Published in final edited form as:
Comput Geosci. 2017 April ; 101: 48–56. doi:10.1016/j.cageo.2016.12.015.

Automated detection of geological landforms on Mars using Convolutional Neural Networks

Leon F. Palafox[a], **Christopher W. Hamilton**[a], **Stephen P. Scheidt**[a], **Alexander M. Alvarez**[b]

[a]Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ, USA

[b]College of Medicine, University of Arizona, Tucson, AZ, USA

Abstract

The large volume of high-resolution images acquired by the Mars Reconnaissance Orbiter has opened a new frontier for developing automated approaches to detecting landforms on the surface of Mars. However, most landform classifiers focus on crater detection, which represents only one of many geological landforms of scientific interest. In this work, we use Convolutional Neural Networks (ConvNets) to detect both volcanic rootless cones and transverse aeolian ridges. Our system, named MarsNet, consists of five networks, each of which is trained to detect landforms of different sizes. We compare our detection algorithm with a widely used method for image recognition, Support Vector Machines (SVMs) using Histogram of Oriented Gradients (HOG) features. We show that ConvNets can detect a wide range of landforms and has better accuracy and recall in testing data than traditional classifiers based on SVMs.

Keywords

Convolutional neural networks; Support vector machines; Volcanic rootless cones; Transverse aeolian ridges; Mars

1. Introduction

During the past ten years, the Mars Reconnaissance Orbiter (MRO) has collected over 30 Terabytes of data. Two of the cameras onboard MRO that are routinely used to study geological landforms include the High Resolution Imaging Science Experiment (HiRISE; 0.3 m/pixel resolution; McEwen et al., 2007) and the Context Camera (CTX; 6 m/pixel resolution; Malin et al., 2007). However, the total data volume of these images poses new challenges for the planetary remote-sensing community. For instance, each image includes limited metadata about its content, and it is time consuming to manually analyze each image to search for non-indexed information. Therefore, there is a need for computational techniques to search the HiRISE and CTX image databases and discover new content.

Many algorithms can classify image content, such as Support Vector Machines (SVMs) and logistic regression. Yet, most of these algorithms require pre-processing steps, like

Correspondence to: Leon F. Palafox.

smoothing filters or Histogram of Oriented Gradients (HOG) methods (Dalal et al., 2005), which are typically tailored to address a specific classification problem. These preprocessing steps extract characteristics of the data, like edges in a picture, or patterns of illumination in a remote sensing scene. The signal processing and computer science communities refer to these characteristics as features. Convolutional Neural Networks (ConvNets) have become an increasingly popular alternative for image classification (LeCun, 2016), and compared with other classifiers, ConvNets have the best performance for recognition of both characters (Ciresan et al., 2012) and images (Graham, 2015). ConvNet architectures are the best performing algorithms in both the Mixed National Institute of Standards and Technology (MNIST) and Canadian Institute for Advanced Research (CIFAR) data sets, which are the standard classification data sets within the computer vision community. ConvNets learn their own input features, which alleviates the need to test different pre-processing algorithms. Furthermore, Graphical Processing Units (GPUs) can significantly increase the speed of training and classification steps in ConvNets. Using GPUs is not unique of ConvNets, and other Deep Learning architectures can also benefit from GPU acceleration.

In this paper, we address the problem of automated landform detection using ConvNets to identify Volcanic Rootless Cones (VRCs) and Transverse Aeolian Ridges (TARs) in two types of Mars satellite imagery by:

1. Training a ConvNet to detect landforms of varying size and shape, using VRCs as an example;
2. Showing that, for VRCs, a ConvNet performs better than optimized SVMs with HOG features; and
3. Showing that ConvNets also have the ability to detect a variety of other landforms, such as TARs.

Although our classifier is designed to detect many geologic features, the scope of the current study focuses on identifying VRCs and TARs as two examples of morphological distinct landforms, which are intended to highlight the broad applicability of our classifier to a wide range of geological classification problems.

2. Background information

2.1. Automated landform detection

Previous applications of machine learning in planetary sciences have typically focused on the automated detection of impact craters (Urbach and Stepinski, 2009; Bandeira et al., 2012; Stepinski et al., 2012; Emami et al., 2015; Cohen et al., 2016). Such Crater Detection Algorithms (CDAs) diminish the need for an operator to delimit manually all craters within a region, which is useful for generating impact crater inventories over large areas; however, manual inspection is still required to validate the results. The most popular CDAs first extract features from the data (e.g., shapes and patterns of light and shadow) and then apply a classifier (Stepinski et al., 2012). For instance, Urbach and Stepinski (2009) proposed a popular and efficient CDA, which applies a series of filters to remove the background noise and then creates a set of features that look for the characteristic crescent-shaped shadow of a crater. Bandeira et al. (2012) used the same approach, but added texture recognition to

improve the precision of the algorithm. Cohen et al. (2016) showed preliminary results using a ConvNet for crater detection and demonstrated that they outperformed previously tested methods in the same dataset.

Aside from detecting impact craters, machine learning methods have only been used to identify a few other landforms in a planetary science context. These efforts include using Self Organizing Feature Maps (SOFMs) to identify VRCs in Mars Global Surveyor (MGS) Mars Orbiter Camera (MOC) imagery (Hamilton and Plug, 2004), applications of SVMs to detect dunes in MOC images (Bandeira et al., 2011), and object-based approaches to estimating the orientation of TARs with HiRISE data (Vaz and Silvestro, 2014). More recently, Palafox et al. (2015) and Scheidt et al. (2015) have also demonstrated the utility of ConvNets for detecting VRCs and TARs in HiRISE images. However, in general, little work has been done to develop generalized classifiers to detect other geological landforms using planetary remote sensing data—with the exception of the hazard navigation and automated rock analysis by robotic rovers on Mars. For instance, Gor and Castano (2001) designed an automated classifier to detect and analyze rocks for both of NASAs Mars Exploration Rovers (MERs) Spirit and Opportunity (Gor et al., 2001). Biesiadecki and Maimone (2006) also designed a self-navigation system using stereo matching and Random Sample Consensus (RANSAC) algorithms, and used these algorithms to estimate the position of the rover by identifying landmarks in the image data (Biesiadecki and Maimone, 2006).

