Crater Detection Using Bayesian Classifiers and LASSO

Ying Wang[1], Wei Ding[2], Kui Yu[1], Hao Wang[1] and Xindong Wu[1,3]

[1]Department of Computer Science, Hefei University of Technology, Hefei, 230009, China
[2]Department of Computer Science, University of Massachusetts Boston, Boston, MA 02125, USA
[3]Department of Computer Science, University of Vermont, Burlington, VT 05405, USA
wangying182@163.com, wei.ding@umb.edu, ykui713@gmail.com, jsjxwangh@hfut.edu.cn, xwu@cs.uvm.edu.

Abstract—surveying a large amount of small sub-kilometer craters in planetary images is a challenging task due to their non-distinguishable features. In this paper, we integrate the LASSO (Least Absolute Shrinkage and Selection Operator) method with the Bayesian network classifier and propose an L1 Regularized Bayesian Network Classifier (L1-BNC) algorithm for this task. The L1-BNC algorithm uses the LASSO method not only to deal with high-dimensional crater features, but also to give a crater feature order for constructing a Bayesian network classifier. Our framework is evaluated on a large Martian image of $37,500 \times 56,250m^2$. Experimental results demonstrate that this proposed method gets higher prediction accuracy than the existing crater detection algorithms.

Keywords-bayesian classification; feature selection; crater detection; Lasso

I. INTRODUCTION

Impact craters, or craters for short, are formed by high velocity collisions of solid celestial bodies onto the planet's surface. Yielding information about the past and present geological processes, craters become the most studied geomorphic features in the solar system. However, during a crater life cycle, they are often eroded, buried, overlaid, and transformed, which lead to various difficulties in crater detection from remote images. Furthermore, geologically more active surfaces contain more degraded craters. Therefore, efficient crater detection in planetary images remains as a challenging task [12, 7, 11].

The existing work on crater detection can be divided into manual detection and automatic detection. Manual detection focuses on large craters of planetary images, for example, 42,283 Martian craters with diameters larger than 5 km [3], and 8,497 named lunar craters with diameters larger than a few kilometers [1]. However, the size distribution of craters follows power-law: large craters that can be easily identified manually are rare and small sub-kilometer craters are abundant [14]. Surveying such small craters is ill-suited for visual detection, due to their shear numbers, but well-suited for an automated technique. Therefore, crater auto-detection techniques are needed, especially to catalog small sub-kilometer craters. In a word, automating the process of small crater detection is the only practical solution to a comprehensive surveying of sub-kilometer craters [7].

Existing research efforts on automatic detection techniques in planetary images can be divided into two general categories: unsupervised and supervised approaches. The unsupervised methods identify crater rims in an image as circular or elliptical features [13, 9, 6, 2, 12]. However, the performance is usually at least one magnitude less accurate than supervised methods. The supervised methods take advantage of domain knowledge in the form of labeled training sets that guide classification algorithms [5, 17, 18]. In addition, those previous efforts relie on inefficient exhaustive search of the entire image using pixel-based approaches which may work for finding a small number of large craters in low resolution images, but not for finding a very large number of small craters in high resolution images. Billions of pixels in a high resolution planetary image inevitably become a bottleneck of scalability of those crater detection methods [7, 15].

Urbach and Stepinski [15] proposed the idea of finding candidate craters instead of inefficient exhaustive search of the entire image. Crater candidates are the regions of an image that can potentially contain craters which can be viewed parts of an image that contain crescent-like pairs of shadows and highlights. The work of Urbach and Stepinski only uses a small set of features to describe the shapes of the shadow and high regions of crater candidates. However, other non-crater landforms in similar shapes make using shape features, an un-ideal choice on crater detection. Recently, Ding et al. [7] proposed a boosting with feature selection framework for crater detection. In this framework, they not only used the idea of candidate crater, but also was the first research team that constructed image gradient texture features from crater candidates for rapid feature extraction. Their framework contains three steps as follows. Firstly, they utilize mathematical morphology on shape detection for efficient identification of regions indicative for craters; Secondly, they extract and select image gradient texture features; at the last, a set of boosting algorithms with feature selection is proposed for crater detection. Texture features effectively enrich the information for crater detection so that we can effectively capture potential gradient structure of remotely sensed images without priori domain knowledge. However, how to efficiently deal with those features in high dimensions is a difficult problem for traditional supervised learning algorithms.

