AUTOMATED CRATER DETECTION USING MACHINE LEARNING

A Dissertation Presented

by

Joseph Paul Cohen

Submitted to the Office of Graduate Studies, University of Massachusetts Boston, in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2016

Computer Science Program

AUTOMATED CRATER DETECTION USING MACHINE LEARNING

A Dissertation Presented

by

Joseph Paul Cohen

Approved as to style and content by:

Wei Ding, Associate Professor
Chairperson of Committee

Dan A. Simovici, Professor
Member

Marc Pomplun, Professor
Member

Zong-Guo Xia, Professor
Member

Dan A. Simovici, Program Director
Computer Science Program

Peter Fejer, Chairperson
Computer Science Department

ABSTRACT

AUTOMATED CRATER DETECTION USING MACHINE LEARNING

May 2016

Joseph Paul Cohen,
A.S., Massachusetts Bay Community College
B.S., University of Massachusetts Boston
M.S., University of Massachusetts Boston

Directed by Associate Professor Wei Ding

Craters are among the most studied geomorphic features in the Solar System because they yield important information about the past and present geological processes and provide information about the relative ages of observed geologic formations. This work focuses on improving the features extracted from satellite imagery in order to more accurately detect craters. The first focus is on improving the accuracy of methods based on semi-automatic Haar image features by only considering subsets of the features. Using feature selection methods for black-box optimization such as genetic algorithms and randomized variable elimination we are able to achieve better performance. The second focus was to learn the optimal filters and features based on training examples and replace the semi-automatic Haar features with full-automatic convolutional filters. For this a Convolutional Neural Network (CNN) called CraterCNN is designed which outperforms all existing methods and achieves up to 90% on the standard crater benchmark dataset. Then the GoogLeNet inception architecture is used to further improve the benchmark and achieve up to 93% F1-Score. In order to decrease the computational cost

of CNN models to make global Martian analysis possible a convolutional feature selection method called RandomOut is proposed. This method identifies convolutional filters which have been abandoned by the network by using the convolutional gradient norm and reinitializes them during training. RandomOut method enables CNNs to increase their accuracy to that of a network containing more filters but without the computational cost of actually adding more filters. This dissertation showcases significant progress in the field of automated crater detection and provides methods that can be applied to many other areas of automated planetary science.

Acknowledgments

This work is partially funded by a grant from the National Science Foundation Graduate Research Fellowship Program (grant number: DGE-1356104), and National Nuclear Security Agency of the U.S. Department of Energy (grant number: DE-NA0001123), and Computing and Communication Foundations Program (grant number: CCF-1062749) as well as by the U.S. National Aeronautics and Space Administration (grant number: NNX09AK86G). This work utilized the supercomputing facilities managed by the Research Computing Department at the University of Massachusetts Boston as well as the resources provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

Table of Contents

1 DESIRE AND CHALLENGES 1

1.1 Applications: Surface Dating 1

1.2 Applications: Query-by-Content and Regionalization 2

1.3 Contributions 3

1.4 Open Problems 4

2 RELATED WORK, TRENDS, AND BACKGROUND 5

2.1 Supervised/Unsupervised Learning 6

2.2 Feature Selection 7

3 FEATURE EXTRACTION 12

3.1 Semi-automatic (Haar) 12

3.2 Fully-automatic (Convolutional Neural Network Layers) 13

4 FEATURE SELECTION AND WEIGHTING ALGORITHMS 17

4.1 Genetic Feature Selection 17

4.1.1 Subset Representation as Chromosome 19

4.1.2 Wrapped Classifier Multiobjective Fitness Function 20

4.1.3 Random Crossover $\bigotimes$ 21

4.1.4 Mutation 22

4.1.5 Greedy Random Search (GRS) 22

4.1.6 Weighted Random Search 24

4.1.7 Weighted Random with Simulated Annealing Search 26
4.1.8 Shrinking Cardinality Chromosome Search (SCCS) 27
4.2 Probabilistic Feature Selection 31
4.2.1 Calculating the Optimal Number of Image Attributes to Remove 33
4.2.2 Cost Function 34
4.2.3 Dynamic Programming to Precompute the Cost Function Offline 36
4.2.4 Feedback Loop Estimation Correction 37
4.3 Neural Network Feature Selection (RandomOut) 40

5 EVALUATION 44
5.1 Convolutional Neural Network 46
5.2 RandomOut 48

6 CONCLUSION 53

REFERENCES 55

List of Tables

2.1 The datasets shown here are well established University of California, Irvine (UCI) datasets. 8

4.1 Descriptions of probabilistic feature selection notation 33

5.1 F1-score via 10-fold cross validation. 48

List of Figures

1.1 An example of two square regions of NLCD data constructed as histograms. The legend to the right shows the name of each category. 3

