

METHODOLOGY

Open Access



Detection of Martian dust storms using mask regional convolutional neural networks

Rasha Alshehhi^{1*} and Claus Gebhardt²

Abstract

Martian dust plays a crucial role in the meteorology and climate of the Martian atmosphere. It heats the atmosphere, enhances the atmospheric general circulation, and affects spacecraft instruments and operations. Compliant with that, studying dust is also essential for future human exploration. In this work, we present a method for the deep-learning-based detection of the areal extent of dust storms in Mars satellite imagery. We use a mask regional convolutional neural network, consisting of a regional-proposal network and a mask network. We apply the detection method to Mars daily global maps of the Mars global surveyor, Mars orbiter camera. We use center coordinates of dust storms from the eight-year Mars dust activity database as ground-truth to train and validate the method. The performance of the regional network is evaluated by the average precision score with 50% overlap (mAP_{50}), which is around 62.1%.

Keywords: Mars, Dust storm, Mask regional convolutional neural networks, Average precision score

1 Introduction

The Martian dust cycle is of fundamental importance to the meteorology and climate of the Martian atmosphere (e.g., Haberle et al. 2017; Kass et al. 2016; Montabone et al. 2015a). Atmospheric dust absorbs and scatters solar and infrared radiation. It thus increases the atmospheric temperature and enhances the atmospheric general circulation (e.g., Gebhardt et al. 2020, 2021; Newman and Richardson 2015). Moreover, dust storms are a very common phenomenon on Mars. Every few Martian years, on average, global dust storm events occur. Hence, the Mars dust cycle has implications for spacecraft engineering parameters, the entry-descent-landing (EDL) operation of spacecraft, the energy production by the solar panels of Mars rovers/landers, etc. Also, it is an essential concern for future human exploration of Mars.

Martian dust storms are evident as frontal features (Wang and Richardson 2015), dust storm texture/convective features (Guzewich et al. 2015), and dust clouds (Cantor et al. 2019). Based on the definition

of Cantor et al. (2001), regional dust storms differ from local dust storms by having an area of $\geq 1.6 \times 10^6 \text{ km}^2$ and a duration of more than two days. Global dust storm events (GDEs) or planet-encircling dust storms start as local/regional dust storms and engulf the entire planet (Forget and Montabone 2017). Still, dust lifting takes place at the regional scale and GDEs have several active dust lifting centers. GDEs have a duration of up to a few months and occur, by average, each few Martian Years (Zurek and Martin 1993). While there may be local and regional dust storms at any time of the Martian year, GDEs occur only during the second half of the Martian year ($L_S = 180^\circ - 360^\circ$). The latter is known as the dust storm season and coincides with the Mars southern hemisphere spring and summer. A yearly repeatable phenomenon is multiple local dust storms at the northern/southern Mars polar cap edge in the respective hemispheric fall to the spring season, known as polar cap edge storms. By contrast, dust devils are another phenomenon. They may have diameters of several hundred meters and durations of several tens of minutes (Reiss et al. 2011).

A comprehensive dust climatology was detailed in Montabone et al. (2015a). The basis for that are data on

*Correspondence: ra130@nyu.edu

¹ Center for Space Science, New York University, Abu Dhabi, United Arab Emirates

Full list of author information is available at the end of the article

the column dust optical depth from the satellite instruments MCS/MRO (Mars climate sounder/MARS reconnaissance orbiter), THEMIS/MO (thermal emission imaging system/Mars Odyssey), and TES/MGS (thermal emission spectrometer/Mars global surveyor). The latter operate at different wavelength ranges, measurement geometries, and spatial and temporal coverage. This dust climatology is made publicly available via the Mars climate database (MCD),¹ together with many other parameters of the Mars atmosphere and surface. It has a moderate spatial resolution of a few degrees latitude and longitude² and was demonstrated to be suitable to follow the evolution of certain regional dust storms by Montabone et al. (2015a). Various studies identified and explored dust storms based on the visual inspection of Mars daily global maps (MDGMs) from the camera system MOC/MGS (Cantor 2007; Hinson et al. 2012). Other studies focused on MDGMs from both MOC/MGS and MARCI/MRO (Battalio and Wang 2021; Wang and Richardson 2015). In this work, we perform a feasibility study on a deep-learning-based approach for dust storm detection from the record of MDGMs by the Mars orbiter camera (MOC) (Malin et al. 2010) aboard the Mars global surveyor (MGS), by applying regional convolutional neural networks (R-CNNs) (Simonyan and Zisserman 2014).

Recently, deep convolutional networks have made significant improvements in the accuracy of object detection, in particular R-CNNs which focus on an object of interest (e.g. face, car, etc.), called reference or ground-truth objects, from images. Object detection is a challenging task because it requires the accurate localization of candidate objects. In this paper, we use mask regional convolutional neural networks (Mask R-CNNs) that jointly learn to classify dust storm candidates and refine their spatial locations. The spatial extent of potential dust storm candidates may be to a certain degree arbitrary because dust storm boundaries are identified based on the subjective perception of individual observers and are interpolated if they intersect gaps in satellite images and/or the edge of the polar night. Here we draw a rectangular box or bounding box around the center coordinates of each dust storm instance and consider it as a reference box or ground-truth box.

In this paper, we detect the presence of dust storms in an image, estimate their edge coordinates and evaluate it based on hand-drawn reference boxes. The main contribution of this work can be summarized as follows:

- It is the first work on the deep-learning-based detection of Mars dust storms which is applied to several Martian-year records of MDGMs. Also, it uses the Mars dust storm database of Battalio and Wang (2021), which is one of the most recent and comprehensive of its kind, as a ground-truth.
- It uses a new architecture that consists of two networks for improving the accuracy of the boundaries of dust storm areas, although the ground-truth includes a certain degree of subjectivity and arbitrariness.
- It uses a dice score as a mask loss function to overcome ambiguous cases at the boundary between a dust storm and non-dust-storm categories with a lower level of uncertainty between the two categories.

The outline of this paper is the following. Section 2 describes the previous work related to automated dust storm detection and the latest R-CNN techniques. In Sect. 3, we explain the observation-based dataset and ground-truth we used. In Sect. 4, we illustrate the methodology used to detect dust storms. We discuss the performance of our method in Sect. 5. In Sect. 6, we summarize the main findings and provide an outlook for the future.

2 Related work

2.1 Automated detection of Martian dust storms

Maeda et al. (2015) proposed an automatic method to detect dust storms. Their method is based on selecting features from Martian images using a minimal redundancy maximal relevance algorithm and classification using the support vector machine (SVM) technique into a dust storm and non-dust storm. It successfully detects around 80% of dust storms, but it did not define the locations of dust storms. Gichu and Ogohara (2019) suggested a segmentation method to classify Martian images into either dust areas or cloud areas. They used principal component analysis (PCA) to reduce the number of Martian image bands and supervised multi-layer perceptron (MLP) neural networks based on subjective ground-truth images. They only focused on the regions (patches) with a high frequency of dust storms revealed by Guzewich et al. (2015) and Kulowski et al. (2017). In this work, we concentrate on non-polar Martian images.

2.2 Regional convolutional neural networks (R-CNNs)

The R-CNN is an extended version of the standard CNN, which is used to identify an object of interest; the presence of an object (e.g., face, dog, car, etc.) in an image, the exact location of an object, and the number of occurrences of an object in an image. This cannot proceed with

¹ <http://www-mars.lmd.jussieu.fr/>

² http://www-mars.lmd.jussieu.fr/mars/dust_climatology/index.html

a standard CNN because the number of occurrences of an object of interest varies from image to image. While an object may be present several times in some image, it may not be included in another image. The R-CNN is usually trained based on subjective reference bounding boxes (ground-truth) around an object in images, which are drawn manually by field experts. Girshick et al. (2013) introduced the first R-CNN. They used the CNN, which consists of convolutional and fully-connected layers to extract regions of interest (RoI), called candidate region proposal, from convolutional layers. These RoIs are fed into a classical support vector machine (SVM) to classify the presence of an object within candidate region proposals (i.e., whether there is an object or not). In addition to predicting the presence of an object, the SVM also predicts four coordinates of vertices of an object to increase the precision of the bounding boxes. Ren et al. (2016) introduced a regional proposal network (RPN) for a better generation of RoI or candidate region proposals within a shorter time. This network is called Fast R-CNN. He et al. (2017) added a new architecture, called segmentation mask network, which works in parallel with the RPN. This network generates a segmentation map i.e., the identification of a bounding box with the input image that indicates the location and extends (boundaries) of the object. This network is called a Mask R-CNN. Cheng et al. (2020) proposed a modified version of the Mask R-CNN architecture to enhance the precision of boundaries of an object of interest. In this work, we use a modified version of the Mask R-CNN to identify dust storms in Martian maps and subjectively estimate its coordinates from RoI.

3 Data

As ground-truth images, we use the Mars dust activity database (MDAD) (Battalio and Wang 2021). This is a dust storm database compiled from eight Martian years (MY) of Mars Daily Global Maps (MDGMs), which means from MY 24, L_S 150° (1999) to MY 32, L_S 171° (2014). The MDAD comprises 14,974 dust storm instances, which are, by definition, enclosed dust storm regions on a single sol (Martian day). The dust storm instances are combined into 7927 dust storm members. These are subdivided further into a total of 228 dust storm sequences (125 originated in the northern hemisphere and 103 in the southern hemisphere).

