

Semi-automatic mapping of linear-trending bedforms using ‘Self-Organizing Maps’ algorithm



M. Foroutan^{a,*}, J.R. Zimbelman^b

^a Department of Geography, University of Calgary, Alberta T2N 1N4, Canada

^b CEPS/NASM MRC 315, Smithsonian Institution, Washington, D.C. 20013-7012, United States

ARTICLE INFO

Keywords:

Semi-automatic mapping
Linear-trending bedforms
Self-organizing maps
Image processing
Transverse Aeolian Ridges

ABSTRACT

Increased application of high resolution spatial data such as high resolution satellite or Unmanned Aerial Vehicle (UAV) images from Earth, as well as High Resolution Imaging Science Experiment (HiRISE) images from Mars, makes it necessary to increase automation techniques capable of extracting detailed geomorphologic elements from such large data sets. Model validation by repeated images in environmental management studies such as climate-related changes as well as increasing access to high-resolution satellite images underline the demand for detailed automatic image-processing techniques in remote sensing. This study presents a methodology based on an unsupervised Artificial Neural Network (ANN) algorithm, known as Self Organizing Maps (SOM), to achieve the semi-automatic extraction of linear features with small footprints on satellite images. SOM is based on competitive learning and is efficient for handling huge data sets. We applied the SOM algorithm to high resolution satellite images of Earth and Mars (Quickbird, Worldview and HiRISE) in order to facilitate and speed up image analysis along with the improvement of the accuracy of results. About 98% overall accuracy and 0.001 quantization error in the recognition of small linear-trending bedforms demonstrate a promising framework.

1. Introduction

Mapping different types of minor linear-like features, such as aeolian ripples and glacial linear landforms, and extracting their morphometric and pattern information on Earth and other planets, is challenging because of their small aerial extent on satellite images and their large number or high density within small areas. In order to spatially analyze small linear features such as ripple-like aeolian bedforms, other studies traced and digitized them manually on Earth and other planets (e.g., Ewing et al., 2010, 2015; Fenton et al., 2015; Pedersen et al., 2015). These manual techniques are time consuming and subject to different types of human error, particularly in closely spaced and aerially extensive bedform fields. Unfortunately, available feature extraction modules require many inputs from human operator or additional pre- and post-processing steps such as Histogram of Oriented Gradients (HOG) and cleaning. For instance, the ENVI feature extraction module (Hölbling and Neubert, 2008), which is one of the most popular object-based modules, is not efficient for very small footprints in huge images. Also, results cannot be acquired automatically and depend heavily on good segmentation, followed by effective rule building (Hölbling and Neubert, 2008).

Other introduced unsupervised feature extraction techniques still

depend on supervised steps, demand human-based pre-processing and post-processing such as manual threshold determination, or are not completely efficient at training time (e.g. Katartzis et al., 2001; Shi et al., 2014; Romero et al., 2016). In addition, current linear feature extraction techniques are merely efficient for man-made structures with homogenous surfaces, such as roads, and cannot identify small features with natural curves and topography (e.g. Miao et al., 2013) or they are suitable for a specific image resolution (e.g. Shi et al., 2014). A supervised digitizing technique, object-based ripple analysis technique (OBRA), has been introduced for aeolian features on Mars by Vaz and Silvestro (2014). Their technique has been tested successfully on ripple patterns on dunes with the same texture and material as their background; however, the methodology sounds difficult to use. A recent paper by Palafox et al. (2017) applied Convolutional Neural Network for automatic detection of Transverse Aeolian Ridges (TARs), enigmatic aeolian bedforms on Mars (Zimbelman, 2010; Berman et al., 2011), but the methodology just identifies areas overlaid by TARs and maps their spatial distribution, and does not map individual crestlines, which is important for wind direction identification in aeolian geomorphology. Generally, during our literature review of automatic and semi-automatic remote sensing techniques for feature extraction, we found many of them rely on human inputs, are vague in defining their evaluation

* Corresponding author.

E-mail address: foroutam@ucalgary.ca (M. Foroutan).

technique and are not entirely reproducible.

The present study introduces a framework based on an unsupervised Artificial Neural Network (ANN) algorithm called Self-Organizing Maps (SOM), to automate outlining bedforms. SOM is generally used for visualization of information that preserves the topological relationships captured in the input (Kohonen, 2001). It translates information relationships of high dimensional input data to a two-dimensional output grid in what is called the map. It was initially applied in engineering tasks, but it has been introduced to the geoscience community as a pattern recognition method used in geophysics (Raiche, 1991). Gradually, several applications of this seminal method attracted scientists in other geoscience disciplines, such as the classification of remote sensing data (Jianwen and Bagan, 2005), class modeling (Marini et al., 2005), and information visualization and knowledge discovery (Koua et al., 2006), with better results acquired compared to traditional approaches based on accuracy and error comparison. Application of this method in geomorphometrical segmentation of bare earth and desert landforms (Ehsani and Quiel, 2008; Foroutan et al., 2013) demonstrated that SOM is an efficient tool for analyzing aeolian landforms in hyper-arid environments, which provides useful information for terrain feature analysis in remote regions. However, these desert studies were based on different order geomorphometric elements derived from digital elevation models (DEMs), for features on the scale of hundreds of meters.

The main objective of our research is to identify a semi-automated unsupervised and reproducible framework for mapping small linear features. This is crucial for further spatial analysis of such features, particularly in dense groups and large areas, and is still a problem in different disciplines. As the SOM algorithm has been shown to be capable of handling large and detailed input data, we tried to test its capability in linear feature identification and extraction. The current study introduces a firm framework in recognizing small linear features, which can be used on different types of satellite images and resolutions, from Earth and other planetary surfaces. Compared to the above mentioned previous techniques, our methodology is unsupervised, with a higher level of automation to reduce human inputs, and is simple to reproduce. The proposed methodology has been tested on both Earth and Mars satellite images for a rigorous evaluation.

2. Data and methods

2.1. Study area

The study area for this research is a unique mega-ripple field in the Lut desert of Iran, one of the hottest and driest deserts on Earth (Foroutan and Zimbelman, 2016). The region hosts millions of granule-covered mega-ripples with various spatial patterns (e.g. saturated or raked patterns) and crest morphologies (e.g. barchan-like, simple) (Fig. S1). These landforms are located in a lowland area about 200 m above sea level in an area of about 250 km² at the fringe of a huge megar-yardang field (Fig. 1). This pristine depression contains a veneer of aeolian sediment over most of the surface, sitting on quaternary alluvial fan deposits or other alluvium surrounded by Dasht deposits (sand, silt, clay and conglomerate) and fine-grained lake deposits (Lut formation) (Foroutan and Zimbelman, 2016). These mega-ripples are typically larger than normal mega-ripples on Earth and of great interest to planetary scientists because they are very similar in horizontal length scale to mysterious aeolian bedforms on Mars known as TARs.

TARs are immobile, small-scale, distinct morphological features with narrow transverse dimensions found mostly throughout the equatorial and mid-latitude regions of Mars. TARs can occur singularly or in groups that number in the thousands; they are abundant but not ubiquitous on Mars (Berman et al., 2011) (Fig. 2). The forms and patterns of the Lut features are essentially identical to many TARs seen in high resolution images of Mars (e.g., Zimbelman, 2010; Berman et al., 2011), which makes them to be one of the best terrestrial analogs for

TARs, here referred to as “TAR-like features” rather than mega-ripples (Foroutan and Zimbelman, 2016). Dimensions of these bedforms in the Lut desert range in width and length from 1 to 30 m and 7 to > 300 m, respectively. Their height varies from < 0.5 m to about 3 m. Although several hypotheses exist regarding TARs formation, they still cannot properly explain the unusual dimensions and characteristics of the bedforms. Considering the very small to stable nature of the Lut desert landforms, which makes them more viable as analogs, we believe that the methodology described in this manuscript, along with future field studies in the Lut desert, may reveal more information about these curious features on both planets. For Mars study areas we focused on images that contain TARs regardless of any other selecting factor or particular region on the planet. Our selected images come from Kaiser Crater, Noctis Labyrinthus and Phlegra Dorsa.