2.2. The characteristics and geological significance of VRCs and TARs

Volcanic Rootless Cones (VRCs) are generated by explosive interactions between lava and external sources of water (Thorarinsson, 1951, 1953), and are commonly associated with the flow of lava into marshes, lacustrian basins, littoral environments, glacial outwash plains, snow, and ice. Terrestrial VRCs cover areas of up to ~150 km^2 and generally include numerous cratered cones ranging from 1 to 35 m in height and ~2–500 m in diameter (Fagents and Thordarson, 2007). VRCs on Mars (Fig. 1) are generally larger, typically ranging from tens of meters to ~1 km in diameter, and can form groups covering thousands of square kilometers (Hamilton et al., 2010a, 2010b, 2011). Rootless cone morphologies and spatial organization strongly depend upon lava emplacement processes (Hamilton et al., 2010a, 2010c) and a balance between the availability and utilization of lava (fuel) and groundwater (coolant) in molten fuel–coolant interactions (MFCIs; Sheridan and Wohletz, 1981, 1983; Wohletz, 1983, 1986, 2002; Zimanowski et al., 1991; Zimanowski, 1998). However, in the presence of excess lava (e.g., in regions inundated by large sheet-like flows of molten lava), it may be assumed that the location of VRC groups will strongly depend on the distribution of near-surface H_2O and that VRCs may be used a proxy for former H_2O deposits (Frey et al., 1979; Frey and Jarosewich, 1982; Greeley and Fagents, 2001; Fagents and Thordarson, 2007; Head and Wilson, 2002; Fagents et al., 2002; Jaeger et al., 2007; Hamilton et al., 2010a, 2010c, 2011). Cratered cones, resembling terrestrial VRCs, have been identified in many regions on Mars (Fagents and Thordarson, 2007) and their widespread occurrence makes them important as a paleo-environmental indicator that can be used to infer the locations of near-surface H_2O at the time of lava flow emplacement.

Wind plays a significant role in shaping the surface of Earth and Mars by moving small particles to generate a variety of depositional and erosional features. Aeolian bedforms include ripples and dunes, as well as a distinct class of bedforms termed Transverse Aeolian Ridges (TARs) (Bourke et al., 2003). TARs occur in the equatorial and mid-latitude regions of Mars (Balme et al., 2008; Berman et al., 2011), but it is uncertain whether or not they form by ripple- or dune-forming processes. It is clear that many martian TARs are constructional landforms, resulting from the transport and deposition of granular material, alternative hypotheses have been proposed for some examples. For instance, Montgomery et al. (2012) explain several TAR-like features on Mars as periodic bedrock ridges, which are erosional landforms with crests that are transverse to the prevailing wind direction (Greeley et al., 1992; Hugenholtz et al., 2015). These contrasting interpretations carry different implications for surface–atmospheric interactions on Mars and the deposition, or erosion, of sedimentary units through time. Mapping the spatial distribution of TARs over regional and global scales could provide important new constraints for their formation processes, but their small size and widespread distribution makes automated approaches to TAR identification preferable to manual mapping efforts.

3. Methods

3.1. Support Vector Machines (SVMs)

In planetary remote sensing, SVMs have been used to detect impact craters on the Moon (Burl, 2000) and to study volcanic landforms on Venus (Burl, 2001; Decoste and Schölkopf, 2002). SVM algorithms use a function, known as a kernel, to create a decision boundary that separates data into distinguishable classes (Boser et al., 1992; Hastie et al., 2009). In remote sensing, these kernels become especially important as objects from different classes may have overlapping characteristics.

Our SVM classifier uses Histogram of Oriented Gradients (HOG) features to accentuate landforms in HiRISE and CTX images. In the HOG transformation, a series of oriented gradients—discrete angles between 0 and 360°—are drawn in small, adjacent non-overlapping units. A histogram representing the number of elements in line with these oriented gradients is created for each unit and depicted as an intensity vector in that unit. An array of HOG features representing the linear landforms of an image can provide additional information beyond the original data set. HOG is very robust to changes in illumination and shadowing, which is a desirable characteristic in a landform detection algorithm (Dalal et al., 2005).

3.2. Convolutional Neural Networks (ConvNets)

Artificial Neural Networks (ANNs) are composed of connected set of linear classifiers, each of which is trained to generate a specific decision boundary and classify simple spaces. Layers within an ANN are connected in sequential order, such that the input of a layer is the output of the previous one. Traditionally, ANNs have an input layer, which receives the input data; a set of hidden layers, which serve as the classifier; and an output layer that provides the result of the classification (Hornik et al., 1989). ConvNets differ from traditional neural networks in that different inputs share weights, rather than each input

having a single weight (LeCun et al., 1998; LeCun et al., 1990). The purpose of sharing weights is to take advantage of local consistency in the data.

The first layer is the input layer, while the last layer is the output layer. Each unit, has the following evaluation function:

$$y_i = f(\sum w_j \cdot x_j), \tag{1}$$

where y_i is the output of the i^{th} unit, w_j refers to the weight of the j^{th} input, and x_j refers to the j^{th} input. The function $f(\cdot)$ is called the activation function and it bounds the output to the range [0,1] or [−1, 1]. This bounding makes the ANN a classifier, with outputs either True (1) or False (0). In the case of multiclass classification, an ANN will use a softmax function, which allows for multiple classes in the output.

After training, the ANN can classify new data points, which may belong to two or more classes. Each of the perceptrons will be sensitive to different features of the dataset (e.g., color, edges, etc.) without the need of adding pre-processing steps. This results in general classifiers that can be used for very different classes, which makes them suitable to detect many kinds of landforms (e.g., linear and sinuous shapes, such as TARs and dust devil tracks, etc).

ConvNets extract features in the images using convolutions. The training of the ConvNet yields good values for the convolutional window for each unit in the network. These convolutions help extract the most descriptive features of an image. ConvNets work with data arranged as an image matrix.

In general, a ConvNet architecture is composed of: convolutional layers, which learn the convolutions that best represent the classes in the data; pooling layers that reduce the number of features from the convolutions to enhance computational performance (Nair and Hinton, 2010), control overfitting and allow for translation invariance; and Rectified Linear Unit (ReLU) layers, which apply the ReLU (Nair and Hinton, 2010) activation function to increase the nonlinear properties of the network. The electronic appendices provide a full description of the three layers used in a traditional ConvNet, basic training paradigms, and a description of how a ConvNets handle training data.

3.2.1. Description of MarsNet—MarsNet consists of five ConvNet architectures running in parallel, each of them tuned for different sliding window sizes. To select the most appropriate window sizes to identify VRCs, we tested more than 20 different windows-size candidates. We used a simple validation scheme, where we tested the error of the different window sizes using 30% of the data for training and 70% for testing. We found five optimal sizes, which are 8, 16, 20, 52, and 100 pixels (Fig. 1). Each pixel size corresponds to a single ConvNet architecture. We employ five ConvNet architectures in parallel—one per pixel size —to search for landforms of different size within a target HiRISE or CTX image (although we can easily adapt our system for use with other image data). The training examples are generated by manual tagging of individual cones in both CTX and HiRISE images (Fig. 1). The output of MarsNet consists of a series of heatmaps that indicate the likelihood of

positive identification for each of the landforms of interest. For instance, if we had three landforms within a scene (e.g., VRCs, TARs, and impact craters), the output of MarsNet will be three heatmaps, indicating the likelihood of each of the landforms.