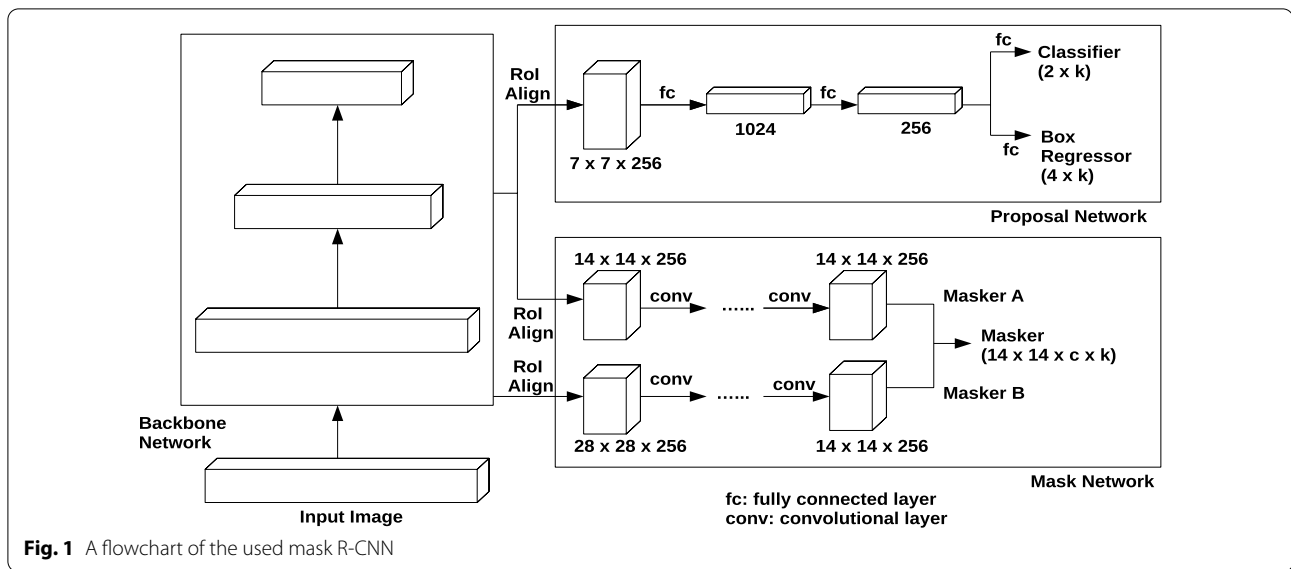
The Mars dust activity database can be found at <https://doi.org/10.7910/DVN/F8R2JX>. It includes the center coordinates (longitude and latitude) and area (in km^2) of individual dust storm instances. We use the center coordinates of each such dust storm instance but, as a simplifying assumption, consider rectangle areas around the center coordinates. The MDAD also includes confidence

levels (CL) of 100, 75, and 50, which are assigned to each dust storm instance based on visual inspection. They rate the accuracy of dust storm boundaries with the highest confidence level of 100 and the lowest confidence level of 25. CL = 100 means the entire perimeter of the dust storm instance is distinct against the background so that the dust storm edge has an error on the order of a few pixels only (which is equivalent with approximately 0.5°). CL = 25 shows rather nebulous boundaries that cannot be exactly discerned from the background within a few degrees of latitude/longitude. The CL is also used to determine how distinct a dust storm instance is from the background atmospheric opacity. Only dust storm instances with CL = 100, 75, 50 are listed in the MDAD.

In the following, we include all Mars Daily Global Maps (MDGMs) based on MOC/MGS, from MY 24, L_S 150° (1999) to MY 28, L_S 121° (2006), as obtained from <https://doi.org/10.7910/DVN/WWRT1V>. We consider the non-polar versions of these MDGMs, which cover latitudes from 60°N–60°S and longitudes from 180°E–180°W and have simple cylindrical map projection. The MDGMs have a resolution of 7.5 km per pixel with 0.1° longitude by 0.1° latitude. They are available as RGB images. Details on the MDGM production process can be found in Wang and Ingersoll (2002). Each MDGM is based on 13 wide-angle global map swath images of the Mars global surveyor (MGS) Mars orbiter camera (MOC). The latter covers the whole sun-lit planet around 2 PM local time each sol. The MDGMs consist of imagery from the two visible bands, red (575–625 nm) which is more sensitive to dust storms, and blue (400–450 nm) which is more sensitive to water ice clouds (Cantor et al. 2001). The green component of the MDGMs is synthesized by combining 1/3 red and 2/3 blue and applying linear stretching.

4 Method

A basic introduction to image classification and object detection by classical convolutional neural networks can be found in Higham and Higham (2019), Zhao et al. (2019). The fast regional convolutional neural network (fast R-CNN) (Girshick 2015) usually includes two classical convolutional neural networks: the base network/backbone network and the detection network. The detection network, in turn, is a regional proposal network (RPN). A modified version of the fast R-CNN is the mask regional convolutional neural network (Mask R-CNN) (He et al. 2017). The detection network in Mask R-CNNs is a combination of a regional proposal network (RPN) and a mask network (called segmentation network). We use a Mask R-CNN to estimate the probabilities of regions to show a dust storm in the Martian map. This Mask R-CNN includes a RPN following Ren



et al. (2016). In addition to that, it consists of a mask network that predicts a dust storm segmentation map on a pixel-to-pixel basis (He et al. 2017). A segmentation map is a bounded region of pixels within the image that have a higher probability of being a dust storm (being classified in the “dust storm category”) than not.

As a backbone network, we use a residual network (ResNet) (He et al. 2016) (see “Appendix” for a description of the ResNet architecture). In principle, the ResNet works better with additional layers. It finds an optimized number of layers to negate the vanishing gradient problem in classic networks. The backbone network takes the initial Martian map (MDGMs) as an input image and outputs convolutional feature maps. The latter are, in turn, the input of the detection network. To extract the region of interest (RoI) from convolutional feature vector maps, we use the feature pyramid network (FPN) (Lin et al. 2017) as a detection network. We use the FPN architecture because it combines the low-resolution, semantically strong features with the high-resolution, semantically weak features, which is sufficient for capturing the difference between dust storm and non-dust-storm regions (see “Appendix” for a description of the FPN architecture).

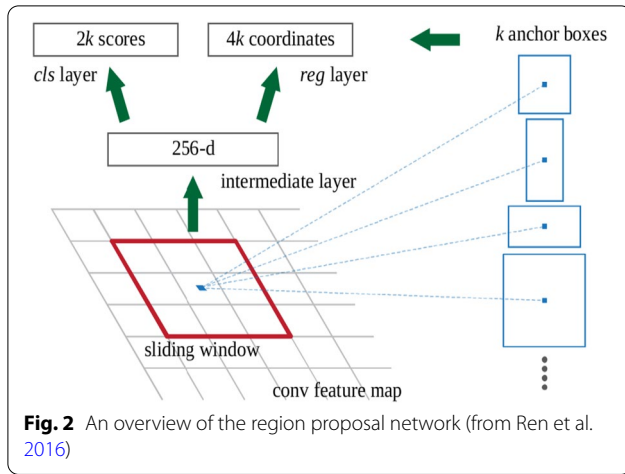
The RPN includes a classifier and a bounding box regressor, for the purpose of object classification and bounding box optimization. The classifier predicts the dust storm probability, called score, of each region; The score is obtained for all regions (collection of pixels), where a high score region refers to a region that has a high probability of being a dust storm, and a low score region is likely not a dust storm. The bounding box regressor predicts four boundary coordinates of the dust

storm regions (Girshick 2015; He et al. 2017). Usually, the RPN use multiple reference boxes, called anchors, to obtain more accurate boundary coordinates of dust storm regions. The reference boxes are used to evaluate the performance of the RPN.

Refining the Mask R-CNN is beyond the scope of this paper and we provide the reader with an overview of this method in this article. Figure 1 shows a flowchart of the current method, and we will discuss different parts in detail in the following sections.

4.1 Region proposal network (RPN)

The RPN takes as input a certain region (called the region of interest, RoI) from the convolutional feature map of the backbone network and outputs a set of rectangular candidate regions, called proposals, over the input region. Each proposal has a dust-storm-probability and non-dust-storm-probability, called score, and four coordinates of the most likely area of a dust storm. In detail, the input region of the RPN is a spatial window with the dimensions $w \times h \times l$, where w , h , l are the width, height and number of feature maps of the convolutional layers. The width, height, and the number of bands of the input region are $7 \times 7 \times 256$ (w and h are determined experimentally). The input region of the RPN is fed into two sibling fully-connected network layers and finally into the box-classification layer (l_c) and the box-regression layer (l_r). These are briefly denoted sibling, classification and regression network in the following, respectively. The outputs of the sibling network layers are classification scores and coordinates of four vertices of k proposals of the input window $w \times h$. k is a pre-defined number of reference boxes with different scales and aspect ratios,



called anchors. The anchors are derived as rectangular regions from the convolutional feature maps of the initial MDGMs, based on the longitude and latitude of the center coordinate of respective dust storms of the MDAD dataset. The region is associated with a scale s and aspect ratio r ($s = 4$ and $r = 3$; these are determined based on computational capacities). The anchors are applied after the re-sampling of any input feature vectors. The output of l_c of window $w \times h$ is $2 \times k$ elements, which are the probabilities of any anchor to contain or not a dust storm. These probabilities are called classification scores. The output of l_r of window $w \times h$ is $4 \times k$ elements, which are the four corner coordinates of each dust storm proposal. These include offsets added by l_r for optimizing the object detection quality. Figure 2 shows an overview of the RPN.