Based on the crater framework proposed by Ding et al. [7], in this paper, we propose a new method on feature selection by integrating the LASSO method [4] into the Bayesian network classifier and propose an L1 Regularized Bayesian Network Classifier (L1-BNC). LASSO (Least Absolute Shrinkage and Selection Operator) was firstly introduced by Tibshirani [4] which is a least-square like problem with the addition of L1 penalty on the parameter vector for shrinkage and feature selection. The advantage of

LASSO is that it is not only for feature selection but also for establishing a good feature order for the selected features. Therefore, in a high-dimensional crater dataset, L1-BNC with LASSO can not only select the strongly relevant crater features, but also establish a good feature order to further construct a Bayesian network classifier. The proposed method is evaluated on a large and high resolution image of Martian ($37,500 \times 56,250$ m^2) surface. Experimental results show that L1-BNC outperforms other crater detection methods with respect to prediction accuracy.

The rest of the paper is organized as follows. Section 2 describes the method for crater detection based on a Bayesian network classifier with LASSO. Section 3 presents our experimental result in a large high resolution planetary image from Mars. Section 4 summarizes our work and discusses future directions.

II. BAYESIAN NETWORK CLASSIFIERS WITH LASSO FOR CRATER DETECTION

In this section, based on the research efforts of Ding et al. [7], a method which integrates Bayesian classifier with LASSSO is proposed for crater detection. The proposed method contains four steps: (1) constructing crater candidates; (2) extracting texture features from candidate craters (3) using LASSSO to select a subset of the strongly relevant features and sorting the selected features by their correlation with the class label, (4) training a Bayesian network classifier based on the sorted features and perform classification.

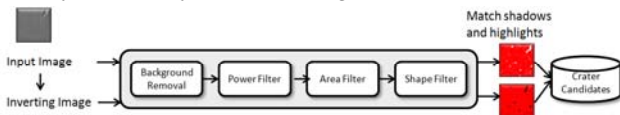### A. A framework for constructing candidate craters

Figure 1. Diagram illustrating individual steps in constructing crater candidates

Fig. 1 shows the framework for constructing candidate craters [7]. A key insight in their method is that pair wise crescent-like shapes (candidate craters, the locations where craters likely reside) are identified from images using a shape detection method based on mathematical morphology [7, 15, 16]. In Fig. 1, the step of background removal deletes shapes, such as mountains, that are too large to be part of the craters; Power Filter, Area Filter and Shape Filter remove those features that lack sufficient contrast and are unreliable for crater detection. In the final step, highlight and shadow regions are matched so that each pair corresponds to a single crater candidate. This framework does not have high enough accuracy to constitute a stand-alone crater detection technique, but is ideal for identification of crater candidates.

B. Feature Extraction and Feature Selection

Wang and Ding et al. [11] constructed the training set and testing set which contains 19 crater geometric shape features from a high resolution panchromatic image h0905_0000. They used three feature selection methods to deal with those 19 features. In their results, when the number

of selected features is up to 14 or 18, the accuracy of crater detection achieves the highest value with SVM using a Greedy and AUC feature selection algorithm. Ding et al. [7] worked on the same image and divided this image into three regions to extract 1,089 image gradient texture features on these regions. And then they iteratively construct a boosting classifier and selected 150 features from the original 1,089 image texture features for crate detection.

Motivated by above research work, we design an L1-BNC to deal with those high dimensional crater features for crate detection which selects features by the LASSO method. The main difference between L1-BNC and the work of Ding and Wang is that L1-BNC can dynamically adjust the number of selected features by the prediction accuracy. In the other words, unlike the previous work, L1-BNC can vary the number of selected features to dynamically assess the prediction accuracy, and then select the feature subset with the highest detection accuracy.