2.2. Satellite data

Small granular mega-ripples cannot be easily identified in images with > 2 m pixel⁻¹ resolution, so these bedforms are best analyzed using satellite images with relatively high resolution. Natural colour panchromatic QuickBird images with 0.63 m pixel⁻¹ resolution, which have been orthorectified, have been used as the main input data in this study, which is the best available for the area. QuickBird satellite collects both multi-spectral and panchromatic images concurrently. The QuickBird panchromatic resolution ranges from 0.61 m at nadir to 0.72 m at 25° off-nadir for a bandwidth which spans from 450 to 900 nm (DigitalGlobe, 2008a). Capability for showing spatial variability at the fine scales represented by these images opened new horizons for field-scale remote sensing applications (Coops et al., 2006; Wu et al., 2008). The boundary of the acquired data for this study is illustrated in Fig. 1, which shows the study area location comparing to a yardang site at the west and sand dunes at the far south. The selected area (Fig. 1) does not have any recognizable topographic variations and contains mega-ripples with diverse patterns, sizes and spatial densities (single, patch, field and sea), which is a challenge to evaluate the usability of the method in all these settings. There are also scattered small yardangs and ventifacts between some of the marginal landform patches.

2.3. SOM (Self Organizing Maps)

Artificial neural networks are structured in layers of neurons as processing units, which are inter-connected. Each output layer neuron is connected to all neurons in the input layer by synaptic weights or weight vectors. By adjusting the weights of an artificial neuron, desired output from the network can be obtained. This adjusting step, which is defined by a certain ANN algorithm, is called a learning process. SOM is an unsupervised and nonparametric artificial neural network algorithm invented and introduced by Kohonen (2001). The SOM characteristics, such as learning ability, abstraction with topology preservation, and visualization, can be applied to complex tasks in different disciplines (e.g. Marini et al., 2005). It has been applied in this study as a powerful classification algorithm with the capability of delineating small linear features in satellite images.

The input layer in the SOM represents the input feature vector and thus has neurons for each measurement dimension. For remote sensing data, this would imply a separate neuron for each reflectance band or image. The output layer is a two-dimensional array of neurons, where each output layer neuron is connected to all neurons in the input layer by weight vectors (Fig. 3). Weights are initialized randomly and the learning procedure modifies the initial weight of connections, or weight vectors, to best describe the spatial pattern of the features presented to them (Kohonen, 2001). The primary aim of the model is to let the weight vectors learn what is presented by the input vectors.

Network learning in SOM is an iterative process. An input vector is represented by the vector $x = \{x_1, x_2, \dots, x_n\}$, where n is the number of



Fig. 1. A part of Lut desert in Iran, which contains large mega-ripples. The white outline shows the boundary of QuickBird satellite images used as the input data for this study. The small black outlie shows Fig. 6's location.

input variables used for the network. During the classification, input vectors are presented to the network and, in each case, the output neuron with the minimum Euclidean distance between its weight vector and the input vector is selected as the winner:

$$winner = \arg \min_j \left(\sqrt{\sum_{i=1}^n (x_i(t) - w_{ij}(t))^2} \right) \tag{1}$$

where $x_i(t)$ is the input to neuron i at iteration t , and $w_{ij}(t)$ is the synaptic weight connecting output neuron j to the input neuron i at iteration t . The weight vector of the winner and output neurons within a neighborhood radius γ of the winner are then adjusted in the direction of the input vector:

$$w_{ij}(t + 1) = w_{ij}(t) + \alpha(t)[x_i(t) - w_{ij}(t)] \tag{2}$$

where $w_{ij}(t + 1)$ is the adjusted weight and $\alpha(t)$ is the learning rate at iteration t . The weights of those neurons outside the neighborhood remain unadjusted. This competitive learning and lateral interaction stage is known as coarse-tuning. The learning rate decreases gradually during the coarse-tuning stage from an initial learning rate (α_{max}) to a final learning rate (α_{min}), after a total number of iterations (t_{max}):

$$\alpha(t) = \left(\frac{\alpha_{min}}{\alpha_{max}} \right)^{\frac{t}{t_{max}}} \tag{3}$$

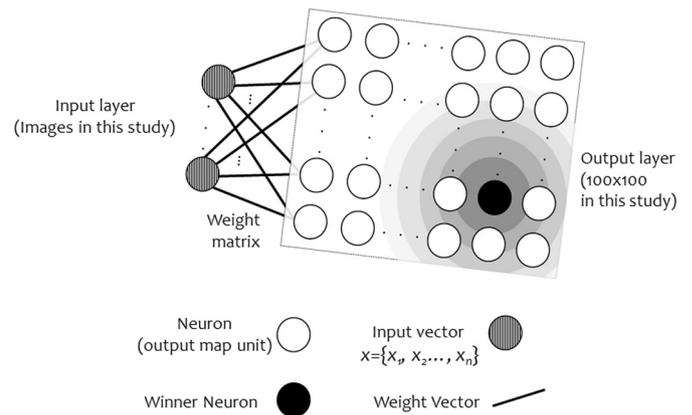


Fig. 3. Simplified structure of the Self Organizing Maps. Lines show the connection between input and output neurons (weight vectors). Each input neuron is connected to all output neurons by weight vectors. Winner neuron update neighboring neurons based on their distance.

Consequently, the radius of the neighborhood decreases steadily during the coarse-tuning stage. A large initial neighborhood radius is usually chosen and it decreases until the weight of only the winning neuron is adjusted in the output layer. Coarse-tuning finally results in a

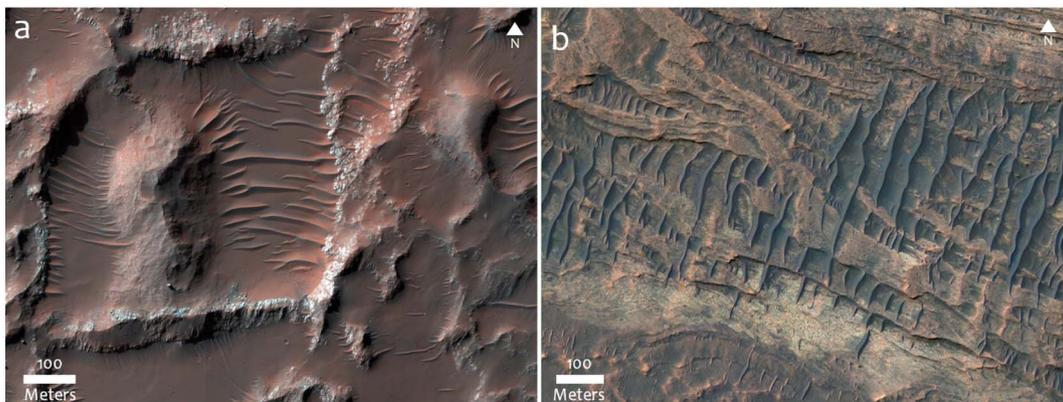


Fig. 2. Examples of TARs on Mars with different ranges of albedo: a) West Holden Crater; TARs in this area have the same albedo and probably the same material as their surrounding area (HiRISE PSP_010197_1540) and b) Miyamoto Crater; here TARs are darker than their surroundings (HiRISE ESP_026378_1730).

fundamental organization (topology) of neuron weights that represent the underlying clusters in the input data. The boundaries between neurons and their distances to their neighbors can be visualized at what is referred to as the unified distance matrix or U-Matrix (Vesanto and Alhoniemi, 2000) (Fig. S2). The final stage in SOM classification is labeling, which determines output neuron classes. An unsupervised weight-based labeling procedure has been used, which labels emergent neuron clusters using statistically significant weights (van Heerden and Engelbrecht, 2013). At this stage the basic structure of the input data has been topologically formed in the SOM output layer topology, which finally defines the class differences and distance (Kohonen, 2001).

