



Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 63 (2019) 3721–3737

**ADVANCES IN
SPACE
RESEARCH**
(*a COSPAR publication*)
www.elsevier.com/locate/asr

Bayesian network-based extraction of lunar impact craters from optical images and DEM data

Juntao Yang, Zhizhong Kang*

Department of Remote Sensing and Geo-Information Engineering, School of Land Science and Technology, China University of Geosciences, Xueyuan Road, Beijing 100083, China

Received 21 August 2018; received in revised form 20 January 2019; accepted 5 February 2019
Available online 13 February 2019

Abstract

Impact craters are among the most noticeable geomorphological features on the planetary surface and yield significant information about terrain evolution and the history of the solar system. Thus, the recognition of impact craters is an important branch of modern planetary studies. Aiming at addressing problems associated with the insufficient and inaccurate detection of lunar impact craters, a decision fusion method within the Bayesian network (BN) framework is developed in this paper to handle multi-source information from both optical images and associated digital elevation model (DEM) data. First, we implement the edge-based method for efficiently searching crater candidates which are the image patches that can potentially contain impact craters. Secondly, the multi-source representations of an impact crater derived from both optical images and DEM data are proposed and constructed to quantitatively describe the two-dimensional (2D) and three-dimensional (3D) morphology, consisting of Histogram of Oriented Gradient (HOG), Histogram of Multi-scale Slope (HMS) and Histogram of Multi-scale Aspect (HMA). Finally, a BN-based framework integrates the multi-source representations of impact craters, which can provide redundant and complementary information, for distinguishing craters from non-craters. To evaluate the effectiveness and robustness of the proposed method, experiments were conducted on three lunar scenes using both orthoimages from the Lunar Reconnaissance Orbiter (LRO) and DEM data acquired by the Lunar Orbiter Laser Altimeter (LOLA). Experimental results demonstrate that integrating optical images with DEM data significantly decreases the number of false positives compared with using optical images alone, with F_1 -score of 84.8% on average. Moreover, compared with other existing fusion methods, our proposed method was quite advantageous especially for the detection of small-scale craters with diameters less than 1000 m.

© 2019 COSPAR. Published by Elsevier Ltd. All rights reserved.

Keywords: Crater detection; Bayesian network fusion; Active learning; Optical images; Digital elevation model

1. Introduction

Impact craters generate significant information primarily regarding relative and absolute surface chronologies, erosional processes and climate history (Cohen and Ding, 2014). The distributions and morphologies of impact

craters are of significance in detecting and analyzing the relative ages of planetary surface units (Tanaka, 1986; Haruyama et al., 2009; Morota et al., 2009). Generally, the heavily cratered surface is relatively older than the less cratered surface (Jin and Zhang, 2014). To date, many crater databases have been built via visual inspection of images (Liu et al., 2012). However, it is time-consuming and laborious to manually build a comprehensive crater catalog (Robbins and Hynek, 2012). Furthermore, considering the known crater databases, we conclude that manual extraction focuses primarily on large craters in planetary

* Corresponding author at: Department of Remote Sensing and Geo-Information Engineering, School of Land Science and Technology, China University of Geosciences, Xueyuan Road 29, Haidian District, Beijing 100083, China.

E-mail address: zzkang@cugb.edu.cn (Z. Kang).

images (Barlow, 1988) and can be greatly affected by the researcher's personal knowledge and skills (Di et al., 2014). Thus, it is necessary to develop an automatic crater detection algorithm (CDA) to recognize craters and obtain much more information.

Due to the importance of impact craters, many CDAs have been proposed in recent years. These methods are generally divided into two categories, i.e., the optical image-based CDAs and the DEM-based CDAs. Many optical image-based CDAs, such as Hough transform (HT) (Honda and Azuma, 2000; Bruzzone et al., 2004; Sawabe et al., 2006; Luo et al., 2011), genetic algorithm (GA) (Honda and Azuma, 2000), random sample consensus (RANSAC) (Kang et al., 2016), radial consistency (Earl et al., 2005), template matching (TM) (Bandeira et al., 2007; Salih et al., 2017) and statistical methods (Ding et al., 2008), were based on the characteristic patterns (i.e., nearly round edges) of impact craters in optical images and had the advantage of high automation level. Troglio et al. (2012) used a set of image processing techniques, including watershed segmentation and generalized Hough transform, for extracting craters and rocks. Marques and Pina (2015) delineated the rims of impact craters using dynamic programming. Liu et al. (2015) introduced a saliency detection idea for capturing crater rims before edge fitting. However, these unsupervised CDAs need to be preprocessed with prior edge extraction and are usually suitable for extracting large-scale craters. Apart from pattern recognition techniques, machine learning-based approaches were also used to distinguish craters from non-craters from optical images by learning a classifier, such as decision tree (Urbach and Stepinski, 2008; Stepinski et al., 2009; Stepinski et al., 2012), neural networks (NN) (Kim et al., 2005), SVM (Stepinski et al., 2007). To enhance the performance of a weak classifier, boosting-based (Jin and Zhang, 2014; Martins et al., 2009; Ding et al., 2011; Stepinski et al., 2012; Bandeira et al., 2012; Wang et al., 2015; Xin et al., 2017) and bagging-based (Yin et al., 2015) techniques were also developed to automatically detect impact craters using shape and texture features extracted from planetary images. Compared with unsupervised methods, current supervised methods have been inspired by face detection and exploited gray-scale texture and shape features such as local binary pattern (LBP) (Ojala et al., 2001), Haar-like (Viola and Jones, 2001), HOG (Dalal and Triggs, 2005) for detecting the small-scale impact craters, which can provide a higher correctness rate.

Nevertheless, the visibility of impact craters in optical images are normally affected by illumination angle, which poses great challenges for the optical image-based CDA applications. With the increasing availability of high-resolution DEM, 3D digital morphology and topography of impact craters were analyzed and used for automatic detection. Like the optical image-based CDAs, Di et al. (2014) extracted texture features from Martian topographic data to detect impact craters using a learned boosting-

based classifier. Stepinski et al. (2009) considered the round and symmetric topographic depressions as crater candidates, which were further fed into a decision tree for distinguishing craters from non-craters. Bue and Stepinski (2006) and Salamuniccar and Loncaric (2010) exploited edge extraction and HT for the large-scale impact crater identification from topographic data. Xie et al. (2013) and Yue et al. (2013) analyzed morphological characteristics for identifying craters from the Chang'E-1 DEM to complete the current crater catalogs. Liu et al. (2017) extracted the raw impact craters using a watershed algorithm from lunar DEM data and used several morphological factors, such as posture ratio and rectangle factor, for the elimination of false impact craters. Zhou et al. (2018) calculated slope of aspect factors from DEM data by terrain analysis for detecting the rims of impact craters. Yamamoto et al. (2017) utilized the rotationally symmetric pattern of slope gradients and slope azimuth of the crater wall to detect the crater shape in DEM/DTM data. Chen et al. (2018) detected lunar impact craters through terrain analysis and mathematical morphology from DEM.

Although many CDAs based on both optical images and DEM data have been studied in recent years, automatic recognition of impact craters still remains a challenge. First, since DEM data contain abundant morphological and topographical characteristics which are insensitive to illumination, they are more suitable for automatic recognition of impact craters through the quantitative geomorphological analysis. To date, many morphological and topographical factors, such as profile line (Zhou et al., 2018), slope (Kang et al., 2016; Zhou et al., 2018), aspect (Bue and Stepinski, 2006; Kang et al., 2016; Zhou et al., 2018), and curvature (Bue and Stepinski, 2006), have been extracted from DEM data and used to reflect the 3D morphology and topography of an impact crater from different perspectives. However, these factors were mainly used in unsupervised methods and fewer machine learning-based studies have been reported due to the lack of the highly descriptive and robust representations of impact craters based on these morphological and topographical factors. Secondly, as mentioned earlier, most of the optical image-based CDAs are based on the visibility of impact craters in optical images, which is easily affected by illumination angle. Although the DEM-based CDAs can effectively avoid the illumination limitations from which the optical image-based CDAs suffer, numerous impact craters, especially small impact craters, might not be detectable. Therefore, multi-source data can provide complementary information and the fusion of multi-source data can acquire more robust extraction results. Kang et al. (2016) and Degirmenci and Ashyralyev (2010) have combined images and DEM data for the impact crater detection and reduced the error extractions. Nevertheless, as a step in the detection, DEM data are only required for verification.

To address these challenges, we first propose Histogram of Multi-scale Slope (HMS) and Histogram of Multi-scale

Aspect (HMA) derived from DEM data to quantitatively describe the 3D morphology and topography of an impact crater since different aspect values present inside and outside impact walls, and slope values can be used to identify the steep crater walls (Zhou et al., 2018). In a HMA (HMS), DEM patch is divided into the regular grids of various scales. Then, the aspect (slope) histogram of each grid in each scale is voted and concatenated in a certain order to form one huge feature vector. At the same time, HOGs are extracted and used to express the impact crater shape in optical images since HOGs can well encode the rims of impact craters. As a result, multi-source representations of impact craters are constructed, consisting of HOG, HMA and HMS. Moreover, to handle these multi-source features of impact craters which provide redundant and complementary information, we use a BN-based decision fusion method for automatically detecting impact craters. In our fusion method, the SVM model is used as the base classifier and built for each type of feature set (i.e., HOG, HMA or HMS) separately because the inherent characteristics of multi-sensor data make different feature sets exhibit different degrees of reliability for classification (Lee, et al., 1987; Waske and Benediktsson, 2007). For the subsequent fusion process, a BN is applied to combine probabilistic evidences generated by multiple SVM models for the final detection results.

The rest of this paper is organized as follows. Section 2 describes our proposed method in detail. Section 3 presents the experimental results and analysis for evaluating the proposed method. This paper concludes with a discussion of future research considerations in Section 4.

2. Methodology

In our proposed method, impact crater sites are detected by integrating optical images with DEM data. Fig. 1 shows the workflow of automatic extraction of impact craters. This workflow consists of the following steps: (1) finding crater candidates from a test image based on both a bilateral filter and the Robert edge extraction operator; (2)

extracting HOG, HMA and HMS from each of the sample pairs, i.e., optical image samples and associated DEM samples; (3) training the optimal SVM models by an active learning strategy (Schohn and Cohn, 2000) using HOG, HMA and HMS respectively for initial detections that are eventually integrated in the BN-based framework; and (4) detecting crater sites inside each crater candidate using the framework constructed in step (3). As a result, the crater candidates that might contain one or more craters are further recognized in our proposed method using a sliding window approach (Lampert et al., 2008). Key algorithms of our proposed method are given in more detail below.