Due to the difference in size of the input images that are passed to each of the parallel ConvNets, these networks each have a different number of convolutional and pooling layers (Table 1) before connecting to the fully-connected softmax output. The softmax output allows MarsNet to output multiple classes instead of only a True/False decision. In Fig. 2, we present a graphical representation of the feature complexity in the different layers in MarsNet. Each convolutional layer increases the number of calculated features, whereas each pooling layer reduces the dimension of the feature map. At the end, the output of the last convolutional layer is passed through to a fully connected ANN with a softmax function to obtain a label for the input patch.

Both HiRISE and CTX images vary in their resolution, depending largely on the initial binning of the data, and training a single ConvNet to detect features in range of image resolutions can lead to an increased number of false positives due to differences in resolution. One alternative is to downsample the HiRISE images to match the resolution of the CTX images. Therefore, we have developed two MarsNet architectures, one to process HiRISE images and another to process CTX images. Both of these networks use independent, but co-registered training and testing images.

3.2.2. Data extraction—We manually labeled examples of VRCs in HiRISE and CTX images, as well as examples of TARs in HiRISE images. To do this, we created a Graphic User Interface (GUI) where the user can tag the landforms of interest directly in a target image. We also tagged other features as a catch-all class for all the features that do not correspond to VRCs. In the case of TARs, we also tagged sand and bedrock, instead of a single catch-all class. For the TARs, we only tagged examples in HiRISE images. We then extract four images surrounding the center pixel of the tagged image (Fig. 1), instead of a single image to train the classifier. These images will serve as training data for partial features instead of only complete landforms. Training on partial landforms allows the classifier to make a positive detection even if an image contains only part of the landform. This extraction creates four $Y \times Y$ images for each training example (where Y can be either 8, 16, 20, 52 or 100 pixels). In the end, the dataset consists of 800 positive examples of rootless cones (from 200 tagged images) and 800 examples of TARs (from 200 tagged images).

Since each image has different illumination parameters, we need to normalize the training examples. Without a normalization, the training examples from one image are not transferable to different images. A normalized image (X') is calculated from a non-normalized image (X) in the following way:

$$X' = \frac{X - \mu}{\sigma}, \tag{2}$$

where the mean (μ) and standard deviation (σ) were calculated over all the possible classification windows in the original HiRISE/CTX image. For instance, if we were classifying 10 HiRISE images, we would have 10 μ parameters and 10 σ parameters.

3.2.3. Data training, metrics, and parameters—The quantitative metrics of interest to evaluate a classification algorithm are accuracy, specificity, and recall (Eqs. (3)–(5)). These metrics are sufficient to test classification algorithms over different datasets as well as their capability to detect false positives and false negatives (Bishop, 2006). Accuracy measures the ratio of mistakes the classifier made on the test dataset. Specificity measures the rate of negative landform detection. And, recall measures the ability to detect positive examples, that is, the occurrence of a landform of interest. For instance, in a dataset with 5 data points of positive examples and 100 negative examples, an algorithm set to detect every example as a negative example would have an accuracy of 95%. On the other hand, that very same algorithm would have a recall of 5%, which is a poor recall measurement. The same case can be done for specificity and few negative examples.

$$\text{Accuracy} = \frac{\text{true negative+truepositive}}{\text{true positive+false negative + true negative + false positive}} \tag{3}$$

$$\text{Specificity} = \frac{\text{true negative}}{\text{true negative+falsepositive}} \tag{4}$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive+false negative}} \tag{5}$$

We can use quantitative evaluation metrics only when the dataset is labeled, which in our example is a very small subset of the real data. For example, a single HiRISE image can have on the order of 100 million windows of 16 × 16, and we already mentioned that our positive samples are 800 for each different landform.

To determine the optimal number of units and layers in the present work, we used the 3-fold Cross-Validation (CV; Bishop, 2006) training paradigm. We used CV for each pixel size to determine the best combination of layers–units. By using CV, we also determined the regularization parameter of each network.

3.2.4. SVM optimization—In our experiments, we used an SVM algorithm enhanced with HOG features. The HOG transformation divided the image into discrete squares four pixels wide and four pixels high and collected information on the size of the oriented gradients in these squares. We concatenated these data to the original data set. Training of the SVM was then performed, utilizing the MATLAB SVM library and a radial-basis function (RBF) kernel (similar to a Gaussian), corrected by ten-fold cross-validation to find the parameters of the decision boundary. SVM optimization is handled very well in Matlab, which allows us to run more folds than in the ConvNet cases, where we only used 3 folds due to the inherent complexity and size of the ConvNet parameter space.

3.2.5. Description of the pipeline—After we finished training MarsNet, we tested the algorithm with previously unseen images. For a new image I, of dimensions $Y \times Z$, we generated a blank image of $Y \times Z$ for each of the possible classes and pixel sizes. We call these images target fields. In our architecture, we have five window sizes and three classes, which makes 15 target fields.

In the new image, we use a sliding window of $S \times S$ pixels with 90% overlap; here S is the window size for each ConvNet architecture in MarsNet. We use the window as an input for the classifier (ConvNet or SVM). In the case of the SVM, we calculated the HOG features before passing the window to the classifier. If the extracted segment evaluates as a positive detection, we add a matrix of ones in the same location in the target field of the class and pixel.

Once we finish processing each patch, we will have five heatmaps per class. The location of the landforms of interest represent hotspots in the heatmap. We then re-scale the heatmap between 0 and 1, which will indicate the likelihood of a landform being located in each individual pixel. Finally, we calculate an average of the results generated by different window sizes to generate a final heatmap. This average will help aggregate the detections of landforms at different scales.

By the end, we have a single heatmap per class, and to complete the mapping, we threshold the heatmap at 0.5 likelihood. Adjusting this threshold to higher or lower values affects the sensitivity of the mapping. Higher thresholds decrease the number detected values, but increases precision, whereas lower thresholds increase detection area, but also increase false positive detection.

3.2.6. Code and Data—The code that implements the pipeline for training and testing is available at: https://github.com/leonpalafox/CNNPlanetaryScience.