4.2 Mask R-CNN

The mask regional convolutional neural network (Mask R-CNN) runs over the convolutional feature vector maps of the ResNet backbone network and produces a segmentation map. The Mask R-CNN works in parallel with the classification and regression network, described in Sect. 4.1. This is also illustrated in Fig. 1. The Mask R-CNN takes a region from convolution feature vector maps with the spatial dimension $x \times y \times z$, where x , y and z are in units of width, height and the number of feature maps ($x \times y \times z$ is $14 \times 14 \times 256$ and $28 \times 28 \times 256$; the window size is determined experimentally). The Mask R-CNN produces a binary map with $x \times y \times 1$. The latter dimension is equal to one because there is only one class (dust storm region or the planetary surface background). The binary map is a classified 2D image, where 1 refers to foreground pixels, i.e., dust storm pixels, and 0 refers to background pixels, i.e., non-dust-storm pixels.

4.3 RoIAlign

RoIAlign (He et al. 2017) (see “Appendix” for details about RoIAlign) is a standard operation for extracting regions of interest (RoI) from convolutional feature maps. We use RoIAlign both within the classification and regression network and segmentation network. The regions extracted by the segmentation network are denoted $\text{RoI}_{\text{Masker}_A}$ and $\text{RoI}_{\text{Masker}_B}$ in the following. The $\text{RoI}_{\text{Masker}_A}$ is based on the output of the second to fifth residual convolutional layer in the first part of the mask segmentation network (the corresponding process is denoted Masker_A in Fig. 1). The $\text{RoI}_{\text{Masker}_B}$ is based on the output of the second part of the segmentation network (Masker_B in Fig. 1). In Masker_A , the $\text{RoI}_{\text{Masker}_A}$ is the result of four consecutive 3×3 convolutions, while in Masker_B , the $\text{RoI}_{\text{Masker}_B}$ is the result of two consecutive 3×3 convolutions and Masker_A . The combination of Masker_A and Masker_B helps obtain a binary map with rough coordinates of dust storm pixels.

4.4 Training the network by learning and optimization

The loss function L is a combination of classification L_c , regression L_r and segmentation mask L_m losses. We use a binary cross-entropy to classify each box and a mean absolute error (MAE) to estimate four coordinates of each box, which are coordinates of the four vertices. To alleviate the class-imbalance problem between positive pixels (dust) and negative pixels (non-dust), we use Dice loss (Milletari et al. 2016) to measure overlapping between prediction and ground-truth.

$$L(y, \tilde{y}) = \frac{1}{N} \sum_{n=0}^N L_c(y_n, \tilde{y}_n) + L_r(y_n, \tilde{y}_n) + L_m(y_n, \tilde{y}_n), \quad (1)$$

$$L_c(y_n, \tilde{y}_n) = -y_n * \log(\tilde{y}_n) + (1 - y_n) * \log(1 - \tilde{y}_n), \quad (2)$$

$$L_r(y_n, \tilde{y}_n) = \sum_{j=1}^{J=4} \|\tilde{y}_n(j) - y_n(j)\|, \quad (3)$$

$$L_m(y_n, \tilde{y}_n) = \frac{2 \sum_i^I y_n(i) * \tilde{y}_n(i)}{\sum_i^I y_n(i) + \sum_i^I \tilde{y}_n(i)}, \quad (4)$$

where N , J and I are the number of reference boxes, the number of four coordinates of each box, the number of pixels of each box, respectively. In Eq. 2, y is the ground-truth value of each reference box n (0 for non-dust storm box and 1 for dust storm box). In Eq. 3, y is (x, y) coordinates of all vertices of each reference box n . In Eq. 4, y is the ground-truth value of each pixel i in each reference box n (0 for non-dust pixel and 1 for dust pixel). In all equations, \tilde{y} is the predicted probability of the box.

5 Performance

In this section, we use various training strategies and present the performance of Mask R-CNNs on MDGMs derived from MGS/MOC observations. This includes comparisons with state-of-the-art methods.

5.1 Evaluation metrics

We use the intersection-over-union (IoU) to evaluate the performance of the convolutional networks. The IOU is a metric that measures the accuracy of the detection method based on comparing the areas of reference and predicted bounding boxes. It is defined as the area of intersection of the predicted bounding box \tilde{Y}_m with the reference box Y_m divided by the area of the union between \tilde{Y}_m and Y_m :

$$\text{IoU} = \frac{\text{area}(Y_m \cap \tilde{Y}_m)}{\text{area}(Y_m \cup \tilde{Y}_m)}, \quad (5)$$

We assign an object, i.e., bounding box, to a value of 1 (dust storm) if the IOU is greater than a certain threshold, while it is assigned to a value of 0 (non-dust storm) if the IOU is smaller than the same threshold (He et al. 2017). We calculate the precision ($P = \text{TP}/(\text{TP} + \text{FP})$) and recall ($R = \text{TP}/(\text{TP} + \text{FN})$) of each object in the testing dataset, where TP, FN and FP are the number of dust storm pixels correctly classified as dust storm pixels, the number of dust storm pixels classified as non-dust storm pixels and the number of non-dust storm pixels classified as dust storm pixels, respectively. We evaluate the performance of the network based on the mean average precision (mAP) score, where AP is the area under the precision-recall curve. We calculate mAP at various intersection-over-union thresholds th_{IoU} (25%, 50% and 75%).

5.2 Implementation details

- Data: we use RGB (red, green, blue) Mars Daily Global Maps (MDGMs) as input of the Mask R-CNN. We define ground-truth for the Mask R-CNN by using boxes of size 120×120 pixels around the center coordinates of dust storms identified in the MDAD database (Battalio and Wang 2021), where the choice of 120 pixels is determined from multiple experiments with different sizes and selecting the value that provided the best agreement and was within available computational GPU capacities.
- Mask R-CNNs: We use four scales for each RPN anchor $s \in \{32 \times 32, 64 \times 64, 128 \times 128, 256 \times 256\}$ on $\{\text{Conv}_2, \text{Conv}_3, \text{Conv}_4, \text{Conv}_5\}$ layers. We use

three aspect ratios $r \in \{1 : 2, 1 : 1, 2 : 1\}$ at each scale. Using four scales and three ratios follows (Lin et al. 2017). Also, this is in line with computational capacities (e.g., memory and speed).

- Training and inference: we use the adam optimization function. The learning rate is assigned the value of 0.0001 and it decreases by a factor of 10 every 1000 iterations. The weight decay is set to 0.001 and the momentum to 0.9, the step per epoch to 1000, and the validation step to 50. We use one GPU with a mini-batch size equal to 32. We train with 1 image per GPU. The highest IoU threshold used is 0.7 and the lowest IoU threshold used is 0.3, in line with previous reported Mask R-CNNs (Ren et al. 2015; He et al. 2017; Lin et al. 2017; Cheng et al. 2020).

5.3 Experimental results in different seasons

Only 607 out of the 2484 MDGM maps (24%) have occurrences of dust storms in them. The MDGM dataset includes maps from spring season only in MY 25, from summer, fall (146 maps) and winter (150 maps) seasons in MY 26, from spring, summer, fall (126 maps) and winter (185 maps) seasons in MY 27, and from spring and summer seasons in MY 28. We use two training strategies. In the first training strategy, the training dataset mainly includes around 296 maps during dustiness seasons from MY 26 and testing includes around 311 maps from MY 27. However, in the second training strategy, the training set includes around 350 maps during dustiness seasons which were randomly selected from MY 25 and MY 28 and the testing set includes 256 maps. Here, we discuss each strategy in detail.

In the first strategy, we use maps from MY 25, $L_S 0^\circ$ to MY 27, $L_S 180^\circ$ as the training dataset (1216 maps; around 49% of the total MDGM maps). We randomly select validation images from MY 25 to MY 27 (614 maps; around 25% of total MDGM maps) which are not used in the training dataset to validate the performance of the convolutional networks during the training process to obtain a lower error. We use maps from the almost full-year period from MY 27, $L_S 180^\circ$ to MY 28, $L_S 121^\circ$ as the testing dataset (659 maps; around 27% of the total MDGM maps). Each dataset includes images from the entire year.