2.1. Finding crater candidates

As shown in Fig. 1, a set of crater candidates is identified efficiently and then fed into a BN-based classifier for the separation of craters from non-craters. Thus, having a dedicated method for finding crater candidates is of significance, which will not only reduce the computational complexity but also decrease the number of false detections. To date, the algorithms used for candidate selection have primarily consisted of the highlight and shadow region-based method (Urbach and Stepinski, 2008; Liu et al., 2012; Wang et al., 2015; Xin et al., 2017), the edge detection-based method (Kang et al., 2016), the feature point-based method (Degirmenci and Ashyralyev, 2010) and the flooding-based method (Stepinski et al., 2007; Stepinski et al., 2009; Stepinski et al., 2012). As introduced by Bandeira et al. (2012), the shadow might be far from regular and even be absent due to the generally low angle of illumination and other factors, such as the occurrence of several processes that degrade the original shape of the impact craters. The feature point-based method is also affected by the noises that might exist in optical images. The flooding-based method can be used to identify single depressions in DEM data. However, Stepinski et al. (2007) has noted that the flooding-based method failed to locate superposed depressions and irregular depressions.

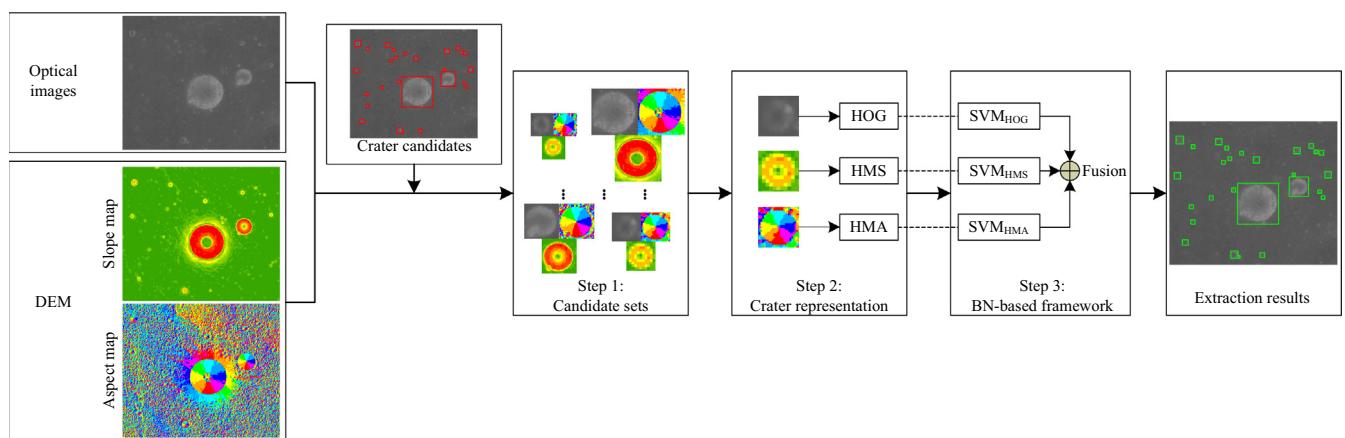


Fig. 1. Workflow of Bayesian-network-based extraction of lunar impact craters.

To reduce the computational complexity of the recognition phase in this paper, we first adopt an edge extraction-based method that was proposed in our previous work (Kang et al., 2016) to find crater candidates in optical images since craters exhibit characteristic patterns in optical images, i.e., nearly round edges. It is a well-known fact that noise patterns may exhibit patterns similar to those of real edges and that the effectiveness of crater extraction is determined by the accurate extraction of edges. Thus, to eliminate the effects of noise, the input images are denoised using a bilateral filter (Paris et al., 2009), which is a nonlinear, self-adaptive filter that considers both spatial information and gray similarity; it also discards noise while retaining edge information, which is effective for edge extraction. Then, we aggregate the edges extracted by the Robert edge extraction operator to obtain the locations that might be craters, namely, crater candidates. Fig. 2 shows an illustration that uses the Robert edge extraction operator. As shown in Fig. 2(b), the extracted edges are aggregated into crater candidates and coloured randomly.

2.2. Multi-source representation learning of a crater based on optical images and DEM data

Prior to machine learning-based step that detects impact crater sites inside a candidate, a set of features is required to describe similarities among impact craters. As aforementioned, both the optical image-based CDAs and the DEM-based CDAs have their own limitations. Hence, constructing multi-source features of impact craters and integrating the multi-source information efficiently to reduce the error extractions have become the key issues in the field of impact crater detection. In this section, we propose and establish the multi-source representations of an impact crater, which includes HOG extracted from optical images and 3D morphological features (i.e., HMA and HMS) extracted from DEM data.

2.2.1. HOG extracted from optical images

It can be noticed that impact craters exhibit characteristic patterns in optical images, i.e., nearly round edges

(Kang et al., 2016). Since HOG well describes the distribution of the gradient intensities and edge directions of impact craters in optical images, we in this paper use HOG to describe the similarities among impact craters in optical images.

HOG (Dalal and Triggs, 2005), which is popular in the domain of pedestrian detection, uses local histograms of the oriented gradients of pixel luminance to characterize a given image. Fig. 5 shows an example of HOG extraction. The detailed process of HOG extraction is as follows:

- (1) All the images are resampled into 64×64 pixels, and the gradient magnitude $G(x, y)$ and gradient directional angle $\theta(x, y)$ at location (x, y) can be evaluated by Eq. (1):

$$\begin{cases} G_x(x, y) = I(x + 1, y) - I(x - 1, y) \\ G_y(x, y) = I(x, y + 1) - I(x, y - 1) \\ G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \\ \theta(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \end{cases} \quad (1)$$

where $G_x(x, y)$, $G_y(x, y)$, $I(x, y)$, $G(x, y)$, and $\theta(x, y)$ represent the central difference across x , the central difference across y , the pixel gray-value, the gradient magnitude and the gradient directional angle at location (x, y) , respectively.

- (2) A 64×64 raw image is partitioned into an array of cells (in this paper, 8×8 cells, as shown in Fig. 5 (b)), and each cell consists of 8×8 gradient direction angles, as shown in Fig. 4(left), which were then weighted and voted into a 9-bin orientation histogram assigned by its gradient magnitudes and Gaussian block-weighted window (Dalal and Triggs, 2005), with a standard deviation equal to one-half the width of the block. In the 9-bin orientation histogram, these orientation bins are evenly spaced over $[0^\circ, 360^\circ]$ (as shown in Fig. 3) as the horizontal axis, and the weighted votes are accumulated

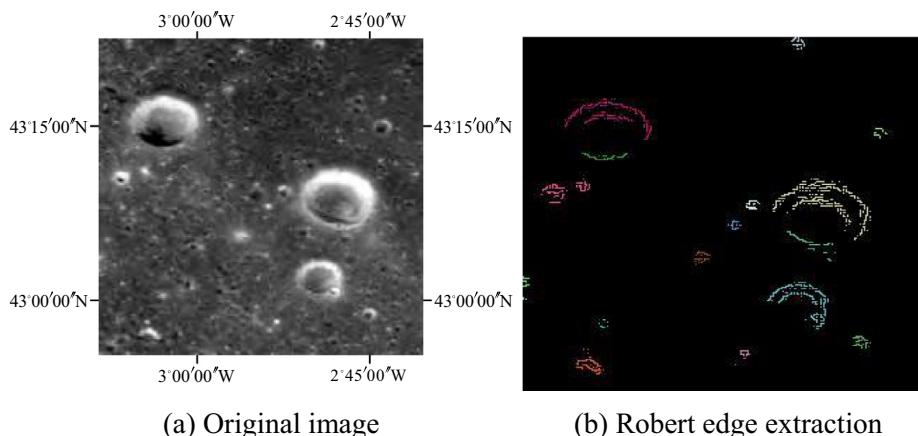


Fig. 2. An illustration of the Robert edge extraction.

into the orientation bins of histograms as the vertical axis, as shown in Fig. 4 (right). To reduce aliasing, the tri-linear interpolation is used to distribute the value of each gradient into adjacent orientation bins in the histogram. Fig. 4 describes the gradient histogram generation of a cell (8×8 gradient direction angles).

- (3) To accumulate a measure of local histogram energy over the larger connected regions, each block contains a group of cells, namely, adjacent 2×2 cells (as shown in Fig. 5(c)). Four orientation histograms of each block are integrated into a 36-dimensional feature vector that is then normalized into an L2 unit length (Lowe, 2004) to compensate for illumination. Moreover, adjacent blocks overlap with each other, which introduces the spatial information to each other.
- (4) The HOG per sliding window contains a concatenated vector of all its blocks (as shown in Fig. 5(d)).

2.2.2. Morphological features extracted from DEM data

According to the procedure described in Section 2.2.1, HOG can well encode the rim of an impact crater in optical images. Nevertheless, it should be noted that the visibility of impact craters in optical images is primarily affected by the illumination angles and the resolutions of images (Zhou et al., 2018). Consequently, the representation ability of the feature vector is limited if the impact craters are expressed from only a single perspective, such as the rim of the impact craters in optical images.

As shown in Fig. 6, different aspect values present inside and outside impact walls whereas slope value reflects the gradient change of impact craters and can be used to identify the steep crater walls. These slope and aspect information exhibit the 3D terrain morphological and topographic characteristics of impact craters for assisting in crater detection and are not affected by illumination. To quantita-

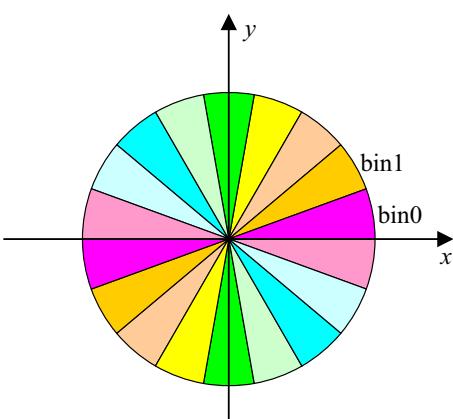


Fig. 3. Quantization of gradient direction θ . In Fig. 3, the angle range of each orientation bin is represented by different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tively describe the topographic structure and the spatial pattern of impact craters in DEM data, we follow the idea of spatial pyramid representations (Lazebnik et al., 2006) with the aim of combining spatial information with the 3D terrain morphological and topographic characteristics of impact craters, such as slope and aspect, into the feature vector. Thus, we in this paper construct HMA and HMS of impact craters derived from the topographic data (as shown in Figs. 7 and 8 respectively). The detailed process of constructing HMA is as follows:

- (1) We calculate the aspect of each pixel, which is voted into an 8-bin orientation histogram to construct the aspect probability histogram at the original spatial scale of the DEM. In the 8-bin orientation histogram, these orientation bins are evenly spaced over $[0^\circ, 360^\circ]$ as the horizontal axis, and the votes are accumulated into orientation bins of histograms as the vertical axis.
- (2) The length and width of the DEM are divided into 2^l units, which means that we can obtain 4^l patches (as shown in Fig. 7 (center left, level 1 and level 2)) of the same size DEM, where l represents the l th layer of the pyramid (in this paper, $l = 0, 1, 2$). The aspect probability histogram of each patch is then counted (as shown in Fig. 7 (center right)) and concatenated as the feature of the l th layer.
- (3) The histograms of all layers are concatenated together as a feature vector called HMA (as shown in Fig. 7(right)). The construction of HMS is akin to that of HMA (as shown in Fig. 8).