In this repository, we provide precise instructions relating to the use of the code and the trained model that we used to detect VRCs and TARs. MarsNet can also be trained to detect other landforms, such as like impact craters, recurring slope lineae, dust devil tracks, etc. The data that we used in this study is freely though the NASA Planetary Data System (PDS) and in the results section, we provide the image identification numbers for all the data used in our analyses. Specifically, for this work we have used HiRISE images PSP_002292_1875 (25 cm/pixel with 1 × 1 binning), ESP_020889_1320 (50 cm/pixel with 2 × 2 binning) and CTX image P03_002147_1865 (6.17 m/pixel).

4. Results

4.1. Comparison between the MarsNet and SVMs classifiers using HiRISE data

We compare the quality of both the MarsNet and SVMs by using the same HiRISE image as input. In both cases, we trained the classifiers, MarsNet and SVMs, using the same training dataset, and we validated the results using the same test dataset. In addition, we used CV to optimize the parameters of the classifiers. The CV segmentation was done using the same hold-out validation sets for both MarsNet and SVMs.

Fig. 3 shows maps generated with MarsNet and SVM for HiRISE image PSP_002292_1875. It also shows the target fields heatmap associated with the VRC class. Both figures show similar mappings and capture the overall cone field. SVMs, however, have a large false detection rate, missing some VRCs. MarsNet detected most of the VRCs in the scene, although it also misclassified some craters as VRCs. Qualitatively, while the SVMs seem to do a good job detecting null areas within a cone field, ConvNets has an overall better performance detecting the cone cluster as a whole.

In Table 2, we present the three metrics used for quantitative analysis: accuracy, specificity and recall. MarsNet outperforms the SVM based classifier in both accuracy in the test dataset and recall. These results indicate that MarsNet works better for classifying previously unseen data and is better at finding the true positives, whereas SVMs work better on training data and are better at detecting true negatives. However, we consider MarsNet to be superior for automated landform detection because the objective of the classifier is to apply it to new data and we consider it more important to identify true positives than true negatives.

We also calculated the total accuracy for both classifiers in the final aggregated map, and it resulted in 93.56% accuracy on the test sets for MarsNet and 91.36% for the SVMs.

4.2. Comparison of the MarsNet and SVMs classifiers using CTX data

Fig. 4, shows maps generated using both MarsNet and SVMs for CTX images. The cone filed represented is the same one we used for the HiRISE images in the previous section, we chose a larger area due to the lower resolution of CTX, which decreases accuracy and processing times.

We can see that for lower pixel sizes (8, 16) MarsNet generates a large number of false positives; however, as we increase the size (52, 100), both mappings look similar to each other, MarsNet is able to do a better delimitation of the mapping area, while the SVM approach does miss a series of cones in the border of the field. While the MarsNet architecture overestimates the field, the SVM architecture underestimates the field.

In Table 3, we see that the metrics obtained from the training data seem to indicate that MarsNet does a better job than the SVMs with labeled data. This performance does not seem to be represented in the figures due to the high imbalance in the labeled and unlabeled data. However, recall in CTX images is clearly better for MarsNet. Relative to HiRISE classifications, both systems exhibit worse performance in all categories, which is understandable given the low resolution of CTX images, and Fig. 1 shows how larger windows actually encompass large areas containing multiple cones rather than small individual cones. As a consequence, the CTX-based classifier recognizes VRC groups instead of individual cratered cones. As with HiRISE images, MarsNet outperforms the SVM in test data, although, for two pixel sizes (16, 20), SVM outperformed MarsNet.

We also calculated the total test set accuracy for both classifiers in the final aggregated map, and it resulted in 91.86% accuracy for MarsNet and 89.97% for the SVMs.

4.3. Applications of MarsNet to other geological landforms

In the following analysis of TARs, we only use HiRISE images because TARs are not well resolved in CTX images given their low spatial resolution. SVMs with the same HOG feature extractor used for VRCs are also unsuitable for TAR classification, which do not exhibit rotational invariance, and so we are unable to directly compare the performance of the SVM to the MarsNet using TARs as an example.

Fig. 5 shows that MarsNet correctly identifies areas which contain TARs, including isolated examples that are located within the bedrock-dominated region in the southern part of the scene. The metrics for TAR detection using MarsNet are consistent with its performance detecting VRCs (Table 2). Nonetheless, Fig. 5 illustrates that the classification includes some false positives that can be attributed to the fact that some structures within bedrock appear very similar to TARs. However, MarsNets performance could be increased by expanding it to include additional output classes and training it to recognize bedrock structures as a separate landform. In this way, the capability of the classifier improves and the number of classes that it is trained to detect also increases.

4.4. Time performance metrics

Using a typical Central Processing Unit (CPU), the processing time for MarsNet operating on a HiRISE image is orders of magnitude slower than the processing time for CTX images. To address this problem, we used a Graphics Processing Unit (GPU) acceleration available in Matlab, via the MatConvNet library (Vedaldi and Lenc, 2015). The library uses the cuDNN library by NVIDIA, which calculates convolutions in the GPU. These calculations decrease the processing time by at least two orders of magnitude. Our processing times are shown in Tables 4, 5. The tables show that even for full resolution HiRISE images, it is reasonable to survey a large area of Mars using MarsNet. As expected, CTX images take considerably less time. Even when they have a larger footprint, their lower resolution makes them faster to process. However, HiRISE images offer better resolution, and for some landforms, like TARs, detection is unreliable using CTX images.

5. Discussion

MarsNet has better landform identification results when we compare it with one of the best off-the-shelf classifiers (e.g., SVMs). SVMs, however, did have higher specificity in most cases, up to 10% compared with ConvNets. A higher specificity from the SVM means that the SVM is very good at detecting the negative cases; that is, the absence of a landform. However, its lower accuracy means that it is not as good as MarsNet in terms of detect the landforms themselves. This higher specificity is reflected in the fact that for most examples, the SVM approach did miss cones in the image. Furthermore, the SVM approach can only detect radially symmetric landforms like VRCs because the HOG feature extractor that we chose is optimized for landforms with rotational invariance and not for linear landforms like TARs. In contrast, MarsNet, calculates its own input features based on the training images, which makes them easily adaptable to a wide range of classification tasks.

Due to their higher spatial resolution, HiRISE images are better than CTX images for detecting small landforms. However, CTX images are 8.55 times faster to process than HiRISE images (Table 5). This means that even if CTX classifications are not as accurate as HiRISE classifications, the challenge of mapping large areas becomes more tractable if we use CTX images. Additionally, individual CTX images have a much larger footprint than HiRISE images, which is a great advantage when conducting regional surveys, and overall CTX has imaged much more of the martian surface (95.3%; Malin, 2007) than HiRISE has to date (2.5% coverage; McEwen et al., 2016). Nonetheless, while CTX images provide the best approach to mapping the distribution of landforms on regional to global scales, HiRISE remains ideally suited for more detailed mapping studies on a local scale.