SOM classification gives a detailed separation between elements of mega-ripples, which means it has the ability to recognize the crestline as a separate class, which is our target class. The quantization error of the SOM algorithm is used as a performance criterion for classification results. The quantization error can be measured by defining the average distance of the input vectors to the cluster centroids by which they are represented.

3. Results

3.1. SOM results

An important issue for documenting crestlines on small bedforms, such as the mega-ripples, with ANN algorithms is finding inputs and adjusting the network to define these features in detail. The choices of suitable inputs and an appropriate kernel size for extracting each input layer, the best dimensions for the output layer, iteration and cluster number, and the neighborhood radius for the SOM classification need to be adjusted with precision. Different network adjustments have been tested to recognize the best resulting design for a framework that captures the crestlines of mega-ripples with the highest accuracy (Fig. S3). The final framework for capturing our target landforms on Quickbird images has the input layers including all bands of the satellite image and two additional layers that were extracted from the panchromatic band; one of these additional layers is the edge detection map, obtained by using a Sobel edge detection filter, and the second additional layer is the result of a high-pass spatial filter (Fig. 4).

The Sobel operator calculates the gradient of the brightness intensity of each pixel in the image, giving areas of abrupt change, horizontal and vertical, relative to those of gradual change (Sobel, 1970). High-pass filtering removes the effect of low frequency components within an image, like ‘geological noise’ such as overlying dust devil tracks or small salt deposits, while retaining the high frequency or dominant local variations (Chavez et al., 1991). These two layers can emphasize study features’ crestlines for better delineation in the next steps. Selection between different filters and extracting kernel sizes were identified by comparing the results between several runs. The best results were acquired by using 3×3 kernel size for Sobel filter and 11×11 kernel size for high-pass filter. The SOM was designed to sample all pixels in the input layers.

The output layer for the SOM was designed to have 100×100 neurons, an initial neighborhood radius of 142.42, the minimum

learning rate, $\alpha(t)$, of 0.5 and the maximum learning rate of 1, have been used for this algorithm. These values were determined as the most efficient and precise settings according to an evaluation of the smallest quantization error through several iterative runs of the network and accuracy assessment (Fig. S3). Post-classification steps are identifying the crestline class and vectorization and conducting a smoothing algorithm to the desired class. For generalizing our methodology, we consider the identification step as a quick recognition by the user, but for all our data in this study we used average aspect ratio (length to width) of classes as an automatic class recognition tool. Based on the linear shape of crestlines, our desired class has the highest aspect ratio compare to other classes. QuickBird image of the study area (white border in Fig. 1) consists of four tiles, which were processed separately with the same structure and training setting of SOM. Using > 1 million iterations for described SOM setting, the final quantization error was ~ 0.001 , which resulted in the anticipated classification. The final extracted shapefile contains almost one million crestlines in the area, which was ready for further spatial analysis. A density map of the extracted crestlines of the whole area is illustrated in Fig. 5. The flowchart in Fig. S4 summarizes the major steps used in this semi-automatic crestline extraction method.

The methodology described above was also applied to a WorldView-1 black and white image of the same region, which has the highest spatial resolution among commercial satellite images with 0.5 m pixel^{-1} resolution at nadir (DigitalGlobe, 2008b). We also used High Resolution Imaging Science Experiment (HiRISE) satellite images of the Martian surface (McEwen et al., 2007), which have a spatial resolution of about 0.50 and 0.25 m per picture element. Both data are single band, but the same set of SOM structure and setting from Quickbird images has been used for them. The results were quite satisfactory with an overall accuracy of about 98%. Considering their higher resolution compared to a Quickbird image, they seem to have almost the same accuracy. We also tested single band for Quickbird, which showed slightly lower accuracy, about 0.3%, in results as compared to using three bands. This implies that application of one band can be sufficient for satisfactory results, while using all bands is recommended for Quickbird images. In the following section the quantitative accuracy of the results and the method challenges are discussed.

3.2. Accuracy assessment and validation

3.2.1. Overall evaluation of results

For comparing our results with the traditional manual technique, about 600 features were randomly selected throughout the study area for manual digitizing as our control points/features. This sample size, which is about 0.05% of the total number of features in the Lut desert, is more than what is typically required for accuracy assessment in the literature (Lillesand et al., 2014). Mega-ripples in this area are not identical all over the study area, so accordingly sample features are also diverse in morphology, such as linear, arcuate, barchan-like, and they also form different patterns, such as closely spaced, widely spaced or singular, topographically confined and influenced features (Figs. S1 and 11). These random features have been digitized manually by photo-



Fig. 4. Input layers: a) The image of bedforms and two resulted layers from: b) high-pass filter with 11×11 kernel size, and c) Sobel filter with 3×3 kernel size. They have been recognized to have the best outcomes for linear features' recognition.

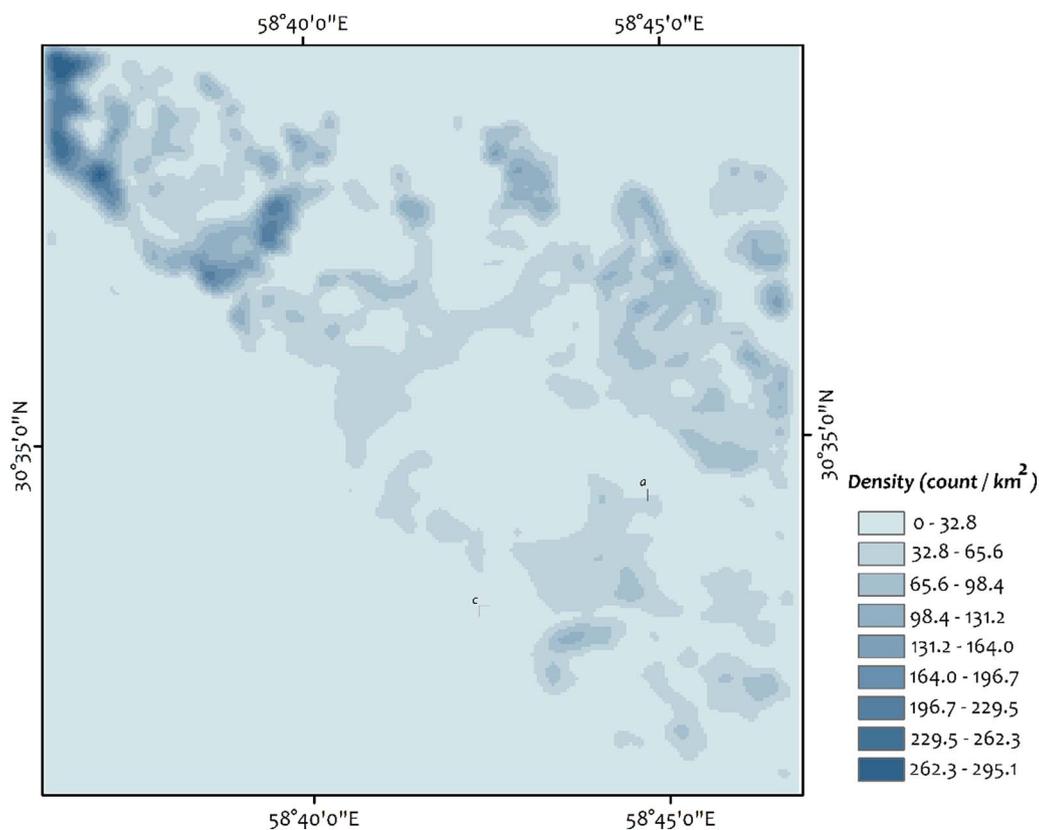


Fig. 5. Density map of > 1 million crestlines of mega-ripples that have been digitized by the framework presented here. Black boxes in the image show locations of areas in Fig. 11.