C. A framework for crater detection

In this section, we first give the main flow for crater detection in Fig. 2. In Fig. 2, at first, crater candidates are identified which contain crescent-like pairs of shadows and highlights from an image. Secondly, the training set and testing set are sampled based on those crater candidates. In the final step, L1-BNC is used to discover craters and non-craters from crater candidates.
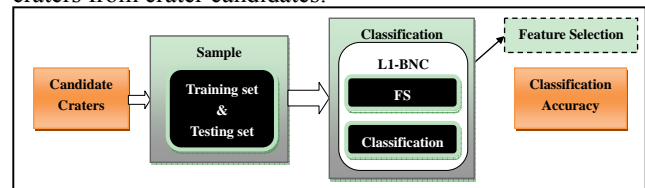

Figure 2. Main flow for carter detection

L1-BNC integrates Bayesian network classifiers with LASSO, so that it can not only choose the most effective feature subsets for classification, but also sort the selected features according to their correlation with the class label. The L1-BNC algorithm in detail is shown in Fig. 3 which is divided into two stages. The first stage is feature subset selection using the LARS algorithm (a very efficient solution to the LASSO problem), which chooses the most relevant features. And the second stage sorts the selected feature subset and uses the feature order to construct an exact Bayesian network classifier for crate detection.

The first step of L1-BNC is mainly used to select the feature subset. We first initialize the number of each feature subset to be selected. And then L1-BNC selects a feature subset according to the correlation with class label by calling the function L1-BNC_LARS (), detailed description of this function is explained in Fig. 4. This function can not only select a strongly relevant feature set, but also give an order of the selected features with the class label. The LARS method is a useful and less greedy version of traditional forward selection methods: given a collection of possible predictors, we select the one having largest absolute correlation with the response y, say xj1, and perform simple linear regression of

y on xj1. This leaves a residual vector orthogonal to xj1, now considered to be the response.

Algorithm:

L1-BNC for crater detection

Require

Input crater candidates $(x_1, y_1),...,(x_n, y_n)$ where $y_i = 0, 1,$ $i = 1,…,n$, for non-crater and crater examples respectively.

1. Initialize Num=number of features to be selected
2. For each Num do
(1)   **Feature selection and Feature Ranking:**
                L1-BNC_LARS(CraterTrainData, Num);
Obtaining a feature subset, including the strongly relevant features with class label, denoted by S, and the feature ordering of this feature subset according to their correlation with the class label.
(2)   **Classification procedure:**
           // CraterTrainData and CraterTestSet are train datasets
// and test datasets.
L1-BNC_Classification(S,CraterTrainData,CraterTestSet);
Return classification accuracy.
End for

Figure 3.   Algorithm L1-BNC for classification

We project the other predictors orthogonally to x_{j1} and repeat the selection process. After k steps, this results in a set of predictors $x_{j1}, x_{j2},..., x_{jk}$ are then used in the usual way to construct a k-parameter linear model. From the function, we can see that LARS selects the most correlated variable with the residual r at each time, while r just presents the information of selected variables and the response variable (the class label).

Feature Selection and feature Ranking:

L1-BNC_LARS (CraterTrainData, Num)

Require
(1) Dataset:CraterTrainData $(x_1, x_2,..., x_n, y)$,
    CraterTestData $(x_1, x_2,..., x_n, y)$, where x_i is texture feature, class
    label $y = 1$ for a crater, otherwise $y = 0$ for a non crater.
(2) Num: the number of selected features.
For i=1:Num do

       Comptute: $\arg\min_{\beta}\{\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2\}$

       β_j is the jth regression coefficient, x_{ij} is crater feature, p is the
       number of selected features, and n is the number of cases, as
       follows:
(1) Start with all coefficients β_i equal to zero. Find predictor x_i most correlated with the response y.
(2) Along with the most possible in the direction of the predictor x_i, stop when some other predictor x_k has as much correlation with the current residual ($r = y - \hat{y}$).
(3) Increase (β_i, β_k) in their joint least squares direction, until some other predictor x_m has as much correlation with the residual r by the inner product.
(4) Continue until the number of selected features achieves num.
End For
    Obtaining a ordering of the selected features.

Figure 4.   Feature Selection and Feature Ranking for crater features

The second step of L1-BNC is to call function L1-BNC_Classification. This function uses the feature order obtained by the first step to learn a Bayesian network classifier by K2 algorithm [8], which learning Bayesian networks from data is often motivated by its ability to find the network structure efficiently, given that a reasonable variable ordering is provided. Then we use this trained classifier to detect the candidate craters, detailed description of this function is discussed in Fig. 5.