In terms of classification accuracy, we note that MarsNet sometimes incorrectly identified impact craters as VRCs within both CTX and HiRISE images, which stems from the fact that we did not use an implicit impact crater class as a training dataset. However, as we increase the number of classes represented within the training data, these false positives will be reduced—thereby resulting in higher accuracy and specificity values for the classification.

Lastly, a remarkable strength of MarsNet is how the same ConvNet architecture was capable of detecting both VRCs and TARs. The configuration of the network and data processing procedure were unchanged in both cases, and this consistency shows that as long as we provide a comprehensive training dataset that contains enough examples of the landforms of interest, we can perform an automatic classification using the same architecture.

6. Conclusion

In this paper, we presented a classifier based on ConvNets called MarsNet, which is capable of outperforming SVMs augmented with HOG feature extractors. Our system is capable of distinguishing between very different landforms, such as VRCs and TARs, and we have shown that the current architecture is fast enough to process HiRISE images, as well as CTX images, using GPU acceleration. We have presented results demonstrating that the same ConvNet architecture is capable of identifying two different landforms types on Mars, without the addition of any extra pre-processing steps in the pipeline. This shows that MarsNet can be used as a generalized classifier using the same architectural system to detect multiple landforms types and at a range of resolutions. For instance, applications to HiRISE imagery are ideally suited to detailed local surveys, whereas CTX images may be used to survey much larger regions of Mars. Furthermore, our code allows any research team to use their own datasets to search for other kind of landforms, which provides a modularity that will enable MarsNet to transition from being a tool for automated landform detection to an automated mapping system for multiple landform types as its training repertoire increases with time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We acknowledge funding support from the National Aeronautics and Space Administration (NASA) Mars Data Analysis Program (MDAP) Grant Number NNX14AN77G.

References

Balme M, Berman DC, Bourke MC, Zimbelman JR. 2008; Transverse aeolian ridges (TARs) on Mars. Geomorphology. 101 :703–720.

Bandeira L, Marques JS, Saraiva J, Pina P. 2011; Automated detection of Martian dune fields. IEEE Geosci Remote Sens Lett. 8 :626–630.

Bandeira L, Ding W, Stepinski TF. 2012; Detection of sub-kilometer craters in high resolution planetary images using shape and texture features. Adv Space Res. 49 :64–74.

Berman DC, Balme MR, Rafkin SCR, Zimbelman JR. 2011; Transverse aeolian ridges (TARs) on Mars II: Distributions, orientations, and ages. Icarus. 213 :116–130.

Biesiadecki, J; Maimone, M. The mars exploration rover surface mobility flight software: driving ambition. Proceedings of the 2006 IEEE Aerospace Conference; 2006. 1–15.

Bishop, CM. Pattern Recognition and Machine Learning. 1. Springer; New York, USA: 2006.

Boser, BE; Guyon, IM; Vapnik, VN. A Training Algorithm for Optimal Margin Classifiers. Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory; 1992. 144–152.

Bourke, MC; Wilson, SA; Zimbelman, JR. The variability of TARs in troughs on Mars. Proceedings of the 34th Lunar and Planetary Science Conference; 2003. Abstract # 2090

Burl, MC. Data Mining for Scientific and Engineering Applications. Springer; US, Boston, MA: 2001. 63–84.

Burl MC. 2000 NASA volcanoes data set at UCI KDD Archive.

Ciresan, D; Meier, U; Schmidhuber, J. Multi-column Deep Neural Networks for Image Classification. Proceedings of the International Conference of Pattern Recognition; 2012. 3642–3649.

Cohen, JP; Lo, HZ; Lu, T; Ding, W. Crater Detection via Convolutional Neural Networks. Proceedings of the 47th Lunar and Planetary Science Conference; 2016. Abstract # 1143

Dalal, N; Bill, T. Histograms of Oriented Gradients for Human Detection. Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2005. 886–893.

Decoste D, Schölkopf B. 2002; Training invariant support vector machines. Mach Learn. 46 :161–190.

Emami E, Bebis G, Nefian A, Fong T. 2015; Automatic crater detection using convex grouping and convolutional neural networks. Adv Vis Comput. 7432 :213–224.

Fagents S, Lanagan P, Greeley R. 2002; Rootless cones on mars: a consequence of lava-ground ice interaction. Geological Society, London, Special Publications. 202 :295–317.

Fagents, SA, Thordarson, T. Rootless volcanic cones in Iceland and on Mars. In: Mars, TG, Chapman, G, editors. The Geology of Mars: Evidence from Earth-Based Analogs. Cambridge Univ, Press; Cambridge, U.K: 2007. 151–177.

Frey H, Jarosewich M. 1982; Subkilometer martian volcanoes - properties and possible terrestrial analogs. J Geophys Res: Solid Earth. 87 :9867.

Frey H, Lowry BL, Chase SA. 1979; Pseudocraters on Mars. J Geophys Res: Solid Earth. 84 :8075.

Gor, V; Castano, R; Manduchi, R. Autonomous rock detection for Mars terrain. Proceedings of the AIAA Space 2001 Conference and Exposition; 2001. 1–14.

Graham, B. Fractional Max-Pooling. Proceedings of the International Conference on Learning Representations; 2015. 1–10.

Greeley R, Fagents SA. 2001; Icelandic pseudocraters as analogs to some volcanic cones on Mars. J Geophys Res. 106 :20527–20546.

Greeley R, Lancaster N, Lee S, Thomas P. 1992; Martian aeolian processes, sediments, and features. Mars. 1 :730–766.

Hamilton, CW; Plug, LJ. Identification of volcanic rootless cones and impact craters using artificial neural networks. Proceedings of the AGU Spring Meeting; 2004.

Hamilton CW, Fagents SA, Thordarson T. 2010a; Explosive lava-water interactions II: self-organization processes among volcanic rootless eruption sites in the 1783–1784 Laki lava flow, Iceland. Bull Volcanol. 72 :469–485.

Hamilton CW, Fagents SA, Wilson L. 2010b Explosive lava-water interactions in Elysium Planitia, Mars: geologic and thermodynamic constraints on the formation of the Tartarus Colles cone groups. J Geophys Res: Planets. :115.

Hamilton CW, Thordarson T, Fagents SA. 2010c; Explosive lava-water interactions I: architecture and emplacement chronology of volcanic rootless cone groups in the 1783–1784 Laki lava flow, Iceland. Bull Volcanol. 72 :449–467.

