$d_c/D = \frac{\overline{hc} - hc_{min}}{D}, \tag{5}$$

$$d_e/A_{maj} = \frac{\overline{he} - he_{min}}{a}, \tag{6}$$

$$d_e/A_{min} = \frac{\overline{he} - he_{min}}{b}, \tag{7}$$

where hc_{min} is the minimum value of all DEM pixels falling inside the fitted circle; D is the diameter of the fitted circle; he_{min} is the minimum value of all DEM pixels falling inside the fitted ellipse; and a and b are the major and minor axes of the fitted ellipse, respectively.

3.2.3. Features Related to Crater Density

Features related to crater density are mainly used to describe crater distribution, as special distribution patterns of secondary craters deviate from the uniform distribution of primary craters. Here, 24 features belonging to 4 groups are designed to express crater patterns, with 12 features aimed at describing chains and 12 describing clusters. A secondary crater chain is a line of regularly spaced rows of 3 or more secondary craters with similar sizes, thus a secondary crater belonging to a chain may have statistically significant increased density in a certain direction. Based on the above considerations, we designed 12 features calculating crater number and density in different areas. Features belonging to chain group I (Chain_I) and chain group II (Chain_II) are defined as follows:

$$Chain\_I = [Ch\_I_1, Ch\_I_2, Ch\_I_3, Ch\_I_4, Ch\_I_5, Ch\_I_6], \tag{8}$$

$$Ch\_I_i = NChs_i, \tag{9}$$

$$Chain\_II = [Ch\_II_1, Ch\_II_2, Ch\_II_3, Ch\_II_4, Ch\_II_5, Ch\_II_6], \tag{10}$$

$$Ch\_II_i = \frac{NChs_i}{NCh_i}, \tag{11}$$

where Ch_I_i and Ch_II_i (i = 1, 2, 3, 4, 5, 6) represent features consisting of Chain_I and Chain_II, respectively; $NChs_i$ (i = 1, 2, 3, 4, 5, 6) represents the number of craters with similar sizes in the corresponding region; and NCh_i (i = 1, 2, 3, 4, 5, 6) represents the number of all craters in the corresponding region (Figure 5a). For each crater to be identified, the total area used for calculating Chain_I and Chain_II covers an area beginning at the center and extending out to a radial distance of 6 R (R is the radius of the crater). The extending distance of the counting area is set as 6 R, as a chain contains at least 3 craters with similar sizes and there may be distance between them. The total area is further divided into 6 regions in which crater counting is conducted, and the counting result in each region is regarded as a feature (Figure 5a).

A secondary crater cluster contains 10 or more secondary craters with similar sizes, and a secondary crater belonging to a cluster may have statistically significant increased density in a certain area. Similar to the features describing crater chains, we also designed 12 features calculating crater number and density. Cluster group I (Cluster_I) and cluster group II (Cluster_II) are defined as follows:

$$\text{Cluster\_I} = [Cl\_I_1 , Cl\_I_2 , Cl\_I_3 , Cl\_I_4 , Cl\_I_5 , Cl\_I_6], \tag{12}$$

$$Cl\_I_i = NCls_i, \tag{13}$$

$$\text{Cluster\_II} = [Cl\_II_1 , Cl\_II_2 , Cl\_II_3 , Cl\_II_4 , Cl\_II_5 , Cl\_II_6], \tag{14}$$

$$Cl\_II_i = \frac{NCls_i}{NCl_i}, \tag{15}$$

where Cl_I_i and Cl_II_i ($i = 1$, 2, 3, 4, 5, 6) represent features consisting of Cluster_I and Cluster_II, respectively; $NCls_i$ ($i = 1$, 2, 3, 4, 5, 6) represents the number of craters with similar sizes in the corresponding region; and NCl_i ($i = 1$, 2, 3, 4, 5, 6) represents the number of all craters in the corresponding region (Figure 5b). For each crater to be identified, the total area used for calculating Cluster_I and Cluster_II is same as that used for calculating Chain_I. The extending distance of counting area is set as 6 R, as a crater cluster contains at least 10 craters with similar sizes. The 6 regions used for crater counting are areas beginning at the center and extending out to different distances. From Region 1 to Region 6, the extending distance ranges from R to 6 R (Figure 5b). In this paper, craters with similar sizes means that the variation of their diameters is within 20%, a little smaller than that set by Wu [28], and all the calculation associated with counting craters is conducted based on Robbins' lunar crater database.
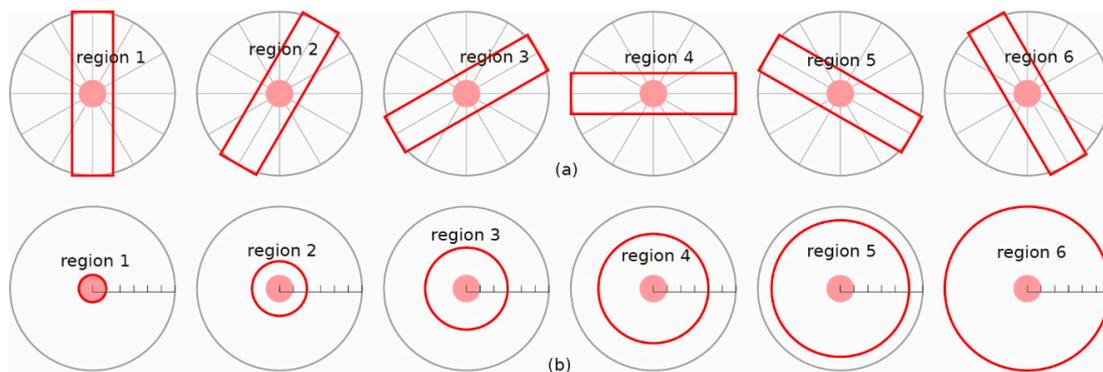
Figure 5. Sketch maps showing corresponding regions used for calculating features related to crater density related to (**a**) crater chain and (**b**) crater cluster. Light red circles represent craters whose features are calculated. Dark red regions labeled 1, 2, 3, 4, 5, 6 mark corresponding regions for calculating Ch_I_i, $Ch\_II_i, Cl\_I_i,$ and Cl_II_i. Gray circles show regions at the centers of craters and extended out to a radial distance of 6 R (R is crater radius).

3.3. Description of Classifiers

Among all the methods in machine learning, the random forest method is thought to have the advantages of simplicity, high accuracy, and an avoidance of overfitting [49]. In addition, the random forest classifier is good at processing data of higher dimensions and can calculate the importance of each feature after the training period. Moreover, the random forest classifier has been applied successfully in many research fields [50] and has been introduced into study related to geomorphology [51]. Besides, we conducted a preliminary test of common machine learning classifiers with our samples, including random forest, support vector machine, and adaptive boosting, and found that among these three, random forest had the best performance. Thus, the random forest classifier was chosen to form the proposed algorithm.

Random forest classifier is a combination of tree predictors [49]. The basic idea of random forest is to combine several weak classifiers (tree predictors) to form a strong classifier (random forest classifier). The process of combining weak classifiers reduces the impact of a single classifier error, thus the classification accuracy and stability of the strong classifier can be improved. Decision trees are the foundation of a random forest. The main components of a decision tree model are nodes and

branches, and the most important steps in building a model are splitting and stopping [52]. Nodes can be further divided into root nodes, internal nodes, and lead nodes according to their different locations in a decision tree. Branches represent chance outcomes or occurrences that emanate from nodes, and this process can be regarded as splitting parent nodes into child nodes. The splitting process would be stopped if stopping rules are met, such as tree depth or minimum number of records in a leaf. Parameters that may influence the performance of a decision tree model include the maximum number of features used for splitting, the maximum depth of the decision tree, the minimum number of samples in a leaf node, and the minimum number of samples to split [53,54]. Among all the parameters, the minimum number of samples to split and the maximum depth of the decision tree are two of the most important that affect the performance of the model and should be seriously considered when adjusting parameters [55–57]. A random forest classifier can model a series of different decision tree models, and each decision tree model would be trained based on its own training samples and features, which are selected randomly from all of the training samples and features. This is different from the training process of a traditional decision tree model, as the traditional process uses all of the training samples and features as input data. The result of a random forest classifier is obtained by calculating the results of each decision tree, and the mode number among the results is regarded as the result. Figure 6 shows a random forest classifier model. Except for the parameters affecting the performance of the decision tree model, the maximum number of decision tree models contained in a random forest model is another parameter that has a great impact on model performance.
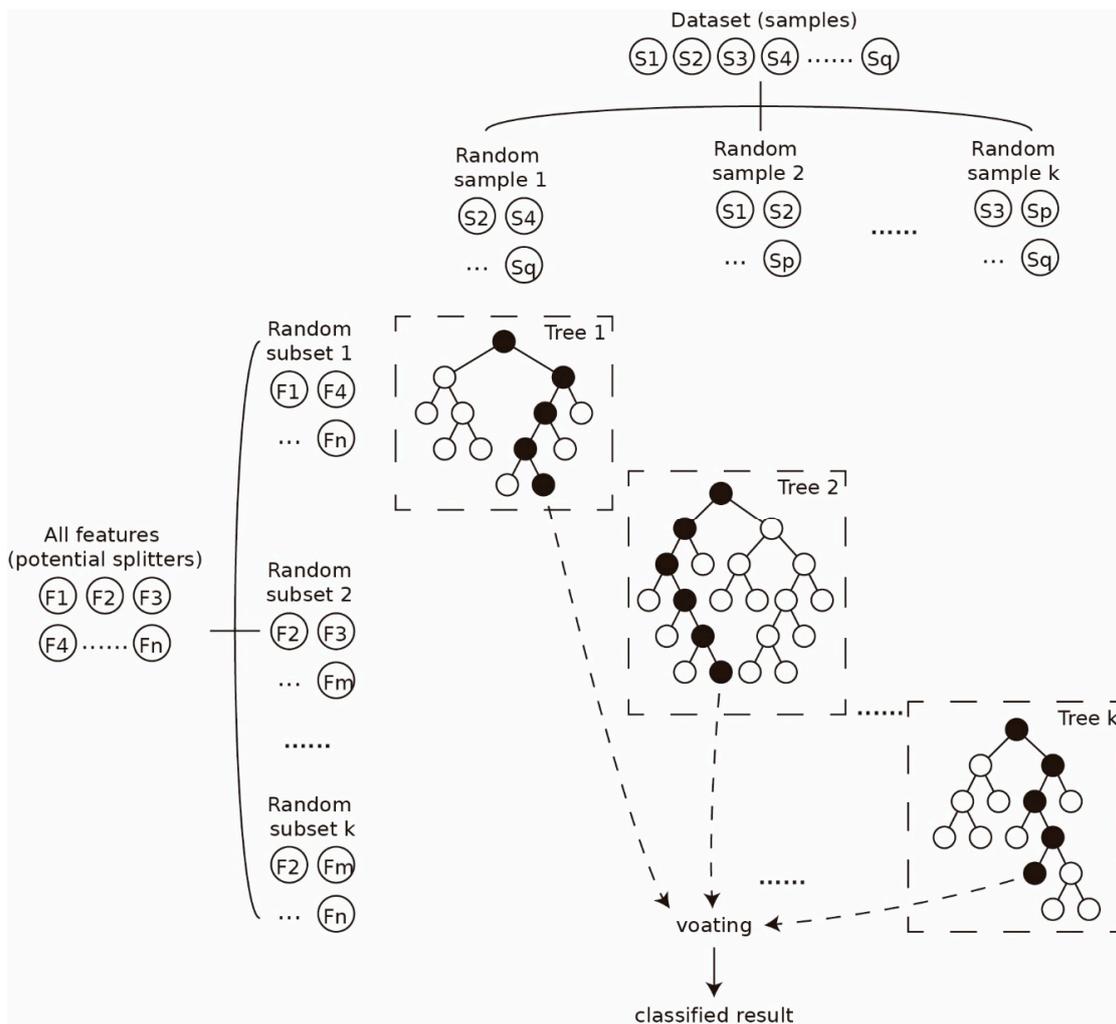
Figure 6. Sketch map showing a random forest classifier.