In this paper, DEM data are divided using regular grids of various scales, and the slope or aspect histogram of each grid in each scale is voted and arranged in a certain order to construct the probability model that contains the spatial information. These spatial relationships between the local regions are extremely crucial for improving the performance of the feature vector, which not only describes the 3D shape of the craters but also well represents the spatial layout. As shown in Figs. 7 and 8 (center right), the size of the feature vector increases with a decrease in the spatial scale, and the multiscale feature vector describes the morphological changes of the craters in 3D space from coarse to fine and well establishes the spatial relationship of the local region.

2.3. BN-based framework

As mentioned in Section 2.2, the shape characteristics of impact craters can be expressed by multi-source features derived from both optical images and DEM data. These multi-source features can usually provide redundant and complementary information from different perspectives for automatically detecting impact craters. Therefore, a combination of these multi-source features over the same

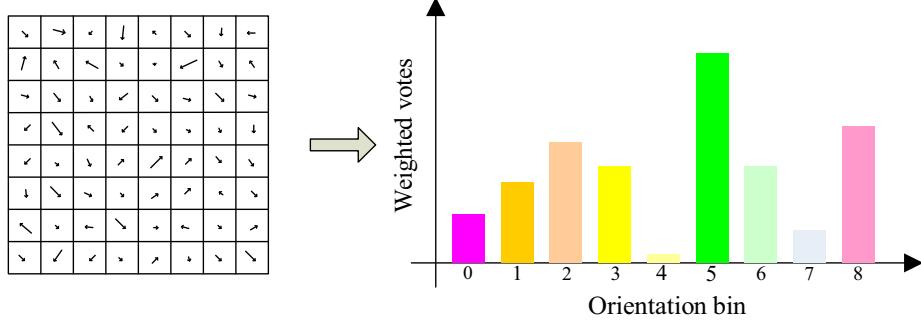


Fig. 4. An example of a cell gradient histogram generation. Each orientation bin is corresponding to angle range in Fig. 3.

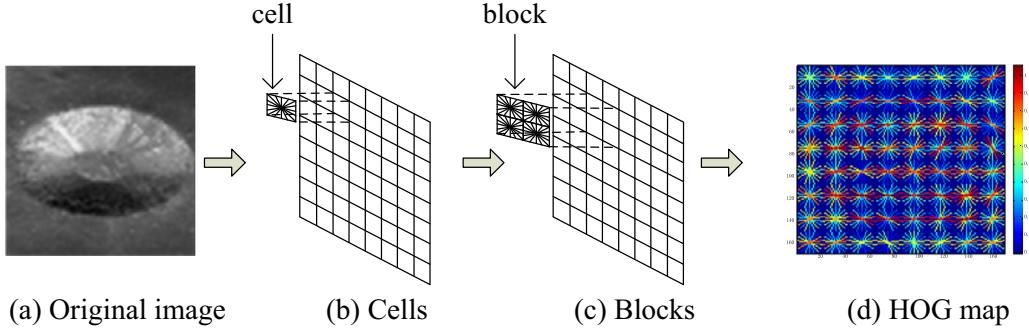


Fig. 5. An example of HOG extraction. (a) Original image; (b) Original image was partitioned into an array of cells and each cell was 8×8 pixels; (c) each block consisted of 2×2 spatial cells; (d) A feature map contained an iconic representation of the array of cells HOG.

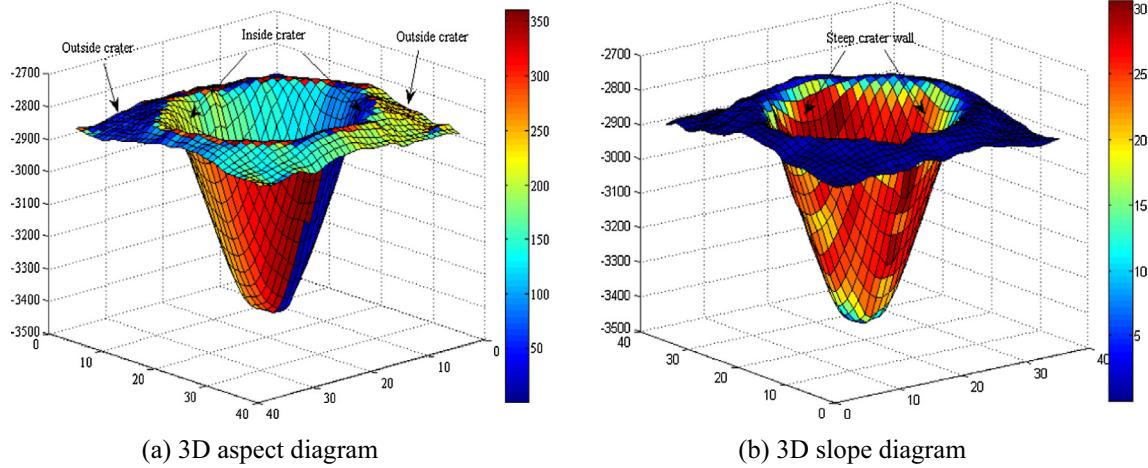


Fig. 6. Terrain analysis of impact craters.

scene will reduce imprecise detections and produce more consistent recognition results. Considering different degrees of reliability among these multi-source features (Lee et al., 1987; Waske and Benediktsson, 2007) for classification, each type of feature set, i.e., HOG, HMA or HMS in this paper, is treated separately. In terms of each type of feature set, a classifier is required to simulate the distribution which best fits the “true” data distribution and its output is an intermediate class posterior probability rather than a class label, which quantitatively measures the degree of “fit” between the instance to be classified and the classifier.

Then, we use a BN to combine all the intermediate class posterior probabilities derived from multiple classifiers for the final determinations. In this section, we build a BN-based decision fusion framework to distinguish craters from non-craters, which includes SVM models trained by an active learning strategy and BN-based fusion.

2.3.1. SVM models trained by an active learning strategy

Collecting positive and negative training samples of impact craters is arduous and time-consuming. Moreover, the dimensions of multi-source representations of an

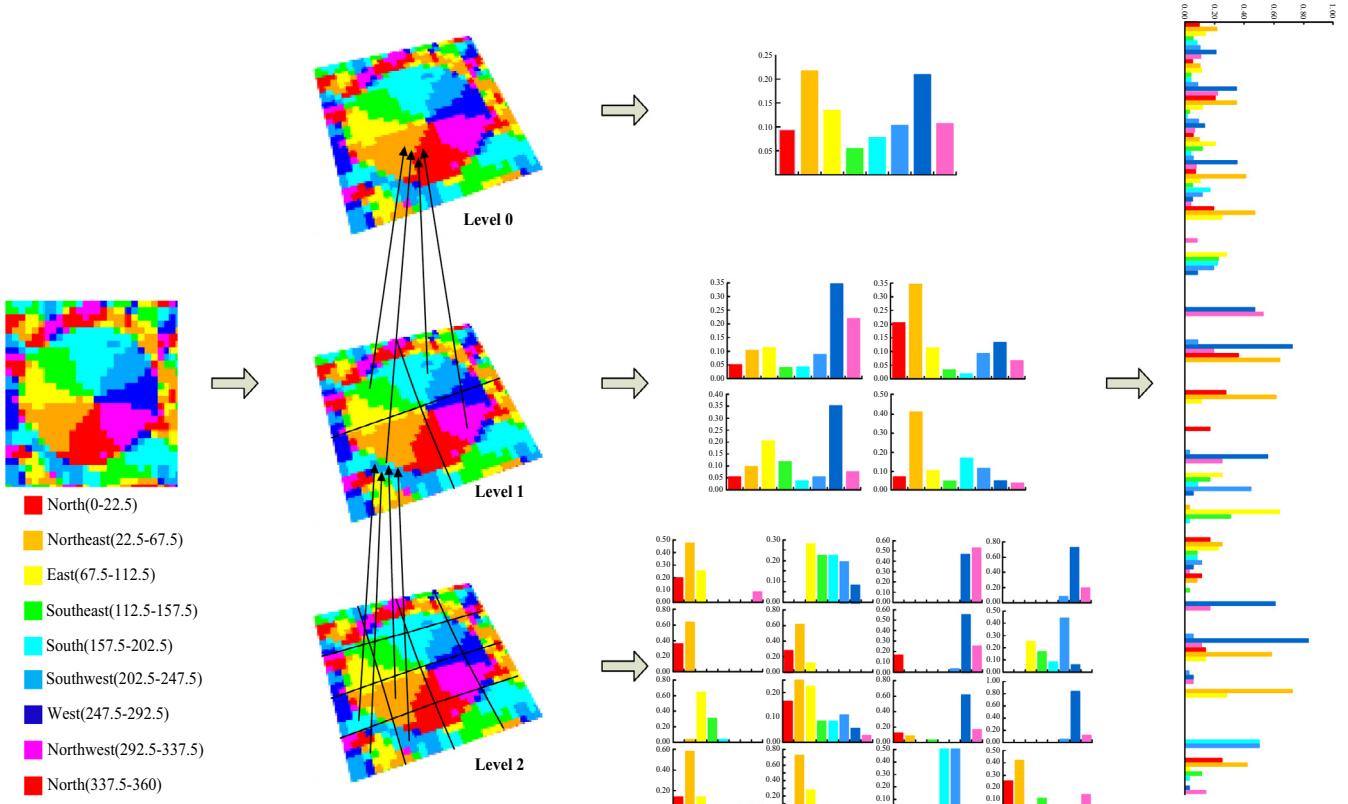


Fig. 7. An example of HMA constructed based on spatial pyramid theory. Different coloured bins in histograms correspond to different aspect ranges. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

impact crater are high. As a supervised method, SVM has an excellent ability in solving small samples, nonlinear and high dimensional classification problems. Therefore, we use the SVM model as the base classifier in the proposed method for the initial detections.

SVM (Cortes and Vapnik, 1995) is a binary classifier and its key idea is to map the nonlinear, separable feature space into a higher dimensional space using a kernel function (e.g., radial basis function) for an optimal decision hyperplane. Fig. 9 illustrates the geometric explanation of SVM. Let the training samples be $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, where n is the number of samples, $x_i (1 \leq i \leq n)$ denotes the feature vector of the training sample i , and $y_i (1 \leq i \leq n)$ ($y_i \in \{-1, 1\}$, in this paper) is the class label of the training sample i . Let the hyperplane function be $f(x) = \omega^T x + b$. The decision hyperplane can be found by solving the following constrained optimization problem:

$$\max \frac{1}{\|\omega\|}, \text{ s.t. } y_i(\omega^T x_i + b) \geq 1, i = 1, 2, \dots, n \quad (2)$$

which is equivalent to

$$\min \frac{1}{2} \|\omega\|^2, \text{ s.t. } y_i(\omega^T x_i + b) \geq 1, i = 1, 2, \dots, n \quad (3)$$

We then need Lagrange multipliers α for each of the constraints. Thus, the new objective function is Eq. (4)

$$\ell(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n \alpha_i (y_i(\omega^T x_i + b) - 1) \quad (4)$$

Let

$$\theta(\omega) = \max_{\alpha_i \geq 0} \ell(\omega, b, \alpha) \quad (5)$$

Given that all constraints are met, minimizing $\frac{1}{2} \|\omega\|^2$ is equivalent to minimizing $\theta(\omega)$. For the details, please refer to (Andrew, 2000).