Hamilton CW, Fagents SA, Thordarson T. 2011 Lava-ground ice interactions in Elysium Planitia, Mars: geomorphological and geospatial analysis of the Tartarus Colles cone groups. J Geophys Res: Planets. :116.

Hastie, T, Tibshirani, R, Friedman, J. Springer series in statistics. Vol. 1. Springer; Berlin: 2009. The elements of statistical learning.

Head, JW, Wilson, L. Mars: a review and synthesis of general environments and geological settings of magma-H_2O interactions. Geological Society; London: 2002. 27–57. Special Publications 202

Hornik K, Stinchcombe M, White H. 1989; Multilayer feedforward networks are universal approximators. Neural Netw. 2 :359–366.

Hugenholtz CH, Barchyn TE, Favaro EA. 2015; Formation of periodic bedrock ridges on Earth. Aeolian Res. 18 :135–144.

Jaeger WL, Keszthelyi LP, McEwen AS, Dundas CM, Russell PS. 2007; Athabasca valles, Mars: a lava-draped channel system. Science. 317 :1709–1711. [PubMed: 17885126]

LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L. 1990 Handwritten Digit Recognition with a Back-Propagation Network. Advances in Neural Information Processing Systems. :396–404.

LeCun Y, Bottou L, Bengio Y, Haffner P. 1998; Gradient-based learning applied to document recognition. Proc IEEE. 86 :2278–2323.

LeCun, Y. [accessed: 01.12.16] MNIST website. 2016. ⟨http://yann.lecun.com/exdb/mnist/⟩

Malin MC, Bell JF, Cantor BA, Caplinger MA, Calvin WM, Clancy RT, Edgett KS, Edwards L, Haberle RM, James PB, Lee SW, Ravine MA, Thomas PC, Wolff MJ. 2007 Context camera investigation on board the mars reconnaissance orbiter. J Geophys Res: Planets. :112.

Malin MC. 2007 MRO Context Camera Experiment Data Record Level 0 V1.0, MRO-M-CTX-2-EDR-L0-V1.0. NASA Planetary Data System.

McEwen AS, Eliason EM, Bergstrom JW, Bridges NT, Hansen CJ, Delamere WA, Grant JA, Gulick VC, Herkenhoff KE, Keszthelyi L, Kirk RL, Mellon MT, Squyres SW, Thomas N, Weitz CM. 2007 Mars reconnaissance orbiter's high resolution imaging science experiment (HiRISE). J Geophys Res: Planets. :112.

McEwen, AS; Sutton, SS; Hansen, CJ; Team, H. The First Decade of HiRISE at Mars. Proceedings of the 47th Lunar and Planetary Science Conference; 2016. Abstract # 1372

Montgomery DR, Bandfield JL, Becker SK. 2012 Periodic bedrock ridges on Mars. J Geophys Res: Planets. :117.

Nair, V; Hinton, GE. Rectified Linear Units Improve Restricted Boltzmann Machines. Proceedings of the 27th International Conference on Machine Learning; 2010. 807–814.

Palafox, LF; Alvarez, AM; Hamilton, CW. Automated Detection of Impact Craters and Volcanic Rootless Cones in Mars Satellite Imagery Using Convolutional Neural Networks and Support Vector Machines. Proceedings of the 46th Lunar and Planetary Science Conference; 2015. Abstract # 2316

Scheidt SP, Palafox LF, Hamilton CW, Zimbelman JR. 2015; Automated detection of transverse aeolian ridges on mars using convolutional neural networks and a field-based terrestrial orthoimage training set. LPI Contrib. 1843 :8047.

Sheridan MF, Wohletz KH. 1981; Hydrovolcanic explosions the systematics of water-tephra equilibration. Science. 212 :1387–1389. [PubMed: 17746257]

Sheridan MF, Wohletz KH. 1983; Implications of large-scale melt-water interactions: particle characteristics and dispersal patterns. Reports of the Geology program. 8512 :1982–1983.

Stepinski TF, Ding W, Vilalta R. 2012; Detecting impact craters in planetary images using machine learning. 1 :146–159.

Thorarinsson S. 1951; Laxárgljúfur and Laxárhraun a tephrochronological study. Geogr Ann. 33 :1–89.

Thorarinsson S. 1953; The crater groups in Iceland. Bull Volcanol. 14 :3–44.

Urbach ER, Stepinski TF. 2009; Automatic detection of sub-km craters in high resolution planetary images. Planet Space Sci. 57 :880–887.

Vaz DA, Silvestro S. 2014; Mapping and characterization of small-scale aeolian structures on Mars: an example from the MSL landing site in gale crater. Icarus. 230 :151–161.

Vedaldi, A; Lenc, K. MatConvNet – Convolutional Neural Networks for MATLAB. Proceedings of the International Conference on Multimedia; 2015. 1–15.

Wohletz KH. 1983; Mechanisms of hydrovolcanic pyroclast formation grain-size, scanning electron microscopy, and experimental studies. J Volcanol Geotherm Res. 17 :31–63.

Wohletz KH. 1986; Explosive magma-water interactions thermodynamics, explosion mechanisms, and field studies. Bull Volcanol. 48 :245–264.

Wohletz K. 2002; Water/magma interaction some theory and experiments on peperite formation. J Volcanol Geotherm Res. 114 :19–35.

Zimanowski B, Fröhlich G, Lorenz V. 1991; Quantitative experiments on phreatomagmatic explosions. J Volcanol Geotherm Res. 48 :341–358.

Zimanowski, B. From magma to tephra. Vol. 4. Elsevier; Amsterdam: 1998. Phreatomagmatic explosions; 25–53.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.cageo.2016.12.015.