A modified fivefold cross-validation was used in training process (Figure 7). The usual fivefold cross-validation means that a dataset is randomly divided into five subsets and each subset is made a verification set, and the remaining four groups of subset data are used as training sets. In this paper, we manually added some negative samples to the subset used for testing, which could help test the performance on a dataset with a biased portion. First, 606 secondary craters in the training dataset were randomly selected and combined with the 606 primary craters forming the preliminary dataset. Then, this dataset and the remaining 2431 secondary craters were randomly divided into five subsets. Each subset combined with a secondary crater subset was made a verification set, and the remaining four groups of subset data were used as training sets.
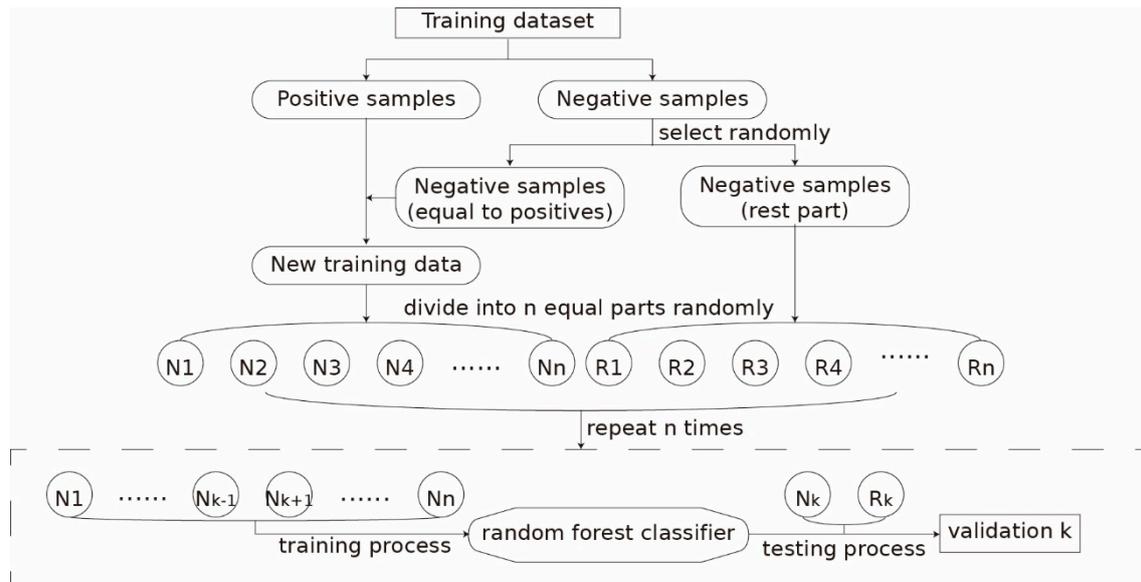
Figure 7. Sketch map showing a modified nfold cross-validation.

The proposed algorithm was implemented based on the scikit-learn package in Python. Model parameter adjustment is a necessary process for better model performance. To find the best parameter settings, we changed the maximum number of decision trees from 100 to 500 in steps of 10, the maximum number of features used for splitting among the total feature number, base 2 logarithm of total feature number, and the square root of the total feature number and minimum number of samples to split from 2 to 10 in steps of 1. The random forest classifier has its best performance when the maximum number of decision trees is set to 350, the maximum number of features used for splitting is set to the square root of the total feature number, and the minimum number of samples to split is set to 2.

A feature selection process was done through recursive feature elimination. The recursive feature elimination can select the model with best performance for a setting model feature number by building a model repeatedly. Each round, the feature having worst importance is eliminated and then the process is repeated on the remaining features until all features are traversed or reach a setting feature number. With the setting feature number changing from 1 to 32, 32 subsets of features were selected and each of them was the best feature combination for a given feature number. After testing model performance, a combination of 29 features was selected as the final features.

3.4. Accuracy Assessment

As an efficient tool, a confusion matrix can describe the relationship between detection results and true values, show the number of correct and incorrect detections directly, and help calculate other quantitative criteria. Metrics used to evaluate the goodness of fit for the proposed approach included sensitivity, precision, accuracy, the F1-score, and the kappa coefficient. They are formulated as follows:

$$Sensivity = \frac{TP}{TP+FN} \text{,} \tag{16}$$

$$Precision = \frac{TP}{TP+FP} \text{,} \tag{17}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \text{,} \tag{18}$$

$$F1-score = 2*\frac{Precision \times Sensivity}{Precision + Sensivity} \text{,} \tag{19}$$

$$Kappa = \frac{p_o - p_e}{1 - p_e} \text{,} \tag{20}$$

$$p_o = \frac{TP+TN}{TP+TN+FP+FN} \text{,} \tag{21}$$

$$p_e = \frac{(TP+FP)\times(TP+FN)+(TN+FN)\times(TN+FP)}{(TP+TN+FP+FN)^2} \text{,} \tag{22}$$

where *TP*, representing true positive, is the number of correct positive detections; *TN*, representing true negative, is the number of correct negative detections; *FN*, representing false negative, is the number of incorrect positive detections; and *FP*, representing false positive, is the number of incorrect negative detections. A higher kappa coefficient means better results. The kappa coefficient is 0.6–0.8 and 0.8–1, representing a substantial and almost perfect agreement between the estimation and observation, respectively [58].

4. Experimental analysis

4.1. Feature Distribution Analysis

4.1.1. Features Related to Crater Shape

The irregularities of the identified craters are shown in Figure 8 and Table 3. The irregularities of primary craters range from 0.020 to 0.057, and of secondary craters range from 0.028 to 0.148. About 75% of primary craters have an irregularity under 0.25, though the max is 0.57. As for secondary craters, nearly half of them are above 0.25, and 10% above 0.5. Figure 8a,b shows the irregularities of identified craters using frequency distribution histograms, in which blue represents primary craters and orange represents secondary craters. These two plots also show that the peaks in the two histograms are all skewed toward lower irregularities. We calculated the mean irregularities of craters in different diameter bins (Figure 8c). It is obvious that secondary craters usually have higher mean irregularities compared with primary craters with similar diameters. Besides, the confidence intervals for the mean irregularity of primary craters and secondary craters also have different change trends. For craters with diameters smaller than 5.7 km, the range of confidence interval for mean irregularity of primary craters is smaller than that of secondary craters. A small number of samples may account for the large confidence interval of craters with diameters larger than 10 km.

The eccentricities of identified craters are shown in Figure 9 and Table 4. Figure 9a,b shows the eccentricities of identified craters using frequency distribution histograms, in which blue represents primary craters and orange represents secondary craters. The eccentricities of primary craters range from 0.06 to 0.62, and of secondary craters range from 0.07 to 0.94. The mean eccentricity of primary craters is 0.36, a little smaller than that of secondary craters (Table 4). The eccentricity distributions of primary and secondary craters are different, which can be seen from the mean and standard deviation, though they both approximate normal distributions. In general, the peak in the histogram of primary craters is skewed toward lower eccentricities. Different from primary craters, the peak in the histogram of secondary craters is skewed toward middle eccentricities. The frequency distribution histogram

results indicate that craters with eccentricity larger than 0.6 are more likely to be identified as secondary craters. Figure 9c shows that secondary craters usually have higher mean eccentricities compared with primary craters with similar diameters, and the eccentricity difference between primary and secondary craters decreases with increased diameter. It also shows that mean eccentricity decreases with increased diameter, and this is in agreement with the results of Guo [1].
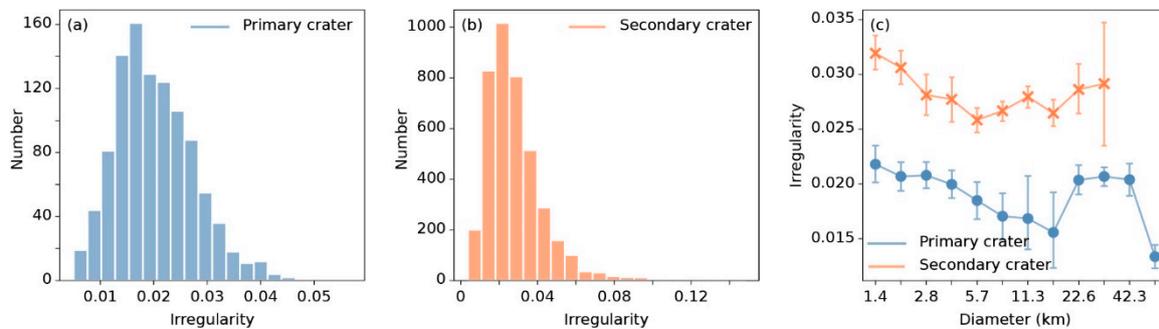
Figure 8. Histograms of irregularity distribution of (**a**) primary craters and (**b**) secondary craters, and (**c**) the relationship between the statistics irregularity and diameter. In (c), the points represent mean irregularity. Blue lines and points represent primary craters, and orange lines and points represent secondary craters.

Table 3. Statistics of irregularity.