Crater detection:

L1-BNC_Classification(S,CraterTrainSet,CraterTestSet)

Require

(1) Feature subset: S.
(2) CraterTrainSet and CraterTestSet: Datasets in the format of $(x_1, x_2,..., x_n, y)$, where x_i is texture feature, class label $y = 1$ for a crater, otherwise $y = 0$ for a non crater.

1.   **K2(S, CraterTrainSet)**; //each CraterTrainSet contains
                             //Num selected features.
   For i:=1:Num
       Search the parents of x_i from set of nodes that precede x_i in the feature set S, which operates by initially assuming that a node has no parents, and then adding incrementally that parent whose addition most increases the probability of the resulting network.
   End For
       We can obtain the topology of the Bayesian network classifier, denoted by BNC.

2.   **Evaluating the BNC**
   For each test set
       Evaluating BNC on the test set.
       Obtaining classification accuracy.
   End For

Figure 5.   L1-BNC for Crater detection

III. EXPERIMENT RESULTS

We have selected a portion of the High Resolution Stereo Camera (HRSC) nadir panchromatic image h0905 [10], taken by the Mars Express spacecraft, to serve as the test set. As illustrated in Fig. 6, the selected image has a resolution of 12.5 meters/pixel and a size of 3,000 by 4,500 pixels (37,500 × 56,250m^2). The image represents a significant challenge to automatic crater detection algorithms because it covers a terrain that has spatially variable morphology and its contrast is rather poor if inspected at a small spatial scale. We divide the image into three sections denoted by the west region, the central region, and the east region (see Fig. 6). The central region is characterized by surface morphology that is distinct from the rest of the image. The west and east regions have similar morphology but the west region is much more heavily cratered than the east region.

The testing set are from the west region, the central region, the east region and the corresponding crater candidate numbers of samples are 6708, 2935, 2026, respectively. The training set consists of 204 true craters and 292 non-crater examples selected randomly from amongst crater candidates located in the northern half of the east region. The number of features is 1,089 and 1 class label for all data sets.
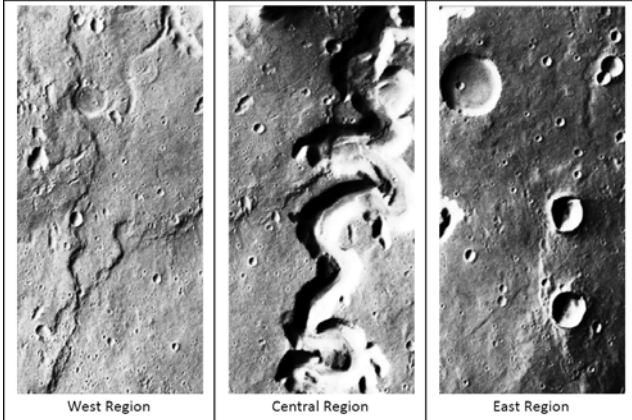
Figure 6.   The Test images

A.   The results with varied number of selected features

To construct classifiers, the numbers of selected features are 2, 5, 10, 14, 18, 20, 25, 50, 100, 120, 150 from the training set (see the analysis of section 2.2), then, we test those classifiers in the west region, the central region and the east region, respectively. Table 1 show that the classification accuracy of each feature subset on the three regions (the highest prediction accuracy marked by the bold face).