Generally speaking, the more training samples are available, the better the performance of the supervised method. To solve the problem of the human annotation cost, an active learning strategy (Schohn and Cohn, 2000; Liu et al., 2012) to learn the optimal SVM model is adopted in this paper. Active learning selects samples from the unlabeled sample pool in a biased manner via query strategies (e.g., margin sampling (Schohn and Cohn, 2000)) that are designed to exploit properties of the classifier and the current labeled and auxiliary unlabeled data. This interactively constructed training set (as shown in Fig. 10) does not necessarily represent the entire sample space but is assumed to consist of samples that are the most reliable and informative for the classifier. Compared with conventional passive learning, in which the training samples are usually chosen randomly without interaction with the classifier, active learning can satisfactorily obtain the classification performance with fewer labeled samples. Thus, an active learning

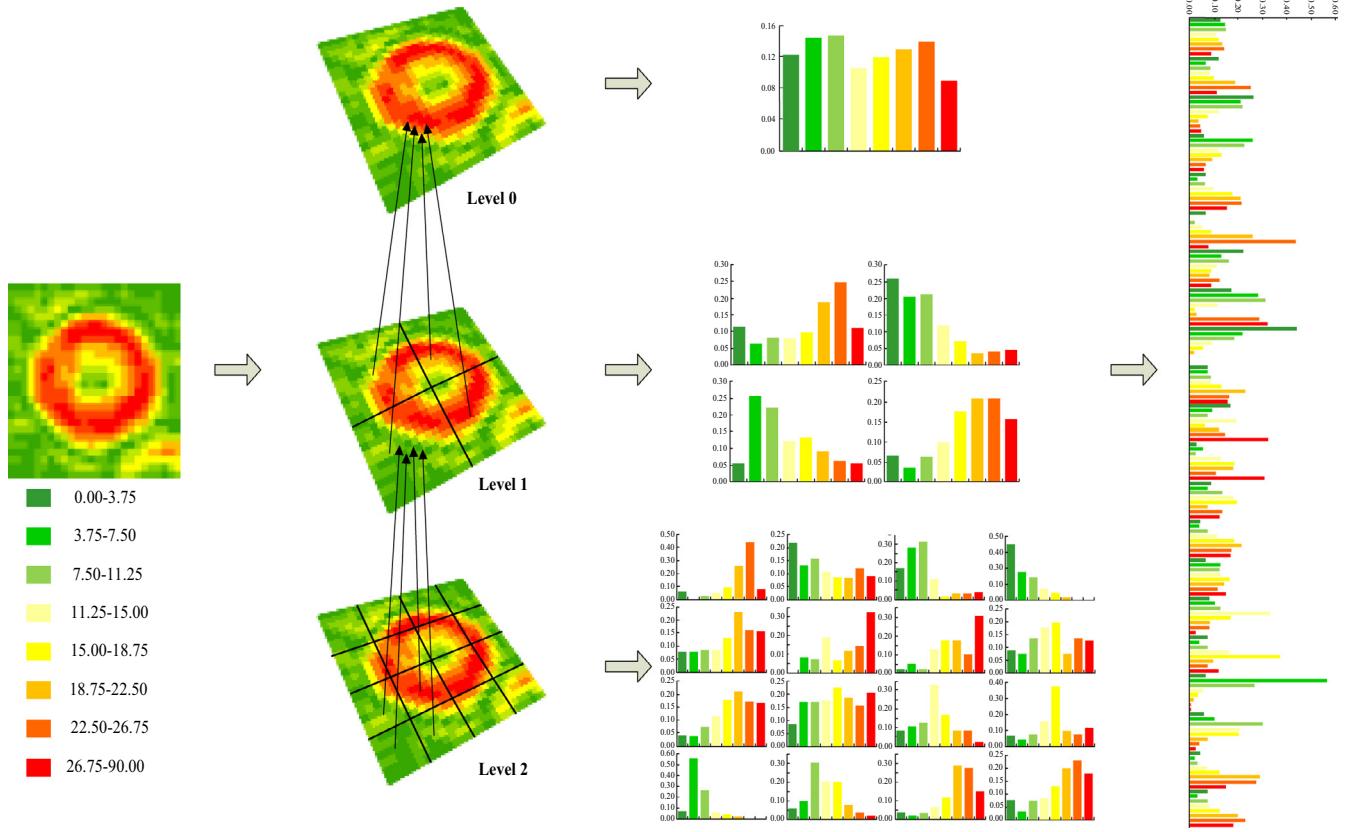


Fig. 8. An example of HMS constructed based on spatial pyramid theory. Different coloured bins in histograms correspond to different slope ranges. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

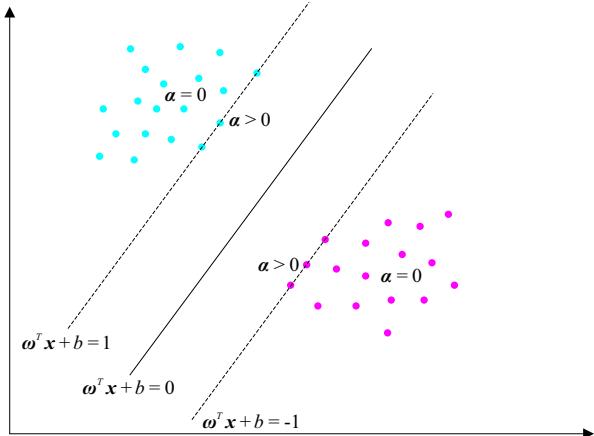


Fig. 9. Geometrical explanation of SVM. The solid line in the middle is the large margin decision hyperplane and its distance to the two dotted lines, where the support vectors are, is equal. In fact, Lagrange multipliers α of all the non-support vectors, i.e., $\omega^T x + b \neq \pm 1$, are zero, which is to say, the optimal decision hyperplane is only defined by the support vectors, i.e., $\omega^T x + b = \pm 1$.

- (1) Let the known sample set be divided into three parts, namely, the initial training sample set U_0 , candidate sample set U_C and validation sample set U_T . Among these sets, the initial training sample set U_0 contains at least one crater sample and one non-crater sample, which ensures that every class is represented.
- (2) The hyperplane f_{i-1} is learned based on the training sample set U_{i-1} . The validation sample set U_T is recognized using f_{i-1} , and the correctness rate of f_{i-1} is recorded. Then, the candidate sample set U_C is detected using f_{i-1} to select the most reliable and informative samples, which are marked as U_N . The training sample set is updated according to the following rules: (i) The sample in U_C is added to the new training sample set U_i if it is wrongly classified. (ii) For a sample, the further the distance between the sample and the hyperplane is, the larger the confidence in the classification result. Therefore, the samples in U_C are selected and added to U_i according to the distance between the samples and the hyperplane. $U_i = U_{i-1} \cup U_N$ and $U_C = U_C - U_N$ are then implemented.
- (3) Repeat step (2), until candidate dataset U_C is empty or correctness rate is unchanged.

The class information acquired from the SVM model can be transformed into the intermediate posterior proba-

strategy potentially exploits a greater amount of information for the data and significantly reduces the human annotation cost. Fig. 11 exhibits an example of hyperplane optimization. The detailed optimization process is as follows (as shown in Fig. 10):

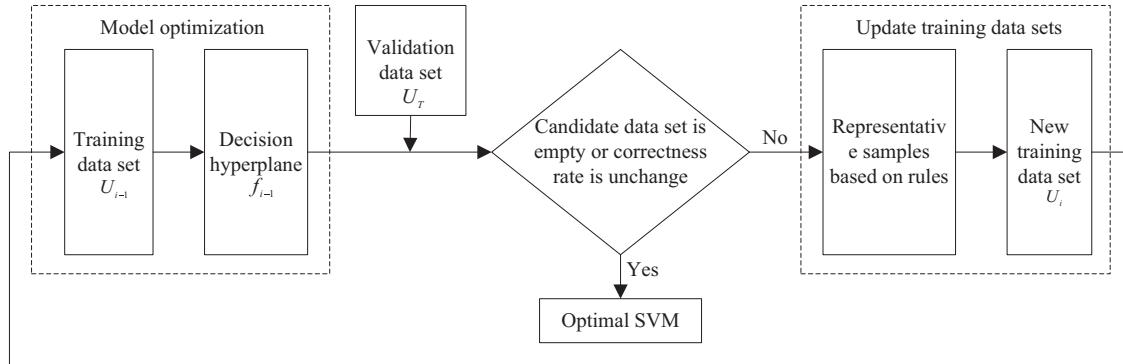


Fig. 10. Optimal SVM learned by active learning strategy.

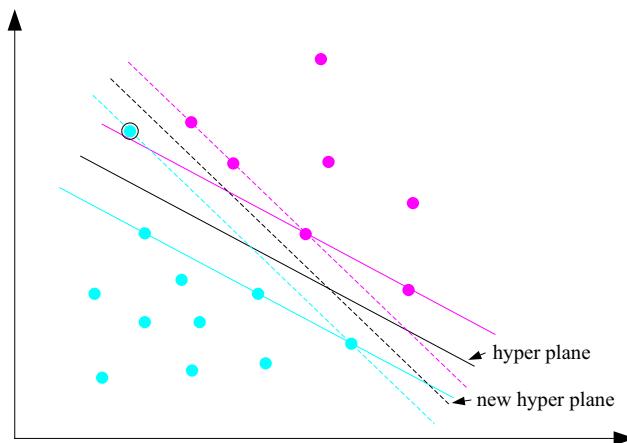


Fig. 11. An example of hyperplane optimization. The blue point with a black box is the sample that is classified wrongly by the current decision hyperplane, and the black full line is the current decision hyperplane while the black dotted line denotes the new optimal decision hyperplane optimized by iteration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

bility by calculating the distance between the unknown sample and the hyperplane after the optimal hyperplane is learned. The output of the standard SVM model, which is the distance between the unknown sample and the optimal hyperplane, is processed and transformed into the posterior probability by using the sigmoid-fitting method (Platt, 2000) as follows:

$$P(y_i = 1|f(\mathbf{x}_i)) = \frac{1}{1 + \exp(Af(\mathbf{x}_i) + B)} \quad (6)$$

where A and B denote the parameters that need to be solved and $f(\mathbf{x}_i)$ denotes the output of the standard SVM model where \mathbf{x}_i is used as the input. The sigmoid-fitting method not only preserves the sparseness of the SVM model but also estimates the posterior probability well.