Parameter	Value		
	Primary Craters	**Secondary Craters**	**Craters**
Count	1032	4041	5073
Mean	0.020	0.028	0.027
Standard deviation	0.007	0.015	0.014
Minimum	0.005	0.004	0.004
25th percentile	0.015	0.018	0.017
Median	0.019	0.025	0.024
75th percentile	0.025	0.035	0.032
Maximum	0.057	0.148	0.148
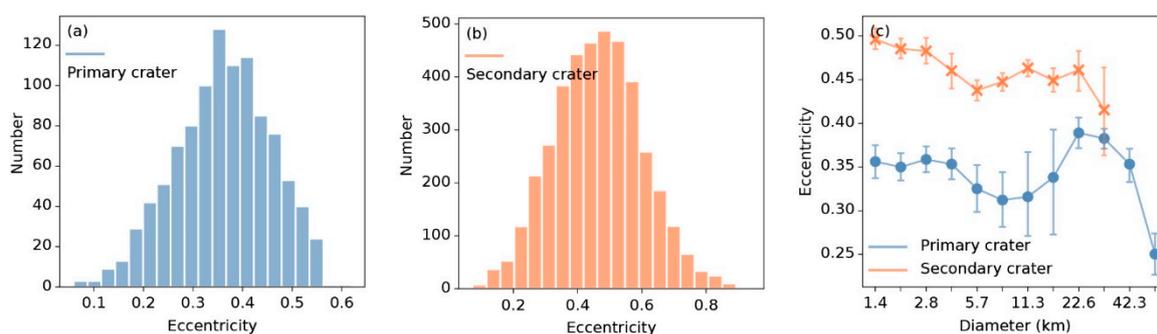
Figure 9. Histograms of eccentricity distribution of (**a**) primary craters and (**b**) secondary craters, and (**c**) the relationship between the statistics eccentricity and diameter. In (c), points represent mean eccentricity. Blue lines and points represent primary craters, and orange lines and points represent secondary craters.

Table 4. Statistics of eccentricity.

Parameter	Value		
	Primary Craters	**Secondary Craters**	**Craters**
Count	1032	4041	5073
Mean	0.36	0.46	0.44
Standard deviation	0.10	0.14	0.14
Minimum	0.06	0.07	0.06
25th percentile	0.30	0.37	0.35
Median	0.37	0.46	0.44
75th percentile	0.43	0.56	0.53
Maximum	0.62	0.94	0.94

The rim integrity (R_i) of identified craters is shown in Figure 10 and Table 5. Figure 10a,b shows the rim integrity of identified craters using frequency distribution histograms, in which blue represents primary craters and orange represents secondary craters. The mean rim integrity of primary craters is 0.87, a little higher than that of secondary craters (Table 5). The rim integrity distributions of primary and secondary craters are different, which can be seen from the mean and 75th percentile rim integrity, though the number of primary and secondary craters increases with increased rim integrity. It can be seen that the distribution of rim integrity for primary and secondary craters generally shows the same trends. Yet the frequency distribution histogram of primary craters indicates that nearly half of them have a rim integrity equal to 1, and only a few have rim integrity smaller than 1. As for secondary craters, though the peak in the histogram is skewed toward higher rim integrity, with increased rim integrity, the number of primary craters increases sharply when rim integrity is higher than 0.7. In other words, in Figure 10b, no significant peak is observed in the interval from 0.7 to 1. Figure 10c shows that compared with secondary craters, primary craters of the same size usually have higher rim integrity. A small number of samples may account for the low rim integrity and large confidence interval of primary craters with diameters larger than 22 km.
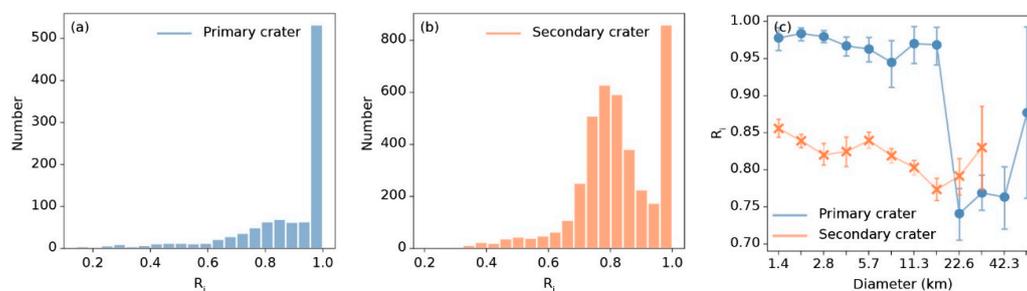
Figure 10. Histograms of rim integrity distribution of (**a**) primary craters and (**b**) secondary craters, and (**c**) the relationship between the statistics boundary integrity and diameter. In (c), points represent mean rim boundaries. Blue lines and points represent primary craters and orange lines and points represent secondary craters.

Table 5. Statistics of boundary integrity.

Parameter	Value		
	Primary Craters	**Secondary Craters**	**Craters**
Count	1032	4041	5073
Mean	0.87	0.82	0.83
Standard deviation	0.18	0.13	0.15
Minimum	0.14	0.20	0.14
25th percentile	0.80	0.75	0.75
Median	0.98	0.81	0.83
75th percentile	1.00	0.92	1.00
Maximum	1.00	1.00	1.00

4.1.2. Features Related to Crater Depth

Features related to crater depth used in this study are the standard deviation of fitted circle height (STD_d_c), the standard deviation of fitted ellipse height (STD_d_e), the fitted circle's depth-to-diameter ratio (d_c/D), the fitted ellipse's depth to major axis (d_e/A_{maj}), and the fitted ellipse's depth to minor axis (d_e/A_{min}), shown in Figure 11. Figure 11a,b shows the average standard deviation of fitted circle height and fitted ellipse. Blue lines and points represent primary craters, and orange lines and points represent secondary craters. In general, the average STD_d_c and STD_d_e increase with increased crater diameter, and secondary craters usually have a slightly higher average STD_d_c and STD_d_e compared with primary craters of the same size. Figure 11c shows the distribution of d_c/D with respect to crater diameter. It indicates that before 20 km, the line of primary craters lies above that of secondary craters, which means at least the average d_c/D of primary craters is higher than that of secondary craters. However, for craters with diameters larger than 20 km, the mean d_c/D of primary craters declines quickly and seems lower than that of secondary craters. Line diagrams of the difference between d_e/A_{maj} and d_c/D of different crater classes as a function of diameter are shown in Figure 11d. Figure 11e shows the difference between d_e/A_{min} and d_c/D of different crater classes as a function of diameter. Both figures show that differences between d_e/A_{maj} ($d_e/A_{\min}$) and d_c/D of primary craters slowly decline with increased diameter, which show an opposite trend to that of secondary craters. The differences between d_e/A_{maj} ($d_e/A_{\min}$) and d_c/D of primary craters are usually smaller than those of secondary craters of the same size, and the diversity between primary and secondary craters enlarges with increased diameter.
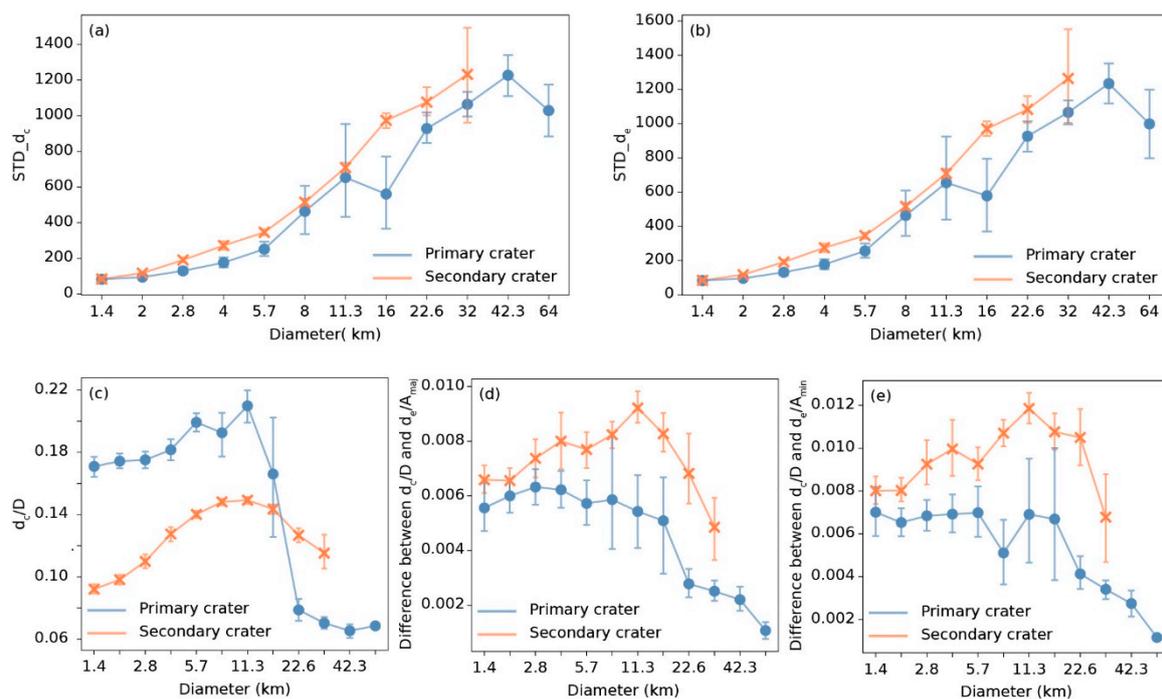
Figure 11. Relationship between diameter and (**a**) standard deviation of fitted circle height (STD_d_c), (**b**) standard deviation of fitted ellipse height (STD_d_e), (**c**) fitted circle's depth-to-diameter ratio (d_c/D), (**d**) difference between fitted circle's depth-to-diameter ratio (d_c/D) and fitted ellipse depth to major axis (d_e/A_{maj}), (**e**) difference between fitted circle depth-to-diameter ratio (d_c/D) and fitted ellipse depth to minor axis ($d_e/A_{\min}$). Blue lines and points represent primary craters, and orange lines and points represent secondary craters.

4.1.3. Features Related to Crater Density

According to the definition of a secondary crater chain, a secondary crater belonging to a chain may have statistically significant increased density in a certain direction, and this means that for

a crater in a chain, the range of its parameters consisting of chain group I (Chain_I) and chain group II (Chain_II) may be larger than that of a primary crater. The mean range of Chain_I for primary craters is 1.18, smaller than that of secondary craters (Table 6). A lower range of crater count in different azimuths means more randomness of crater distribution, thus it less likely to consist of a secondary crater chain. Figure 12a,b shows the distribution of the range of parameters in Chain_I for different diameters and crater classes. Figure 12c,d shows the distribution of the range of parameters in Chain_II. Blue boxes represent primary craters and orange boxes represent secondary craters. For craters with diameters larger than 10 km, the range of Chain_I of secondary craters is seldom smaller than 1.5, but for primary craters, nearly 25% have a range larger than 1.5. Compared with primary craters, the box diagram of secondary craters shows more abnormal values, which means that the ranges of Chain_I and Chain_II of secondary craters present a more dispersed distribution, and this is also consistent with our perception.