2.1 An overview of the last 16 years of automated crater detection research. Major concepts in the field are shown in colored boxes with their citations along with lines showing where these publications appear on the timeline. 6

2.2 Visualizations of the wrapper subset search space for multiple datasets. The top node in each panel represents the subset containing no features and the bottom node represents the subset containing all features. The colors in b, c, and d represent the accuracy and are scaled such that blue represents the minimum and red represents the maximum. 9

3.1 A visualization of a Haar feature to demonstrate the alignment of it's values. 13

3.2 An example convolutional layer 14

3.3 Crater Convolutional Neural Network (CNN) architecture computation graph. Each layer is identified with a letter and lines show processing from left to right. 15

3.4 A crater and non-crater candidate are processed by the first convolutional layer. Eight filters with interesting activation patterns are shown to the right of each candidate image in false color. Values are scaled and then colored to make these values visible and to maximize contrast within each square with blue = low and red = high. 16

4.1 Shown here is diagram based on one used by Kohavi to represent the search space for 4 features [KJ97]. Each vertex is a possible subset of features and it's size represents the result of it's evaluation function in relation to other verticies. Vertices that have been visited by the algorithm are colored blue and unvisited in gray. 19

4.2 Sample F1 values of subsets given proportional value in the pie chart based on their computed *norm* values. The edge of the pie chart is taken up based on the F1 score of the subset. 25

4.3 This brick chart details the GetWeightedSubset function in Algorithm 4. It shows the subsets depicted in Figure 4.2 but as rectangles side by side. Each block contains it's fitness score and also it's normalized fitness value. 26

4.4 Subset size distributions and their F1 score on the datasets. These are the resulting subsets from over 1200 runs of the algorithm. It is clear that the smaller the subset size the better the F1 is. 28

4.5 The SCCS algorithm restricts the search space by eliminating higher cardinality vertices from selection. This upper bound is constructed with the mean of the highest performing vertices. Each vertex is a possible subset of features and it's size represents the result of it's evaluation function in relation to other vertices. 29

4.6 Pipeline of attribute selection process 32

4.7 The testing accuracy of the network is plotted while varying nothing but the random seeds used to initialize the network. 41

5.1 Labels of each crater dataset tile. This dataset is composed of 6 tiles (1700x1700 pixels each) with resolution 12.5m/pixel shown in Figure. The imagery is nadir panchromatic footprint h0905_0000 that was captured by the HRSC aboard the Mars Express spacecraft. 45

5.2 The computation graph of the CraterCNN [CLL16] is shown split between two columns. The data input is shown in the upper left and the softmax output is shown in the lower right corner. 47

5.3 The computation graph of the Inception BN network [IS15] is shown split between two columns. The data input is shown in the upper left and the softmax output is shown in the lower right corner. . . . 49

5.4 Resulting accuracy gain of when using two hyperparameters of extscRandomOut threshold and % of epochs are varied. Each cell value in the heatmap is the mean gain of of 50 different random seeds when using RandomOut. 50

5.5 Evaluation of 50 random seeds sorted by accuracy and shown side by side when using RandomOut and when not. Left: the East region with RandomOut was run with $\tau = 1\times10^{-8}$ with $\mathcal{P} = 60\%$. Right: the West region with RandomOut was run with $\tau = 1\times10^{-12}$ with $\mathcal{P} = 90\%$. 51

5.6 Here the number of filters used in the network is varied between 1 to 10 with and without RandomOut enabled. This plot shows the mean accuracy score from 50 different random seeds of CNNs with RandomOut lead those without it by around 1 to 2 filters consistently. This means the RandomOut method enables CNNs to increase their accuracy to that of a network containing more filters but without the computational cost of actually adding more filters. 52

CHAPTER 1

DESIRE AND CHALLENGES

Craters are among the most studied geomorphic features in the Solar System because they yield important information about the past and present geological processes and provide information about the relative ages of observed geologic formations. We systematically study and present a method for automatic crater detection using advanced machine learning to deal with the large amount of satellite imagery collected.

The challenge of automatically detecting craters comes from their is complex surface because their shape erodes over time to blend into the surface. The limiting factor in existing work is the use of hand crafted filters on the image such as Gabor, Sobel filters, or Haar features. These hand crafted methods rely on domain knowledge to construct. We would like to learn the optimal filters and features based on training examples.

1.1 Applications: Surface Dating

The most popular usage of crater counts is to estimate the age of a specific area of a planetary body. Studies involving crater counts have discovered evidence for recent volcanism on Mars by studying the distribution of Martian craters in relation to craters on the Moon [HMM99]. This is possible because of the observation that old