2.3.2. BN-based fusion

Owing to external space environments, such as solar wind and internal geological deformations, impact craters on the lunar surface generate various degrees of degener-

ation (Bandeira et al., 2012). For example, some craters have distinctive edges and bright splash lines while others exhibit blurry edges and are merged into the surrounding background, which results in uncertainty of the initial detections generated from the SVM models. BN is a powerful model for uncertainty reasoning, which combines probability and graph theory (Koller and Friedman, 2009). Thus, we construct a decision fusion model using BN to combine all the intermediate class posterior probabilities derived from multiple SVMs for the final determinations.

BN is an easily understood and increasingly popular model that represents domain variables and their conditional dependencies as nodes in a directed acyclic graph. Given a database, a node or variable is conditionally dependent upon its parents. This model partitions the joint probability distribution over all the investigated variables into a set of conditional probabilities. Hence, for a BN with K nodes, the joint distribution is given by

$$p(\mathbf{X}) = \sum_{k=1}^K p(X_k|\text{pa}(X_k)) \quad (7)$$

where X_k is a random variable, $\text{pa}(X_k)$ denotes the set of parent node of X_k , and $\mathbf{X} = \{X_i | 1 \leq i \leq K\}$.

The calculation of posterior probabilities and decision-making by the BN takes place in Eq. (8). Assume that $\chi_1, \chi_2, \dots, \chi_w$ are the outputs of the SVM models, w is the number of SVM models (in this paper, $w = 3$), and c is a class node, and the following formula can be derived according to the Bayes formula:

$$\begin{aligned} P(c|\chi_1, \chi_2, \dots, \chi_w) &= \frac{P(\chi_1, \chi_2, \dots, \chi_w|c)P(c)}{P(\chi_1, \chi_2, \dots, \chi_w)} \\ &= P(c)\sum_{i=1}^w \frac{P(\chi_i|c)}{P(\chi_i)} \end{aligned} \quad (8)$$

where $P(c)$ is a flat prior (Li et al., 2016), $P(\chi_i)$ is constant for all the possible states of the class node c , and $P(\chi_i|c)$ is obtained from the SVM model. The category for each instance to be classified (as a crater or non-crater) is determined based on maximum a posterior probability.

Fig. 12 demonstrates a simple illustration of detecting impact craters inside each crater candidate. Following the idea of face or pedestrian detection, we scan each of the

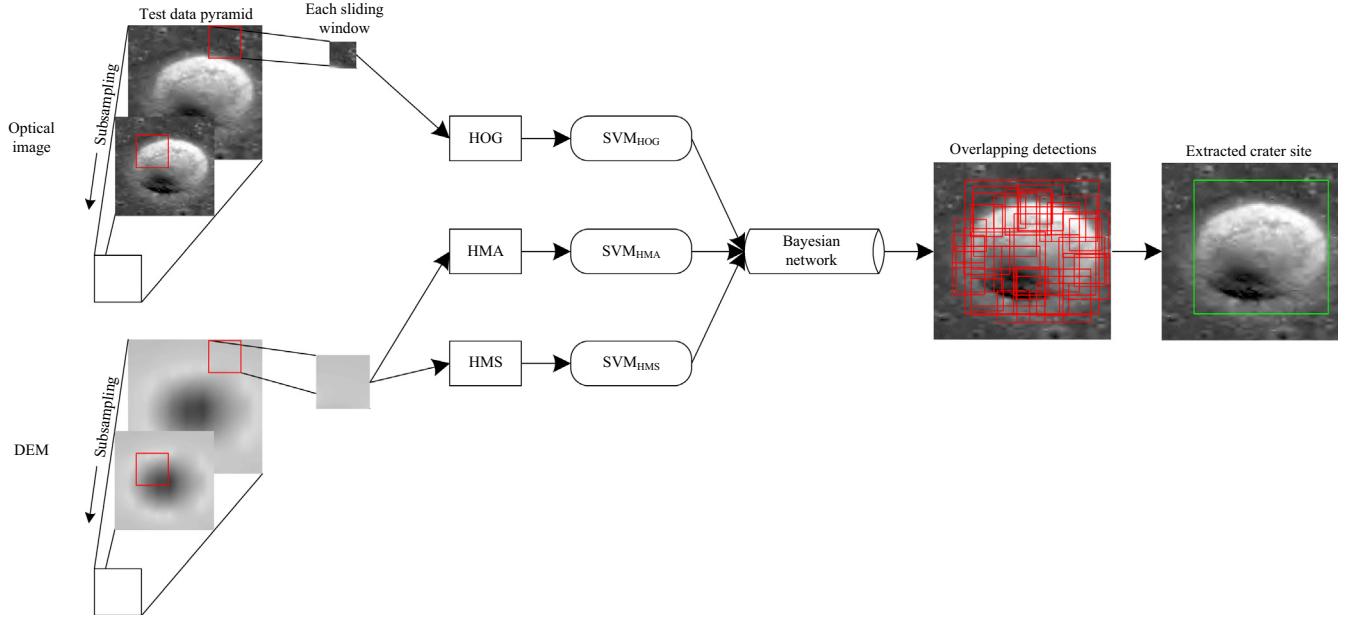


Fig. 12. An illustration of detecting crater site inside the crater candidate.

crater candidates found in the new and unseen test scenes using a sliding window (Lampert et al., 2008). In the procedure of recognition; HOG, HMA and HMS of each sliding window are calculated and fed into the corresponding SVM models respectively for the intermediate posterior probability matrix p_{ij} ($i = 1, 2, \dots, n$, $j = 1, 2, \dots, k$; in this paper, $k = 2$, i.e., crater and non-crater). These posterior probabilities serve as the input of BN and the posterior probability of the class node is further inferred. As a result, each sliding window inside candidate is classified to the class (i.e., crater or non-crater) that holds the highest posterior probability. As a matter of fact, impact craters on the lunar surface might have different sizes. To detect impact craters of different size, we resize the test data iteratively and fix the size of sliding window (set to 10×10 pixels). During iteration, the test data is decreased to 0.9 times the size of previous one. The iteration is not terminated until the size of query candidate is decreased less than the size of sliding window. In our experiment, to insure the feature vector have the same size, we resize the image of the sliding window into 64×64 pixels to calculate HOG, and DEM of the sliding window into 128×128 pixels to construct HMA and HMS. Due to the sliding window in the crater detection, an impact crater inside each candidate may be detected multiple times by different sliding windows. Thus, we use the non-maximum suppression method based on the mean-shift algorithm (Dala, 2006) to check the position of the detected impact crater regions and delete the repeated ones.

3. Experimentation and analysis

To evaluate the effectiveness and robustness of the proposed method, we performed both qualitative and quanti-

tative evaluations using both orthoimages by LRO and DEM data acquired by LOLA from three lunar test scenes. The spatial resolution of the orthoimages was approximately 50 m. The LOLA, which is one instrument that the orbiter of LRO mission carries, provided a precise global lunar topographic model and its spatial resolution was approximately 30 m. Table 1 lists the detailed information from the three lunar test scenes, primarily including latitude range, longitude range, the approximate area and the number of craters. The three lunar test scenes were situated in the middle and eastern part of the Mare Imbrium, which were relatively flat area. Training and test data were chosen by taking into account different erosion rates of the craters, craters with various diameters and densities, different gray intensity contrast changes, in order to verify the effectiveness of combining the optical images with DEM data since two types of data could provide complementary information from different perspectives. The three test scenes were marked manually, depending on the following criteria: (i) impact craters exhibit circular or elliptical characteristics (Wetzler et al., 2005), (ii) impact craters are considered as a pair of crescent-like highlight and shadow regions in a gray-level image (Urbach and Stepinski, 2008), and used as the ground truth data to which the results of our proposed method were compared. The dimension of these craters varies between 500 and 17,000 m in diameter. To measure the detection performance, we use the region overlap criterion (Mikolajczyk et al., 2005) between the region A and the ground truth region A_{gt} to determine whether an impact crater is correctly detected. The area of the union and the intersection of the regions are computed numerically.

$$\frac{A \cap A_{gt}}{A \cup A_{gt}} > 0.5 \quad (9)$$

Table 1
Information from the three lunar test scenes.

	Top	Bottom	Left	Right	Approximate area (km ²)	Crater count manually
Scene I	44°0'58"N	42°2'17"N	11°42'35"W	9°36'11"W	3900	124
Scene II	43°29'11"N	40°26'23"N	4°24'35"W	1°21'16"W	8600	303
Scene III	33°26'53"N	30°58'8"N	16°27'19"W	13°43'7"W	6200	367

where $A \cap A_{gt}$ denotes the intersection of the regions, and $A \cup A_{gt}$ denotes their union.

This paper uses the recall (Re), precision (Pr) and F_1 -score as the evaluation metrics. The recall represents a measure of completeness or quantity, the precision denotes a measure of exactness or quality, and the F_1 -score combines the precision and recall with equal weights. Hence, we have

$$Re(\%) = \frac{TP}{TP + FN} \times 100 \quad (10)$$

$$Pr(\%) = \frac{TP}{TP + FP} \times 100 \quad (11)$$

$$F_1 - \text{score}(\%) = \frac{2 \times Pr \times Re}{Pr + Re} \times 100 \quad (12)$$

where TP denotes the number of craters that are correctly determined to be craters, FN represents the number of craters that are incorrectly classified as non-craters, FP denotes the number of non-craters that are incorrectly determined to be craters.

To learn our proposed method, we in this paper randomly selected 2000 non-crater sample pairs and 2068 crater sample pairs (517 original crater sample pairs and 1551 rotated or symmetric ones (Krizhevsky et al., 2012) that were used to avoid the effect of illumination direction) from optical images and associated DEM data. In these sample pairs, 100 non-crater sample pairs and 100 crater sample pairs were chosen randomly to establish a test sample set U_T , 10 non-crater sample pairs and 10 crater sample pairs were selected randomly to build an initial training sample set U_0 , and the rest of the sample pairs were the candidate training sample set U_C . Part of the different scale crater sample pairs are shown in Fig. 13(a) and (b).

3.1. Impact crater candidate construction and detection

As aforementioned, the edge extraction-based method for finding a crater candidate (as mentioned in Section 2.1) was conducted prior to the BN-based fusion framework. As shown in Fig. 14(a)–(c), crater candidates in the test scenes (I–III) were highlighted with red polygons. In the procedure for selecting crater candidates, regions with a sharp edge were detected and marked from the images and a large part of image background was removed from being recognized. These crater candidates were then fed into the BN-based framework (as mentioned in Section 2.3) to distinguish craters from non-craters.