Table 6. Statistics of the range of Chain_I.

Parameter	Value		
	Primary Craters	**Secondary Craters**	**Craters**
Count	1032	4041	5073
Mean	1.18	1.84	1.71
Standard deviation	1.24	1.33	1.34
Minimum	0.00	0.00	0.00
25th percentile	0.00	1.00	1.00
Median	1.00	2.00	2.00
75th percentile	2.00	3.00	3.00
Maximum	7.00	9.00	9.00
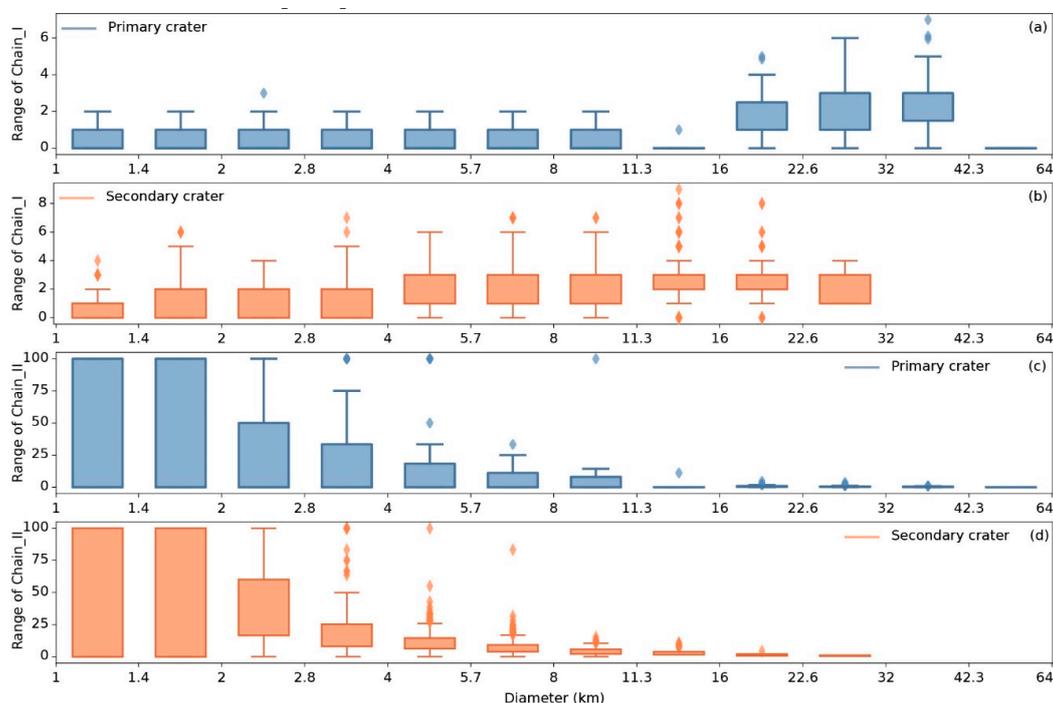
Figure 12. Relationship between diameters and (**a**) ranges of Chain_I of primary craters, (**b**) ranges of Chain_I of secondary craters, (**c**) ranges of Chain_II of primary craters, (**d**) ranges of Chain_II of secondary craters. Blue boxes represent primary craters and orange boxes represent secondary craters.

Cluster group I (Cluster_I) and cluster group II (Cluster_II) are used to describe the aggregation level of craters with a similar degree in a given area and can represent the possibility that they make up a secondary cluster. According to the definition of a secondary crater cluster, a secondary crater

belonging to a cluster may have statistically significant increased density in a certain area, and this means that for a crater in a cluster, the standard deviation of its parameters consisting of cluster group I (Cluster_I) and cluster group II (Cluster_II) may be larger than that of a primary crater. In general, the standard deviation of Cluster_I of primary craters is smaller than that of secondary craters, with the former ranging from 0 to 9.29 and the latter ranging from 0 to 10.02 (Table 7). Box diagrams of standard deviations of Cluster_I and Cluster_II as a function of crater diameter are shown in Figure 13a,b. Among primary craters, the standard deviation of Cluster_I changes slowly with increased diameter. For secondary craters, the crater number first keeps stable with increased standard deviation, and then increases quickly. By contrasting Figure 13a and 13b, we can see that with diameters ranging from 2.8 to 16 km, secondary craters generally have a higher standard deviation of the six parameters in Cluster_I. Figure 13c,d shows that for craters with diameters smaller than 4 km, the standard deviation of Cluster_II of primary craters ranges from 0 to 40, while that of secondary craters ranges from 15 to 40. This can imply that small craters with a standard deviation of vector 2 Cluster_II of more than 15 are more likely to be primary craters.

Table 7. Statistics of standard deviation for Cluster_I.

Parameter	Value		
	Primary Craters	**Secondary Craters**	**Craters**
Count	1032	4041	5073
Mean	1.69	2.59	2.40
Standard deviation	1.72	1.86	1.87
Minimum	0.00	0.00	0.00
25th percentile	0.37	1.00	0.76
Median	1.07	2.34	2.11
75th percentile	2.63	3.80	3.64
Maximum	9.29	10.02	10.02
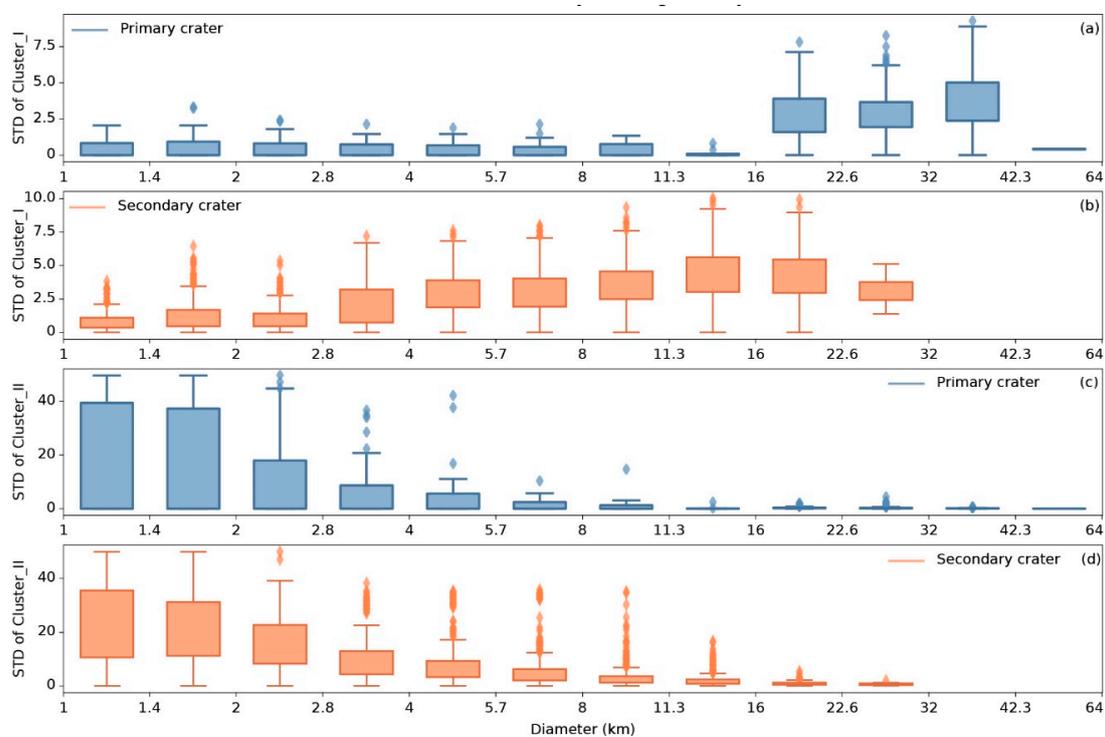
Figure 13. Relationship between diameter and standard deviation (STD) of (**a**) Cluster_I 1 for primary craters, (**b**) Cluster_I for secondary craters, (**c**) Cluster_II for primary craters, (**d**) Cluster_II for secondary craters. Blue boxes represent primary craters and orange boxes represent secondary craters.

4.2. Model Validation

Modified fivefold cross-validation on the training dataset was conducted. The performance of a model is given by the statistical parameters. Figure 14 and Table 8 show the training and fivefold cross-validation testing results. A comparison of the training and testing metrics indicates a clear decrease in accuracy and sensitivity, which indicates overfitting of the models with the training data and that further model validation is necessary. The cross-validation testing dataset has 727 craters, containing 121 primary and 606 secondary craters. Among all the craters, 115 primary and 568 secondary craters were identified correctly, 38 secondary craters were wrongly regarded as primary craters, and 6 primary craters were incorrectly regarded as secondary craters. The evaluation results of the fivefold cross-validation are listed in Table 8. The results show that the trained model had higher sensitivity compared by precision (0.950 and 0.752, respectively). This means that through this approach, most primary craters can be truly predicted but some miscalculation of secondary craters leads to decreased precision. This may be because nearly five times the number of secondary craters than primary craters were used in testing. But we think this is in line with the actual situation, as secondary craters may usually have a larger population than primary craters, especially among small craters. Additionally, this model had a high kappa coefficient of 0.803, signifying substantial consistency between prediction and observation. Figure 14 shows the classification results by enlarging three regions in the cross-validation testing dataset.

Table 8. Statistics of model performance.

	Precision	Sensitivity	Accuracy	F1-score	Kappa
Training dataset	1	1	1	1	1
Testing dataset	0.752	0.950	0.939	0.839	0.803

The evaluation results of Testing Dataset I are listed in Table 9. Testing dataset I consists of 151 craters, 62 primary and 89 secondary craters. Among all the craters, 55 primary and 82 secondary craters were identified correctly, 7 primary craters were wrongly regarded as secondary craters, and 7 secondary craters were mistaken as primary craters (Figure 15). Unlike the results of fivefold cross-validation, the precision of Testing Dataset I was same as sensitivity. This means that nearly 88% of craters predicted as primary craters by this approach are actual primary craters. Although there may be some primary craters mistaken as secondary craters, the accuracy indicates that nearly 90 percent of craters can be correctly identified, which proves that this approach also performs well in lunar mare.