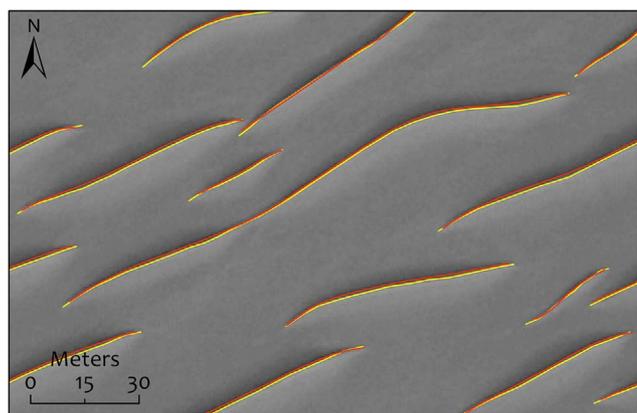


Fig. 6. A portion of the study area, overlaid by two different crestline layers resulted from manual (yellow) and automatic (red) procedures. The location of this area is shown in Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interpretation and independently of the results of the automatic method (Fig. 6). The same procedure has been conducted on Martian images and about 600 control points or TARs have been selected for manual digitizing from three different HiRISE images from Kaiser Crater, Noctis Labyrinthus and Phlegra Dorsa. This is about 0.1% of TARs in these areas. Finally, the two results were quantitatively and qualitatively compared in order to validate the described methodology.

Length and orientation are the first-hand metrics obtained from the crestline map of mega-ripples. Plots in Fig. 7 show the comparison between cumulative probability distribution of length resulted from manual and SOM semi-automatic methodologies on Earth and Mars. The length-weighted trend circular plot of Lut bedforms in Fig. 8 shows the orientation of the crestlines, weighted by length, for both manual and SOM results. Based on data statistics, on average a difference of 0.42° and 2.4° exists between the mean azimuth computed for datasets

extracted from automatic and manual methods for terrestrial and Martian study bedforms, respectively. In addition, an average length difference of 1.14 m was obtained between the traditional or automatic mapping on Earth and about 0.64 m average length difference for both methods on Martian TARs. These values demonstrate the high accuracy for the presented SOM methodology on Earth and Mars satellite images. Table 1 shows the statistical values of these two datasets for orientation, sinuosity and length. It has been determined in some studies that such small differences between results, or even higher differences, might happen in traditionally produced maps conducted by different users (e.g. Vaz and Silvestro, 2014). A visual qualitative comparison between features shows that highest differences occur for the study bedforms with the smallest footprints on the image.

3.2.2. Overall accuracy assessment

For calculating the overall accuracy of the resultant crestline maps, > 1000 points have been selected as accuracy points throughout both Earth and Mars study areas. Results demonstrate an overall accuracy of about 98% for all our data. While the highest accuracy for Quickbird image was 98.3% using three bands, and the lowest was the Quickbird using single band, with 97.3% as the overall accuracy. HiRISE and WorldView-1 images overall accuracy was 98.6% and 98.0%, respectively. While there is a slight difference between the results, all results are still quite satisfactory and demonstrate the precision of the present methodology. Comparing the false positive and false negative of the results shows, for our three types of data, there is a positive correlation between the resolution of the image and false positive of the resulting map; the highest false positive was for HiRISE images, with about 2%. This result for HiRISE images can rise from the existence of some morphologies on Mars that might be mistaken for TARs, such as periodic bedrock ridges (Montgomery et al., 2012) and small linear yardangs, as well as other similar patterns such as polar lineated terrains in other HiRISE images. These features are hard to interpret and recognize from TARs even for manual users. Some salt accumulations between mega-ripples on Earth that are shaped along

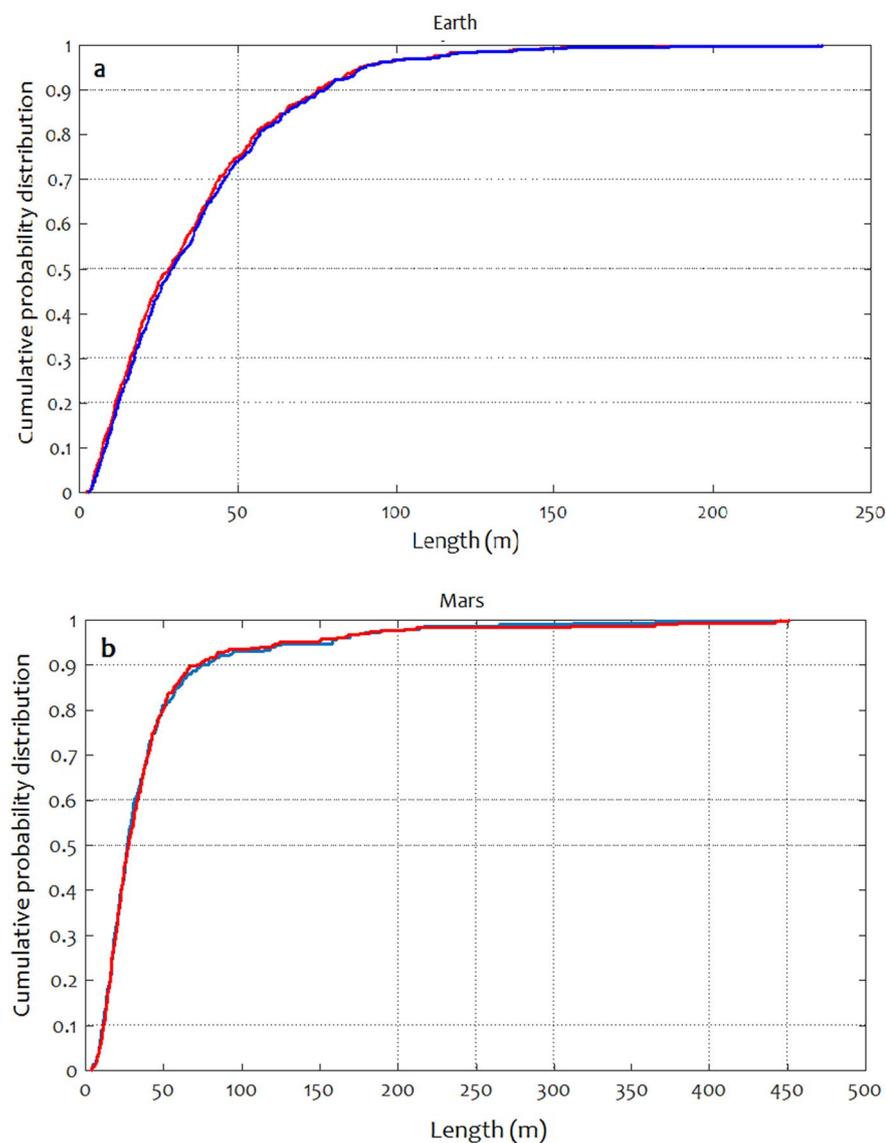


Fig. 7. Cumulative probability distribution graphs of length the crestline on Earth (a) and Mars (b) results from two different methods of digitizing; the red line and the blue line show the automatic and the manual results (reference data), respectively. The comparison between these two methodologies in this figure indicates slight difference between manual and automatic results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bedforms also included in false positives on Earth. The highest false negative percentage was for QuickBird images with more than about 1%; it seems that for our data there is a negative correlation between resolution and false negative percentage. Our survey for false negative cases showed most TARs and mega-ripples that are not captured by SOM were among very small curve-shaped features and the ones that are buried by sediments. Features with less than six pixels in length and three pixels in width, as well as having < 0.5 m altitude, could not be captured by this method. These small features are not $> 0.1\%$ of this mega-ripple field in the Lut desert. Fig. 9 indicates some of the false positive and false negative cases on satellite images of the Lut desert as examples.