As shown in Fig. 15(a)–(c), impact craters detected in the lunar test scene (I–III) using our proposed method in

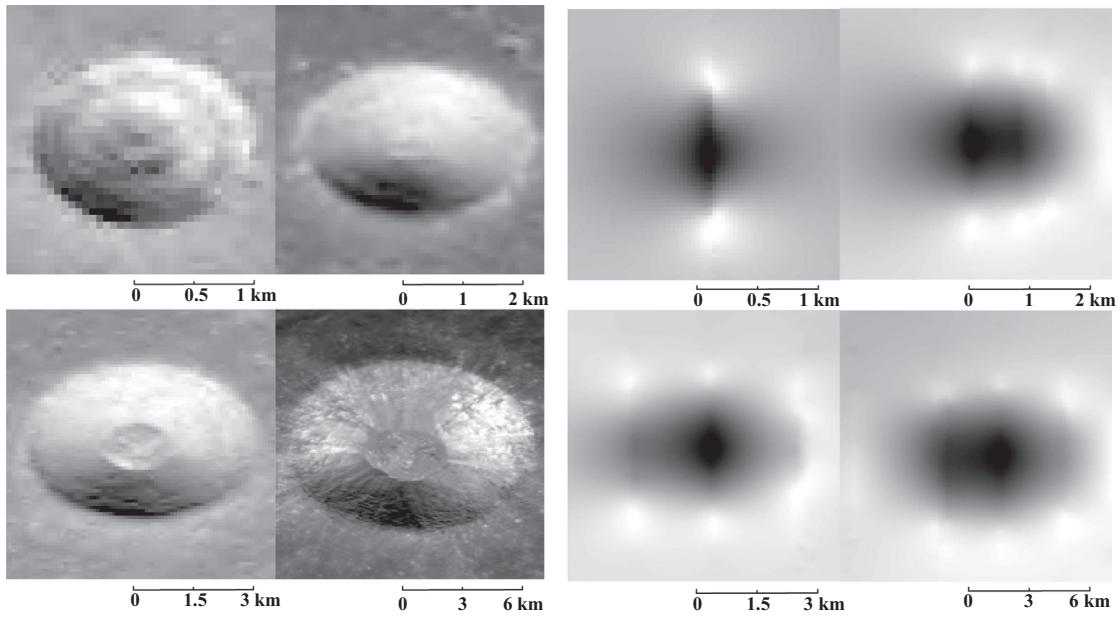
this paper were highlighted with green polygons. In the step for finding crater candidates, multiple impact craters might be aggregated into the same crater candidate (as shown in Fig. 14(a)–(c), the sites that light blue¹ arrows point to) since they were close to each other. In the crater recognition stage based on the BN-based framework, a plurality of craters belonging to the same crater candidate area could be effectively separated (as shown in Fig. 15(a)–(c), the sites that light blue arrows point to). Nevertheless, at the stage of eliminating repeated identifications, the small craters near the large craters might also be falsely rejected from the identification results (as shown in Fig. 15(b), the sites that orange arrows point to, as shown in Fig. 15(a), the sites that yellow arrows point to).

Computational efficiency is an important factor for crater detection algorithms. In the training stage, the processing time mainly depends on the number of both positive and negative samples. In our experiments, 2000 non-crater sample pairs and 2068 crater sample pairs were used to establish the optimal BN-based framework. The total training time is approximately 1 h 41 m 20.3 s. The computational time of the proposed method for processing the three lunar test scenes was shown in Table 2.

3.2. Effectiveness of combining optical image with DEM

As described in Section 2.2.2, inclusion of 3D morphological information of impact craters aimed to reduce wrong extraction. To compare the performance of different feature combinations, the influences of combining different types of features used in the detection were evaluated in this paper. The HOG-based method, the HMA-based method and the HMS-based method, which used associated feature set respectively, directly built a SVM classifier for detecting impact craters. For the HOG-HMA-based method, which used a combination of HOG and HMA, we trained SVM classifiers based on HOG and HMA respectively and implemented the decision-making within the BN framework to fuse the two SVM classifiers for the final detections. With regards to the HOG-HMS-based method, which used a combination of HOG and HMS, the SVM classifiers were learned based on HOG and HMS respectively, which were embedded into the BN framework as child nodes and the extraction results were finally determined by integrating the two SVM classifiers within the BN framework. The proposed method in

¹ For interpretation of color in Figs. 14 and 15, the reader is referred to the web version of this article.



(a) Impact crater samples from optical images (b) Impact crater samples from DEM

Fig. 13. Different scale crater sample pairs from optical images and corresponding DEM (partial). DEM data are rendered as gray-scale images.

this paper exploited a combination of HOG, HMA and HMS (as described in Section 2). The receiver operating characteristic (ROC) curves (Fawcett, 2003) corresponding to different feature combinations are generated and shown in Fig. 16. The ROC curve is a parametric curve that compares two characteristics of the classifier, i.e. true positive rate and false positive rate. Experimental results suggested that compared with the methods that only used one type of feature set (i.e., the HOG-based method, the HMA-based method and the HMS-based method), the methods based on a combination of multi-source features (i.e., the HOG-HMA-based method, the HOG-HMS-based method and the proposed method) exhibited a better recognition performance. Furthermore, the proposed method achieved a higher true positive rate and a lower false positive rate than both the HOG-HMA-based method and the HOG-HMS-based method.

3.3. Comparison with other combination techniques

As mentioned earlier, combining optical image with DEM was used for reducing the number of false positives and improving the detection performance. In recent years, numerous methods have been developed to address multi-source data fusion issues. To prove the superiority of the proposed method, the Vector Stacking (VS) method using SVM (Bruzzone and Carlin, 2006), Majority Voting (MV) method (Benediktsson and Kanellopoulos, 1999) and logarithmic opinion pool (LOGP) method (Benediktsson and Sveinsson, 2003) were compared to our proposed method. For the VS method using SVM, HOG, HMA and HMS were concatenated together as a feature vector that was used to train a SVM classifier for detecting impact craters.

In terms of the MV method, different SVM classifiers were trained and then the final decision results were aggregated according to the majority voting rule. The LOGP method used the weighted product of the posterior probabilities obtained by the individual classifier. Performance comparison between the proposed method and other combination techniques was summarized in Table 3, which demonstrated that compared with other combination techniques, the proposed method showed higher F_1 -score, with average difference of 4.25%. Experimental comparison results illustrated that the proposed method succeeded in addressing multi-source data fusion issues for improving the crater detection.

3.4. Comparison with other existing methods

To further evaluate the performance of our proposed method in this paper, the frequently-used boosting-based method (Martins et al., 2009; Ding et al., 2011; Bandeira et al., 2012; Xin et al., 2017) and the MPGA-based method (Degirmenci and Ashyralyev, 2010) were compared to our proposed method using the three scenes. For the boosting-based method, three different methods, including Boosting-based method (image), Boosting-based method (DEM) and Boosting-based method (DEM + image), were performed. Boosting-based method (image) used optical images alone, where Haar-like and LBP features were extracted and then used to train an AdaBoost classifier for distinguishing craters from non-craters. Boosting-based method (DEM) used DEM data alone, where DEM was simply considered as an image, Haar-like and LBP features were extracted and then used to train an AdaBoost classifier for distinguishing craters from non-craters. Boosting-based method (DEM +

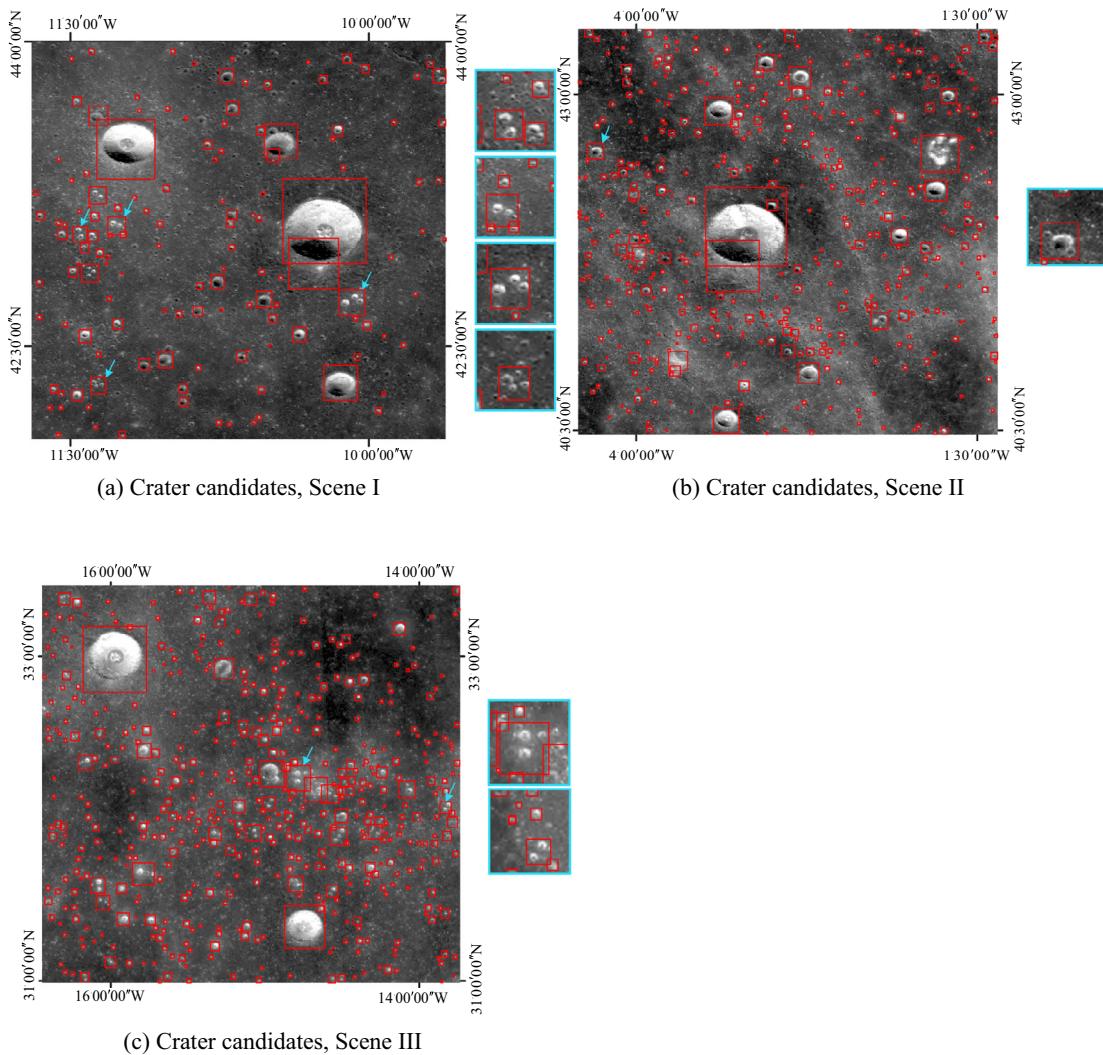


Fig. 14. Crater candidates marked in different scenes.