Table 9. Statistics of Testing Dataset I.

	Precision	Sensitivity	Accuracy	F1-score	Kappa
Testing Dataset I	0.887	0.887	0.907	0.887	0.808

The evaluation results of Testing Dataset II are listed in Table 10. Testing dataset II has 1280 craters, 364 primary and 915 secondary craters. Among all the craters, 340 primary and 898 secondary craters were identified correctly, 24 primary craters were wrongly regarded as secondary craters, and 17 secondary craters were mistaken as primary craters. Figure 16 shows the classification results by enlarging two regions in Testing Dataset II. The results of this dataset are the best among the three testing results. Though the kappa coefficients of the three testing datasets were all within the range of 0.8–1, which indicates almost perfect agreement between estimation and observation, the kappa coefficient of Testing Dataset II was 0.921, much higher than that of the other two datasets.

Table 10. Statistics of Testing Dataset II.

	Precision	Sensitivity	Accuracy	F1-score	Kappa
Testing dataset II	0.952	0.934	0.968	0.943	0.921
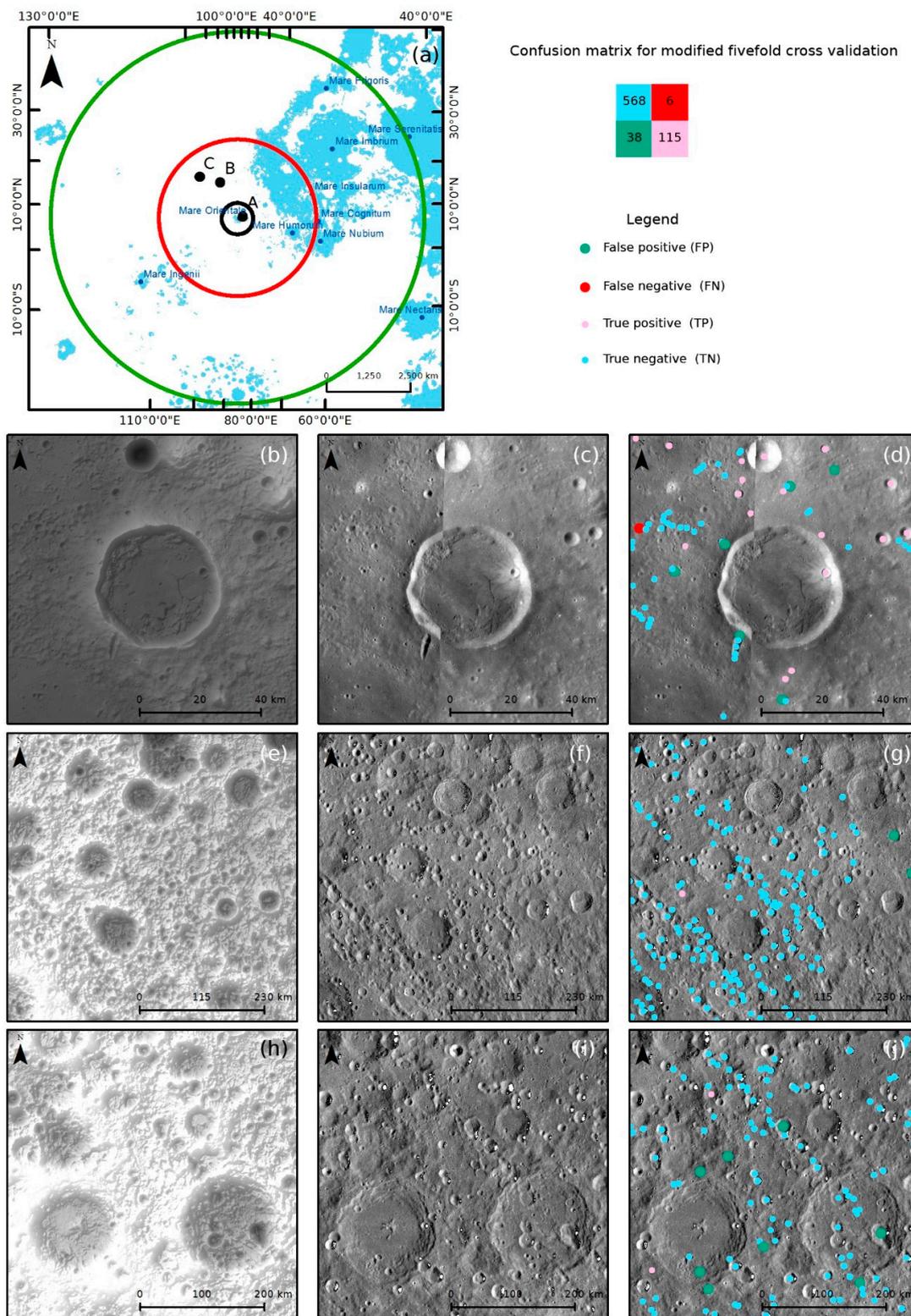
Figure 14. Fivefold cross-validation results: (**a**) sketched map showing locations of regions A, B, and C; (**b**–**d**) DEM and LROC-WAC (Lunar Reconnaissance Orbiter Camera-wide angle camera) images, and classification results for region A; (**e**–**g**) DEM and LROC-WAC images, and classification results for region B; (**h**–**j**) DEM and LROC-WAC images, and classification results for region C. Green points represent false positive (FP), red points represent false negative (FN), pink points represent true positive (TP), and blue points represent true negative (TN).
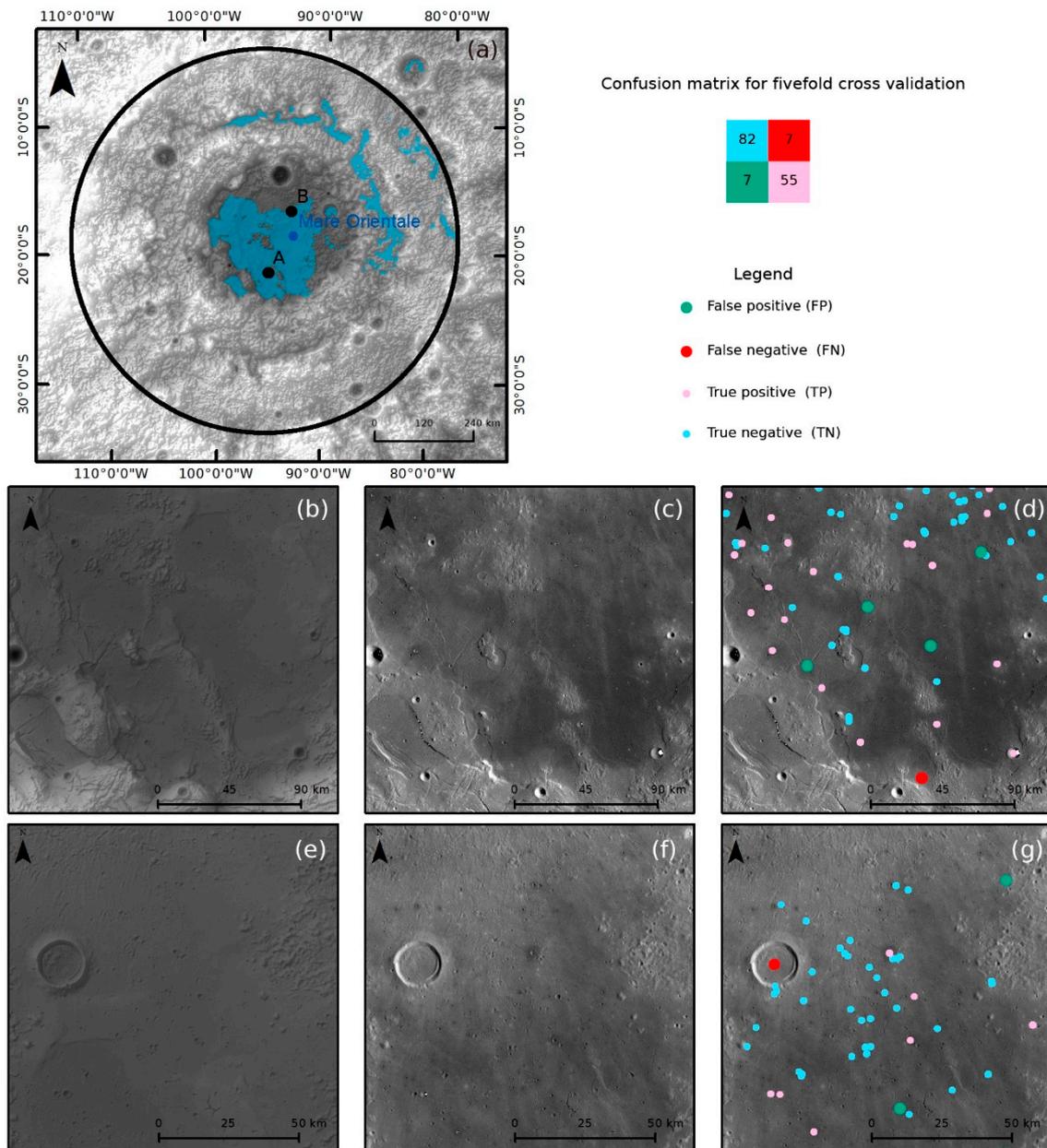
Figure 15. Testing Dataset I validation results showing classification results by enlarging two regions in dataset I: (**a**) sketched map showing locations of Regions A and B; (**b–d**) DEM and LROC-WAC images, and classification results for region A; (**e–g**) DEM and LROC-WAC images, and classification results for Region B; Green points represent FP, red points represent FN, pink points represent TP, and blue points represent TN.
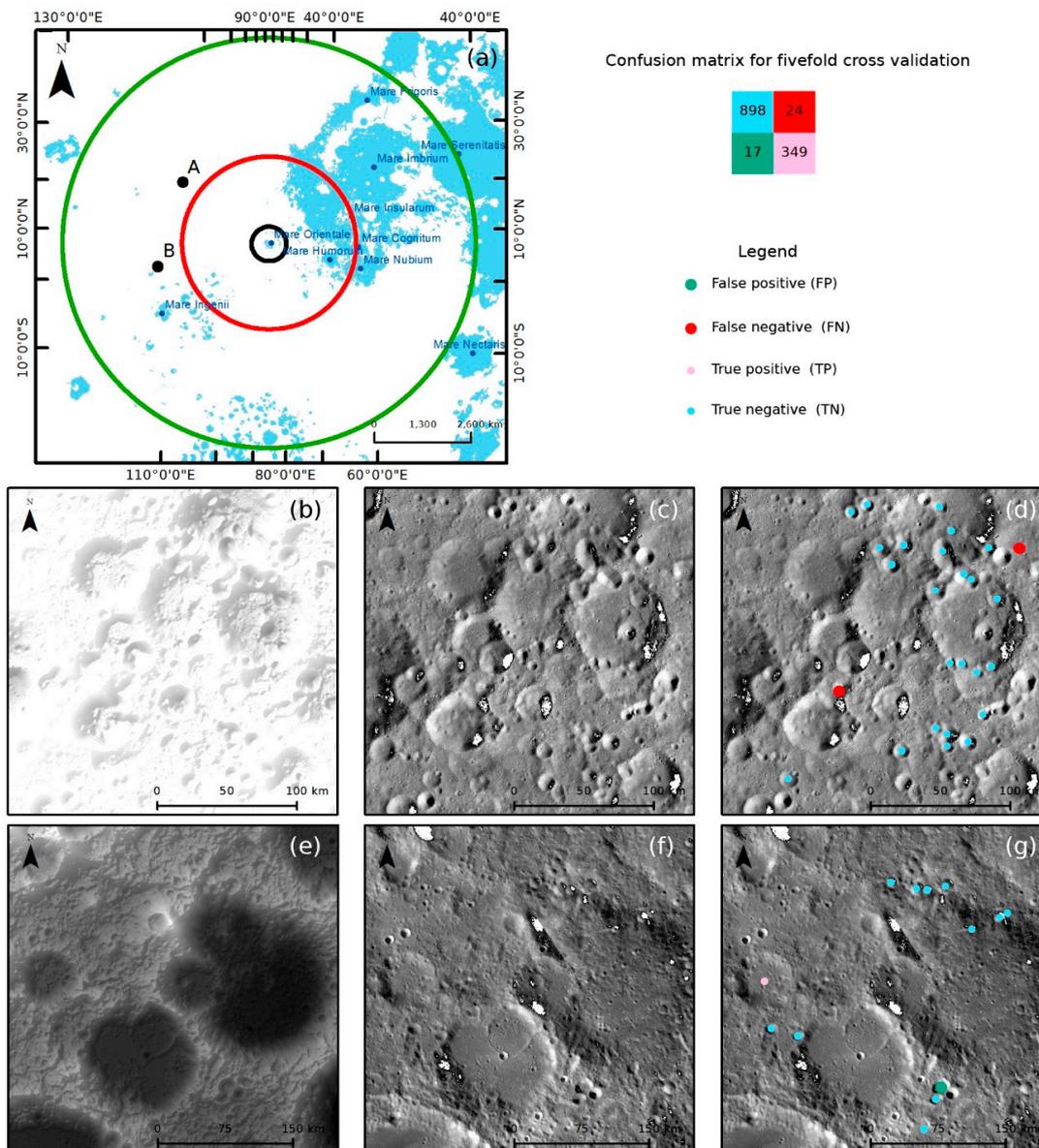
Figure 16. Testing dataset II validation results: (**a**) sketched map showing locations of Regions A and B; (**b**–**d**) DEM and LROC-WAC images, and classification results for region A; (**e**–**g**) DEM and LROC-WAC images, and classification results for Region B. Green points represent FP, red points represent FN, pink points represent TP, and blue points represent TN.