3.2.3. Position and shape accuracy and evaluation

The small and linear natures of our resulting maps cannot be properly evaluated by an overall accuracy. So apart from final overall accuracy, three other approaches of line-specific accuracy assessment measurements have been tried for verification of the introduced methodology. We think each one of these accuracy measures does not necessarily imply the full degree of accuracy since each can have some pitfalls for comparing the matching degree. However, when they all come together, each can add a valuable insight to the matching evaluation between resultant and reference data. For evaluating the positional accuracy, we used Longitudinal Root Mean Squared Error

(LRMSE), which is the quantitative assessment of the horizontal variance between two lines computed between a number of paired sets of points located along both automatic and manual extracted crestlines (Anderson et al., 2014). For assigning the number of points on each line, intervals along the lines have been shortened until we reached to a constant RMSE value, then root mean square errors or differences of these points along each pair of lines have been extracted for all lines. In addition, for better defining the position accuracy we followed a buffer approach, which provides a simple and comprehensive measurement (Goodchild and Hunter, 1997), whereby the manual or reference linear features have been buffered consecutively and the created buffer zones (with different widths) are intersected with the tested linear features or our automatic line results. The percent of total length of the tested feature inside the buffered reference feature (manual line) indicates the agreement between the two lines. Finally, for shape accuracy, relative sinuosity has been calculated by dividing the sinuosity of the resultant lines over our reference lines (Anderson et al., 2014) and the mean and the standard deviation of the results have been calculated for the overall view. This indicates to what extent the curvature and the complexity of the resultant line and reference line are similar. In addition, length differences and azimuth differences have been considered as two other factors for comparing results. Table 2 shows the results; in all these analyses, we used over 600 sample features from our initial manual digitized map (by the first author) as a reference map and a

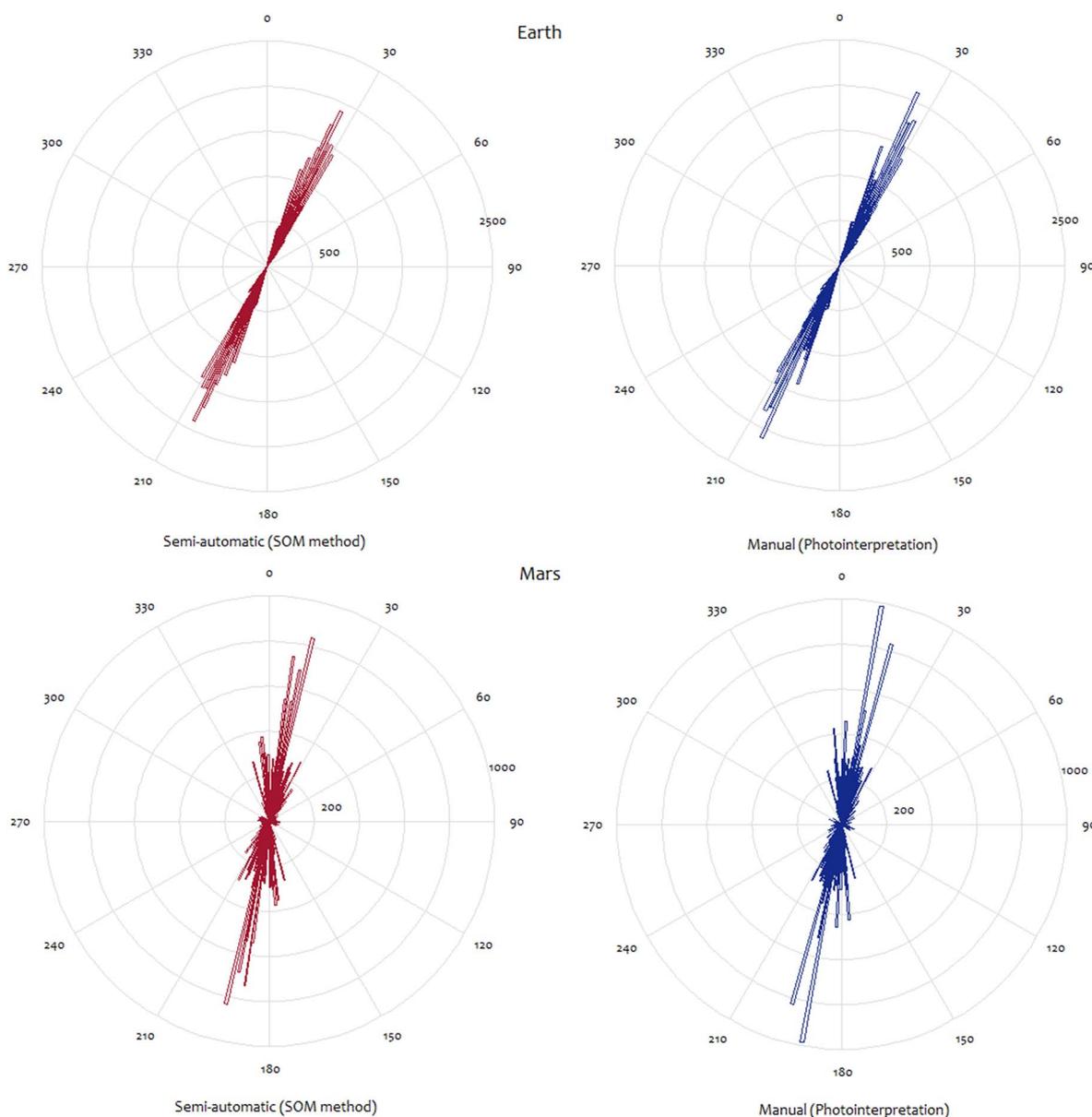


Fig. 8. Features' length-weighted orientation circular plot. Comparing two different methods on Earth satellite images from the Lut desert, red is the automatic result and blue is the manual result. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
The summary statistics comparing automatic versus manual mapping methods.

	Manual digitizing (reference)		Automatic digitizing	
	QuickBird	HiRISE	QuickBird	HiRISE
Input data				
Number of bedforms	610	580	610	580
Mean length (m)	36.94	41.52	35.80	42.16
Maximum length (m)	234.62	445.21	233.15	451.02
Minimum length (m)	2.77	4.20	2.23	3.87
Standard deviation of length (m)	29.79	51.07	29.63	56.08
Mean azimuth (°)	25.24	55.99	24.82	58.39
Maximum azimuth (°)	39.14	179.25	41.57	179.51
Minimum azimuth (°)	6.78	0	6.71	0
Standard deviation of azimuth (°)	5.86	60.51	5.89	61.90
Mean sinuosity	0.99	0.97	0.98	0.96
Maximum sinuosity	1.00	1.00	1.00	1.00
Minimum sinuosity	0.92	0.86	0.62	0.82
Standard deviation of sinuosity	0.008	0.02	0.018	0.03

second manual map from an operator, who was unfamiliar with the purpose of the measurement, as a testifier.

Considering the relatively high length differences in Quickbird images (~seven pixels) as compared to HiRISE images may come from the lower spatial resolution of the former, which can affect length accuracy. Other accuracy criteria in the table have a sound agreement with the traditional method, which validates the precise mapping ability of the presented methodology.

3.2.4. Orientation and illumination condition

Bedforms in the Lut desert seem to be all unidirectional and this may raise doubts about the method to be directionally biased. Interestingly, one of the preliminary results of the overall crestline map is that they are not all unidirectional and they have slightly different trends that seems to be affected by concentration effects of yardangs on local winds, or different wind regimes within the study area. Fig. 10 shows the classified crestline orientation map of the whole area, which demonstrates the features in this area are not completely unidirectional and the methodology captured even small divergence from the whole

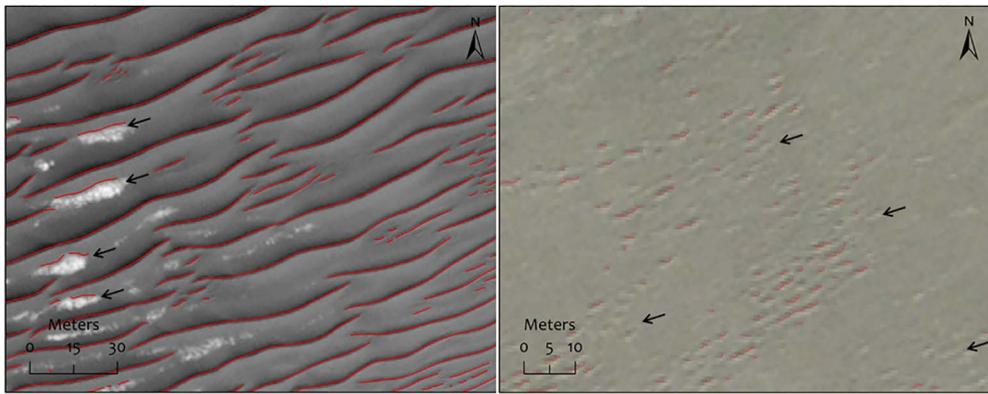


Fig. 9. Example errors that have been detected in presented methodology on satellite images of the Lut desert a) false positive. Arrows indicate large salt accumulations between large mega-ripples that can be digitized as bedforms when they are accumulated along with the bedforms. b) False negative for very small bedforms (indicated by arrows) with small dimensions that can be missed by this method.