+ image) used both DEM data and optical images, where HOG, HMA and HMS were concatenated together as a feature vector that was used to train an AdaBoost classifier for detecting impact craters. The MPGA-based method was a data fusion approach, where impact craters were extracted from optical images and these extracted impact craters were further verified and eliminated by the basins extracted from DEM for reducing the number of false positives. Table 4 presents the performance of the proposed method and other existing methods according to their recall, precision and F_1 -score values. Compared with other existing methods, the proposed method achieved the highest detection rates in terms of F_1 -score with average difference of 3.14% and recall with average difference of 4.75%. Fig. 17 shows comparison of F_1 -score on different size range among different methods, which suggested that the proposed method was quite advantageous for detecting the impact craters with diameters less than 1000 m, between 1000 m and 2000 m and larger than 2000 m. We also concluded that for detecting impact craters larger than

2000 m, all of these methods showed a good extraction performance with F_1 -score of over 90%. However, both the proposed method and the boosting-based method significantly outperformed the MPGA-based method for extracting impact craters less than 2000 m, with average difference in the value of F_1 -score of 3.02%.

4. Conclusions

This paper developed a decision fusion method within the BN framework for automatically detecting impact craters using both optical images and DEM data. The edge extraction-based method was first implemented to find impact crater candidates, which remarkably reduced the computational complexity. Secondly, the multi-source representations of an impact crater were proposed and established based on both optical images and DEM data to provide redundant information for reducing imprecise extractions and complementary information for generating a more consistent recognition. With regards to multi-

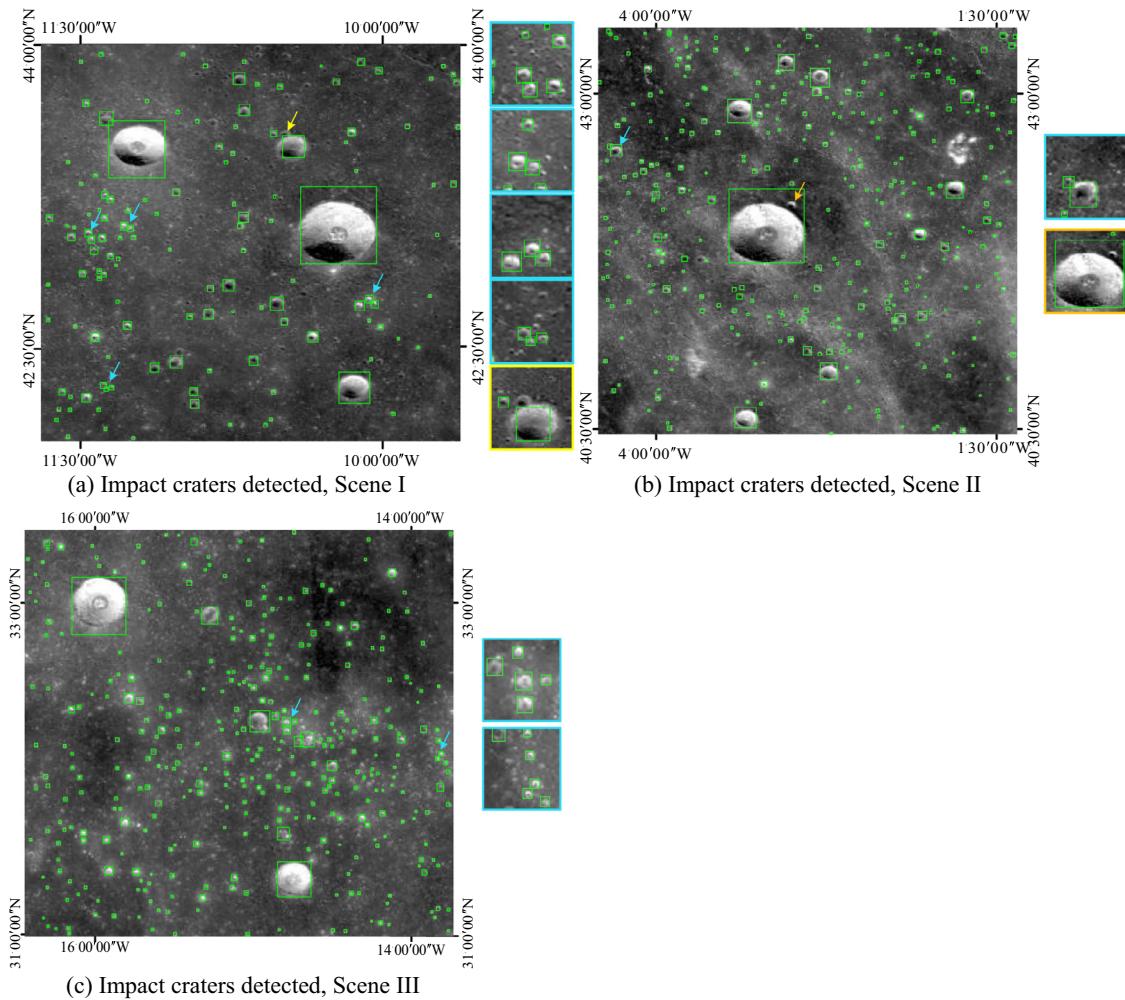


Fig. 15. Detection results of the proposed method in different scenes.

Table 2

Computational time for processing the three lunar test scenes (m for minute, s for second).

Test scenes	Number of candidates	Processing time	
		Candidate construction	Crater detection
Scene I	151	1 m 16.8 s	4 m 11.2 s
Scene II	490	2 m 37.1 s	10 m 48.2 s
Scene III	509	3 m 7.3 s	11 m 50.7 s

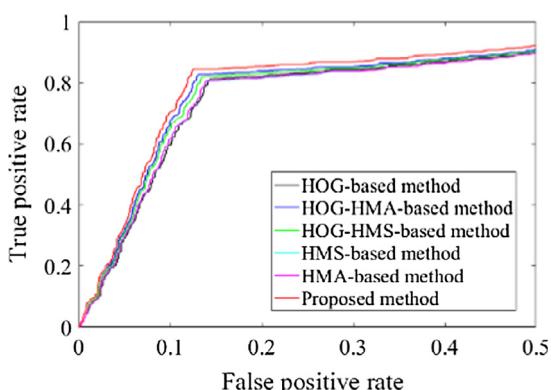


Fig. 16. Receiver operating characteristic curve for different feature combinations.

Table 3

Performance comparison between the proposed method and other combination techniques.

Test scenes	Methods	Re (%)	Pr (%)	F ₁ -score (%)
Scene I	Proposed method	84.68	82.68	83.68
	VS method using SVM	79.03	77.78	78.40
	MV method	80.65	81.30	80.97
	LOGP method	78.23	79.51	78.86

source features, HMA and HMS derived from DEM data were built based on aspect and slope information respectively to quantitatively describe the 3D morphology and topography of impact craters while HOG was used to

Table 4
Performance comparisons between different methods.

Test scenes	Methods	<i>Re</i> (%)	<i>Pr</i> (%)	<i>F</i> ₁ -score (%)
Scene I	Proposed method	84.68	82.68	83.68
	Boosting-based method (image)	83.87	80.62	82.21
	Boosting-based method (DEM)	80.65	78.13	79.37
	Boosting-based method (DEM + image)	84.68	81.40	83.00
	MPGA-based method	71.77	86.41	78.41
Scene II	Proposed method	89.11	82.32	85.58
	Boosting-based method (image)	87.30	80.24	83.54
	Boosting-based method (DEM)	87.79	79.17	83.26
	Boosting-based method (DEM + image)	88.78	81.27	84.86
	MPGA-based method	73.93	83.90	78.60
Scene III	Proposed method	83.38	86.69	85.00
	Boosting-based method (image)	81.30	87.39	84.18
	Boosting-based method (DEM)	79.56	83.43	81.45
	Boosting-based method (DEM + image)	82.29	86.29	84.24
	MPGA-based method	70.03	83.44	76.15

The best performance is marked with Bold fonts.

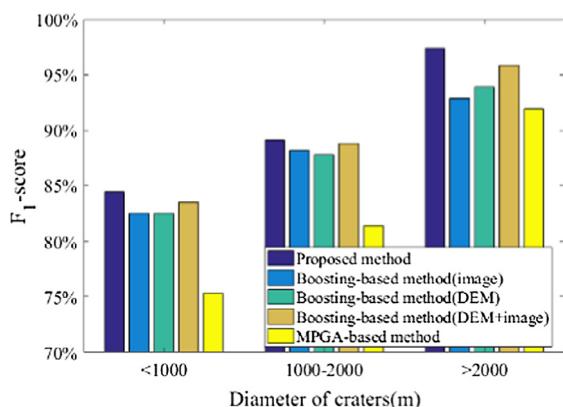


Fig. 17. Comparison of *F*₁-score on different size range among different methods.

encode the rim of impact craters in optical images. In the subsequent detection procedure, each type of feature set (i.e., HOG, HMA and HMS) was treated separately and classified by an SVM model. Then, the intermediate class posterior probabilities estimated by multiple SVM models were integrated under the BN framework for the uncertainty reasoning, which reduced the error extractions obtained by the single-source features. Experimental results indicated that integrating optical images with DEM data significantly decreased the number of false positives compared with using optical images alone, with *F*₁-score of 84.8% on average. Moreover, our proposed method exhibited notable advantages over the other existing methods for craters of all size ranges.

With the improvement of planetary image resolution, it is feasible to automatically extract the small-scale impact craters. However, serious noise problems from which the high resolution planetary images suffer also pose challenges for representing these small-scale impact craters. Thus, representation learning of the small-scale impact craters

using convolutional neural network will be the future research since convolutional neural network can generate the high-level semantic representation for image classification. Moreover, most existing supervised methods for detecting impact craters use the sliding window as primitive and therefore produce a large number of repeated and redundant computations for detecting the impact craters with different sizes. In the future, we attempt to use a pixel-level method, such as fully convolutional neural network, for addressing the repeated and redundant detection issues.

Acknowledgement

This work was supported in part by the Natural Science Foundation of China under Grant 41872207, 41602215 and in part by the National Key Research and Development Program of China under Grant 2016YFE0104400.

References

- Andrew, A.M., 2000. An introduction to support vector machines and other kernel-based learning methods. Printed in the United Kingdom at the University Press.
- Bandeira, L., Ding, W., Stepinski, T.F., 2012. Detection of sub-kilometer craters in high resolution planetary images using shape and texture features. *Adv. Space Res.* 49 (1), 64–74.
- Bandeira, L., Saraiva, J., Pina, P., 2007. Impact crater recognition on mars based on a probability volume created by template matching. *IEEE T. Geosci. Remote* 45 (2), 4008–4015.
- Barlow, N.G., 1988. Crater size-frequency distributions and a revised Martian relative chronology. *Icarus* 5 (2), 285–305.
- Benediktsson, J.A., Kanellopoulos, I., 1999. Classification of multisource and hyperspectral data based on decision fusion. *IEEE T. Geosci. Remote* 37 (3), 1367–1377.
- Benediktsson, J.A., Sveinsson, J.R., 2003. Multisource remote sensing data classification based on consensus and pruning. *IEEE T. Geosci. Remote* 41 (4), 932–936.
- Bruzzone, L., Carlini, L., 2006. A multilevel context-based system for classification of very high spatial resolution images. *IEEE T. Geosci. Remote* 44 (9), 2587–2600.