4.3. Feature Sensitivity Analysis

The importance of each feature can be evaluated based on the worsening of the prediction if the parameter is randomly permuted. The feature importance of each model was calculated during the training procedure with fivefold cross-validation. Figure 17 shows the importance of the features according to category. Features related to crater shape, depth, and density are marked in yellow, blue, and dark pink and dark green, respectively.

Among all features involved in this model, Ch_II_4 and Cl_II_2 are the two most important. The importance value of eccentricity is 4.72, nearly double the importance value of irregularity (Irr), and the importance value of fitted ellipse depth to major axis is 6.89, which is the highest of the three similar features. This means that compared with fitted circle depth-to-diameter ratio (d_c/D) and fitted ellipse depth to minor axis (d_e/A_{min}), fitted ellipse depth to major axis is more useful in distinguishing primary craters from secondary craters. Besides, by comparing the two groups of features describing

crater chains, we found that features in Chain_II are generally more important than those in Chain_I. This may be partly because features in Chain_II are composite indicators that not only contain the information of features in Chain_I but also include information such as crater diameter and so on. Moreover, we can conclude that compared with other statistical regions used in calculating features related to crater clusters, the region beginning at the crater center and extending out to 2 radius is more suitable for distinguishing primary craters from secondary craters, as its corresponding feature has a higher importance value.
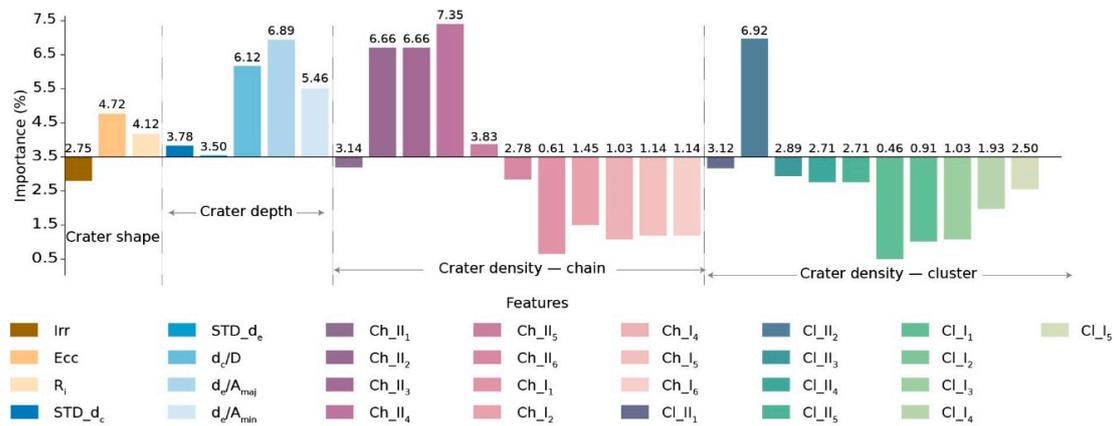
Figure 17. Relative importance values of features.

4.4. Comparision with Previous Work

A comparison with previous studies could better explain the differences and advantages of the method proposed in this paper. However, none of the previous works expressed the performance of their methods in a statistical way. Although the method proposed by Wu et al. was more related to ours concerning the features involved, it was aimed at detecting secondary craters belonging to a certain parent crater and some settings of the key thresholds, such as the degree used to define a crater chain, were not revealed. So here, to contrast the performance of our method with a previous study, we conducted an experiment in a small region using a traditional rule-based method [29]. The key point of this method is to calculate the nonrandom degree of craters. By comparing the results of a Voronoi diagram and the average and standard deviation of ideally random spatial distributions, one can determine whether a crater is nonrandom with high significance, and craters with a significance of nonrandomness higher than a certain value could be regarded as secondary craters [29].

The difference between this previous rule-based method and the method in this paper lies in two aspects. From the feature aspect, this previous rule-based method is only concerned with the difference in spatial distribution between primary and secondary craters. In this paper, we synthesized the literature explaining the difference between primary and secondary craters, extracted features that may be considered when the work is conducted via visual identification, and preliminarily selected 32 features concerning crater shape, depth, and density. After seriously testing the importance of these features, a series of 29 features were finally selected as the input features of the classifier. Although both methods use crater density as a key point to distinguish primary and secondary craters, the expression of this feature in the two methods is quite different. As we explained in the introduction section, this previous rule-based method uses the degree of spatial aggregation as the expression of the difference in crater density realized through a Voronoi diagram. In our method, we describe crater density in a group of features and the description is based on the definitions of crater chains and clusters. On the one hand, this expression fails to express the difference of degree in spatial aggregation between crater chains and clusters and may introduce uncertainty in the detection result, as this method only uses one threshold. One the other hand, a certain fraction of spatially random distributions are in fact secondary, and this expression can be used to solve such a situation. In our method, we describe

a crater chain by counting the number of craters with approximate size in certain directions, counting the number of craters with approximate size in certain areas. These descriptions can help to better express the characteristics of crater chains and clusters and convert them to a classifier more accurately. From the method aspect, the previous method is rule-based and mainly solves this problem from a statistical perspective. A rule-based method is relatively flexible and easy to understand. It is suitable for simple questions or questions that have been carefully studied and can be accurately described by rules. When talking about distinguishing primary and secondary craters, researchers already know in which aspects their differences exist, and this means that researchers know the objects that rules should describe. However, most of the thresholds of the rules are unknown. For example, a secondary crater cluster consists of some clustered craters, and to distinguish them, a rule describing the degree of clustering should be created, but researchers are unable to quantify the extent to which spatial aggregation could be considered as a crater cluster. To solve this problem, the method proposed in this paper is a machine learning–based method. Choosing random forest as the classifier, one can replace the process of setting thresholds with selecting positive and negative samples. The random forest classifier can automatically learn the rules hidden in samples through the training process. A rule-based method is more like a decision tree, and a random forest classifier is a combination of decision trees, which reduces the impact of single decision tree classifier error and improves the accuracy and stability of the strong classifier.

Ground truth data are important when assessing the performance of a certain method. Compared with craters classified by us, a classified crater database proposed by others would be more suitable and impartial when used for contrast. There are many lunar crater databases, but no database containing classified secondary and primary craters has been published. Thus, we used part of the secondary craters in training dataset II identified by Guo [1] as ground truth data and the number of correctively identified secondary craters as a metric to assess the methods. The area used for comparison contained 146 identified secondary craters belonging to the Orientale Basin. Using the typical rule-based method, 66 of them were identified correctly, corresponding with 144 craters identified through the machine-learning based method. The identification result shows that the proposed method has a better performance than the typical rule-based method. Two reasons may account for the rule-based model's slightly worse performance in this comparison. Based on crater density, the typical rule-based method has a better performance in the whole testing region, and an assessment on only parts of secondary craters may be a little biased. The setting of the thresholds also affects the model performance. Though the performance of the rule-based model may be worse than its ideal one, the identified ratio of the two models can still prove that the proposed model has a better performance.

5. Conclusions and Discussions

A machine learning approach is proposed for distinguishing primary and secondary craters automatically and with good performance. The developed approach was evaluated with actual datasets collected on the moon. The theoretical analysis and experimental validation lead to the following conclusions:

- The evaluation process was conducted with manually labeled primary and secondary craters. The whole dataset contains 1032 primary craters and 4041 secondary craters, with 1974 craters identified for this research, 2621 referenced from other research, and 478 large primary craters selected.
- The proposed machine learning shows favorable performance. The accuracy and F1-score for fivefold cross-validation were 0.939 and 0.839, respectively, for Testing Dataset I, 0.907 and 0.887, for Testing Dataset II, 0.968 and 0.943.
- The experimental results of Testing Datasets I and II show that the proposed machine learning method produces accurate crater classification in other areas. This means that the method has good generalization ability. Also, as Testing Dataset II consists of secondary craters identified by other researchers and large primary craters, its good performance shows homogeneity with others.