Table 2
Position and shape accuracy of resultant crestlines comparing to the initial manual digitizing map as a reference map and a second manual digitized map as a testifier.

	SOM method (Quickbird-Earth)	Testifier (Earth) 2nd operator	SOM method (HiRISE-Mars)	Testifier (Mars) 2nd operator
1 LRMSE	0.0851	0.0802	0.065	0.053
2 Relative sinuosity	0.9989	0.9987	0.989	0.999
3 Buffer accuracy width (m)	1.2	0.45	0.25	0.23
4 Maximum azimuth differences (°)	3.2°	2.4°	3.0°	0.7°
Maximum length differences (m)	4.6	1.5	0.56	0.5

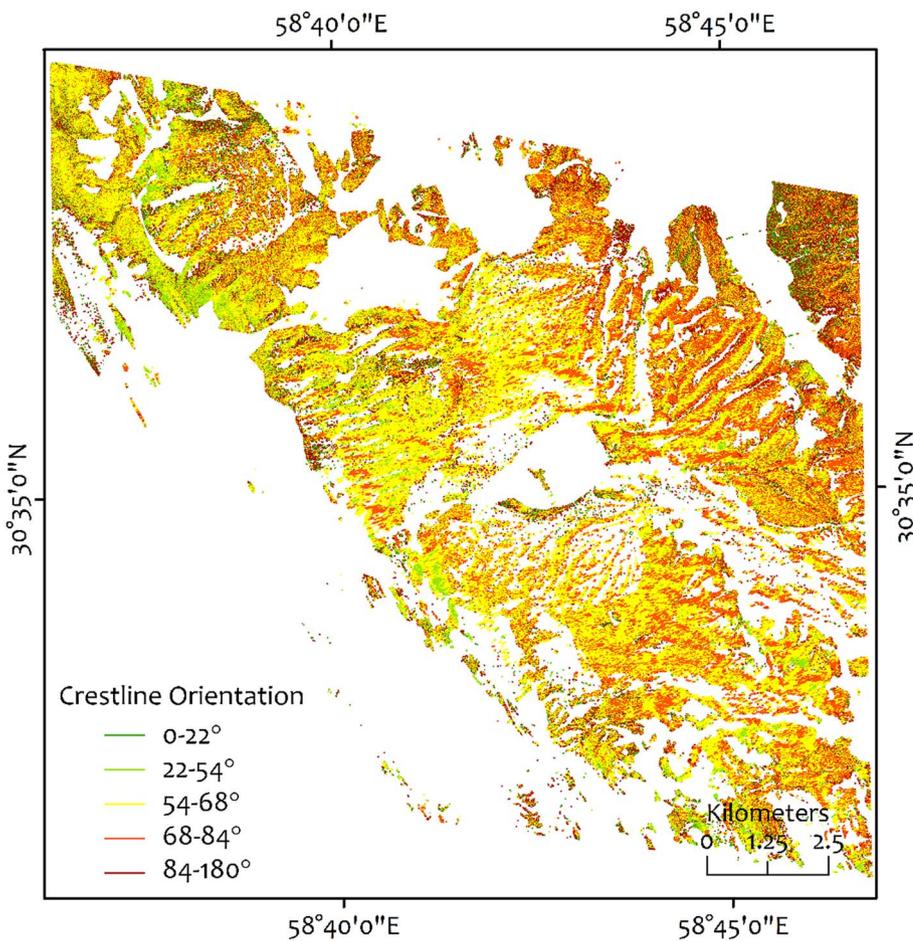


Fig. 10. The resultant overall crestline map, which has been classified based on the orientation of the crestlines. The map indicates that all features in this area are not unidirectional and some factors particularly in marginal regions effect their trends.

trend. Fig. 11e,f shows a closer view of the mega-ripples that have deviated by influence of nearby topography and all have been captured by the SOM method. The trend-based classified TAR crestlines in Kaiser crater (Fig. 12b), also demonstrates larger trend differences for Martian features, which have been perfectly captured by our method. This

evidences indicate that the presented methodology is not directionally biased.

Illumination conditions in the images may play an important role in accuracy of the results, and other studies have not evaluated its effects. Generally, these linear bedforms tend to show their distinctive form

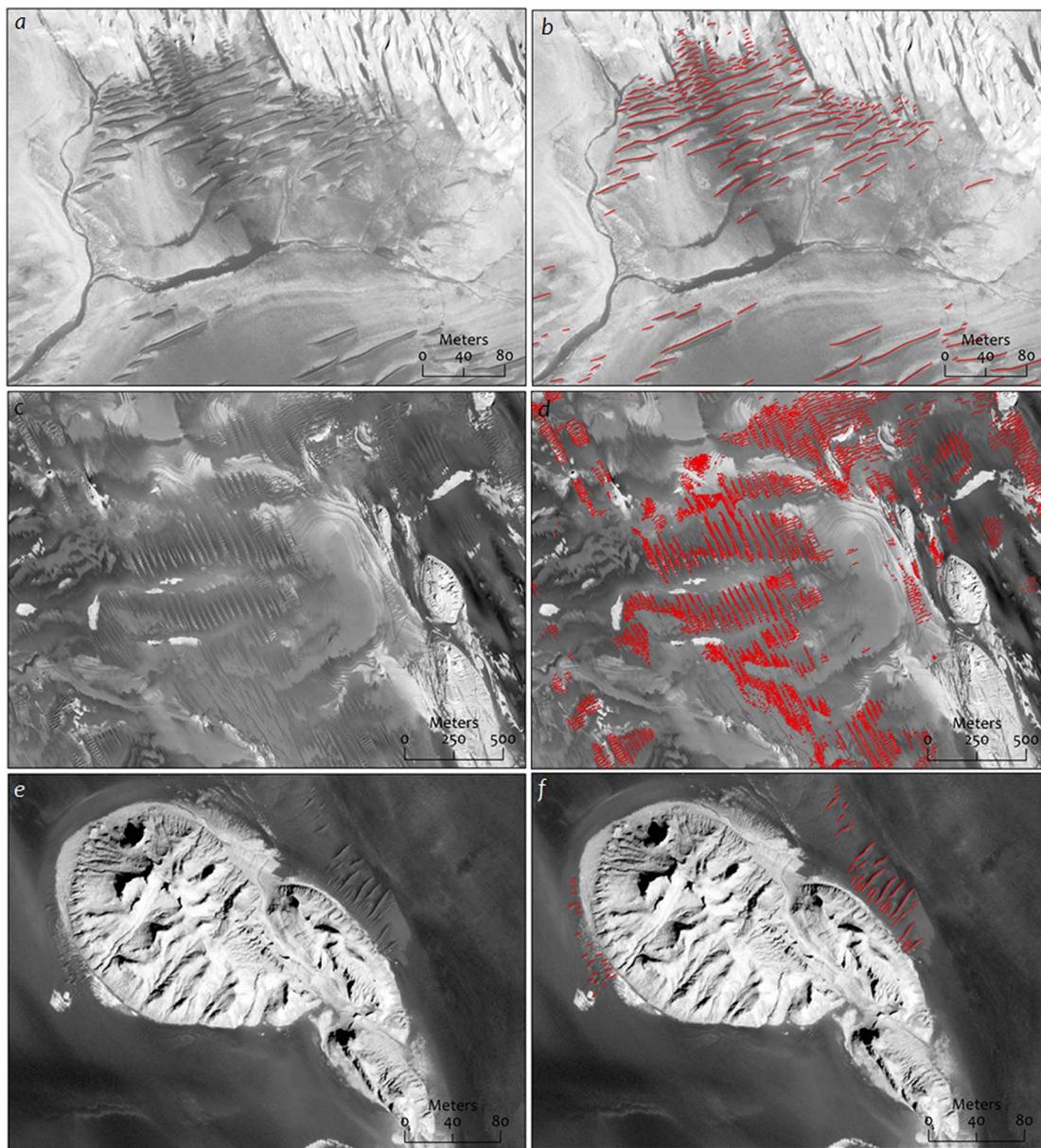


Fig. 11. Different forms and patterns of TAR-like features with diverse background albedo and geology in the Lut desert of Iran that have been overlaid by digitized crestlines from the SOM semi-automatic method. a and b) Confined TAR-like bedforms on lacustrine sediments surrounded by light ventifacts and fluvial features. c and d) Raked TAR-like patterns surrounded by different colors of sediments. e and f) Widely spaced and topographically influenced features around a yardang.