As depicted in Table 1, when the number of selected features is 2, the classification accuracy on three areas is rather low, which reveals that although a crater can be consider as containing crescent-like pairs of shadows and highlights, only using two features to detect a crater is far from enough because of the eroding, burying, overlaying, and transforming problem during a crater life cycle. Therefore, it is necessary to increase more features to improving the accuracy of crater detection.
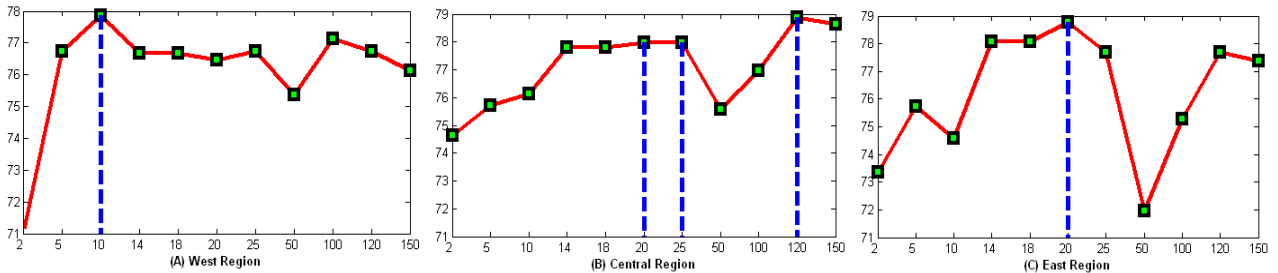
TABLE I.   THE CLASSIFICATION ACCURACY FOR EACH FEATURE SUBSET

Features	Training Set	TestSet1	TestSet2	TestSet3
2	crater_train2	71.17	74.62	73.35
5	crater_train5	76.73	75.72	75.72
10	crater_train10	**77.85**	76.12	74.58
14	crater_train14	76.68	77.79	78.08
18	crater_train18	76.68	77.79	78.08
20	crater_train20	76.45	**77.96**	**78.78**
25	crater_train25	76.73	**77.96**	77.69
50	crater_train50	75.37	75.57	71.96
100	crater_train100	77.12	76.97	75.27
120	crater_train120	76.73	**78.88**	77.69
150	crater_train150	76.13	78.64	77.39

In order to clearly illustrate the results of Table 1, Fig. 7 gives how the classification accuracy changes with varied number of selected features on three regions where x-axis is the number of selected features: 2, 5, 10, 14, 18, 20, 25, 50, 100, 120, and 150, y-axis is the classification accuracy and dashed represents the number of selected features when the classification accuracy achieve the highest value.

Fig. 7(A) shows that when the number of selected features is up to 10, the prediction accuracy achieves the highest value 77.85 in west region. In the central region, when the number of selected features is up to 120, the prediction accuracy obtain the best value 78.88, while the number of selected features is up to 20 or 25, the prediction accuracy only decrease 0.0092 in Fig. 7(B). In Fig. 7(C), when the number of selected features is up to 20, the prediction accuracy achieves the highest value 78.78 in east region, which is slightly higher than the west region. But when the number of selected features is equal to 50, L1-BNC gets the lowest accuracy. Therefore, when the number of selected features from three areas is 10, 120, and 20 respectively, the corresponding prediction accuracy is highest, up to 77.85, 78.88 and 78.78, respectively.

Figure 7.   Prediction accuracy changes with the varied number of selected features

B.   Comparative Performance with the Existing Algorithms

In this section, we compare L1-BNC with the existing algorithms for crater detection, such as Naïve Boosting, J48, and SVM. A Naïve Boosting algorithm (NBoost for short) was proposed by Ding et al. [7] which integrated the boosting algorithm and greedy feature selection algorithm while J48 and SVM perform without feature selection in our experiments. Table 2 shows the prediction accuracy of the four algorithms, where 10/120/20 denotes the number of selected features by L1-BNC in the west, central and east regions respectively (the highest prediction accuracy in Table 2 marked in the bold face).

TABLE II.   THE PREDICTIVE ACCURACY OF COMPARING L1-BNC WITH NAÏVE BOOSTING, J48 AND SVM

Algorithm		Features	West Region	Central Region	East Region
Feature Selection	L1-BNC	10/120/20	**77.85**	**78.88**	**78.78**
	NBoost	150	76.61	**78.88**	77.49
No Feature Selection	J48	1090	76.91	77.24	76.36
	SVM	1090	76.83	77.10	77.54

Table 2 shows that the prediction accuracies of both L1-BNC and Naïve Boosting are higher than those of J48 and SVM without feature selection. L1-BNC outperforms the Naïve Boosting algorithm on the west and east regions. On the central region, the two algorithms have the same performance.