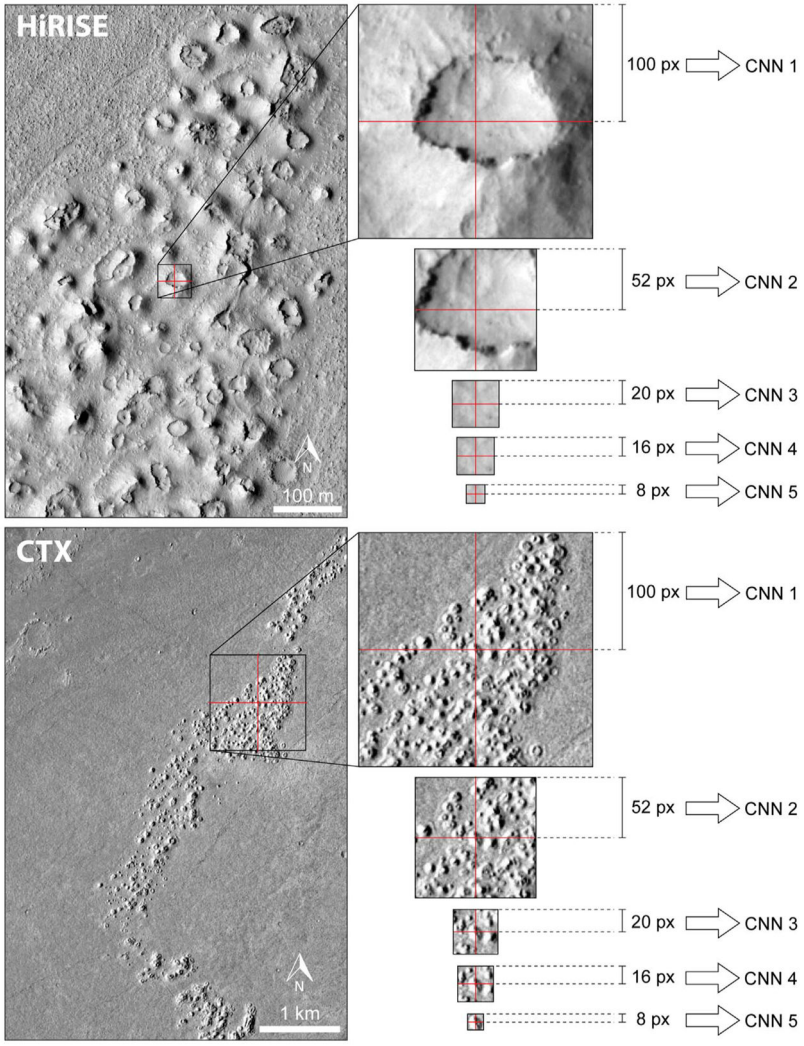

Fig. 1.

Top: Section of HiRISE image PSP_002292_1875 along with VRC training samples extracted at the different sizes for each of the CNNs in the MarsNet architecture. All the images are centered in the same feature, and five images sizes are created out of each target, each of the images feeds one of the five different ConvNets in MarsNet. Each image is segmented in four sub-images. Bottom: Target examples from a section of CTX image P03_002147_1865. Both images are illuminated from the left.

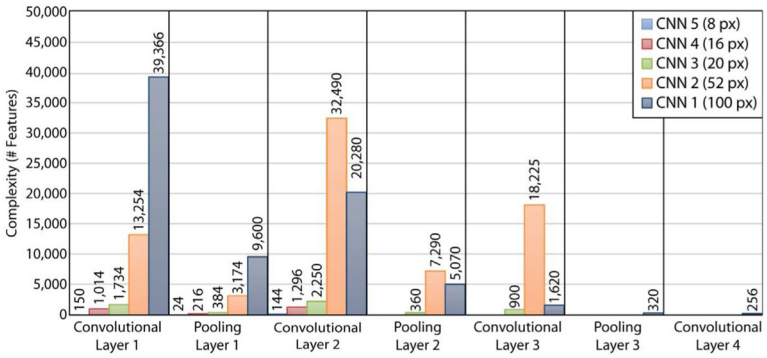

Fig. 2.

Description of the MarsNets complexity illustrating the number of features within each of the five parallel CNNs. The x-axis shows the depth within each CNN and the y-axis shows the number of features calculated in each layer. Each feature indicates a weight in the ConvNet. CNN layers operating on larger inputs (i.e., images subsets with a larger number of pixels) require more layers to obtain the best results.

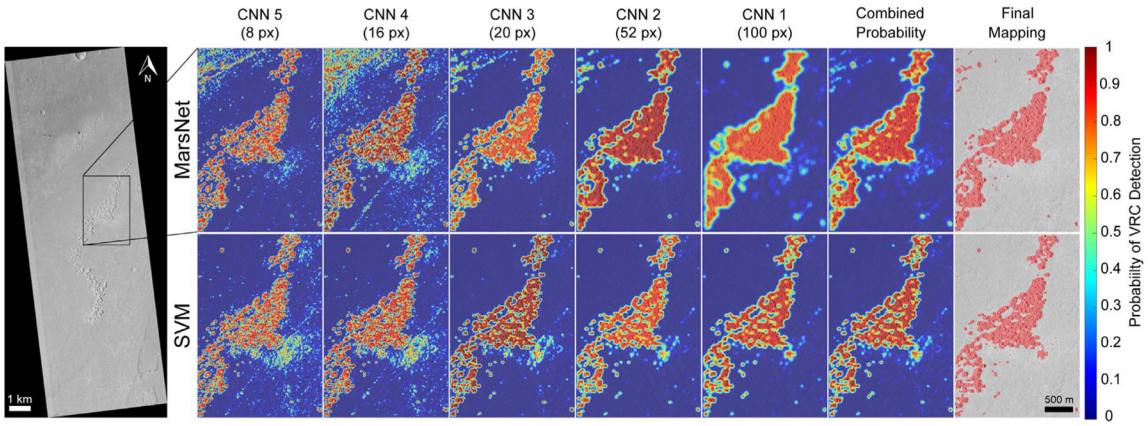

Fig. 3.

Visual comparison maps obtained for HiRISE image PSP_002292_1875. The map in the top was obtained using the MarsNet architecture, and the map in the bottom was created using the SVM architecture. All the figures without a scale bar share the same scale bar of 500 m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