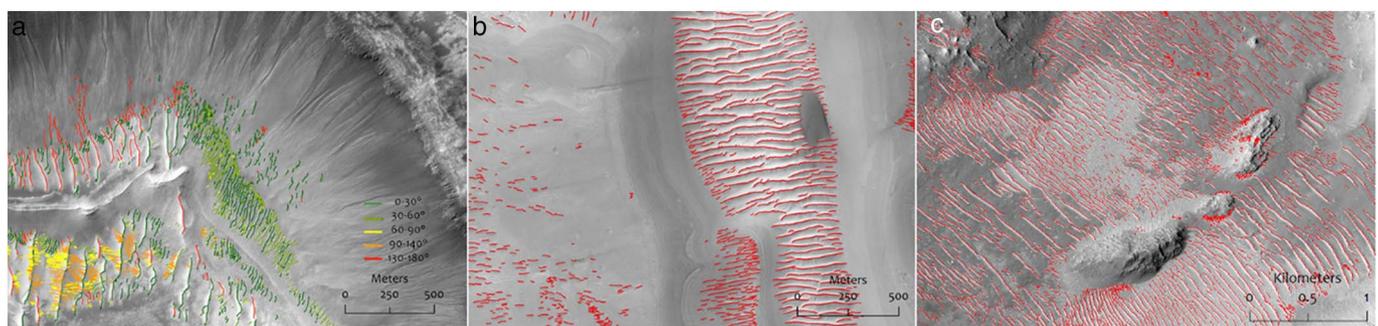


Fig. 12. Crestline maps extracted by the introduced mapping framework for TARs on Mars observed in diverse topographic settings. a) TARs with variable sizes in Kaiser Crater (PSP_006820_1325). Classified based on orientation. b) TARs in Noctis Labyrinthus, superimposed by a large dark dune (LDD) (ESP_044009_1730). c) TARs on the floor of an impact crater in Phlegra Dorsa (ESP_045503_2000).

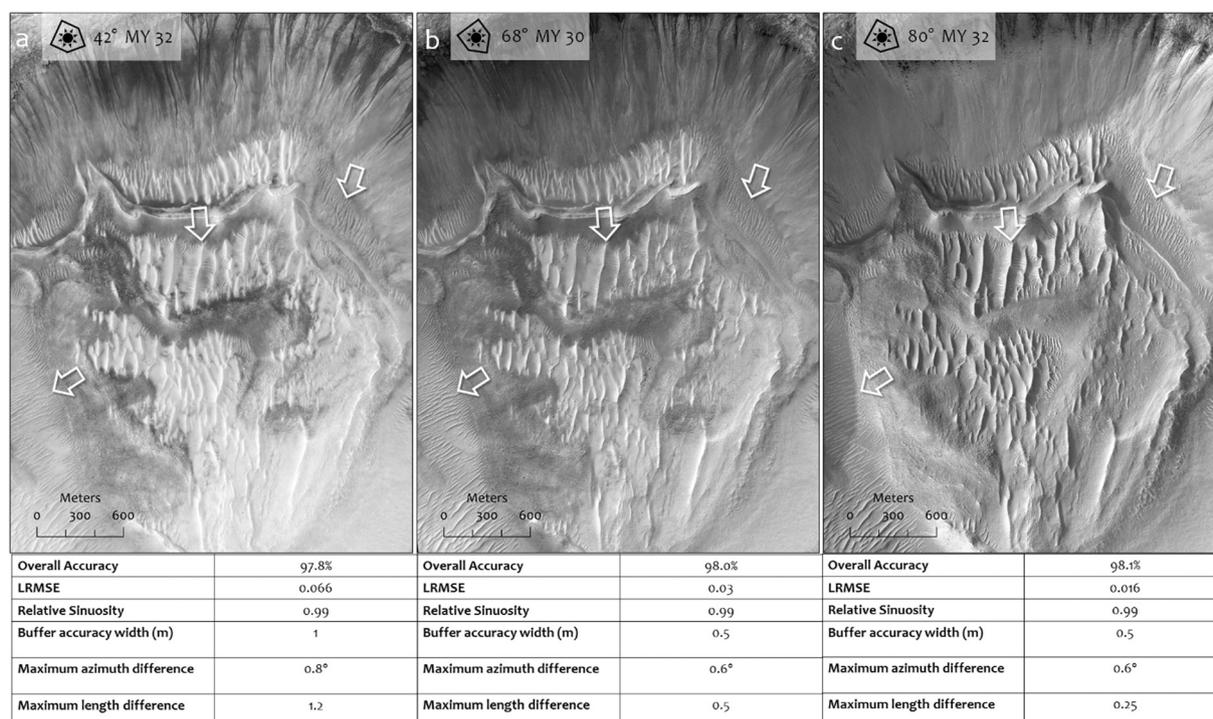


Fig. 13. Kaiser Crater with different illumination conditions. Illumination direction in all images is indicated by a sun symbol, followed by the solar incidence angle and Mars Year. Results indicate minor differences in position accuracy of the crestlines. (a) HiRISE ESP_039944_1325, (b) HiRISE ESP_015918_1325, (c) HiRISE ESP_036871_1325).

clearly in the satellite images that have one bright side and one shadowed side. All our Earth data had almost uniform light conditions with one dark and one bright face. For evaluating the uncertainty that could result from illumination condition, we processed three different images from Kaiser crater with clearly different illumination conditions (Fig. 13). We examined each TAR one by one for the position and shape accuracy with the resultant manual one (as our reference map). Regarding the diverse feature orientation in this crater (Fig. 12a) the reference data for each TAR came from digitizing the image with the best illumination condition for that TAR. This means the reference data created from all images and indicated the highest accuracy for the crestline position and shape. Arrows in Fig. 13 indicate the same regions, with different TAR clarity in the images. The overall accuracy assessment of the results showed only slightly different accuracy, < 0.3%, which could be ignored. However, the main differences in the results were in buffer width, length and azimuth accuracy. This means the methodology is capable of capturing the features in images with different illumination conditions; however, the position accuracy of the crestlines are affected for features with almost the same brightness on both faces or under the shadow of other nearby features. Considering the 0.25 m pixel⁻¹ resolution of the Fig. 13a comparing to 0.50 m pixel⁻¹ in two other images may emphasize the importance of illumination conditions for small linear features, even more than pixel resolution. Fig. S5 indicates the rose plot of orientation distribution of TARs in the Kaiser crater.

4. Conclusion

The unsupervised semi-automatic ANN method known as SOM within a simple framework was conducted using high resolution satellite images of a large mega-ripple field in Iran and TARs on Mars, for aeolian bedforms digitizing. The proposed semi-automatic methodology has been evaluated for accuracy and the results are consistent with traditional photo-interpretation, with about 98% overall accuracy. This method is capable of providing a precise delineation between small linear features, such as TARs and mega-ripples, oriented in all directions, from other morphologic linear features present in the same

satellite image. The network settings were examined in order to define ones that most accurately captured the mega-ripple crestlines and outlines from the images.