- Bruzzone, L., Lizzi, L., Marchetti, P.G., Earl, J., Milnes, M., 2004. Recognition and detection of impact craters from EO products. Proceedings of ESA-EUSC.
- Bue, B.D., Stepinski, T.F., 2006. Machine detection of Martian impact craters from digital topography data. *IEEE T. Geosci. Remote* 45 (1), 265–274.
- Chen, M., Liu, D., Qian, K., Li, J., Lei, M., Zhou, Y., 2018. Lunar crater detection based on terrain analysis and mathematical morphology methods using digital elevation models. *IEEE T. Geosci. Remote* 56 (7), 3681–3692.
- Cohen, J.P., Ding, W., 2014. Crater detection via genetic search methods to reduce image features. *Adv. Space Res.* 53 (12), 1768–1782.
- Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Kluwer Academic Publishers.
- Dalal, N., 2006. Finding People in Images and Videos. Institut National Polytechnique de Grenoble.
- Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. In: Computer Vision and Pattern Recognition, IEEE Computer Society Conference, pp. 886–893.
- Degirmenci, M., Ashyralyyev, S., 2010. Impact Crater Detection on Mars Digital Elevation and Image Model. Middle East Technical University.
- Di, K., Li, W., Yue, Z., Sun, Y., Liu, Y., 2014. A machine learning approach to crater detection from topographic data. *Adv. Space Res.* 54 (11), 2419–2429.
- Ding, M., Cao, Y.F., Wu, Q.X., 2008. Autonomous craters detection from planetary image. In: International Conference on Innovative Computing Information and Control. IEEE, p. 443.
- Ding, W., Stepinski, T.F., Mu, Y., Bandeira, L., Ricardo, R., Wu, Y., Lu, Z., Cao, T., Wu, X., 2011. Subkilometer crater discovery with boosting and transfer learning. *ACM T. Intel. Syst. Tec.* 2 (4), 39.
- Earl, J., Chicarro, A.F., Koeberl, C., Marchetti, P.G., Milnes, M., 2005. Automatic recognition of crater-like structures in terrestrial and planetary images. In: 36th Annual Lunar and Planetary Science Conference, p. 36.
- Fawcett, T., 2003. ROC graphs: notes and practical considerations for researchers. *Pattern Recogn. Lett.* 31 (8), 1–38.
- Haruyama, J., Ohtake, M., Matsunaga, T., Morota, T., Honda, C., Yokota, Y., Abe, M., Ogawa, Y., Miyamoto, H., Iwasaki, A., Pieters, C.M., Asada, N., Demura, H., Hirata, N., Terazono, J., Sasaki, S., Saiki, K., Yamaji, A., Torii, M., Josset, J., 2009. Long-lived volcanism on the lunar farside revealed by SELENE Terrain Camera. *Sci.* 323 (5916), 905.
- Honda, R., Azuma, R., 2000. Crater extraction and classification system for lunar images. *Mem. Fac. Sci. Kochi Univ.* 21, 13–22.
- Jin, S., Zhang, T., 2014. Automatic detection of impact craters on Mars using a modified adaboosting method. *Planet. Space Sci.* 99 (1), 112–117.
- Kang, Z., Luo, Z., Hu, T., Gamba, P., 2016. Automatic extraction and identification of lunar impact craters based on optical data and DEMs acquired by the Chang'e satellites. *IEEE J-Stars* 8 (10), 4751–4761.
- Kim, J.R., Muller, J.P., Gasselt, S.V., 2005. Automated crater detection, a new tool for mars cartography and chronology. *Photogramm. Eng. Rem. S.* 71 (10), 1205–1218.
- Koller, D., Friedman, N., 2009. Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning. MIT Press.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. In: International Conference on Neural Information Processing Systems. Curran Associates Inc., pp. 1097–1105.
- Lampert, C.H., Blaschko, M.B., Hofmann, T., 2008. Beyond sliding windows: object localization by efficient subwindow search. *Comput. Vis. Pattern Recogn.*, 1–8.
- Lazebnik, S., Schmid, C., Ponce, J., 2006. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. *IEEE CVPR*, 2169–2178.
- Lee, T., Richards, J.A., Swain, P.H., 1987. Probabilistic and evidential approaches for multisource data analysis. *IEEE T. Geosci. Remote* 3, 283–293.
- Li, M., Stein, A., Bikker, W., Zhan, Q., 2016. Urban land use extraction from very high resolution remote sensing imagery using a Bayesian network. *Isprs J. Photogramm.* 122, 192–205.
- Liu, A., Zhou, D., Chen, L., Chen, M., 2015. Saliency detection and edge feature matching approach for crater extraction. *J. Syst. Eng. Electron.* 26 (6), 1291–1300.
- Liu, D., Chen, M., Qian, K., Lu, G., Lin, H., 2017. Boundary detection of dispersal impact craters based on morphological characteristics using lunar digital elevation model. *IEEE J-Stars* 99, 1–15.
- Liu, S., Ding, W., Gao, F., Stepinski, T.F., 2012. Adaptive selective learning for automatic identification of sub-kilometer craters. *Neurocomputing* 92, 78–87.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60 (2), 91–110.
- Luo, L., Wang, X.Y., Ji, W., Li, C., 2011. Automated detection of lunar craters based on Chang'E-1 CCD data. In: International Congress on Image and Signal Processing. IEEE, pp. 883–887.
- Marques, J.S., Pina, P., 2015. Crater delineation by dynamic programming. *IEEE Geosci. Remote S.* 12 (7), 1581–1585.
- Martins, R., Pina, P., Marques, J.S., Silveira, M., Silveira, M., 2009. Crater detection by a boosting approach. *IEEE Geosci. Remote S.* 6 (1), 127–131.
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van, G.L., 2005. A comparison of affine region detectors. *Int. J. Comput. Vision* 65 (1–2), 43–72.
- Morota, T., Haruyama, J., Honda, C., Ohtake, M., Yokota, Y., Kimura, J., Matsunaga, T., Ogawa, Y., Hirata, N., Demura, H., Iwasaki, A., Miyamoto, H., Nakamura, R., Takeda, H., Ishihara, Y., Sasaki, S., 2009. Mare volcanism in the lunar farside Moscovense region: implication for lateral variation in magma production of the Moon. *Geophys. Res. Lett.* 36 (21), 272–277.
- Ojala, T., Pietikäinen, M., Mäenpää, T., 2001. A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. *24(7)*, 397–406.
- Paris, S., Kornprobst, P., Tumblin, J., 2009. Bilateral filtering. *Int. J. Numer. Meth. Eng.* 63 (13), 1911–1938.
- Platt, J.C., 2000. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Adv. Large Margin Class.* 10 (4), 61–74.
- Robbins, S.J., Hynek, B.M., 2012. A new global database of Mars impact craters ≥ 1 km: 1. database creation, properties, and parameters. *J. Geophys. Res. Planet.* 117 (E5), 5004.
- Salamuniccar, G., Loncaric, S., 2010. Method for crater detection from Martian digital topography data using gradient value/orientation, morphometry, vote analysis, slip tuning, and calibration. *IEEE T. Geosci. Remote* 48 (5), 2317–2329.
- Salih, A.L., Schulte, P., Grumpe, A., Wohler, C., Hiesinger, H., 2017. Automatic crater detection and age estimation for mare regions on the lunar surface. 2017 25th European Signal Processing Conference.
- Sawabe, Y., Matsunaga, T., Rokugawa, S., 2006. Automated detection and classification of lunar craters using multiple approaches. *Adv. Space Res.* 37 (1), 21–27.
- Schohn, G., Cohn, D., 2000. Less is more: active learning with support vector machines. In: Seventeenth International Conference on Machine Learning. Morgan Kaufmann Publishers Inc., pp. 839–846.
- Stepinski, T.F., Ding, W., Vilalta, R., 2012. Detecting impact craters in planetary images using machine learning. *Intell. Data Anal. Real-Life Appl.: Theory Pract.*
- Stepinski, T.F., Ghosh, S., Vilalta, R., 2007a. Machine learning for automatic mapping of planetary surfaces. In: National Conference on Innovative Applications of Artificial Intelligence. AAAI Press, pp. 1807–1812.
- Stepinski, T.F., Mendenhall, M.P., Bue, B.D., 2007b. Robust automated identification of Martian impact craters. In: Lunar and Planetary Science Conference, Lunar and Planetary Science Conference, p. 1202.
- Stepinski, T.F., Mendenhall, M.P., Bue, B.D., 2009. Machine cataloging of impact craters on Mars. *Icarus* 203 (1), 77–87.

- Tanaka, K.L., 1986. The stratigraphy of Mars. *J. Geophys. Res. Sol. Ea.* 91 (B13), 139–E158.
- Troglia, G., Moigne, J.L., Benediktsson, J.A., Moser, G., Serpico, S.B., 2012. Automatic extraction of ellipsoidal features for planetary image registration. *IEEE Geosci. Remote S.* 9 (1), 95–99.
- Urbach, E.R., Stepinski, T.F., 2008. Automatic detection of sub-kilometer craters in high resolution images of Mars. In: *Lunar and Planetary Science Conference*, p. 2184.
- Viola, P., Jones, M., 2001. Robust object detection using a boosted cascade of simple features. *Proc. Cvpr*, 1–511.
- Wang, Y., Yang, G., Guo, L., 2015. A novel sparse boosting method for crater detection in the high resolution planetary image. *Adv. Space Res.* 56 (5), 982–991.
- Waske, B., Benediktsson, J.A., 2007. Fusion of support vector machines for classification of multisensor data. *IEEE T. Geosci. Remote* 45 (2), 3858–3866.
- Wetzel, P.G., Honda, R., Enke, B., Merline, W.J., Chapman, C.R., Burl, M.C., 2005. Learning to detect small impact craters. *Appl. Comput. Vision* 1, 178–184.
- Xie, Y., Tang, G., Yan, S., Lin, H., 2013. Crater detection using the morphological characteristics of Chang'E-1 digital elevation models. *IEEE Geosci. Remote S.* 10 (4), 885–889.
- Xin, X., Di, K., Wang, Y., Wan, W., Yue, Z., 2017. Automated detection of new impact sites on Martian surface from HiRISE images. *Adv. Space Res.* 60 (7), 1557–1569.
- Yamamoto, S., Matsunaga, T., Nakamura, R., Sekine, Y., Hirata, N., Yamaguchi, Y., 2017. An automated method for crater counting using rotational pixel swapping method. *IEEE T. Geosci. Remote* 55 (8), 4384–4397.
- Yin, J., Li, H., Jia, X., 2015. Crater detection based on gist features. *IEEE J-stars* 8 (1), 23–29.
- Yue, S., He, L., Wen, Y., Lu, G., Lin, H., 2013. Shape characteristics-based extraction of lunar impact craters: using DEM from the Chang'E-1 satellite as a data source. *Ann. Gis* 19 (1), 53–62.
- Zhou, Y., Zhao, H., Chen, M., Tu, J., Yan, L., 2018. Automatic detection of lunar craters based on DEM data with the terrain analysis method. *Planet. Space Sci.*