Figures 3 and 4 show reference dust storm regions (a) and (c) and predicted dust storm regions by the R-CNN (b) and (d) for selected MDGMs of the testing dataset. In Fig. 3a, the reference bounding boxes for MY 28, $L_S 83.04^\circ$ is given by dust storm instances at the coordinates $(89.25^\circ\text{W}, 26.7^\circ\text{S})$ and $(132.05^\circ\text{E}, 30.2^\circ\text{N})$. The dust storm instances are classified with $\text{CL} = 75$ and $\text{CL} = 50$, respectively, which implies that their subjective edges/

Table 1 Comparison between R-CNNs based on various mAP

Networks	mAP_{25}		mAP_{50}		mAP_{75}		Time (ms)	
	strategy 1	strategy 2	strategy 1	strategy 2	strategy 1	strategy 2	strategy 1	strategy 2
Fast R-CNN	64.1	64.8	59.1	58.1	32.2	32.1	350	320
Mask R-CNN	67.8	68.1	61.2	60.5	38.1	33.1	290	300
SPPnet	62.5	63.2	50.1	53.1	22.5	20.1	370	360
Current R-CNN	68.2	67.5	62.1	61.2	39.0	37.7	320	310

The highest mAP and the shortest time are presented in bold font. The mAP_{25} , mAP_{50} , mAP_{75} are mAP s with th_{IoU} of 25%, 50% and 75%

boundaries are less clearly identifiable. As follows from Fig. 3b, the R-CNN identifies a dust storm instance close to the reference dust storm instance (89.25°W, 26.7°S). However, the R-CNN predicts two bounding boxes in the same region that overlaps with the ground-truth reference box. This may be because there are 12 arbitrary reference boxes/anchors ($s = 4$ and $r = 3$), and thus be considered in the same region of interest. Figure 3c, d presents accurate results for MY 28, L_S 110.25° (summer). The ground-truth is given by dust storm instances at the coordinates (96.75°W, 26°S) and (73.05°W, 34.3°S) with $CL = 100$ and $CL = 75$ near to southern polar ice cap. The detection accuracy is approximately 0.99 for both, i.e., high overlapping areas with the ground-truth. Figure 4a, b at MY 27, L_S 222.83°, i.e., during the dust storm season, show that the R-CNN identifies dust storm instances in different regions. However, it mismatches some of the center coordinates and has a certain overlap with surrounding areas. That is the case around the reference bounding boxes at (148.85°W, 39.5°N), (54.35°W, 4.7°N), (14.65°W, 46.2°N) with $CL = 50$, $CL = 75$ and $CL = 75$, respectively. This may be at least partly due to the fact that $CL = 100$ means the dust storm instances still have an error of a few pixels, or approximately around 0.5°, and $CL = 75$ and $CL = 50$ have an error greater than 0.5°, accordingly. Also, it fails to distinguish the dust storm instances with $CL = 100$ and $CL = 50$ at the coordinates (147.75°E, 33.7°N) and (46.55°W, 17.9°S) from the background. A potential explanation for that is increased atmospheric background dustiness during the dust storm season. Figure 4c, d shows accurate results at L_S 305.93° with references bounding boxes at (32.55°W, 0.90°N) and (158.55°W, 36.1°N) and $CL = 75$.

In the second strategy, we apply the network to images from MY 25, L_S 0° to MY 28, L_S 121°, which are randomly divided into a training dataset (1300 maps; around 52% of total MDGM maps), a validation dataset (586 maps; around 24% of total MDGM maps) and a testing dataset (600 maps; around 24% of total MDGM maps) and analyze the performance in all seasons. Figures 5 and 6 show examples from all four seasons at L_S 53.47°, L_S 105.36°, L_S 238.51° and L_S 313.57° in MY 26 and MY 27, respectively.

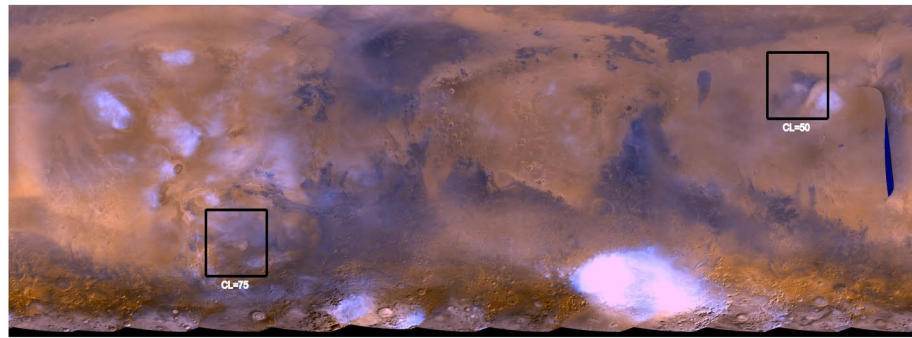
Figure 5a–d shows detected dust storms at L_S 53.47° (spring) and L_S 105.36° (summer). Figure 6a–d presents results at L_S 238.51° (fall) and L_S 313.57° (winter). Here, our method successfully identifies most of the dust storm instances; however, it misses some center coordinates of dust storms (error of a few pixels). Thus, they are to a certain extent subjective and the dust storm instances may even extend over a larger area. If so, our method may have identified nearby regions because they have similar spatial and spectral characteristics. Among others, our method may also have produced some false-negative and false-positive cases due to water ice clouds, large background dustiness, or image gaps in the MDGMs, as in Figs. 5d and 6b, respectively. In line with that, we may integrate some additional processes in the future (e.g., filling missing data, cloud detection, etc.).

5.4 Distribution of longitude-latitude coordinates

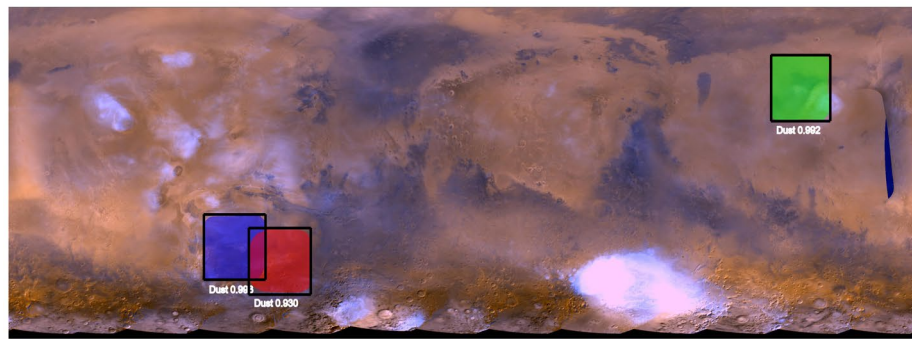
Figure 7a–d shows the distribution of longitude-latitude coordinates of the predicted bounding boxes compared to the subjective coordinates of the reference bounding boxes (delta-longitude dx and delta-latitude dy in pixel) using the first and the second training strategies. We note dx and dy are approximately between -4 and 4 in both strategies, but variations in longitude value are larger compared to variations in latitude value.

5.5 Comparison with state-of-the-art methods

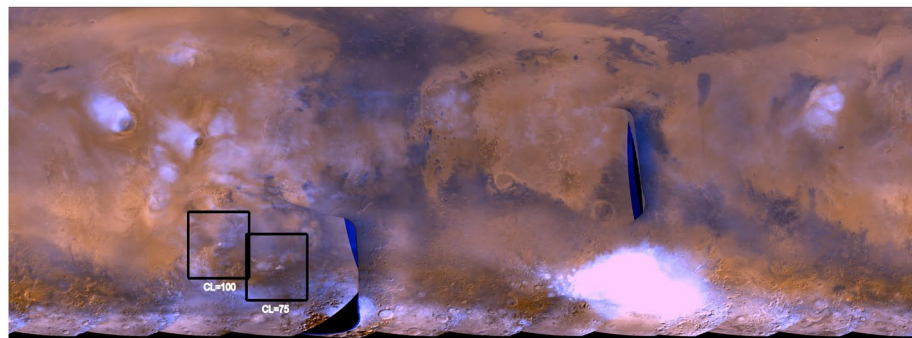
We compare the performance of the regional networks based on the first and second training strategies. The first strategy identifies locations of dust storms for coming years (from middle MY 27 to MY 28) based on old years (from MY 25 to the middle MY 27) and the second strategy identifies locations based on random dust storms (from MY 25 to MY 28). In Table 1, we compare results among the fast R-CNN, Mask R-CNN, SPPnet and the current R-CNN. We compare mAP with th_{IoU} equal to 25%, 50% and 75%. As expected, selecting higher thresholds reduces the effectiveness of all R-CNNs. In addition, the inference times required for each image lie between 300 and 370 milliseconds (ms) for all networks. Mask R-CNNs have higher mAP and are faster



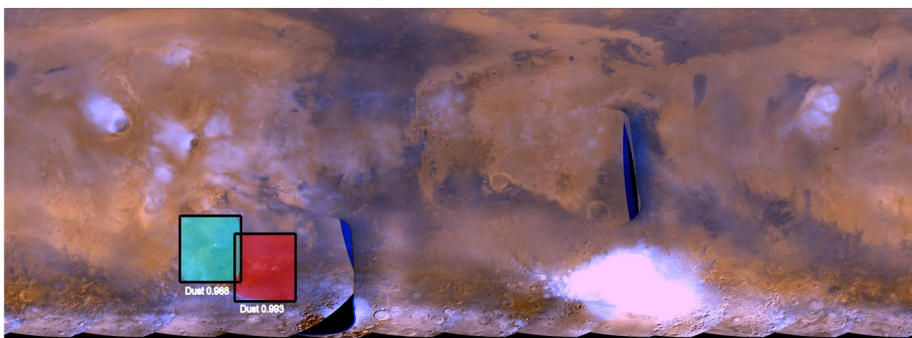
(a)