The overall accuracy of the mapping for these small linear features could not properly explain the accuracy of the results. We suggest both positional and shape accuracy, along with overall accuracy, may shed light on different challenges and shortcomings of these kinds of methods. Our evaluation shows illumination condition presents small uncertainties in the position accuracy of the resulted crestlines. So, this should be taken into consideration in change recognition maps of small features using our or other semi-automatic methodology. In addition, results indicate that features with small altitudes (shadow length) and < 2-pixel aspect ratio cannot properly be captured by this method.

The framework introduced in this work could probably be applicable and adjustable to many types of images which contain linear features. It can facilitate and speed up image analysis along with improving the accuracy of the results. Furthermore, model validation based on multi-temporal images in different environmental management issues and geophysical problems, such as climate change effects and motion tracking in aeolian geomorphology, could greatly benefit from this precise feature extraction technique.

Acknowledgement

Sincere appreciation is expressed to the two anonymous reviewers, who carefully read the manuscript and provided us with constructive suggestions and comments.

Appendix A. Supplementary Material

Supplementary material to this article can be found online at <http://dx.doi.org/10.1016/j.geomorph.2017.05.016>.

References

Anderson, D.L., Ames, D.P., Yang, P., 2014. Quantitative methods for comparing different polyline stream network models. *J. Geogr. Inf. Syst.* 6, 88–98.

- Berman, D.C., Balme, M.R., Rafkin, S.C., Zimbelman, J.R., 2011. Transverse aeolian ridges (TARs) on Mars II: distributions, orientations, and ages. *Icarus* 213 (1), 116–130.
- Chavez, P., Sides, S.C., Anderson, J.A., 1991. Comparison of three different methods to merge multiresolution and multispectral data-Landsat TM and SPOT panchromatic. *Photogramm. Eng. Remote Sens.* 57 (3), 295–303.
- Coops, N.C., Johnson, M., Wulder, M.A., White, J.C., 2006. Assessment of QuickBird high spatial resolution imagery to detect red attack damage due to mountain pine beetle infestation. *Remote Sens. Environ.* 103 (1), 67–80.
- DigitalGlobe, 2008a. QuickBird imagery products: Product guide. Longmont, CO (2008) [Online]. <http://digitalglobe.com>.
- DigitalGlobe, 2008b. In: Longmont, CO (2008) [Online (Ed.), WorldView-1 products quick reference guide, . <http://digitalglobe.com>.
- Ehsani, A.H., Quiel, F., 2008. Application of self organizing map and SRTM data to characterize yardangs in the Lut desert, Iran. *Remote Sens. Environ.* 112 (7), 3284–3294.
- Ewing, R.C., Peyret, A.P.B., Kocurek, G., Bourke, M., 2010. Dune field pattern formation and recent transporting winds in the Olympia Undae Dune Field, north polar region of Mars. *Journal of Geophysical Research: Planets* 115 (E8).
- Ewing, R.C., Hayes, A.G., Lucas, A., 2015. Sand dune patterns on Titan controlled by long-term climate cycles. *Nat. Geosci.* 8 (1), 15–19.
- Fenton, L.K., Michaels, T.I., Chojnacki, M., 2015. Late Amazonian aeolian features, gradation, wind regimes, and sediment state in the vicinity of the Mars exploration rover opportunity, Meridiani Planum. *Mars. Aeolian Research.* 16, 75–99.
- Foroutan, M., Zimbelman, J.R., 2016. Mega-ripples in Iran: a new analog for transverse aeolian ridges on Mars. *Icarus* 274, 99–105.
- Foroutan, M., Kompanizare, M., Ehsani, A.H., 2013. Semiautomatic morphometric land surface segmentation of an arid mountainous area using DEM and self-organizing maps. *Arab. J. Geosci.* 6 (12), 4795–4810.
- Goodchild, M.F., Hunter, G.J., 1997. A simple positional accuracy measure for linear features. *Int. J. Geogr. Inf. Sci.* 11 (3), 299–306.
- Hölbling, D., Neubert, M., 2008. ENVI Feature Extraction 4.5. Snapshot. In: *GIS Business*, 7/2008, pp. 48–51.
- Jianwen, M., Bagan, H., 2005. Land-use classification using ASTER data and self-organized neural networks. *Int. J. Appl. Earth Obs. Geoinf.* 7, 183–188.
- Katartzis, A., Sahli, H., Pizurica, V., Cornelis, J., 2001. A model-based approach to the automatic extraction of linear features from airborne images. *IEEE Trans. Geosci. Remote Sens.* 39 (9), 2073–2079.
- Kohonen, T., 2001. *Self Organizing Maps*, third ed. Springer, New York.
- Koua, E.L., MacEachren, A., Kraak, M.J., 2006. Evaluating the usability of visualization methods in an exploratory geovisualization environment. *Int. J. Geogr. Inf. Sci.* 20, 425–448.
- Lillesand, T., Kiefer, R.W., Chipman, J., 2014. *Remote Sensing and Image Interpretation*. John Wiley & Sons.
- Marini, F., Zupan, J., Mageri, A.L., 2005. Class-modeling using Kohonen artificial neural networks. *Anal. Chim. Acta* 544, 306–314.
- McEwen, A.S., et al., 2007. Mars Reconnaissance Orbiter's High Resolution Imaging Science Experiment (HiRISE). *J. Geophys. Res.* 112, E05S02. <http://dx.doi.org/10.1029/2205JE002605>.
- Miao, Z., Shi, W., Zhang, H., Wang, X., 2013. Road centerline extraction from high-resolution imagery based on shape features and multivariate adaptive regression splines. *IEEE Geosci. Remote Sens. Lett.* 10 (3), 583–587.
- Montgomery, D.R., Bandfield, J.L., Becker, S.K., 2012. Periodic bedrock ridges on Mars. *Journal of Geophysical Research: Planets* 117 (E3).
- Palafox, L.F., Hamilton, C.W., Scheidt, S.P., Alvarez, A.M., 2017. Automated detection of geological landforms on Mars using convolutional neural networks. *Comput. Geosci.* 101, 48–56.
- Pedersen, A., Kocurek, G., Mohrig, D., Smith, V., 2015. Dune deformation in a multi-directional wind regime: White Sands Dune Field, New Mexico. *Earth Surf. Process. Landf.* 40 (7), 925–941.
- Raiche, A., 1991. A pattern recognition approach to geophysical inversion using neural nets. *Geophys. J. Int.* 105 (3), 629–648.
- Romero, A., Gatta, C., Camps-Valls, G., 2016. Unsupervised deep feature extraction for remote sensing image classification. *IEEE Trans. Geosci. Remote Sens.* 54 (3), 1349–1362.
- Shi, W., Miao, Z., Wang, Q., Zhang, H., 2014. Spectral-spatial classification and shape features for urban road centerline extraction. *IEEE Geosci. Remote Sens. Lett.* 11 (4), 788–792.
- Sobel, I.E., 1970. *Camera models and machine perception*. Ph.D. Thesis, Electrical Engineering Department, Stanford University, Stanford, CA.
- van Heerden, W.S., Engelbrecht, A.P., 2013. Unsupervised Weight-Based Cluster Labeling for Self-Organizing Maps. In *Advances in Self-Organizing Maps* (pp. 45–54) Springer, Berlin Heidelberg.
- Vaz, D.A., 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.
- Vesanto, J., Alhoniemi, E., 2000. Clustering of the self-organizing map. *IEEE Trans. Neural Netw.* 11 (3), 586–600.
- Wu, J., Bauer, M.E., Wang, D., Manson, S.M., 2008. A comparison of illumination geometry-based methods for topographic correction of QuickBird images of an undulant area. *ISPRS J. Photogramm. Remote Sens.* 63 (2), 223–236.
- Zimbelman, J.R., 2010. Transverse aeolian ridges on Mars: first results from HiRISE images. *Geomorphology* 121 (1), 22–29.