IV.   CONCLUSIONS

The aim of this paper is to present a reliable method for auto-detection of sub-kilometer craters from high resolution images of planetary surfaces. In this framework, we integrate the LASSO method with the Bayesian network classifiers and propose the L1 Regularized Bayesian Network Classifier (L1-BNC) for crater detection. Different from previous work, L1-BNC can vary the number of selected features to dynamically assess the prediction accuracy, and then select the feature subset with the highest detection accuracy. Experiments show that L1-BNC outperforms other classification algorithms.

REFERENCES

[1] Andersson, L. E. and Whitaker, E. A. 1982. Nasa catalog of Lunar nomenclature. In NASA Reference Publication 1097.

[2] Bandeira, L., Saraiva, J., and Pina, P. 2007. Impact crater recognition on Mars based on a probability volume created by template matching. In IEEE Transactions on Geoscience and Remote Sensing. 4008 - 4015.

[3] Barlow, N. G. 1988. Crater size-frequency distributions and a revised Martian relative chronol-ogy. Icarus 75, 285-305.

[4] Bradley Efron, Trevor Hastie, Iain Johnstone and Robert Tibshirani, Least Angle Regression, Annals of Statistics (with discussion) (2004) 32(2), 407-499.

[5] Burl, M. C., Stough, T., Colwell, W., Bierhaus, E. B., Merline, W. J., and Chapman, C. 2001. Automated detection of craters and other geological features. In Int. Symp. Artif. Intell. Robot. and Autom. Space. Montreal, QC, Canada.

[6] Cheng, Y., Johnson, A. E., Matthies, L. H., and Olson, C. F. 2002. Optical landmark detection for spacecraft navigation. In the 13th Annual AAS/AIAA Space Flight Mechanics Meeting. AAS 03-224. Puerto Rico, 1785-1803.

[7] W. Ding, T. Stepinski, Y. Mu, L. Bandeira, R. Vilalta, Y. Wu, Z. Lu, T. Cao, X. Wu. Sub-Kilometer Crater Discovery with Boosting and Transfer Learning, ACM Transactions on Intelligent Systems and Technology, 2011, accepted to appear.

[8] G. F. Cooper and E. Herskovits (1992) A Bayesian method for the induction of probabilistic networks from data. Machine Learning, 9(4), 309–347

[9] Honda, R., Iijima, Y., and Konishi, O. 2002. Mining of topographic feature from heterogeneous imagery and its application to lunar craters. In Progress in Discovery Science, Final Report of the Japanese Discovery Science Project. Springer-Verlag, London, UK, 395407.

[10] HRSC Data Browser. 2011.
http://europlanet.dlr.de/node/index.php?id=209.

[11] Jue Wang, Wei Ding, Effective classification for crater detection: A case study on Mars, Cognitive Informatics (ICCI), 2010 9th IEEE International Conference on ,7-9 July 2010 , Pages:688 – 695

[12] Kim, J., Muller, J.-P., Van Gasselt, S., Morley, J., and Neukum, G. 2005. Automated crater detection: a new tool for Mars cartography and chronology. Photogrammetric Engineering and Remote Sensing 71, 10 (October), 1205-1217.

[13] Leroy, B., Medioni, G., and Matthies, E. J. L. 2001. Crater detection for autonomous landing on asteroids. Image and Vision Computing 19, 787-792.

[14] Tanaka, K. L. 1986. The stratigraphy of Mars. Journal of Geophysical Research 91, E139-E158.

[15] Urbach, E. R. and Stepinski, T. F. 2009. Automatic detection of sub-km craters in high resolution planetary images. Planetary and Space Science 57, 880-887.

[16] Urbach, E. R., Roerdink, J. B. T. M., and Wilkinson, M. H. F. 2007. Connected shape size pattern spectra for rotation and scale-invariant classi_cation of gray-scale images. IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 272-285.

[17] Vinogradova, T., Burl, M., and Mjolsness, E. 2002. Training of a crater detection algorithm for Mars crater imagery. In IEEE Aerospace Conference Proceedings. Vol. 7. 7-3201-7-3211.

[18] Wetzler, P., Honda, R., Enke, B., Merline, W., Chapman, C., and Burl, M. 2005. Learning to detect small impact craters. In Seventh IEEE Workshops on Application of Computer Vision. Vol. 1. Breckenridge, CO, 178-184.