(b)

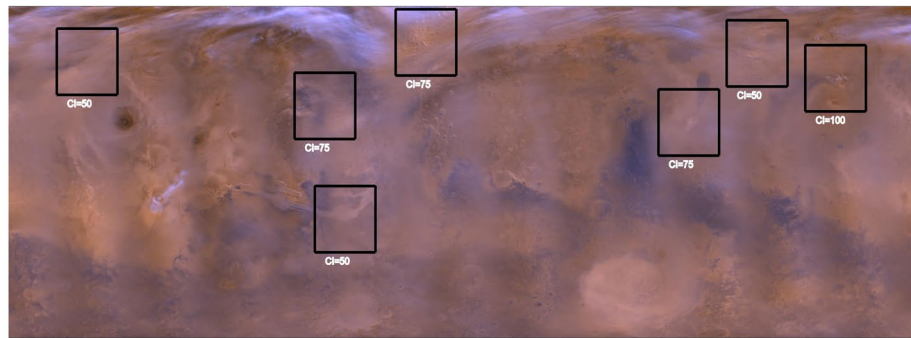


(c)

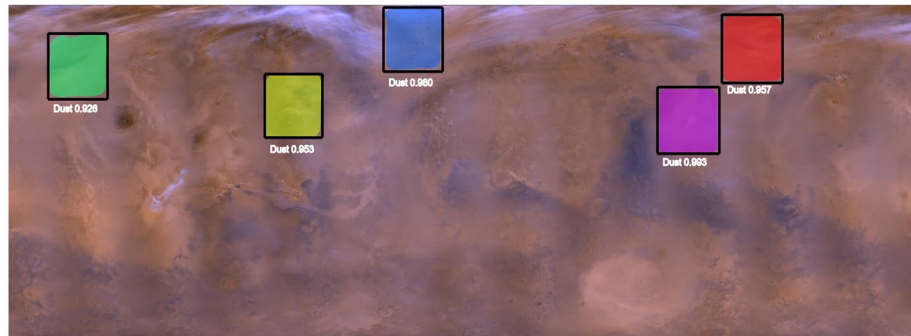


(d)

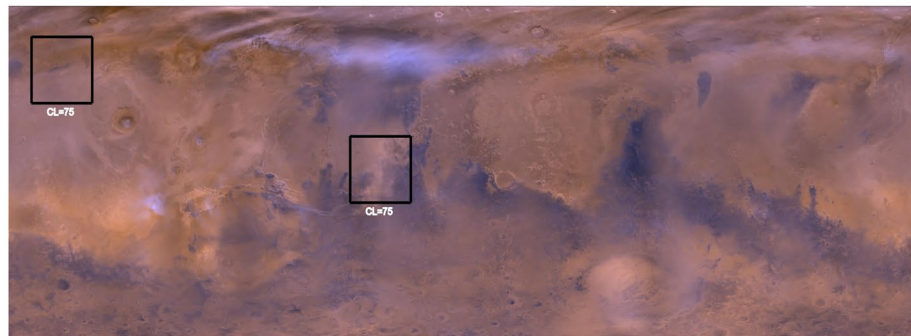
Fig. 3 1st strategy: the training and validation images are selected randomly from MY 25 to middle of MY 27. The testing images are selected randomly from middle of MY 27 to MY 28. **a** and **c** are ground-truth images from spring (s20_23) and summer (s22_22) seasons at $L_S = 83.04^\circ$ and $L_S = 110.25^\circ$, respectively. **b** and **d** are their predicted dust maps



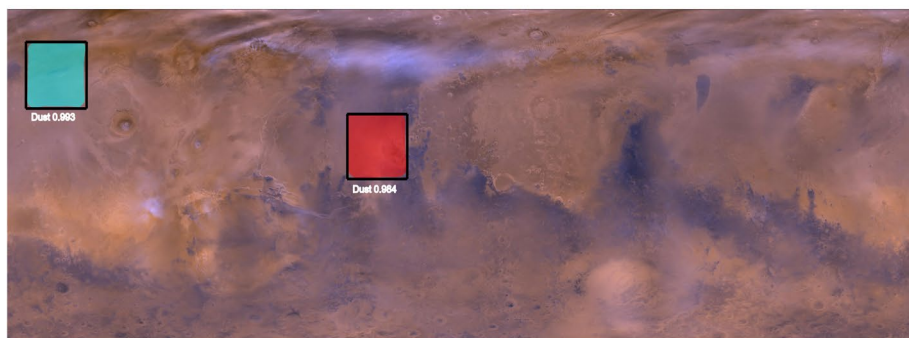
(a)



(b)



(c)



(d)

Fig. 4 1st strategy: the training and validation images are selected randomly from MY 25 to middle of MY 27. The testing images are selected randomly from middle of MY 27 to MY 28. **a** and **c** are ground-truth images from fall season (s20_011) at $L_S = 222.83^\circ$ and winter season (s11_018) at $L_S = 305.93^\circ$, respectively. **b** and **d** are their predicted dust maps

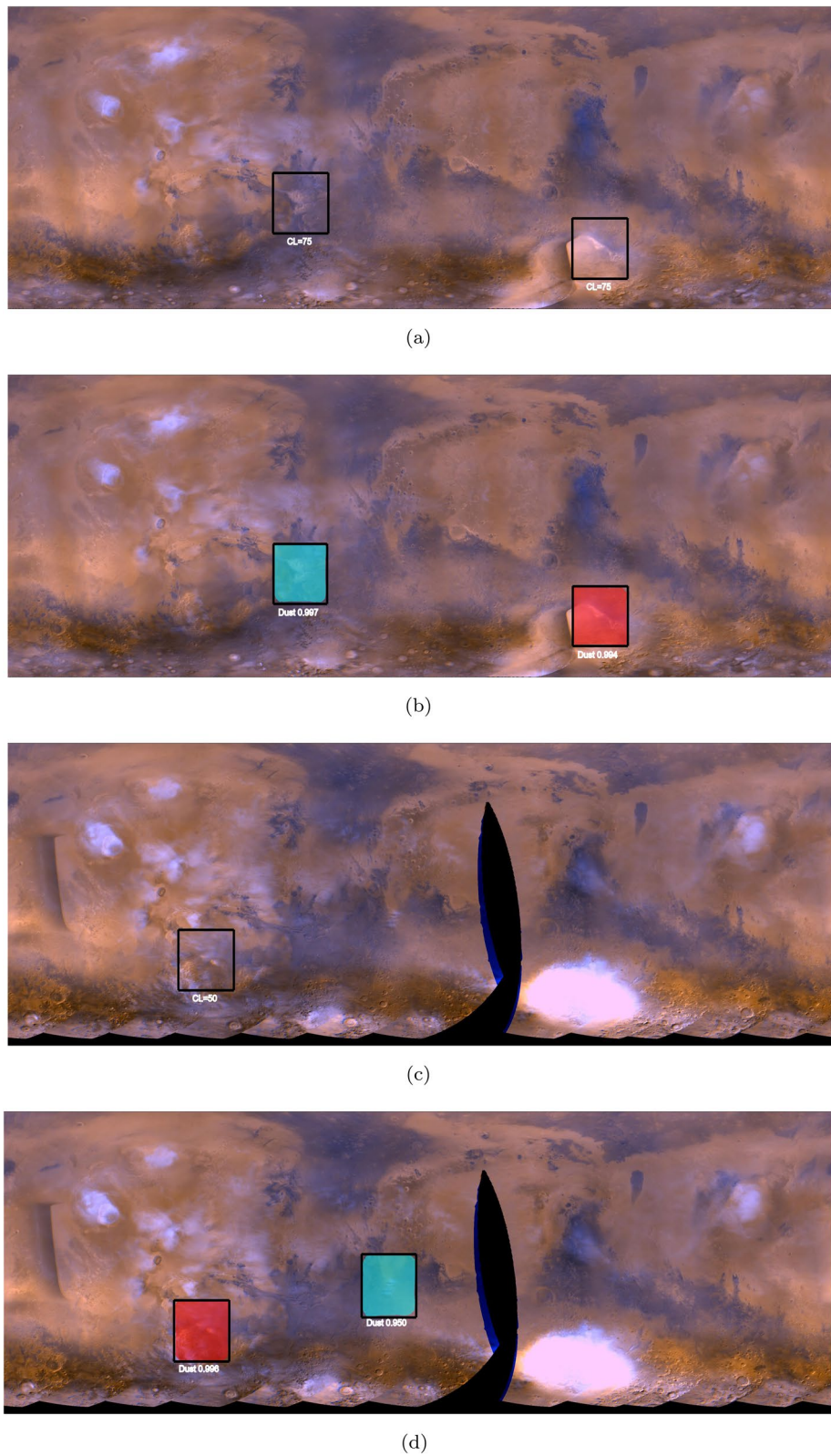
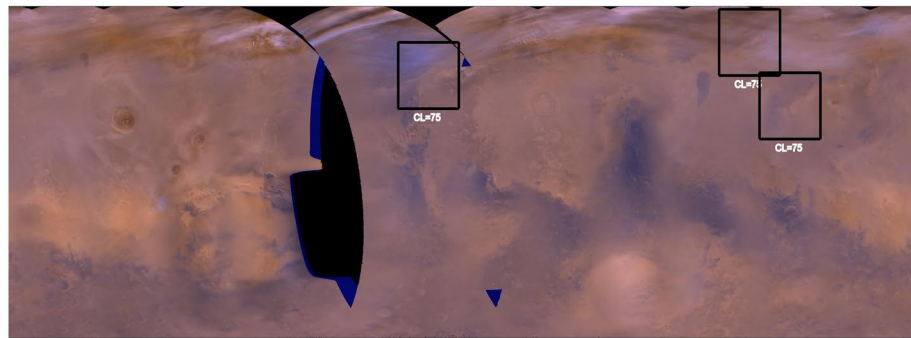
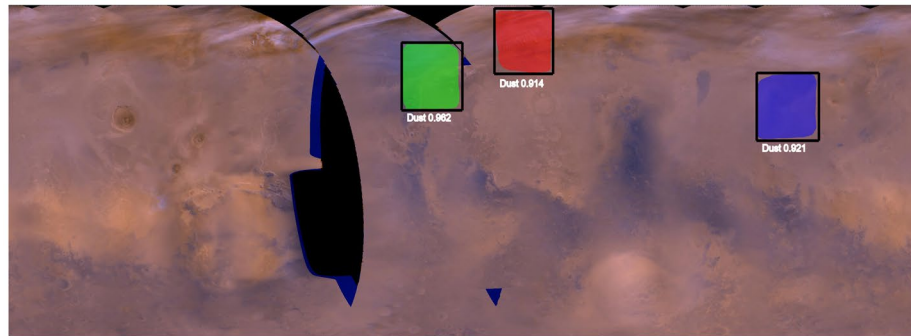