There exist great differences between diameters of primary and secondary craters in Testing Dataset II. The diameter of a primary crater is usually larger than 20 km, but a secondary crater is usually smaller than 20 km; only 16 craters had a diameter larger than 15 km. The classification result of Testing Dataset II had higher precision (0.952) and kappa value (0.921) compared with the training dataset (0.752 and 0.803) and Testing Dataset I (0.887 and 0.808), and this great difference may have led to the extremely good performance of the model. Though the parameters of the model do not directly contain diameter, some of them may be greatly affected by diameter, such as parameters in Chain_II, and the difference in diameter between primary and secondary craters may reduce the difficulty of differentiation to some degree. However, this does not totally negate the good performance of our model on Testing Dataset II. This is in part because the model was trained with the training dataset and there is not a great difference between the diameters of primary and secondary craters, which means that the trained model did not focus much on distinguishing primary craters from secondary craters by diameter, but also because the trained model performed well on Testing Dataset I and the diameters of wrongly identified primary and secondary craters of Testing Dataset II range from 20 km to 29 km, and from 12 km to 20 km, respectively, both not fully concentrated around 20 km. It should be noted that the automatic classification process mainly concerns craters smaller than 20 km in diameter, due to the lack of large secondary samples. The proposed approach is able to detect craters with diameters larger than 20 km on images; however, the detection rate might be decreased. To improve accuracy, all craters were projected again before feature extraction, setting the new projection centered at the crater center. If the approach is only applied in a small region, projection has little effect on the final classification result. The most time-consuming part of the whole method was feature extraction. Calculating features for about 5000 craters needs about 2–3 days based on computer performance and this time could be shorten if the whole process is conducted using GPU or the research region is small and the process of reprojection could be left out. Except the feature extraction part, all remain processes could be finished in within 10 minutes.

Researchers have tried to create reliable methods for automatically distinguishing primary craters from secondary craters [5,10,27,28]. These rule-based methods were either mainly aimed at detecting secondary crater chains and clusters and failed to distinguish primary craters located within or near a cluster, or needed a threshold defined before use. The proposed method mainly differs from previous studies in feature selection and base method, and these two points guarantee the performance of the proposed method. Primary craters differ from secondary craters in a lot of aspects, such as depth-to-diameter ratio and irregularity, and manual identification is a process that takes many parameters into consideration. However, previous studies using parameters related to crater density failed to consider parameters related to shape and depth. Not considering crater shape or morphologic features will certainly lead to negative effects on performance and a full consideration of these differences helps improve the proposed model's performance. In this paper, we take full consideration of all features that can be used; those related to crater shape and depth, such as Irr, Ecc, STD_d_c, and d_c/D are used together with those related to crater density. A full use of these features help the proposed method better study the difference between primary and secondary craters and imitate the process of manual identification. Taking advantage of the random forest classifier also guarantees the precision of the model. Previous rule-based methods are more like a decision tree, and a random forest classifier is a combination of decision trees. By combination a serious of decision trees, a random forest classifier could reduce the impact of single decision tree classifier error and improve the accuracy and stability. Besides, experts can conclude whether a crater is a primary crater but cannot give the threshold for distinguishing, and a little difference in setting thresholds may lead to a great difference in the results. Previous approaches usually needed to define a threshold to distinguish primary craters from secondary craters. In the proposed method, the threshold setting is included in the learning process and is done by random forest classifier which can weaken the influence of improper setting of thresholds.

Our future work will try to enlarge the sample dataset with more distant secondary craters and to find primary craters that the identified secondary craters belong to, using a crater degradation model and the relationship between primary craters and their secondary craters. The proposed machine learning approach enables an automated method of classifying primary and secondary craters, which results in better performance.

Author Contributions: conceptualization, Q.L. and W.C.; methodology, Q.Y. and W.C.; software, Q.L.; validation, Q.L., W.C., and Y.Z.; formal analysis, Q.L.; writing—original draft preparation, Q.L.; writing—review and editing, Q.L., W.C., G.Y., and J.L.; visualization, Q.L.; supervision, Q.L. and W.C.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Pseudocode of the Whole Algorithm

Input identified positive and negative sample database (target dataset), Robbins' crater database, DEM
num ← the total number of positive and negative samples
-------------------------------------- **Feature extraction**--------------------------------------
for i ← 1 to num
    extract circular center longitude, circular center latitude, ellipse center longitude, ellipse center latitude, diameter,
        ellipse major axis, ellipse minor axis, ellipse angle, eccentricity and irregularity from database
    ------------------------Calculation: features related to crater shape --------------------------
    irr ← irregularity / diameter
    ------------------------ Calculation: features related to crater depth --------------------------
    circle ← a circle generate according circular center longitude, circular center latitude and diameter
    overlay DEM with circle according to their location
    circle _polyline ← extract value from DEM for every pixel located on circle
    circle _polygon ← extract value from DEM for every pixel within the circle
    std _cl ← standard deviation of circle _ polyline
    dc/d ← (the average of circle _polyline – the minimum of circle _polygon) / diameter
    ellipse ← generate according ellipse center longitude, ellipse center latitude, ellipse major axis,
        ellipse minor axis and ellipse angle
    overlay DEM with ellipse according to their location
    ellipse _polyline ← extract value from DEM for every pixel located on ellipse
    ellipse _polygon ← extract value from DEM for every pixel within the ellipse
    std _el ← standard deviation of ellipse _polyline
    de/Amaj ← (the average of ellipse _polyline – the minimum of ellipse _polygon) / ellipse major axis
    de/Amin ← (the average of ellipse _polyline – the minimum of ellipse _polygon) / ellipse minor axis
    -----------------------Calculation: features related to crater density --------------------------
    radius ← diameter / 2
    region ← a circle generate according circular center longitude, circular center latitude and 6 times radius
    degree ← 0
    rectangle ← a rectangle with a width of 1.5 times radius, a length of 6 times radius facing north
    for j ← 1 to 6
        area ← the overlap part of region and rectangle
        ch _I[j] ← the number of crater of similar size in whole crater database located in area
        count ← the number of crater in whole crater database located in area
        ch _II[j] ← ch _I[j] / count
        rectangle ← rectangle rotate 30 degrees clockwise at center

for j ← 1 to 6

    area ← a circle generate according circular center longitude, circular center latitude and j times radius

    cl _I[j] ← the number of crater of similar size in whole crater database located in area

    count ← the number of crater in whole crater database located in area

    cl _II[j] ← cl _I[j] / count

    feature dataset← restore eccentricity, irr, std _cl, std _el, dc /d, de _Amaj, de _Amin, ch _I, ch _II, cl _I, cl _II

Divide feature dataset and target dataset into Training Dataset, Testing Dataset I and Testing Dataset II

------------------------------------- **Training Process** - - -------------------------------------

Training Dataset I ← positive samples in Training Dataset and the same number of negative samples selected randomly

Training Dataset II ← the rest part of Training Dataset

Classifier ← Training: setting parameters, selecting features, modified fivefold cross-validation

Analysis feature importance

------------------------------------- **Testing Process**- - -------------------------------------

Test the performance of classifier in Testing Dataset I and Testing Dataset II

References

1. Guo, D.; Liu, J.; Head, J.W.; Kreslavsky, M.A. Lunar Orientale Impact Basin Secondary Craters: Spatial Distribution, Size-Frequency Distribution, and Estimation of Fragment Size. *J. Geophys. Res. Planets* **2018**, *123*, 1344–1367. [CrossRef]

2. Sun, S.; Yue, Z.; Di, K. Investigation of the depth and diameter relationship of subkilometer-diameter lunar craters. *Icarus* **2018**, *309*, 61–68. [CrossRef]

3. Head, J.W.; Fassett, C.I.; Kadish, S.J.; Smith, D.E.; Zuber, M.T.; Neumann, G.A.; Mazarico, E. Global Distribution of Large Lunar Craters: Implications for Resurfacing and Impactor Populations. *Science* **2010**, *329*, 1504–1507. [CrossRef] [PubMed]

4. Ivanov, B.A. Size-Frequency Distribution of Small Lunar Craters: Widening with Degradation and Crater Lifetime. *Sol. Syst. Res.* **2018**, *52*, 1–25. [CrossRef]

5. Bierhaus, E.B.; Chapman, C.R.; Merline, W.J. Secondary craters on Europa and implications for cratered surfaces. *Nature* **2005**, *437*, 1125–1127. [CrossRef]

6. Melosh, H.J. *Impact Cratering: A Geologic Process*; Oxford Universitr Press: New York, NY, USA, 1989.

7. Bierhaus, E.B.; Merline, W.J.; Chapman, C.R. Variation in Size-Distributions between Adjacent and Distant Secondary Craters. *Lunar Planet. Sci. Conf.* **2005**, *Abstract#238*.

8. Xiao, Z. On the importance of self-secondaries. *Geosci. Lett.* **2018**, *5*, 17. [CrossRef]

9. Xiao, Z.; Werner, S.C. Size-frequency distribution of crater populations in equilibrium on the Moon. *J. Geophys. Res. Planets* **2015**, *120*, 2277–2292. [CrossRef]

10. Michael, G.G.; Platz, T.; Kneissl, T.; Schmedemann, N. Planetary surface dating from crater size–frequency distribution measurements: Spatial randomness and clustering. *Icarus* **2012**, *218*, 169–177. [CrossRef]

11. Michael, G.G.; Neukum, G. Planetary surface dating from crater size–frequency distribution measurements: Partial resurfacing events and statistical age uncertainty. *Earth Planet. Sci. Lett.* **2010**, *294*, 223–229. [CrossRef]

12. Wilhelms, D.E. Secondary impact craters of lunar basins. In Proceedings of the 7th Lunar Science Conference, Houston, TX, USA, 15–19 March 1976; pp. 2883–2901.