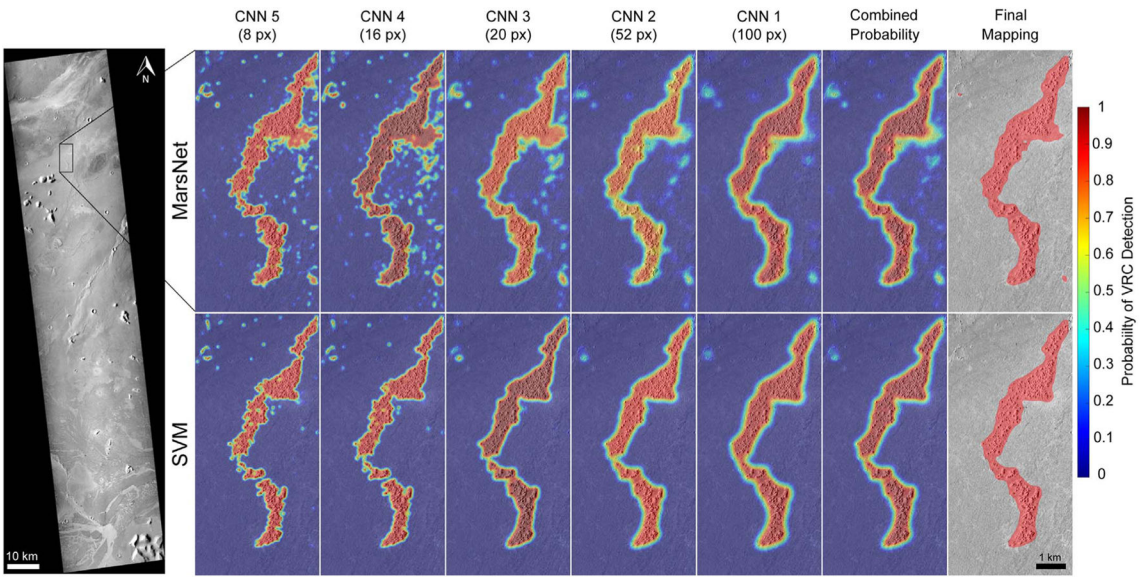

Fig. 4.

Visual comparison maps obtained for CTX image P03_002147_1865. The map in the top was obtained using the MarsNet architecture, and the map in the bottom was created using the SVM architecture. All the figures without a scale bar share the same scale bar of 1 km. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

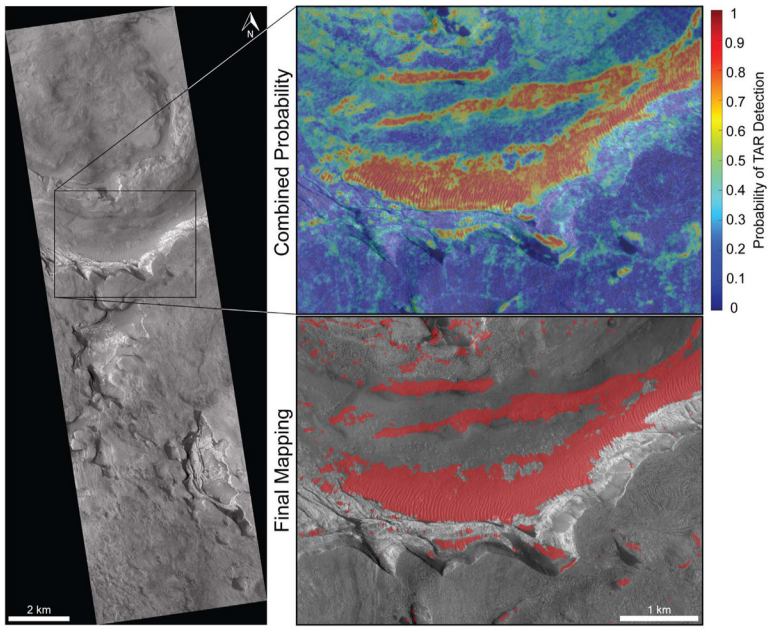

Fig. 5.

Left: HiRISE image ESP_020889_1320, includes a large number of TARs along with bedrock and sand. Upper Right: Probability of TAR detection using MarsNet after combining the results of its five classifiers. Lower Right: Final mapping of TARs (shown in red) based on a detection threshold of 0.5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

MarsNet includes layers with differing sizes to handle inputs with a range of sizes. Larger-sized inputs require more layers to process the larger amount of data contained with the scenes, as well as larger convolutional filters to maintain a small number of layers within the overall network.

	Input layer				1st hidden layer				2nd hidden layer				3rd hidden layer	
Pixel Size	**Conv.**	**Feat.**	**Pool.**	**Feat.**	**Conv.**	**Feat.**	**Pool.**	**Feat.**	**Conv.**	**Feat.**	**Pool.**	**Feat.**	**Conv.**	**Feat.**
8	4	6	2	6	2	6	0	0	0	0	0	0	0	0
16	4	6	2	6	2	6	0	0	0	0	0	0	0	0
20	4	6	2	6	4	15	2	15	2	15	0	0	0	0
52	6	6	2	6	5	15	2	15	9	15	0	0	0	0
100	20	6	2	6	15	5	2	5	5	4	2	4	4	4

Table 2

Quantitative results for the evaluation of MarsNet and SVM maps over the same area shown in Fig. 3. Final values refer to the performance of the classifier when the results of all five windows are combined.

	MarsNet						**SVM**					
Windows size (pixels)	8	16	20	52	100	Final	8	16	20	52	100	Final
Accuracy on training data	92.70	96.45	99.58	99.10	92.40	95.85	95.75	96.00	96.70	97.10	97.65	97.23
Accuracy on test data	88.54	92.36	94.44	95.83	92.01	93.56	96.25	90.50	92.30	90.70	87.50	91.36
Specificity	87.43	90.96	92.05	96.95	90.48	92.34	99.75	99.75	99.50	99.60	100.00	99.65
Recall	90.27	94.26	97.08	94.35	94.17	94.86	91.75	94.75	94.50	94.60	95.25	94.23

Table 3

Quantitative results for the evaluation of MarsNet and SVM maps over the same area shown in Fig. 4. Final values refer to the performance of the classifier when the results of all five windows are combined.

	MarsNet						**SVM**					
Windows size (pixels)	8	16	20	52	100	Final	8	16	20	52	100	Final
Accuracy on training data	95.62	97.99	100.00	87.70	85.10	93.02	92.90	96.45	99.58	85.10	83.01	91.04
Accuracy on test data	90.23	87.89	91.01	78.70	75.01	91.86	88.54	88.01	94.44	75.01	72.01	89.97
Specificity	91.60	93.71	91.90	74.01	72.01	85.02	87.43	90.96	92.05	75.01	71.01	83.02
Recall	87.50	78.35	90.72	82.01	81.01	85.98	90.27	76.01	97.08	80.01	78.01	86.08

Table 4

Estimated times for processing a HiRISE image with an average size (~1 GB).

Windows size (pixels)	8 × 8	16 × 16	20 × 20	52 × 52	100 × 100
Training time (windows/sec.)	8000	6000	5000	4000	4000
Classifying time (windows/sec.)	16,000	16,000	11,000	7000	7000
Total processing Time (min.)	196.80	196.80	160.80	161.40	112.20

Table 5

Processing times for a CTX image with an average size (~40 MB).

Windows size (pixels)	8 × 8	16 × 16	20 × 20	52 × 52	100 × 100
Training time (windows/sec.)	8000	6000	5000	4000	4000
Classifying time (windows/sec.)	16,000	16,000	11,000	7000	7000
Total processing time (min.)	23.01	23.09	19.30	19.30	4.20