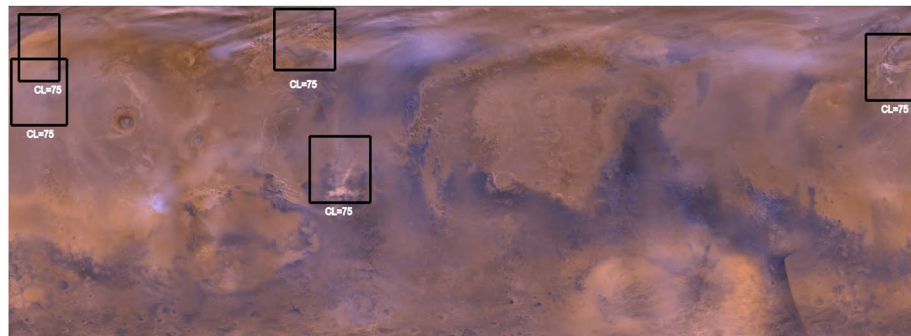
Fig. 5 2nd strategy: the training, validation and testing images are selected randomly from MY 25 to MY 28. **a** and **c** are ground-truth images from spring season (r18_040) at $L_S = 53.47^\circ$ and summer seasons (r22_008) at $L_S = 105.36^\circ$, respectively. **b** and **d** are their predicted dust maps



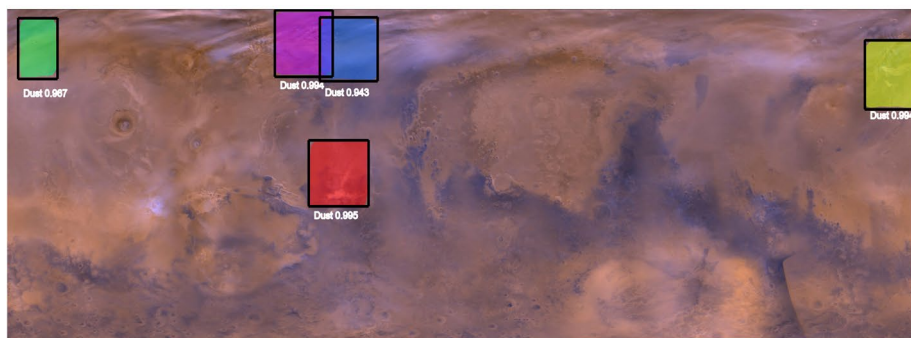
(a)



(b)

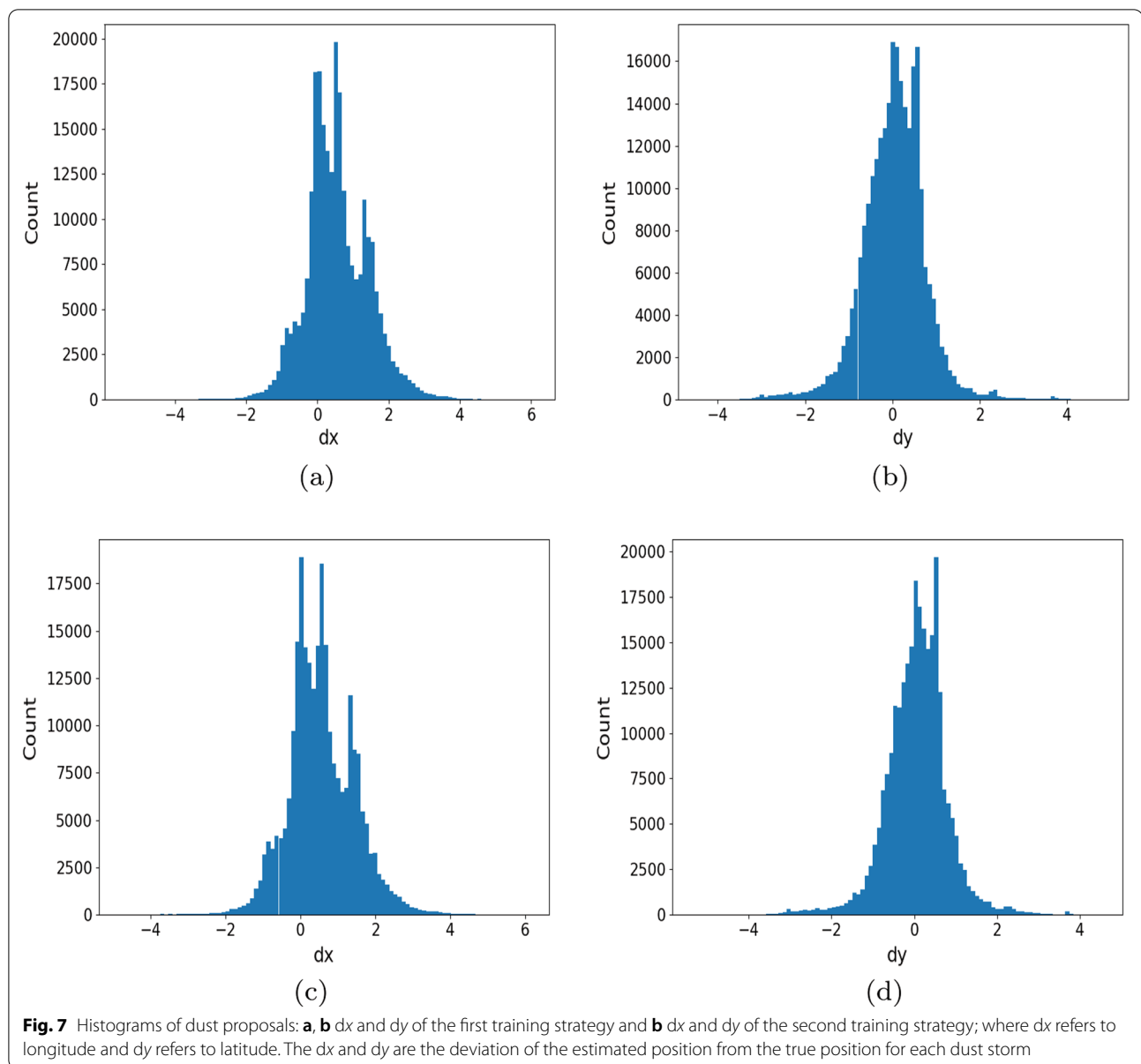


(c)



(d)

Fig. 6 2nd strategy: the training, validation and testing images are selected randomly from MY 25 to MY 28. Left panel shows ground-truth maps, and right panel shows predicted maps. **a** and **c** are ground-truth images from fall season (r08_005) at $L_S = 238.51^\circ$ and winter season (r12_017) at $L_S = 313.57^\circ$ and **b** and **d** are their predicted dust maps



compared to non-mask networks. However, the current method has a slightly higher score because the mask network has an additional component (Masker_B) that focuses on dust pixels and edges or boundaries to refine the mask with minor improvement. The current method achieves around 2/3 precision in less than half second in all MDGMs in the testing dataset, which makes it a more efficient solution with Martian larger datasets of higher dimensions. The Mask R-CNNs in both strategies succeed in non-dustiness seasons and they miss some cases when there are multiple dust events at the same time. Although, the number of maps during dustiness seasons in both strategies are small, the current Mask

R-CNN achieves higher than 60% precision using both training strategies and can achieve better with further observations.