13. Williams, J.-P.; Bandfield, J.L.; Paige, D.A.; Powell, T.M.; Greenhagen, B.T.; Taylor, S.; Hayne, P.O.; Speyerer, E.J.; Ghent, R.R.; Costello, E.S. Lunar Cold Spots and Crater Production on the Moon. *J. Geophys. Res. Planets* **2018**, *123*, 2380–2392. [CrossRef]

14. McEwen, A.S.; Bierhaus, E.B. The importance of secondary cratering to age constraints on planetary surfaces. *Annu. Rev. Earth Planet. Sci.* **2006**, *34*, 535–567. [CrossRef]

15. Mcewen, A.; Preblich, B.; Turtle, E.; Artemieva, N.; Golombek, M.; Hurst, M.; Kirk, R.; Burr, D.; Christensen, P. The rayed crater Zunil and interpretations of small impact craters on Mars. *Icarus* **2005**, *176*, 351–381. [CrossRef]

16. Werner, S.C.; Ivanov, B.A.; Neukum, G. Theoretical analysis of secondary cratering on Mars and an image-based study on the Cerberus Plains. *Icarus* **2009**, *200*, 406–417. [CrossRef]

17. Oberbeck, V.R.; Morrison, R.H. Laboratory simulation of the herringbone pattern associated with lunar secondary crater chains. *Moon* **1974**, *9*, 415–455. [CrossRef]

18. Pike, R.J.; Wilhelms, D.E. Secondary-Impact Craters on the Moon: Topographic Form and Geologic Process. *Lunar Planet. Sci. Conf.* **1978**, *9*, 907–909.

19. Bierhaus, E.B.; Schenk, P.M. Constraints on Europa's surface properties from primary and secondary crater morphology. *J. Geophys. Res.* **2010**, *115*, E12004. [CrossRef]

20. Senthil Kumar, P.; Senthil Kumar, A.; Keerthi, V.; Goswami, J.N.; Gopala Krishna, B.; Kiran Kumar, A.S. Chandrayaan-1 observation of distant secondary craters of Copernicus exhibiting central mound morphology: Evidence for low velocity clustered impacts on the Moon. *Planet. Space Sci.* **2011**, *59*, 870–879. [CrossRef]

21. Wells, K.S.; Campbell, D.B.; Campbell, B.A.; Carter, L.M. Detection of small lunar secondary craters in circular polarization ratio radar images. *J. Geophys. Res.* **2010**, *115*, E06008. [CrossRef]

22. Bart, G.D.; Melosh, H.J. Using lunar boulders to distinguish primary from distant secondary impact craters. *Geophys. Res. Lett.* **2007**, *34*, L07203. [CrossRef]

23. Basilevsky, A.T.; Kozlova, N.A.; Zavyalov, I.Y.; Karachevtseva, I.P.; Kreslavsky, M.A. Morphometric studies of the Copernicus and Tycho secondary craters on the moon: Dependence of crater degradation rate on crater size. *Planet. Space Sci.* **2018**, *162*, 31–40. [CrossRef]

24. Calef, F.J.; Herrick, R.R.; Sharpton, V.L. Geomorphic analysis of small rayed craters on Mars: Examining primary versus secondary impacts: Analysis of small rayed craters on mars. *J. Geophys. Res.* **2009**, *114*, E10007. [CrossRef]

25. Grant, J.A.; Arvidson, R.E.; Crumpler, L.S.; Golombek, M.P.; Hahn, B.; Haldemann, A.F.C.; Li, R.; Soderblom, L.A.; Squyres, S.W.; Wright, S.P.; et al. Crater gradation in Gusev crater and Meridiani Planum, Mars. *J. Geophys. Res.* **2006**, *111*, E02S08. [CrossRef]

26. Nagumo, K.; Nakamura, A.M. Reconsideration of crater size-frequency distribution on the moon: Effect of projectile population and secondary craters. *Adv. Space Res.* **2001**, *28*, 1181–1186. [CrossRef]

27. Salih, A.L.; Lompart, A.; Grumpe, A.; Wöhler, C.; Hiesinger, H. Automatic detection of secondary craters and mapping of planetary surface age based on lunar orbital images. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *XLII-3/W1*, 125–132. [CrossRef]

28. Wu, B.; Wang, Y.; Lin, T.J.; Hu, H.; Werner, S.C. Impact cratering in and around the Orientale Basin: Results from recent high-resolution remote sensing datasets. *Icarus* **2019**, *333*, 343–355. [CrossRef]

29. Honda, C.; Kinoshita, T.; Hirata, N.; Morota, T. Detection abilities of secondary craters based on the clustering analysis and Voronoi diagram. In Proceedings of the European Planetary Science Congress, Cascais, Portugal, 7–12 September 2014.

30. Kreslavsky, M.A. Statistical Characterization of Spatial Distribution of Impact Craters: Implications to Present-Day Cratering Rate on Mars. *LPI Contrib.* **2007**, *1353*, 3325–3328.

31. Wang, Y.; Wu, B. Active Machine Learning Approach for Crater Detection from Planetary Imagery and Digital Elevation Models. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5777–5789. [CrossRef]

32. Chen, M.; Liu, D.; Qian, K.; Li, J.; Lei, M.; Zhou, Y. Lunar Crater Detection Based on Terrain Analysis and Mathematical Morphology Methods Using Digital Elevation Models. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3681–3692. [CrossRef]

33. Zhou, Y.; Zhao, H.; Chen, M.; Tu, J.; Yan, L. Automatic detection of lunar craters based on DEM data with the terrain analysis method. *Planet. Space Sci.* **2018**, *160*, 1–11. [CrossRef]

34. Zuo, W.; Zhang, Z.; Li, C.; Wang, R.; Yu, L.; Geng, L. Contour-based automatic crater recognition using digital elevation models from Chang'E missions. *Comput. Geosci.* **2016**, *97*, 79–88. [CrossRef]

35. Di, K.; Li, W.; Yue, Z.; Sun, Y.; Liu, Y. A machine learning approach to crater detection from topographic data. *Adv. Space Res.* **2014**, *54*, 2419–2429. [CrossRef]

36. Xie, Y.; Tang, G.; Yan, S.; Lin, H. Crater Detection Using the Morphological Characteristics of Chang'E-1 Digital Elevation Models. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 885–889.

37. Stepinski, T.F.; Mendenhall, M.P.; Bue, B.D. Machine cataloging of impact craters on Mars. *Icarus* **2009**, *203*, 77–87. [CrossRef]

38. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

39. Robbins, S.J. A New Global Database of Lunar Impact Craters > 1–2 km: 1. Crater Locations and Sizes, Comparisons with Published Databases, and Global Analysis. *J. Geophys. Res. Planets* **2019**, *124*, 871–892. [CrossRef]

40. Barker, M.K.; Mazarico, E.; Neumann, G.A.; Zuber, M.T.; Haruyama, J.; Smith, D.E. A new lunar digital elevation model from the Lunar Orbiter Laser Altimeter and SELENE Terrain Camera. *Icarus* **2016**, *273*, 346–355. [CrossRef]

41. Smith, D.E.; Zuber, M.T.; Jackson, G.B.; Cavanaugh, J.F.; Neumann, G.A.; Riris, H.; Sun, X.; Zellar, R.S.; Coltharp, C.; Connelly, J.; et al. The Lunar Orbiter Laser Altimeter Investigation on the Lunar Reconnaissance Orbiter Mission. *Space Sci. Rev.* **2010**, *150*, 209–241. [CrossRef]

42. Robinson, M.S.; Brylow, S.M.; Tschimmel, M.; Humm, D.; Lawrence, S.J.; Thomas, P.C.; Denevi, B.W.; Bowman-Cisneros, E.; Zerr, J.; Ravine, M.A.; et al. Lunar Reconnaissance Orbiter Camera (LROC) Instrument Overview. *Space Sci. Rev.* **2010**, *150*, 81–124. [CrossRef]

43. Robbins, S.J.; Antonenko, I.; Kirchoff, M.R.; Chapman, C.R.; Fassett, C.I.; Herrick, R.R.; Singer, K.; Zanetti, M.; Lehan, C.; Di, H. The variability of crater identification among expert and community crater analysts. *Icarus* **2014**, *234*, 109–131. [CrossRef]

44. Hirata, N.; Nakamura, M.A. Secondary craters of Tycho: Size-frequency distributions and estimated fragment size–velocity relationships. *J. Geophys. Res.* **2006**, *111*, E03005. [CrossRef]

45. Preblich, B.S.; Mcewen, A.S.; Studer, D.M. Mapping rays and secondary craters from the Martian crater Zunil. *J. Geophys. Res.* **2017**, *112*, E05006. [CrossRef]

46. Wilhelms, D.E.; Mccauley, J.F.; Trask, N.J. The Geologic History of the Moon, USGS Professional Paper 1348. US Government Printing Office: Washington, DC, USA, 1987.

47. Zhou, S.; Xiao, Z.; Zeng, Z. Impact Craters with Circular and Isolated Secondary Craters on the Continuous Secondaries Facies on the Moon. *J. Earth Sci.* **2015**, *26*, 740–745. [CrossRef]

48. Barnouin, O.S.; Zuber, M.T.; Smith, D.E.; Neumann, G.A.; Herrick, R.R.; Chappelow, J.E.; Murchie, S.L.; Prockter, L.M. The morphology of craters on Mercury: Results from MESSENGER flybys. *Icarus* **2012**, *219*, 414–427. [CrossRef]

49. Sammut, C.; Webb, G.I. *Encyclopedia of Machine Learning*; Springer: Berlin, Germany, 2010; Springer reference.

50. Bassa, Z.; Bob, U.; Szantoi, Z.; Ismail, R. Land cover and land use mapping of the iSimangaliso Wetland Park, South Africa: Comparison of oblique and orthogonal random forest algorithms. *J. Appl. Remote Sens.* **2016**, *10*, 015017. [CrossRef]

51. Veronesi, F.; Hurni, L. Random Forest with semantic tie points for classifying landforms and creating rigorous shaded relief representations. *Geomorphology* **2014**, *224*, 152–160. [CrossRef]

52. Song, Y.Y.; Lu, Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch. Psychiatry* **2015**, *27*, 130–135. [PubMed]

53. Quinlan, J.R. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. [CrossRef]

54. Subramanian, A.A.B.; Pramala, S.; Rajalakshmi, B.; Rajaram, R. Improving Decision Tree Performance by Exception Handling. *Int. J. Autom. Comput.* **2010**, *7*, 372–380. [CrossRef]

55. Grömping, U. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *Am. Stat.* **2009**, *63*, 308–319. [CrossRef]

56. Strobl, C.; Malley, J.; Tutz, G. An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol. Methods* **2009**, *14*, 323–348. [CrossRef] [PubMed]

57. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. Variable selection using random forests. *Pattern Recognit. Lett.* **2010**, *31*, 2225–2236. [CrossRef]

58. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33*, 159–174. [CrossRef] [PubMed]