6 Conclusion and outlook

We use a Mask R-CNN for the automated localization of dust storms in Mars daily global maps (MDGMs) derived from MGS/MOC observations. We evaluate the performance of the network by calculating the area under the ROC curves from the dust storm probability images by using various th_{IoU} and obtain the best performance at AP₂₅. One of the main strengths of this method is its speed and ease of use after training. The proposed Mask

R-CNN has been applied to a several-Martian-Year record of satellite images and demonstrated to provide reasonable results at various seasons. Thus, our deep-learning method is interesting for seeing the existing climatology of Martian dust storms from MOC/MGS observations, in particular during the non-dust storm season, from another perspective. This may become even more interesting given that observations of column-dust-optical-depth (CDOD) from satellite observations, and their subsequent use as a means of dust storm detection, are not a straightforward alternative to the detection of dust storms by visual inspection. For instance, the CDOD climatology of Montabone et al. (2015b) is based on the instruments MGS/TES, MO/THEMIS and MRO/MCS, and identification of dust storms from MGS/MOC and MRO/MARCI observations can provide complimentary information on CDOD climatology.

Potential challenges are due to the MGS/MOC derived MDGMs occasionally having data gaps (missing observations), and the R-CNN fails to detect dust storms that intersect these regions of missing data. Moreover, it is possible that R-CNN confuses between dust storms and enhanced atmospheric background dustiness in the dust storm season and/or different dust storms that are near to each other. It is self-explaining that such limiting factors are less critical during the first half of the year. It is worthwhile noting that also during the Mars cloud season, dust storms need to be accounted for. That is in line with the fact that local/regional dust storms may occur during any time of the Martian year and there is the frequent occurrence of polar cap edge dust storms in both hemispheres in the respective fall to the spring season. As demonstrated in our Figs. 3 and 5, the R-CNN method used here detects dust storms sufficiently well during the $L_S = 0^\circ - 180^\circ$ period or first half of the year. False-positive detection sporadically occurs as in Fig. 5d, but they are generally limited to a small area compared to the spatial extent of water clouds, which occur over large parts of the planet during the Mars cloud season. Based on that, the R-CNN method used here is scientifically interesting because it can potentially distinguish dust storms and water clouds. Thus, it has the potential to prepare satellite images for further automated image analysis methods which may possibly emerge in the future and are beyond the scope of this paper. The latter might be given by deep learning-methods for retrieving data on water-ice-cloud and dust characteristics from Mars satellite images.

We may refine our current R-CNN method and results further and thus obtain more accurate dust storm characteristics (location, size, shape, texture, etc.) as follows. It is widely known that Mars dust storms are bright in the red band and dark in the blue

band. By contrast, Martian water and CO₂ clouds are bright in the red and blue bands and much brighter than the surface in the blue band (Gichu and Ogohara 2019). In future, we aim to include surface albedo and/or cloudiness when preparing the ground-truth to avoid confusion between dust storms, clouds, and albedo features. Also, we aim to predict the probability of accurate contours of dust-storm regions based on reference polygon areas from the MDAD dataset by converting dimensions of dust-storm areas from km to pixels to avoid a small error in pixels in estimating dust-storm regions. In addition, we aim to classify each dust storm based on class (main, continuous, sequential, etc.), type (flushing, turning, GDE, etc.) and K16 class (A, C, GDE, etc.).

We attempted to apply the proposed R-CNN on the Mars reconnaissance orbiter (MRO) Mars color imager (MARCI) MDGMs from MY 28, L_S 133° (2006) to MY 32, L_S 171° (2014). However, we have not succeeded so far. A potential limiting factor is that adjacent global map swath images typically do not overlap and have gaps in between. Apart from that, the MDGMs derived from MRO/MARCI observations only utilize three spectral bands (red, green and blue) of the full seven bands that are available in the original data (five visible bands and two ultraviolet bands). In future, we are also considering using the proposed technique with all seven spectral band information, and using feature reduction techniques to define the most significant bands for dust storm detection and testing the effectiveness of Mask R-CNNs with multiple (more than three) spectral bands. By implication, our method is particularly interesting for upcoming/future Mars satellite missions/instruments that provide imagery without inherent gaps.

Appendix

ResNet

The description of residual blocks is behind the scope of this paper, but here we describe them briefly. The residual network (ResNet) is one of the most successful deep architectures. The shortest version of ResNet consists of 16 local residual blocks (Fig. 8a); each consists of convolution, batch normalization, ReLU activation followed by convolution and batch normalization. The local residual blocks are followed by a fully connected layer. The reason behind the success of ResNet models is the skip connection component. There is a direct connection in each block that skips layers. It solves the problem of vanishing gradient in deep architecture by allowing alternate shortcut paths for the gradient to flow through (He et al. 2016).

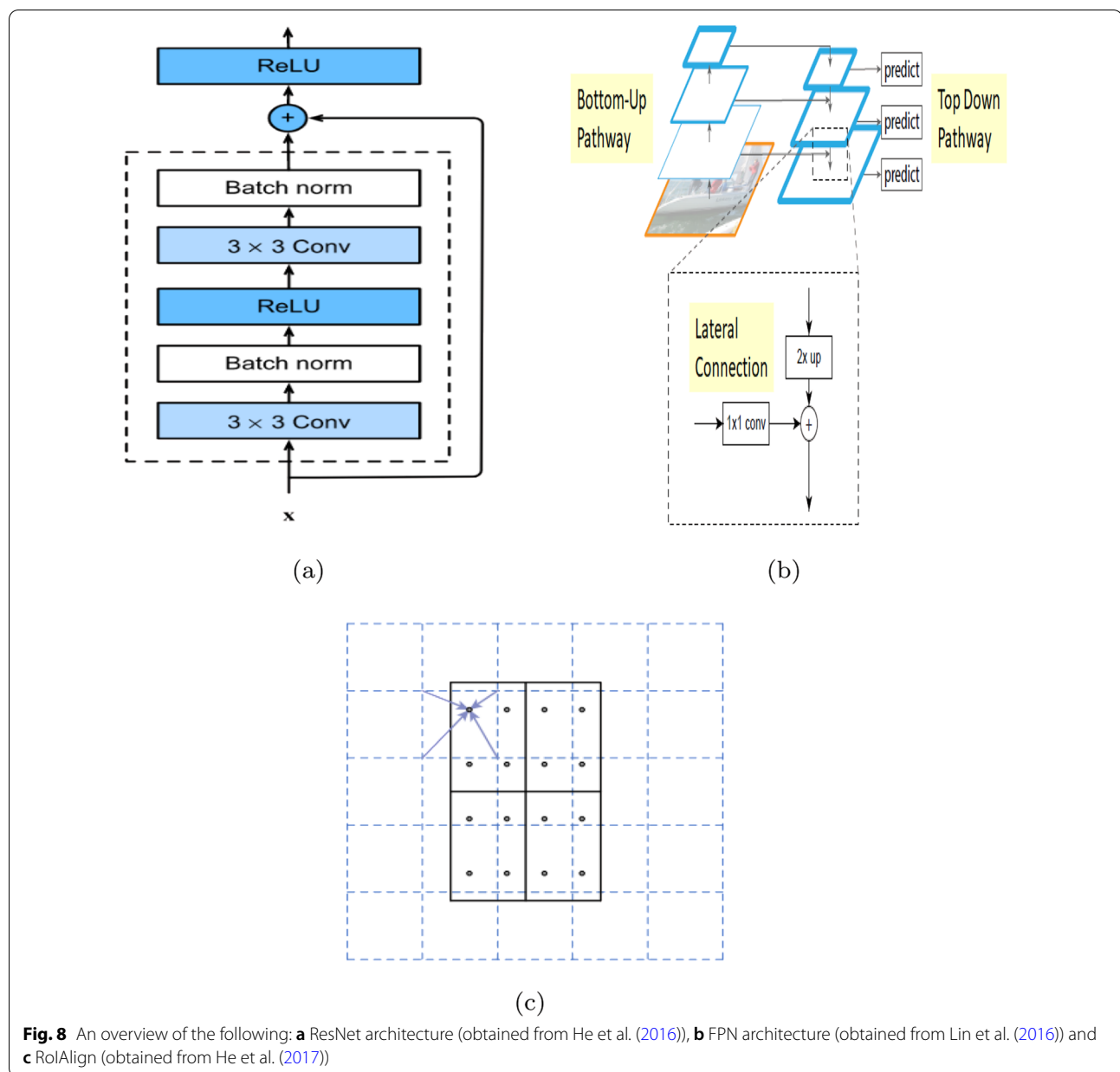


Fig. 8 An overview of the following: **a** ResNet architecture (obtained from He et al. (2016)), **b** FPN architecture (obtained from Lin et al. (2016)) and **c** RoIAlign (obtained from He et al. (2017))

FPN

The feature pyramid network (FPN) (Lin et al. 2016) consists of two main components: bottom-up and top-down (Fig. 8b). The bottom-up is the feed-forward computation of the backbone network, while the top-down is the backward computation. Each hierarchical convolutional layer is a pyramid level; it starts from the largest to the smallest convolutional layers in the bottom-up component and from the smallest to the largest in the top-down component. The convolutional feature maps from the higher pyramid level in top-down are spatial up-sampled by a certain factor using the nearest neighbor technique to the next

coarse pyramid level. The output of each pyramid level in the bottom-up is merged with the pyramid level with the same spatial size in the top-down using lateral connection. The lateral connection is mainly based on applying a 1×1 convolution process to the output of each pyramid level in the bottom-up component and merging the pyramid levels both components by element-wise addition.

RoIAlign

Region of Interest Align, called RoIAlign, is an operation for extracting a small region from a convolutional feature

map using the aligning operation. It divides the input feature map into spatial bins based on the ratio of the width of the input image to the width of the convolutional feature map ($r = W/w$). Then, it uses bilinear interpolation to compute the exact values of the feature map at four regularly sampled locations in each bin, and produces an average or maximum value (He et al. 2017). For instance, Fig. 8c shows an example of the input feature map and RoI, presented in dashed grid and solid line squares. The RoI is within 2×2 bins with four sampling points in each. The RoIAlign computes the value of each point using bilinear interpolation of nearby grid points in the feature map.

Abbreviations

MDAD: Mars dust activity database; MDGMs: Mars daily global maps; MCS: Mars climate sounder; MRO: Mars reconnaissance orbiter; THEMIS: Thermal emission imaging system; MO: Mars Odyssey; TES: Thermal emission spectrometer; MGS: Mars global surveyor; MOC: Mars orbiter camera; MARCI: Mars color imager; MCD: Mars climate database; MY: Martian year; GDEs: Global dust storm events; EDL: Entry descent landing; R-CNN: Regional convolutional neural network; PRN: Regional proposal network; PCA: Principal component analysis; MLP: Multi layer perceptron; SVM: Support vector machine; MAE: Mean absolute error; IoU: Intersection-over-union; mAP: Mean average precision.

Acknowledgements

The authors acknowledge the effort devoted to developing open-source python packages. The authors acknowledge support from the New York University Abu Dhabi, Emirati Research Program.

Authors' contributions

RA proposed the idea, downloaded all global images, trained the R-CNN and evaluated its performance on Martian global images. CG collaborated with the corresponding author in the modification of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the New York University Abu Dhabi Institute Grant (ADH01-73-71210-G1502-ADHPG) and Emirati Research Grant (ADH01-76-71202-EMIRP-ADHPG).

Availability of data and materials

The dataset supporting this article is available in the <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/F8R2JX>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

¹Center for Space Science, New York University, Abu Dhabi, United Arab Emirates. ²National Space Science and Technology Center, United Arab Emirates University, Al Ain, United Arab Emirates.

Received: 25 July 2021 Accepted: 16 December 2021

Published online: 09 January 2022

References

Battalio M, Wang H (2021) The Mars dust activity database (MDAD): a comprehensive statistical study of dust storm sequences. *Icarus* 354(114):059

- Cantor BA (2007) MOC observations of the 2001 mars planet-encircling dust storm. *Icarus* 186(1):60–96
- Cantor BA, James PB, Caplinger M, Wolff MJ (2001) Martian dust storms: 1999 mars orbiter camera observations. *J Geophys Res Planets* 106(E10):23,653–23,687
- Cantor BA, Pickett NB, Malin MC, Lee SW, Wolff MJ, Caplinger MA (2019) Martian dust storm activity near the mars 2020 candidate landing sites: Mro-marci observations from mars years 28–34. *Icarus* 321:161–170
- Cheng T, Wang X, Huang L, Liu W (2020) Boundary-preserving mask R-CNN. In: European conference on computer, pp 660–676
- Forget F, Montabone L (2017) Atmospheric dust on Mars: a review. In: International conference on environmental systems
- Gebhardt C, Abuelgasim A, Fonseca RM, Martín-Torres J, Zorzano MP (2020) Fully interactive and refined resolution simulations of the martian dust cycle by the marswrf model. *J Geophys Res Planets* 125(9):e2019JE006,253
- Gebhardt C, Abuelgasim A, Fonseca RM, Martín-Torres J, Zorzano MP (2021) Characterizing dust-radiation feedback and refining the horizontal resolution of the marswrf model down to 0.5 degree. *J Geophys Res Planets* 126(3):e2020JE006,672
- Gichu R, Ogohara K (2019) Segmentation of dust storm areas on mars images using principal component analysis and neural network. *Prog Earth Planet Sci* 6(12):1–2
- Girshick R (2015) Fast R-CNN. In: IEEE international conference on computer vision, pp 1440–1448
- Girshick RB, Donahue J, Darrell T, Malik J (2013) Rich feature hierarchies for accurate object detection and semantic segmentation. In: IEEE conference on computer vision and pattern recognition
- Guzewich SD, Toigo AD, Kulowski L, Wang H (2015) Mars orbiter camera climatology of textured dust storms. *Icarus* 258:1–13
- Haberle R, Clancy R, Forget F, Smith M, Zurek R (2017) The atmosphere and climate of Mars. Cambridge planetary science. Cambridge University Press, Cambridge
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, pp 770–778
- He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: IEEE international conference on computer vision, pp 2980–2988
- Higham CF, Higham DJ (2019) Deep learning: an introduction for applied mathematicians. [arXiv:abs/1801.05894](https://arxiv.org/abs/1801.05894)
- Hinson DP, Wang H, Smith MD (2012) A multi-year survey of dynamics near the surface in the northern hemisphere of mars: short-period baroclinic waves and dust storms. *Icarus* 219(1):307–320
- Kass DM, Kleinböhl A, McCleese DJ, Schofield JT, Smith MD (2016) Interannual similarity in the Martian atmosphere during the dust storm season. *Geophys Res Lett* 43(12):6111–6118
- Kulowski L, Wang H, Toigo AD (2017) The seasonal and spatial distribution of textured dust storms observed by Mars global surveyor Mars orbiter camera. *Adv Space Res* 59(2):715–721
- Lin T, Dollár P, Girshick RB, He K, Hariharan B, Belongie SJ (2016) Feature pyramid networks for object detection. *CoRR* [arXiv:abs/1612.03144](https://arxiv.org/abs/1612.03144)
- Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: IEEE conference on computer vision and pattern recognition, pp 936–944
- Maeda K, Ogawa T, Haseyama M (2015) Automatic Martian dust storm detection from multiple wavelength data based on decision level fusion. *IPSP Trans Comput Vis Appl* 7:79–83
- Malin MC, Edgett KS, Cantor BA, Caplinger MA, Danielson GE, Jensen EH, Ravine MA, Sandoval JL, Supulver KD (2010) An overview of the 1985–2006 Mars orbiter camera science investigation. *Int J Mars Sci Explor* 4:1–60
- Milletari F, Navab N, Ahmadi SA (2016) V-net: fully convolutional neural networks for volumetric medical image segmentation. In: International conference on 3D vision (3DV), pp 565–571
- Montabone L, Forget F, Millour E, Wilson R, Lewis S, Cantor B, Kass D, Kleinböhl A, Lemmon M, Smith M, Wolff M (2015a) Eight-year climatology of dust optical depth on Mars. *Icarus* 251:65–95
- Montabone L, Forget F, Millour E, Wilson R, Lewis S, Cantor B, Kass D, Kleinböhl A, Lemmon M, Smith M, Wolff M (2015b) Eight-year climatology of dust optical depth on Mars. *Icarus* 251:65–95

- Newman CE, Richardson MI (2015) The impact of surface dust source exhaustion on the Martian dust cycle, dust storms and interannual variability, as simulated by the Marswrf general circulation model. *Icarus* 257:47–87
- Reiss D, Zanetti M, Neukum G (2011) Multitemporal observations of identical active dust devils on mars with the high resolution stereo camera (HRSC) and mars orbiter camera (MOC). *Icarus* 215(1):358–369
- Ren S, He K, Girshick R, Sun J (2015) Faster r-cnn: towards real-time object detection with region proposal networks. In: *Advances in neural information processing systems*, vol 28
- Ren S, He K, Girshick R, Sun J (2016) Faster r-cnn: towards real-time object detection with region proposal networks. In: *Advances in neural information processing systems*, vol 28
- Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. In: *International conference on learning representations*
- Wang H, Ingersoll AP (2002) Martian clouds observed by Mars global surveyor Mars orbiter camera. *J Geophys Res Planets* 107(E10):8-1-8–16
- Wang H, Richardson MI (2015) The origin, evolution, and trajectory of large dust storms on Mars during Mars years 24–30 (1999–2011). *Icarus* 251:112–127
- Zhao ZQ, Zheng P, Xu ST, Wu X (2019) Object detection with deep learning: a review. *IEEE Trans Neural Netw Learn Syst* 30:3212–3232
- Zurek RW, Martin LJ (1993) Interannual variability of planet-encircling dust storms on Mars. *J Geophys Res Planets* 98(E2):3247–3259

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
